

В.С. Григорків, М.В. Григорків, О.І. Ярошенко



ОПТИМІЗАЦІЙНІ МЕТОДИ ТА МОДЕЛІ

Міністерство освіти і науки України
Чернівецький національний університет
імені Юрія Федьковича

В.С. Григорків, М.В. Григорків, О.І. Ярошенко

ОПТИМІЗАЦІЙНІ МЕТОДИ ТА МОДЕЛІ

Підручник
Видання 2-ге



Чернівці
Чернівецький національний університет
імені Юрія Федьковича
2024

УДК 519.863 (075.8)

Г 831

Друкується за ухвалою Вченої ради
Чернівецького національного університету імені Юрія Федьковича
(протокол №13 від 28.12.2021 року)

Рецензенти:

Капустян В.О.

– професор кафедри економічної кібернетики
Національного технічного університету України
«Київський політехнічний інститут імені Ігоря
Сікорського», доктор фізико-математичних наук,
професор

Петрик М.Р.

– завідувач кафедри програмної інженерії
Тернопільського національного технічного університету
імені Івана Пулюя, доктор фізико-математичних наук,
професор

Григорків В.С.

Г 831 Оптимізаційні методи та моделі : підручник, 2-ге вид., випр. і
доп. / В.С. Григорків, М.В. Григорків, О.І. Ярошенко. Чернівці :
Чернівецький нац. ун-т, 2024. 464 с.

ISBN 978-966-423-899-8

У підручнику викладено теоретичні основи моделей та методів оптимізації, які формують класичну базу сучасної теорії оптимізації. Матеріал підручника структурований на три частини, що відповідно присвячені безумовній та умовній оптимізації функцій однієї і багатьох змінних, спеціальним класам моделей та методів умовної оптимізації, які широко використовуються у різних галузях знань та практичної діяльності, а також прикладним аспектам сучасного комп'ютерного забезпечення оптимізаційних методів і моделей.

Для студентів економічних, математичних і технічних спеціальностей та спеціалізацій, а також аспірантів, викладачів і широке коло фахівців.

УДК 519.863 (075.8)

Навчальне видання

Григорків Василь Степанович, **Григорків** Марія Василівна, **Ярошенко** Олена Іванівна

ОПТИМІЗАЦІЙНІ МЕТОДИ ТА МОДЕЛІ

Підручник

Літературний редактор *О.В. Лукул*

Підписано до друку 24.10.2024. Формат 60х84/16

Папір офсетний. Друк різнографічний. Ум.-друк. арк. 25,4.

Обл.-вид. арк. 27,3. Тираж 30. Зам. Н-085.

Видавництво Чернівецького національного університету

58002, Чернівці, вул. Коцюбинського, 2

e-mail: ruta@chnu.edu.ua

Свідоцтво суб'єкта видавничої справи ДК №891 від 08.04.2002 р.

© Чернівецький національний університет
імені Юрія Федьковича, 2024

© В.С. Григорків, 2024

© М.В. Григорків, 2024

© О.І. Ярошенко, 2024

ISBN 978-966-423-899-8

ЗМІСТ

ПЕРЕДМОВА.....	8
ВСТУП	11
ЧАСТИНА І. ОДНОВИМІРНА ОПТИМІЗАЦІЯ	30
Розділ 1. Основи теорії мінімізації функції однієї	
змінної	30
1.1. Поняття мінімуму функції однієї змінної	30
1.2. Класичний метод мінімізації функції однієї	
змінної	33
1.3. Унімодальні та опуклі функції	45
Розділ 2. Прямі числові методи	53
2.1. Метод перебору (рівномірного пошуку)	53
2.2. Методи поділу відрізка пополам	56
2.3. Метод золотого перерізу	64
2.4. Метод Фібоначчі	69
2.5. Метод квадратичної інтерполяції	74
Розділ 3. Методи мінімізації першого та другого	
порядків	79
3.1. Метод середньої точки.....	79
3.2. Метод хорд.....	81
3.3. Метод Ньютона	84
3.4. Метод кубічної апроксимації.....	87
Розділ 4. Методи мінімізації багатомодальних	
функцій	89
4.1. Умова Ліпшиця	89
4.2. Метод перебору.....	90
4.3. Метод ламаних	93

ЧАСТИНА II. БАГАТОВИМІРНА ОПТИМІЗАЦІЯ.....	102
Розділ 5. Теоретичні основи та методи безумовної мінімізації функції багатьох змінних.....	102
5.1. Постановка задачі безумовної мінімізації функції багатьох змінних	102
5.2. Класичний метод мінімізації функції багатьох змінних	104
5.3. Поняття опуклої задачі оптимізації	118
5.4. Деякі прямі методи безумовної мінімізації.....	125
5.4.1. Метод циклічного покоординатного спуску.....	125
5.4.2. Метод Хука-Джівса	131
5.4.3. Методи випадкового пошуку	133
5.4.4. Метод спряжених напрямків	136
5.5. Методи безумовної мінімізації першого та другого порядків	139
5.5.1. Метод градієнтного спуску	140
5.5.2. Метод найшвидшого градієнтного спуску.....	143
5.5.3. Метод Ньютона	145
Розділ 6. Основи теорії та методи умовної мінімізації функції багатьох змінних	150
6.1. Постановка та класифікація задач умовної мінімізації функції багатьох змінних.....	150
6.2. Класична задача на умовний екстремум.....	154
6.3. Необхідні та достатні умови екстремуму у випадку задачі математичного програмування із функціональними обмеженнями-нерівностями.....	168
6.4. Необхідні та достатні умови екстремуму у випадку задачі математичного програмування	

зі змішаними функціональними обмеженнями	186
6.5. Задача умовної оптимізації з прямим обмеженням.....	195
6.6. Методи можливих напрямків	201
6.6.1. Метод можливих напрямків у випадку лінійних обмежень	202
6.6.2. Метод можливих напрямків у випадку нелінійних обмежень	209
6.7. Градієнтні методи умовної мінімізації	215
6.7.1. Метод проекції градієнта	215
6.7.2. Метод умовного градієнта.....	220
6.8. Методи послідовної безумовної мінімізації.....	224
6.8.1. Метод штрафних функцій	224
6.8.2. Метод бар'єрних функцій	227

ЧАСТИНА III. СПЕЦІАЛЬНІ КЛАСИ МОДЕЛЕЙ ТА МЕТОДІВ УМОВНОЇ ОПТИМІЗАЦІЇ.....

Розділ 7. Лінійне програмування.....

7.1. Лінійне програмування як ефективний інструментарій дослідження лінійних моделей	230
7.2. Форми запису задачі лінійного програмування та їх еквівалентність.....	232
7.3. Основні властивості задач лінійного програмування.....	236
7.4. Графічний метод розв'язування задач лінійного програмування.....	246
7.5. Симплексний метод розв'язування задач лінійного програмування	256
7.5.1. Базисні та опорні розв'язки систем лінійних алгебраїчних рівнянь.....	257

7.5.2. Основи симплекс-методу.....	271
7.5.3. Знаходження початкового опорного розв'язку канонічної задачі лінійного програмування.....	282
7.5.4. Вироджені задачі лінійного програмування. Антициклін	295
7.6. Двоїсті задачі лінійного програмування	300
7.6.1. Правила побудови двоїстих задач	300
7.6.2. Основні властивості пари взаємно двоїстих задач.....	304
7.6.3. Сумісне розв'язування взаємно двоїстих задач.....	310
7.6.4. Економічна інтерпретація двоїстих задач.....	317
7.7. Транспортні задачі (Т-задачі).....	318
7.7.1. Теоретичні основи Т-задач.....	318
7.7.2. Методи розв'язування Т-задач.....	323
7.8. Задачі цілочислового лінійного програмування	332
Розділ 8. Нелінійне програмування	340
8.1. Нелінійні оптимізаційні моделі та їх формалізація задачами нелінійного програмування	340
8.2. Графічний метод розв'язування задач нелінійного програмування.....	342
8.3. Задачі дробово-лінійного програмування	350
8.4. Опукле та квадратичне програмування	355
8.5. Поняття двоїстої задачі опуклого програмування	370
Розділ 9. Динамічне програмування.....	373
9.1. Вступ до динамічного програмування	373

9.2. Постановка задачі керування дискретним процесом зі скінченним числом кроків	375
9.3. Принцип оптимальності Беллмана й алгоритм методу динамічного програмування	378
9.4. Приклади розв’язування оптимізаційних задач методом динамічного програмування	382

Розділ 10. Комп’ютерні технології розв’язування оптимізаційних задач	399
10.1. Стандартна задача лінійного програмування та двоїста до неї	399
10.2. Транспортні задачі (Т-задачі)	413
10.2.1. Стандартна Т-задача	413
10.2.2. Цілочислова задача транспортного типу	424
10.3. Стандартна задача цілочислового лінійного програмування	432
10.4. Деякі задачі нелінійного програмування	439
10.4.1. Задача дробово-лінійного програмування	439
10.4.2. Стандартна задача квадратичного програмування	445
10.5. Задача динамічного програмування	452
СПИСОК ЛІТЕРАТУРИ	462

ПЕРЕДМОВА

Сучасні освітні програми зацентовані на підготовку висококваліфікованих фахівців, які б володіли ґрунтовними теоретичними знаннями та практичними навиками щодо вибору оптимальних рішень у будь-якій сфері людської діяльності. Тому оптимізаційні методи та моделі складають основу базових дисциплін, які внесені до стандартів вищої освіти, зокрема вищої економічної, та спрямовані на формування теоретичного інструментарію для побудови, аналізу і знаходження розв'язків оптимізаційних моделей з метою прийняття відповідних оптимальних рішень.

Незважаючи на наявність деякої навчально-методичної літератури з теорії та практичного застосування методів і моделей оптимізації, її використання для студентів багатьох спеціальностей (наприклад, економічних спеціальностей), як правило, потребує певного посередництва фахівців з метою адаптації до конкретних навчальних планів і професійної специфіки студентів. Ця обставина зумовлює актуальність розробки відповідних підручників з даної дисципліни, що, крім уже зазначених особливостей, повинні також відрізнятися підбором матеріалу, який би адекватно репрезентував узгоджені з відповідними компетенціями основи базових знань з теорії оптимізаційних методів і моделей.

Усе вищесказане стосується також відповідних завдань для практичних занять, самостійної та індивідуальної роботи студентів. Вони повинні задовольняти той рівень адекватності, який максимально підпорядкований сучасним вимогам до компетенцій спеціаліста того чи іншого профілю.

Очевидно, що написання навчально-методичної літератури, яка б ідеально поєднала зміст і форму викладення матеріалу, наукову строгість і методичну досконалість, теоретичну завершеність і прикладну спрямованість, – архіскладне завдання, яке далеко не завжди можна виконати і з яким практично зустрічається кожний автор, особливо, коли йдеться про літературу для досить широкого кола читачів, які не належать до професіоналів у відповідній галузі знань, але формують особливу категорію, яка потребує таких знань для практичного застосування у своїй теперішній чи, можливо, майбутній професійній діяльності. До такої категорії належать, зокрема, студенти різних економічних напрямків підготовки. У запропонованому підручнику авторами враховано вищезазначені особливості та певні можливості для їх удосконалення.

Структура матеріалу даного підручника відображає класичні підходи до його впорядкування у літературі, присвяченій оптимізаційним методам і моделям. Перша частина традиційно пов'язана з одновимірною оптимізацією. У ній викладено основи теорії оптимізації функції однієї змінної, прямі числові методи, методи мінімізації першого та другого порядків, а також методи мінімізації багатомодальних функцій. Друга частина – це матеріал з багатовимірної оптимізації, який містить теоретичні основи та числові методи безумовної та умовної мінімізації функції багатьох змінних. Спеціальним класам методів і моделей умовної оптимізації присвячена третя частина посібника, зміст якої складає лінійне, нелінійне та динамічне програмування, а також приклади застосування сучасних комп'ютерних систем для розв'язування деяких оптимізаційних задач цього класу.

Обсяг будь-якого підручника зазвичай обмежений не авторськими бажаннями, а іншими обставинами. Це повною мірою стосується і даного підручника. Але як за змістом, так і за структурою він достатньо інформаційний, тобто такий, що дозволяє отримати базові знання з теорії оптимізації функцій та навички їх практичного застосування. Передбачається, що матеріал зі скінченновимірної оптимізації, який не внесений до підручника, але може поглибити ерудицію читача у галузі оптимізаційних методів та моделей, є предметом самостійної роботи як з рекомендованою літературою [1-22], так і з іншими навчально-методичними та науковими працями з теорії оптимізації. Також зазначимо, що основний матеріал пропонованого підручника є розширеним варіантом матеріалу виданого раніше однойменного підручника перших двох співавторів, тому є достатньо апробованим у плані змісту, форми викладення та цільового використання.

Підручник адресований студентам не тільки економічних, а й математичних та інших спеціальностей вишів, навчальні плани яких передбачають вивчення дисциплін, пов'язаних з методами та моделями оптимізації. Він може бути корисним для всіх, хто викладає відповідні дисципліни або використовує інструментарій оптимізаційних методів і моделей у своїй професійній діяльності.

ВСТУП

В.1. Зміст, структура та формалізація оптимізаційної моделі

У реальній ситуації побудова оптимізаційної моделі, тобто математичної моделі об'єкта оптимізації (зокрема об'єкта оптимізації в економіці), складна задача, сутність якої полягає у створенні наближеної, але адекватної системи уявлень (знань) про об'єкт і математичній формалізації, якість якої залежить від правильно вибраних змінних, параметрів, обмежень, критеріїв тощо [4-11].

При побудові математичної моделі необхідно врахувати найбільш важливі характеристики та фактори досліджуваного об'єкта, які можуть бути описані на мові математики і в кінцевому результаті привести до коректно поставленої оптимізаційної задачі. При цьому надзвичайно важливо, щоб невраховані у моделі фактори суттєво не впливали на її розв'язок.

Процес математичного моделювання оптимізаційного об'єкта можна умовно розбити на кілька етапів. Коротко розкриємо їх зміст.

1. Установлення меж об'єкта оптимізації. Наявність так званих меж об'єкта оптимізації пов'язана з тим, що у більшості випадків неможливо встановити та врахувати всі його змістові показники та властивості. Інакше кажучи, система характеристик об'єкта моделювання практично не може бути всесторонньою, тобто його повним еквівалентом. Ця система обов'язково обмежена у плані опису об'єкта, тобто вона має певні межі, які переважно визначаються

метою моделювання й оптимізації. Наприклад, на практиці оптимізація процесу функціонування конкретного виробничого підприємства далеко не завжди враховує його зв'язки з іншими подібними підприємствами, поведінку на ринку продукції та ресурсів, зовнішньоекономічну діяльність тощо. Зрозуміло, що неврахування всіх цих факторів формує межі об'єкта оптимізації. Цілком очевидно, що ці межі можуть бути вибрані невдало, тому їх доведеться коректувати. Це можна виявити, аналізуючи систему врахованих чинників, її математичну модель і розв'язок моделі. Тоді в одних випадках ці межі треба розширити, в інших – звужити, причому цей процес може бути неодноразово повторений.

2. Вибір змінних і параметрів моделі. На цьому етапі необхідно визначитись із тими величинами, які змінюють свої значення і якими можна керувати з метою досягнення оптимального результату, а також з постійними величинами, які залежать від зовнішніх факторів. Задача оптимізації полягає у виборі таких значень змінних величин (керованих змінних), яким відповідає оптимальна ситуація. Зазначимо, що залежно від рівня деталізації системи оптимізації, тобто від рівня опису концептуальної моделі об'єкта оптимізації одні і ті ж величини можуть бути як керованими змінними, так і постійними.

3. Визначення допустимої множини значень керованих змінних. Кожна математична модель описує відповідний об'єкт (процес, систему, явище), тому керовані змінні, наприклад економічні змінні, відображають реальні ресурси, потужності, обсяги продукції тощо, які завжди є обмеженими. У математичній моделі ці обмеження описуються відповідними рівняннями і нерівностями відносно керованих змінних або в деяких випадках

задаються спеціальними множинами. Сукупність усіх обмежень на керовані змінні визначає так звану допустиму множину оптимізаційної моделі.

4. Вибір критерію оптимізації. Коли йдеться про оптимізацію, то завжди мається на увазі вибір оптимального у певному сенсі варіанта. Відповідь на те, який із варіантів оптимальний, може бути отримана лише тоді, коли відомий критерій оптимізації. Це означає, що у структурі оптимізаційної моделі, крім допустимої множини, обов'язково повинен бути критерій оптимізації, здебільшого числовий. Тоді оптимальним варіантом буде той допустимий варіант, якому відповідає максимальне або мінімальне значення (залежно від змісту поставленої задачі) числового критерію оптимізації, який формалізується деякою функцією (функціоналом) від керованих змінних, яку традиційно називають *цільовою функцією*. В оптимізаційних економіко-математичних моделях у ролі критерію оптимізації може бути дохід, прибуток, витрати, собівартість, рентабельність, продуктивність тощо.

Вибір і формалізація критерію оптимізації – це часто складна задача, пов'язана з неоднозначністю багатокритеріальністю, невизначеністю. Такі ситуації приводять до різних класів оптимізаційних моделей, зокрема до моделей багатокритеріальної та стохастичної оптимізації, хоча на практиці найбільш простими і часто використовуваними є детерміновані моделі однокритеріальної оптимізації.

5. Узагальнена формалізація оптимізаційної моделі. Цей етап завершує процес побудови та формалізації оптимізаційної моделі, яка у математичному плані є задачею оптимізації, тобто задачею максимізації (мінімізації) цільової функції на допустимій множині значень керованих

змінних. Нехай $x_1, \dots, x_n$ – керовані змінні, а X – множина точок $x = (x_1, \dots, x_n)^T$ (T – транспонування) із n -вимірного точкового (векторного) простору $\mathbb{R}^n$ ($X \subseteq \mathbb{R}^n$). Якщо $f(x) \equiv f(x_1, \dots, x_n)$ – деяка цільова функція, визначена на множині X , то оптимізаційна модель у найбільш узагальненій формі формалізується так:

$$\begin{cases} f(x) \rightarrow \min(\max), \\ x \in X. \end{cases} \quad (\text{B.1})$$

Зміст оптимізаційної моделі (задачі) (B.1) – полягає у знаходженні принаймні однієї точки $x^* \in X$, у якій значення функції $f(x)$ буде мінімальним (максимальним). При цьому знаходиться також значення $f(x^*)$. Більш глибоке математичне обґрунтування даної моделі буде здійснене пізніше. Крім того, будуть також досліджені більш конкретизовані її форми. Ще одне важливе зауваження стосується того, що всі описані нижче моделі та методи їх розв'язування будуть скінченновимірними, тобто їх допустимі множини будуть належати простору $\mathbb{R}^n$ ($n \geq 1$), що і відображено вище.

6. Інформаційне забезпечення моделі. Щоб побудована математична модель мала прикладну значущість, тобто була використана при розв'язанні задач реальної практики, потрібно мати необхідну інформацію про її параметри та змінні. Якщо таку інформацію не вдається здобути, то насправді побудова моделі повинна бути продовжена і всі попередні етапи цієї побудови мають бути максимально підпорядковані тому інформаційному забезпеченню, яке є у розпорядженні дослідника. Цілком можливі також ситуації, коли потрібна нам інформація може бути отримана як результат окремого дослідження, наприклад інформаційного

моделювання, оскільки дуже часто статистичних даних чи даних спостережень не тільки немає, але їх і неможливо отримати. У цих випадках існують відповідні експертні дослідження, у тому числі ті, які використовують імітаційні експерименти, зокрема машинну імітацію.

Отже, інформаційне забезпечення моделі – це одна із принципових проблем моделювання. Без її розв'язання неможливо побудувати адекватну модель.

В.2. Приклади найпростіших оптимізаційних моделей в економіці

У попередньому пункті була здійснена узагальнена формалізація оптимізаційної моделі як математичної задачі з теорії оптимізації, тому надалі терміни «модель» і «задача» будуть розумітись як еквівалентні поняття. Крім того, модель (задача) (В.1) у кожному випадку буде мати свій конкретний зміст і відповідну математичну конкретизацію, що буде продемонстровано, зокрема, і в даному пункті, де наводяться найпростіші приклади різних економічних задач, які належать до різних класів задач (моделей) оптимізації, що насправді складають предмет дослідження у наступних розділах даної книги.

В2.1 (максимізація випуску продукції). При дослідженні діяльності виробника деякої продукції з'ясувалося, що залежність її випуску f від витрат x описується виробничою функцією $f(x) = ax^\alpha - bx$, де a, b, α – додатні числові параметри, причому $\alpha \in (0, 1)$. Необхідно визначити на множині $X \subset (0, +\infty)$ обсяг витрат, який максимізує випуск продукції.

◀ Очевидно, що у даній ситуації потрібно розв'язати задачу

$$\begin{cases} ax^\alpha - bx \rightarrow \max, \\ x \in X. \end{cases} \quad (\text{B.2})$$

Задача (B.2) належить до класу найпростіших задач максимізації функції однієї змінної, що має конкретний економічний зміст і є частинним випадком (B.1). У таких задачах більш складний етап побудови виробничої функції. ►

B2.2 (максимізація прибутку однопродуктової фірми). Для виробництва продукції фірма-виробник використовує n видів ресурсів. Якщо $x = (x_1, \dots, x_n)^T \in \mathbb{R}_+^n$ – вектор обсягів використаних ресурсів ($\mathbb{R}_+^n = \{x \in \mathbb{R}^n / x_j \geq 0, j = \overline{1, n}\}$ – невід’ємний ортант n -вимірного точкового (векторного) простору $\mathbb{R}^n$), $y = f(x)$ – обсяг випуску продукції ($f(x)$ – задана виробнича функція), p – ціна продукції, $w = (w_1, \dots, w_n)^T \in \mathbb{R}_+^n$ – вектор цін на ресурси, то прибуток фірми визначається величиною

$$\Pi(x) = pf(x) - \langle w, x \rangle = pf(x) - \sum_{j=1}^n w_j x_j,$$

де $\langle \cdot, \cdot \rangle$ – операція скалярного добутку векторів.

Потрібно визначити такий набір ресурсів, при якому прибуток фірми досягає свого максимального значення.

◄ Нагадаємо, що прибуток фірми визначається як різниця між доходом і витратами. Максимізація прибутку зводиться до задачі

$$\begin{cases} \Pi(x) \rightarrow \max, \\ x \in X, \end{cases} \quad (\text{B.3})$$

де допустима множина X може збігатися з $\mathbb{R}^n$, $\mathbb{R}_+^n$ або бути деякою власною підмножиною із $\mathbb{R}_+^n$. Якщо $X = \mathbb{R}^n$, то задача (B.3) є задачею безумовної багатовимірної

оптимізації, розв'язок якої має реальний зміст лише у тому випадку, коли він невід'ємний. Якщо $X = \mathbb{R}_+^n$ або $X \subset \mathbb{R}_+^n$, то (В.3) буде належати до класу задач умовної багатовимірної оптимізації. У кожному з цих випадків задача (В.3) вкладається у рамки узагальненої моделі оптимізації (В.1). При $X = \mathbb{R}_+^n$ задача (В.3) формалізує так звану довгострокову модель фірми [5]. ►

В2.3 (максимізація корисності споживача). Поведінка споживача на ринку n товарів визначається його функцією корисності

$$u(x_1, \dots, x_n) = Ax_1^{\alpha_1} x_2^{\alpha_2} \dots x_n^{\alpha_n},$$

де A, α_i ($i = \overline{1, n}$) – додатні параметри ($\alpha_i < 1, i = \overline{1, n}$),

$x = (x_1, \dots, x_n)^T$ – придбаний набір або вектор товарів (x_j – кількість одиниць товару j), вектором цін на товари $p = (p_1, \dots, p_n)^T$ і його бюджетом h (наявною сумою грошей) (h, p_i ($i = \overline{1, n}$) – додатні числові параметри). Споживач повністю витрачає свій бюджет. Знайти такий набір товарів, який максимізує корисність споживача.

◄ Математична модель поведінки споживача у даному випадку має вигляд

$$\begin{cases} u(x_1, \dots, x_n) \rightarrow \max, \\ \sum_{j=1}^n p_j x_j = h. \end{cases} \quad (\text{В.4})$$

Модель (В.4) формалізована як задача на умовний максимум (у загальному випадку – на умовний екстремум). Її допустима множина $X \subset \mathbb{R}^n$ задається одним обмеженням – рівністю. Специфіка задачі (В.4) дозволяє використати для її розв'язування методи, відмінні від методів розв'язування

задачі (В.3). Зауважимо також, що реальний розв'язок задачі (В.4) повинен бути невід'ємним, оскільки кількості придбаних товарів не можуть бути від'ємними. Зрозуміло, що умову невід'ємності набору товарів (вектора x) можна вказати у постановці задачі максимізації корисності споживача й отримати модель максимізації корисності з умовою невід'ємності, що приведе до іншого класу оптимізаційних задач. ►

В2.4 (оптимальне планування виробництва). Для виробництва n ($n > 1$) видів продукції підприємство використовує m ($m > 1$) видів ресурсів. Технологія виробництва задається матрицею $A = (a_{ij})_{i,j=1}^{m,n}$ (елемент a_{ij} – це кількість одиниць ресурсу i , яка необхідна для виробництва однієї одиниці продукції j). Крім того, задані вектор $b = (b_1, \dots, b_m)^T$ обмежень на ресурси та вектор $c = (c_1, \dots, c_n)^T$ цін на продукцію. Потрібно визначити такий план виробництва $x = (x_1, \dots, x_n)^T$ (x_j – шукана кількість одиниць продукції j), який би при заданих умовах виробництва максимізував сумарну вартість (дохід) виготовленої продукції.

◄ Математична модель задачі про оптимальне планування виробництва формалізується так:

$$\begin{cases} \sum_{j=1}^n c_j x_j \rightarrow \max, \\ \sum_{j=1}^n a_{ij} x_j \leq b_i, \quad i = \overline{1, m}, \\ x_j \geq 0, \quad j = \overline{1, n}. \end{cases} \quad (\text{В.5})$$

Очевидно, допустимою множиною оптимізаційної задачі

(B.5) є множина $X = \left\{ x = (x_1, \dots, x_n)^T \in \mathbb{R}_+^n / \sum_{j=1}^n a_{ij} x_j \leq b_i, i = \overline{1, m} \right\}$, а

цільовою функцією – функція $f(x) = f(x_1, \dots, x_n) = \sum_{j=1}^n c_j x_j$.

Оскільки цільова функція та обмеження допустимої множини лінійні, то (B.5) належить до класу так званих задач лінійного програмування, які є частинним випадком задач математичного програмування. ►

В2.5 (оптимізація раціону). Для набору з n видів продуктів харчування, кожний з яких містить m видів різних поживних речовин (білки, жири, вуглеводи, вітаміни тощо), відома матриця $A = (a_{ij})_{i,j=1}^{m,n}$, де a_{ij} – це кількість одиниць поживної речовини i у одній одиниці продукту j , вектор $b = (b_1, \dots, b_m)^T$ мінімальних потреб поживних речовин у раціоні та вектор $c = (c_1, \dots, c_n)^T$ цін на продукти. Визначити такий конкретний набір продуктів (такий раціон) $x = (x_1, \dots, x_n)^T$ (x_j – шукана кількість продукту j), який задовольняє відповідні потреби у поживних речовинах і мінімізує витрати на його придбання.

◄ Це класична задача планування найбільш економного (тобто найбільш дешевого) раціону, який відповідає всім медичним вимогам щодо його повноцінності. Така задача актуальна, наприклад там, де потрібно нагодувати велику кількість людей (в армії, санаторії тощо). Інколи її ще називають задачею про дієту. У прийнятих позначеннях математична модель оптимізації раціону має вигляд

$$\begin{cases} \sum_{j=1}^n c_j x_j \rightarrow \min, \\ \sum_{j=1}^n a_{ij} x_j \geq b_i, \quad i = \overline{1, m}, \\ x_j \geq 0, \quad j = \overline{1, n}. \end{cases} \quad (\text{B.6})$$

Зрозуміло, що модель (B.6) відображає не всі аспекти реальної ситуації, адже у ній враховані лише вимоги медицини й економіки, але зовсім не враховані інші чинники, наприклад смаки людини, для якої і складається раціон. Проте модель (B.6) є основою для розуміння сутності оптимального планування раціонів, а також різних сумішей та з'єднань. Як і (B.5) модель (B.6) є задачею лінійного програмування. ►

В2.6 (транспортна задача (Т-задача)). Нехай m постачальників (виробників) виробляють деяку однорідну продукцію, яка споживається n споживачами. При цьому обсяги виробництва продукції та обсяги її споживання відповідно дорівнюють $a_i (i = \overline{1, m})$ та $b_j (j = \overline{1, n})$ одиниць. Крім того, відома матриця $C = (c_{ij})_{i,j=1}^{m,n}$, де c_{ij} – вартість транспортних перевезень однієї одиниці продукції (кг, тонни тощо) від виробника i до споживача j . Потрібно побудувати такий план перевезень продукції $X = (x_{ij})_{i,j=1}^{m,n}$ (x_{ij} – шукана кількість одиниць продукції, що перевозиться від постачальника i до споживача j), при якому:

- 1) вся продукція буде вивезена від постачальників;
- 2) всі споживачі будуть повністю забезпечені необхідною кількістю продукції;
- 3) сумарна вартість перевезень буде мінімальною.

◀ Транспортна задача або Т-задача ще називається задачею оптимізації транспортних перевезень, а інколи задачею оптимізації маршруту переміщень, що безпосередньо пов'язане із її змістом. Математична модель описаної вище Т-задачі така:

$$\left\{ \begin{array}{l} \sum_{i=1}^m \sum_{j=1}^n c_{ij} x_{ij} \rightarrow \min, \\ \sum_{j=1}^n x_{ij} = a_i, \quad i = \overline{1, m}, \\ \sum_{i=1}^m x_{ij} = b_j, \quad j = \overline{1, n}, \\ x_{ij} \geq 0, \quad i = \overline{1, m}, \quad j = \overline{1, n}. \end{array} \right. \quad (B.7)$$

Модель (B.7) як оптимізаційна задача відноситься до класу спеціальних задач лінійного програмування. Якщо у ній перейти від двоіндексних змінних до одноіндексних, то її можна записати у формі, аналогічній до (B.5) і (B.6), про що йтиметься пізніше. Зауважимо, що транспортна задача та її модифікації мають широке практичне застосування. ►

B2.7 (оптимізація виробничої потужності підприємства). Підприємству потрібно придбати n видів обладнання з відповідним вектором виробничих потужностей $c = (c_1, \dots, c_n)^T$ за зміну. На придбання цього обладнання підприємство може виділити m видів виробничих ресурсів (гроші, виробничі площі тощо) з відповідним вектором обмежень $b = (b_1, \dots, b_m)^T$. Відома матриця $A = (a_{ij})_{i,j=1}^{m,n}$, де a_{ij} – це кількість одиниць ресурсу i , що витрачається на придбання одиниці обладнання j . Визначити такий набір придбаного обладнання $x = (x_1, \dots, x_n)^T$ (x_j – шукана кількість одиниць обладнання j), який забезпечує максимум

виробничої потужності підприємства за зміну.

◀ Математична модель задачі оптимізації виробничої потужності підприємства формалізується у вигляді

$$\begin{cases} \sum_{j=1}^n c_j x_j \rightarrow \max, \\ \sum_{j=1}^n a_{ij} x_j \leq b_i, \quad i = \overline{1, m}, \\ x_j \geq 0, \quad j = \overline{1, n}, \\ x_j \in Z, \quad j = \overline{1, n}, \end{cases} \quad (B.8)$$

де Z – множина цілих чисел. Як задача оптимізації, (B.8) є також спеціальною задачею лінійного програмування, а більш конкретно – задачею повністю цілочислового лінійного програмування. Коли йдеться про специфіку задачі з відповідного класу задач, то передбачається, що методи її розв’язування обов’язково повинні враховувати цю специфіку. Як побачимо пізніше, методи розв’язування Т-задач і задач цілочислового лінійного програмування справді відображають їх специфіку, що суттєво відрізняє їх від методів розв’язування типових задач лінійного програмування. ►

В2.8 (максимізація рентабельності виробництва).

Нехай виконуються умови задачі про оптимальне планування виробництва В2.4, але, крім технологічної матриці $A = (a_{ij})_{i,j=1}^{m,n}$, вектора обмежень на ресурси $b = (b_1, \dots, b_m)^T$ та вектора цін на продукцію $c = (c_1, \dots, c_n)^T$, заданий також вектор $d = (d_1, \dots, d_n)^T$ витрат виробничих фондів на одиничні випуски кожного з видів продукції (компоненти векторів c та d вимірюються у грошових одиницях). Необхідно визначити такий план виробництва n

($n > 1$) видів продукції $x = (x_1, \dots, x_n)^T$, який би забезпечив його максимальну рентабельність.

◀ Модель максимізації рентабельності виробництва описується задачею

$$\left\{ \begin{array}{l} \frac{\sum_{j=1}^n c_j x_j}{\sum_{j=1}^n d_j x_j} \rightarrow \max, \\ \sum_{j=1}^n a_{ij} x_j \leq b_i, \quad i = \overline{1, m}, \\ x_j \geq 0, \quad j = \overline{1, n}. \end{array} \right. \quad (B.9)$$

Оскільки цільова функція у даному випадку дробово-лінійна, то задачу (B.9) називають задачею дробово-лінійного програмування. Особливість цієї задачі у тому, що вона належить до класу задач нелінійного програмування (цільова функція нелінійна), але зводиться до задачі лінійного програмування, тому може бути розв'язана його методами. Зауважимо, що задача B2.8 та її модифікації мають широке практичне застосування. ►

B2.9 (максимізація випуску продукції одно-продуктової фірми). Однопродуктова фірма-виробник використовує n ($n > 1$) ресурсів (витрат) $x_1, \dots, x_n$, технологічний зв'язок яких з випуском продукції у задається неокласичною виробничою функцією [4, 5] $y = f(x_1, \dots, x_n)$. Ринок продукції та ресурсів досконалий, тому вектор цін на ресурси $w = (w_1, \dots, w_n)^T$ вважається також заданим. Бюджет фірми визначається величиною C . Потрібно знайти такий

набір ресурсів, вартість якого не перевищує бюджету і який максимізує випуск продукції.

◀ У математичному плані модель максимізації випуску продукції формалізується задачею

$$\begin{cases} f(x_1, \dots, x_n) \rightarrow \max, \\ \sum_{j=1}^n w_j x_j \leq C, \\ x_j \geq 0, \quad j = \overline{1, n}. \end{cases} \quad (\text{B.10})$$

Модель (B.10) ще називають короткостроковою моделлю фірми (відносно можливості придбання ресурсів). У сенсі математичної класифікації модель (B.10) є задачею нелінійного програмування, оскільки виробнича функція нелінійна. У неокласичній теорії виробництва [5] функція $f(x_1, \dots, x_n)$ вважається двічі неперервно диференційовною та угнутою (опуклою вгору), у зв'язку з чим задачу (B.10) ще називають задачею угнутого (чи опуклого) програмування. ►

В2.10 (максимізація розподілу капітальних вкладень). Для збільшення потужності виробництва продукції, яка користується підвищеним попитом, m підприємствам виділяються капітальні вкладень (КВ) загальним обсягом s одиниць (од.). Використання i -м підприємством u_i од. КВ приносить йому прибуток $k_i(u_i)$ од. Необхідно визначити такий варіант розподілу фінансових ресурсів, при якому сумарний прибуток усіх підприємств буде максимальним.

◀ Математична модель задачі про оптимальний розподіл капіталовкладень має вигляд

$$\begin{cases} \sum_{i=1}^m k_i(u_i) \rightarrow \max, \\ \sum_{i=1}^m u_i = s, \\ u_i \geq 0, \quad i = \overline{1, m}. \end{cases} \quad (B.11)$$

У загальному випадку модель (B.11) є задачею нелінійного програмування, але далеко не завжди її можна розв'язати методами нелінійного програмування (все залежить від класу функцій $k_i(u_i), i = \overline{1, m}$). Однак задачу (B.11) можна ефективно розв'язати методом динамічного програмування (термін «динамічне програмування» стосується обчислювального методу, а не певного класу задач). Для цього задачу (B.11) потрібно розглядати як багатетапну (чи багатокрокову) задачу оптимізації, коли оцінюється ефективність не допустимих варіантів розподілу КВ між m підприємствами, а ефективність укладених КВ на одному підприємстві, двох підприємствах і т.д. – m підприємствах. Це приводить нас до m етапів, на кожному з яких стан виробничої системи описується обсягом x_i КВ, виділених для освоєння i підприємствам з номерами від 1 до i включно. Величина u_i є керуванням (зовнішнім впливом), яке визначає процес зміни станів системи. Отже, задача (B.11) зведеться до багатокрокової задачі оптимізації

$$\begin{cases} \sum_{i=1}^m k_i(u_i) \rightarrow \max, \\ x_i = x_{i-1} + u_i, \quad i = \overline{1, m}, \\ u_i \in [0, s - x_{i-1}], \quad i = \overline{1, m}, \\ x_0 = 0, \quad x_m = s. \end{cases} \quad (B.12)$$

Задача (В.12) записана у формі задачі керування, для якої і розроблений метод динамічного програмування. ►

В.3. Класифікація моделей оптимізації

Як уже зазначено у попередньому пункті, формалізована мовою математики модель є математичною задачею, тому словосполучення «модель оптимізації» та «задача оптимізації» у цьому випадку рівносильні. Отже, перейдемо до класифікації оптимізаційних задач (моделей), яка залежить від таких критеріїв:

1. За кількістю керованих (шуканих) змінних:

- *одновимірні* задачі оптимізації, якщо x – скалярна змінна ($n = 1$);
- *багатовимірні* задачі оптимізації, якщо x – n -вимірний вектор ($n > 1$).

2. За наявністю обмежень:

- *задачі безумовної оптимізації*, якщо множина допустимих розв'язків $X = \mathbb{R}^n$ ($n \geq 1$) або співпадає з відкритою областю існування функції і не має ніяких додаткових обмежень;
- *задачі умовної оптимізації*, якщо допустима множина $X \subset \mathbb{R}^n$, тобто є власною підмножиною простору $\mathbb{R}^n$ ($n \geq 1$), яка задається прямими обмеженнями або функціональними обмеженнями у вигляді рівностей чи нерівностей, кількість і вигляд яких залежить від конкретної задачі.

3. За характером обмежень:

- *задачі умовної оптимізації з прямим обмеженням*;
- *задачі умовної оптимізації з обмеженнями – рівностями*;

- задачі умовної оптимізації з обмеженнями-нерівностями;
- задачі умовної оптимізації зі змішаними обмеженнями (типу рівності та нерівності);
- загальні задачі умовної оптимізації (із прямим обмеженням і обмеженнями-рівностями та обмеженнями-нерівностями).

4. За типом критерію оптимізації та обмежень:

- задачі лінійного програмування, якщо цільова функція й обмеження є лінійними функціями;
- задачі нелінійного програмування, якщо цільова функція нелінійна або обмеження містять нелінійні функції;
- задачі цілочислового програмування, якщо всі або деякі змінні задовольняють умову цілочисловості;
- задачі дробово-лінійного програмування, якщо цільова функція – це відношення двох лінійних функцій, а обмеження містять лінійні функції;
- задачі опуклого програмування, якщо цільова функція – опукла функція, а допустима множина – опукла множина;
- задачі квадратичного програмування, якщо цільова функція є квадратичною, а функції обмежень є лінійними;
- задачі сепарабельного програмування, якщо цільова функція і функції обмежень є лінійними комбінаціями функцій однієї змінної, тобто сепарабельними;
- задачі геометричного програмування, якщо цільова функція і функції обмежень є поліномами;
- задачі параметричного програмування, якщо цільова функція або коефіцієнти системи обмежень залежать від деяких параметрів;
- задачі стохастичного програмування, якщо цільова функція або система обмежень містять випадкові величини;

- *задачі динамічного програмування, якщо вони є багатоетапними задачами оптимізації.*

5. За кількістю екстремальних точок:

- *одноекстремальні задачі оптимізації, якщо цільова функція має в області допустимих значень єдиний екстремум;*
- *багatoesкстремальні задачі оптимізації, якщо цільова функція має в області допустимих значень більше одного екстремуму.*

В.4. Поняття про методи оптимізації

Методи розв'язування математичних задач, зокрема задач оптимізації, можуть бути *точними* (методи знаходження точних розв'язків) і *наближеними* (методи знаходження наближених розв'язків). А точні та наближені методи розрізняються як *аналітичні*, *числові* та *графічні*. Наприклад, якщо розв'язком задачі є функція, то вона може бути знайдена у вигляді точної чи наближеної аналітичної залежності, точної чи наближеної таблиці числових значень, точного чи наближеного графіка. Найчастіше на практиці використовуються аналітичні та числові методи. Графічні методи мають обмежену сферу застосування, оскільки вони суттєво залежать від розмірності задачі. Вибір методу розв'язування задачі (якщо він є) повинен бути глибоко обґрунтованим і привести у певному сенсі до оптимального варіанта.

Серед перелічених вище класів методів особливе місце займають числові методи. Більшість прикладних задач усе-таки розв'язуються цими методами. Числові методи поділяються на *скінченні* та *ітераційні*. Сутність скінченного методу полягає у тому, що розв'язок задачі знаходиться за скінченне число кроків (ітерацій). До таких методів

належать, наприклад метод Гаусса для розв'язування систем лінійних алгебраїчних рівнянь і симплекс-метод розв'язування задач лінійного програмування. Коли йдеться про ітераційні методи (або алгоритми), то мається на увазі, що ці методи дозволяють будувати нескінченну послідовність точок у відповідному точковому просторі, яка збігається до шуканого розв'язку. Зауважимо, що ітераційні методи в задачах економіки отримують чітку економічну інтерпретацію як процеси узгодження економічних рішень на шляху досягнення відповідної мети.

Будь-яка класифікація числових методів розв'язування оптимізаційних задач пов'язана з відповідним класом задач і його характеристиками.

Методи розв'язування задач оптимізації поділяються також на методи, які не використовують похідні від функцій, та методи, які базуються на цих похідних. Перші називаються *прямими методами*, або *методами нульового порядку*, а другі – *методами першого порядку*, чи *градієнтними методами* (використовуються перші похідні), або *методами другого порядку*, чи *методами Ньютонa* (використовуються другі похідні).

Зазначимо, що методи розв'язування задач умовної оптимізації у багатьох випадках є модифікаціями методів безумовної оптимізації або зводять задачу умовної оптимізації до розв'язування послідовності задач безумовної оптимізації.

На завершення даного пункту зазначимо, що спектр оптимізаційних моделей та методів їх розв'язування надзвичайно широкий, тому їх детальний огляд і аналіз навіть на неформалізованому рівні у рамках даного підручника неможливий і недоцільний.

ЧАСТИНА I. ОДНОВИМІРНА ОПТИМІЗАЦІЯ

Розділ 1. Основи теорії мінімізації функції однієї змінної

1.1. Поняття мінімуму функції однієї змінної

Нехай функція $f(x)$ визначена на множині X дійсної числової осі $\mathbb{R}^1 = \{x / -\infty < x < +\infty\}$ ($X \subseteq \mathbb{R}^1$) і у всіх точках $x \in X$ набуває скінченні значення (змінна x змінюється неперервно). Уточнимо постановку оптимізаційної моделі (задачі) (В.1) у даному випадку, конкретизувавши деякі базові поняття з теорії оптимізації функції однієї змінної.

Оскільки задача максимізації функції $f(x)$ еквівалентна задачі мінімізації функції $(-f(x))$, то надалі достатньо розглядати лише одну із задач (мінімізації чи максимізації), наприклад задачу мінімізації

$$\begin{cases} f(x) \rightarrow \min, \\ x \in X. \end{cases} \quad (1.1)$$

Зауважимо, що задачі мінімізації та максимізації функції у загальному випадку ще називаються задачами на екстремум функції.

Означення 1.1. Точку $x^* \in X$ називають точкою глобального (абсолютного) мінімуму функції $f(x)$ на множині X , якщо $f(x^*) \leq f(x)$ для всіх $x \in X$.

Значення $f(x^*)$ називають мінімальним значенням $f(x)$ на X і позначають $f^* = f(x^*) = \min_x f(x)$.

Множину всіх точок мінімуму надалі будемо позначати X^* . Множина X^* може містити одну, декілька або навіть нескінченно багато точок, а також може бути порожньою. Наприклад: 1) якщо $f(x) = \ln x$, $X = (0, 1]$, то множина точок мінімуму $X^* = \emptyset$; 2) якщо $f(x) = x^2$, $X = [-1, 1]$, то $X^* = \{0\}$, тобто складається з однієї точки мінімуму; 3) якщо $f(x) = \sin x$, $X = \mathbb{R}^1$, то множина точок мінімуму X^* – нескінченна множина; 4) якщо $f(x) = |x-1| + |x| - 1$, то на $X = [-1, 1]$ множина $X^* = [0, 1]$, на $X = [1, 2]$ – $X^* = \{1\}$, а на $X = (1, 2)$ – $X^* = \emptyset$; 5) якщо $f(x) = x$ при $x \in (0, 1]$ і $f(0) = 1$, то на множині $X = [0, 1]$ чи $X = (0, 1]$ вона не має найменшого значення, тому $X^* = \emptyset$.

Означення 1.2. Точка $x^* \in X$ називається точкою локального мінімуму функції $f(x)$ на множині X , якщо існує таке число $\varepsilon > 0$, що $f(x^*) \leq f(x)$ для $x \in O(x^*, \varepsilon) = \{x \in X / |x - x^*| < \varepsilon\}$.

Точка глобального мінімуму завжди є точкою локального мінімуму, але не навпаки.

Означення 1.3. Функція $f(x)$ називається обмеженою знизу на множині X , якщо існує таке число $m > -\infty$, що $f(x) \geq m$ для всіх $x \in X$. Функція $f(x)$ необмежена знизу на X , якщо існує послідовність точок $\{x_k\}_1^\infty \in X$, для якої $\lim_{k \rightarrow \infty} f(x_k) = -\infty$.

Якщо $X^* = \emptyset$, то узагальненням поняття найменшого значення функції є поняття точної нижньої межі функції.

Означення 1.4. Нехай функція $f(x)$ обмежена знизу на множині X . Тоді число f_* називається точною нижньою

межею функції $f(x)$ на множині X і позначається $f_* = \inf_X f(x)$, якщо:

- 1) $f(x) \geq f_*$ для всіх $x \in X$;
- 2) для довільного як завгодно малого числа $\varepsilon > 0$ існує точка $x_* \in X$, для якої $f(x_*) < f_* + \varepsilon$ (тобто серед значень $f(x)$ на множині X знайдуться як завгодно близькі до f_*).

Точна нижня (як і точна верхня) межа функції $f(x)$ не є досяжною, що демонструють приклади функцій $f(x) = 1/x$ на $X = (0, +\infty)$ та

$$f(x) = \begin{cases} \sin x, & x \in (0, \pi/2) \\ 1/3 & \text{при } x = 0 \text{ і } x = \pi/2. \end{cases}$$

Для необмежених знизу функцій $f(x)$ точна нижня межа $f_* = -\infty$.

Якщо $X^* = \emptyset$, то задача (1.1) втрачає зміст. Тоді, якщо існує $f_* = \inf_X f(x)$, будують мінімізуючу послідовність $\{x_k\}_1^\infty$ для функції $f(x)$ на X і за наближений розв'язок беруть таке x_k , для якого $f(x_k)$ із заданою похибкою ε наближає f_* ($f(x_k) - f_* < \varepsilon$).

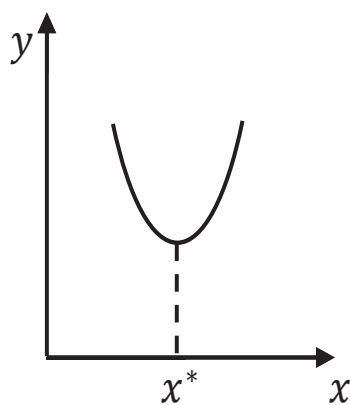
Якщо множина X – замкнена й обмежена (компакт), тобто відрізок, а функція $f(x)$ – неперервна на X , то, згідно з відомою із математичного аналізу теоремою Вейерштраса [1, 14, 20], функція $f(x)$ досягає на X своїх мінімального і максимального значень. Отже, існує досить широкий клас задач (1.1), які завжди мають розв'язок, і довільна мінімізуюча послідовність збігається до X^* .

1.2. Класичний метод мінімізації функції однієї змінної

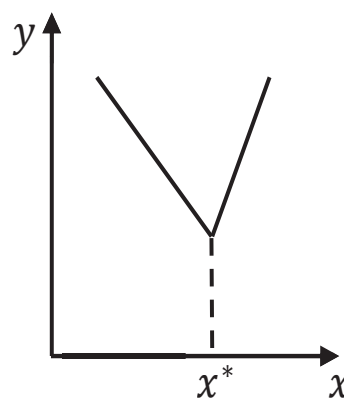
Під класичним методом будемо розуміти той підхід до пошуку екстремуму функції, який базується на диференціальному численні та детально викладений у літературі з математичного аналізу. Нижче сформулюємо основні результати, які формують основи даного методу.

Як відомо, точками екстремуму (наприклад, мінімуму) функції $f(x)$ на X можуть бути лише ті точки x^* , у яких виконується одна з умов:

- 1) похідна $f'(x^*)$ існує і дорівнює нулю (рис. 1.1, а));
- 2) $f(x)$ – неперервна функція, але похідної $f'(x^*)$ не існує (рис. 1.1, б));
- 3) функція $f(x)$ має розрив у точці x^* (рис. 1.1, в), г), д));
- 4) точка є крайньою для проміжку X , наприклад точка $x^* = a$ чи $x^* = b$ для відрізка $X = [a, b]$ (a, b – конкретні числові значення) (рис. 1.1, е)).



а)



б)

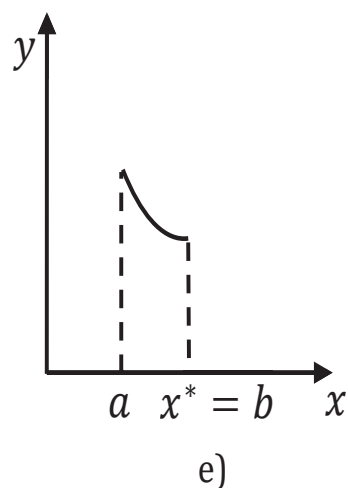
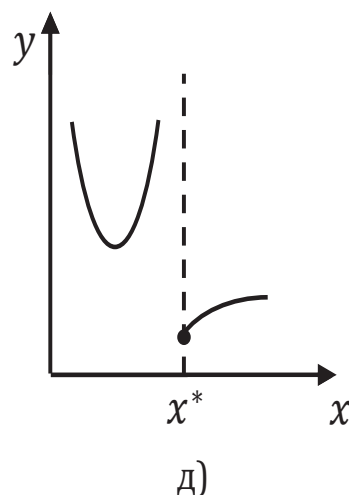
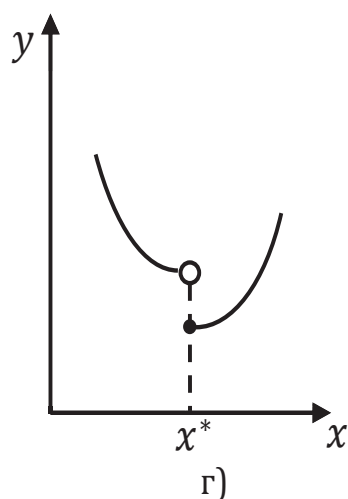
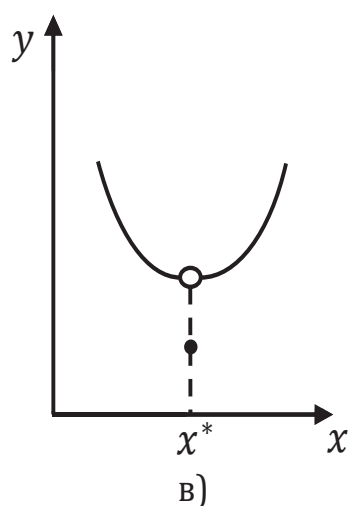


Рис. 1.1. Ілюстрація умов існування точки мінімуму
функції $y = f(x)$ (а) – е))

Підкреслимо, що у випадках в), г) і д) (рис. 1.1) функція $f(x)$ має розриви різних типів, а саме: в) – усувний розрив (функція має у точці x^* скінченне значення $f(x^*)$, але $\lim_{x \rightarrow x^*} f(x) \neq f(x^*)$); г) – розрив 1-го роду (функція задана у точці x^* , причому $f(x^*) = \lim_{x \rightarrow x^* + 0} f(x) = f(x^* + 0)$, але

$f(x^* - 0) = \lim_{x \rightarrow x^* - 0} f(x) \neq f(x^* + 0)$); д) – розрив 2-го роду (функція задана у точці x^* , причому $f(x^*) = f(x^* + 0)$, але ліва границя $f(x^* - 0) = +\infty$, оскільки $x = x^*$ – вертикальна асимптота). Приклади реальних функцій, які у точці $x^* = 0$ відображають ситуації а)–е) (рис. 1.1) відповідно є: $f(x) = x^2$, $f(x) = |x|$,

$$f(x) = \begin{cases} x^2, & x \neq 0, \\ -1, & x = 0, \end{cases} \quad f(x) = \begin{cases} x^2, & x < 0, \\ x^2 - 1, & x \geq 0, \end{cases}$$

$$f(x) = \begin{cases} -\frac{1}{x}, & x < 0, \\ \sqrt{x}, & x \geq 0, \end{cases} \quad f(x) = x^2 \quad (X = (-\infty, 0]).$$

Пошук точок екстремуму починають зі знаходження всіх підозрілих на екстремум точок, а потім, провівши допоміжні дослідження, відбирають ті, які справді є точками мінімуму чи максимуму.

Для з'ясування умов, які повинні виконуватися у точках локальних екстремумів функції $f(x)$, розглянемо задачу (1.1) лише для випадків, коли $X = \mathbb{R}^1$ або X є відкритою множиною числової осі $\mathbb{R}^1$, тобто складається тільки із внутрішніх точок (точок, що належать X разом із деяким оточенням (інтервалом)).

Перш ніж сформулювати необхідні й достатні умови екстремуму функції $f(x)$ у деякій точці $x^* \in X$, нагадаємо, що функція $f(x)$ диференційовна у точці x^* тоді і лише тоді, коли вона має у цій точці скінченну похідну (тоді $\Delta f(x^*) = f(x^* + \Delta x) - f(x^*) = f'(x^*)\Delta x + \alpha(\Delta x)\Delta x = f'(x^*)\Delta x + o(\Delta x)$, де Δx – приріст аргумента, $\alpha(\Delta x)$ – нескінченно мала функція

при $\Delta x \rightarrow 0$). Інакше кажучи, поняття диференційовності функції у точці та поняття існування похідної від функції у цій точці тотожні. Диференційовна у точці функція буде неперервною у цій точці. Коли кажуть, що функція $f(x)$ n ($n \geq 2$) разів диференційовна у точці x^* , то мається на увазі, що її похідна $(n-1)$ -го порядку, тобто функція $f^{(n-1)}(x)$, диференційовна у точці x^* .

Важливу роль у теорії функцій, у тому числі у теорії екстремумів відіграє так звана формула Тейлора [14, 20].

Теорема 1.1. *Нехай функція $f(x)$ диференційовна $(n-1)$ разів у деякому околі точки x^* і n разів диференційовна у самій точці x^* . Тоді для кожного x із цього окола ($x \rightarrow x^*$) справджується рівність*

$$f(x) = f(x^*) + f'(x^*)(x - x^*) + \frac{1}{2!} f''(x^*)(x - x^*)^2 + \dots + \frac{f^{(n)}(x^*)(x - x^*)^n}{n!} + o((x - x^*)^n) = P_n(x, x^*) + R_{n+1}(x, x^*), \quad (1.2)$$

де $o((x - x^*)^n) = o((\Delta x)^n)$ - означає нескінченну малу величину при $\Delta x \rightarrow 0$ ($x \rightarrow x^*$) більш високого порядку, ніж $(\Delta x)^n$.

Формула (1.2) називається формулою Тейлора з центром розкладу у точці x^* . При цьому $P_n(x, x^*)$ та $R_{n+1}(x, x^*)$ - відповідно многочлен Тейлора порядку n функції $f(x)$ у точці x^* та залишковий член формули Тейлора у формі Пеано.

При більш сильних (жорстких) вимогах до функції $f(x)$ має місце таке твердження:

якщо функція $f(x)$ диференційовна n разів у деякому околі точки x^ , то для кожної точки x із цього окола*

знайдеться точка x' , розміщена між точками x і x^* , що

$$f(x) = P_{n-1}(x, x^*) + \tilde{R}_n(x, x^*),$$

де $P_{n-1}(x, x^*)$ – многочлен Тейлора порядку $(n-1)$ функції

$f(x)$ у точці x^* , а $\tilde{R}_n(x, x^*) = \frac{f^{(n)}(x')}{n!}(x - x^*)^n$ – залишковий

член у формі Лагранжа.

Перейдемо до конкретизації необхідних і достатніх умов екстремуму.

Теорема 1.2 (необхідна умова екстремуму першого порядку). Нехай:

1) x^* – точка локального екстремуму функції $f(x)$;

2) функція $f(x)$ диференційовна у точці x^* .

Тоді

$$f'(x^*) = 0. \quad (1.3)$$

◀ Умова (1.3) доводиться елементарно. Наприклад, якщо x^* – точка мінімуму, то у деякому її околі $f(x^*) \leq f(x)$, а для лівої та правої похідних у точці x^* виконуються нерівності

$$f'(x^* - 0) = \lim_{\Delta x \rightarrow -0} \frac{f(x^* + \Delta x) - f(x^*)}{\Delta x} \leq 0,$$

$$f'(x^* + 0) = \lim_{\Delta x \rightarrow +0} \frac{f(x^* + \Delta x) - f(x^*)}{\Delta x} \geq 0,$$

тобто $f'(x^*) = f'(x^* - 0) = f'(x^* + 0) = 0$. ▶

Теорему 1.2 ще називають теоремою Ферма.

Точки $x^* \in X$, для яких виконується необхідна умова екстремуму (1.3), тобто точки підозрілі на екстремум, ще називаються *стаціонарними*, або *критичними* точками. До таких точок, як уже сказано раніше, належать також точки, у яких похідна не існує, наприклад для функції $f(x) = |x|$ цією точкою є $x^* = 0$.

Крім необхідної умови екстремуму першого порядку, для виявлення «зайвих» стаціонарних точок використовується також необхідна умова екстремуму другого порядку, яка конкретизується у наступному твердженні.

Теорема 1.3 (необхідна умова екстремуму другого порядку). *Нехай:*

1) x^* – точка локального мінімуму (максимуму) функції $f(x)$;

2) функція $f(x)$ диференційовна у деякому околі точки x^* і двічі диференційовна у самій точці x^* .

Тоді друга похідна у цій точці невід’ємна (недодатна):
 $f''(x^*) \geq 0$ ($f''(x^*) \leq 0$).

◀ Для доведення теореми 1.3 достатньо скористатися формулою Тейлора (1.2), згідно з якою у зазначеному околі точки x^*

$$\begin{aligned} f(x) - f(x^*) &= f'(x^*)\Delta x + \frac{1}{2} f''(x^*)(\Delta x)^2 + o((\Delta x)^2) = \\ &= \frac{1}{2} f''(x^*)(\Delta x)^2 + o((\Delta x)^2). \end{aligned} \quad (1.4)$$

Якщо припустити, наприклад, що у точці мінімуму $f''(x^*) < 0$, то при достатньо малому Δx права частина рівності (1.4) буде від’ємною, тобто її знак повністю визначається першим доданком і $f(x) < f(x^*)$, що суперечить тому, що точка x^* є точкою локального мінімуму. Отже, у точці мінімуму $f''(x^*) \geq 0$. Аналогічно доводиться друга частина теореми. ►

Але не всі зі стаціонарних точок є точками мінімуму, тобто умова (1.3) виділяє стаціонарні точки, але не визначає їх характер (не вказує є вони точками мінімуму, максимуму чи перегину). Очевидно, для визначення характеру

стаціонарної точки потрібно мати деякі допоміжні умови, які виконуються лише у точках мінімуму чи максимуму.

Теорема 1.4 (перша достатня умова екстремуму).

Нехай:

- 1) x^* – стаціонарна точка функції $f(x)$;
- 2) функція $f(x)$ диференційовна у деякому околі точки x^* , причому зліва і справа від неї похідна $f'(x)$ має протилежні знаки.

Тоді у точці x^ функція $f(x)$ має:*

- 1) максимум, якщо похідна змінює свій знак з «+» на «-»;
- 2) мінімум, якщо похідна змінює свій знак з «-» на «+».

◀ Оскільки x^* – стаціонарна точка, то $f'(x^*) = 0$. Функція $f(x)$ – неперервна у точці x^* , тобто має скінченне значення $f(x^*)$. Якщо $f'(x) < 0$ при $x < x^*$ і $f'(x) > 0$ при $x > x^*$, то це означає, що зліва від точки x^* функція спадає, а справа – зростає, тому $f(x^*)$ є найменшим значенням функції $f(x)$ у деякому околі точки x^* . Аналогічно доводиться, що у точці x^* функція має максимум, якщо $f'(x)$ при переході через цю точку змінює свій знак з «+» на «-». ►

Зауважимо: а) якщо похідна $f'(x)$ при переході через стаціонарну точку x^* не змінює знак, то в цій точці функція екстремуму не має; б) якщо функція $f(x)$ неперервна на інтервалі X і диференційовна у деякому околі точки $x^* \in X$, крім, можливо, самої точки x^* , то твердження теореми 1.4 залишаються справедливими.

Приклад 1.1. Знайти екстремум функції $f(x) = x^4 - x^3$.

◀ Оскільки $f'(x) = 4x^3 - 3x^2$, то, розв'язавши рівняння

$f'(x)=0$, отримаємо дві стаціонарні точки $x_1^*=0$, $x_2^*=3/4$. Легко перевірити, що $f'(x)<0$ при $x<x_1^*=0$ і при $x_1^*<x<x_2^*=3/4$, а при $x>x_2^*$ похідна $f'(x)>0$. Отже, у точці x_1^* функція не має екстремуму, а у точці x_2^* вона має мінімум, причому $f_{\min}=f\left(\frac{3}{4}\right)=\left(\frac{3}{4}\right)^4-\left(\frac{3}{4}\right)^3=-\frac{27}{256}$. ►

Теорема 1.5 (друга достатня умова екстремуму). Нехай:

- 1) x^* – стаціонарна точка функції $f(x)$;
- 2) функція $f(x)$ диференційовна у деякому околі точки x^* і двічі диференційовна у самій точці x^* .

Тоді у точці x^* функція $f(x)$ має:

- 1) мінімум, якщо $f''(x^*)>0$;
- 2) максимум, якщо $f''(x^*)<0$.

◀ Наприклад, якщо $f''(x^*)>0$, то існує окіл точки x^* , у якому функція $f'(x)$ зростає, причому, враховуючи, що $f'(x^*)=0$, змінює свій знак із «–» на «+» при переході через точку x^* . Отже, якщо $f''(x^*)>0$, то, згідно з теоремою 1.4, точка x^* є точкою мінімуму. Аналогічно встановлюється, що при $f''(x^*)<0$ точка x^* є точкою максимуму. ►

Зауважимо, що спосіб доведення теореми 1.5 можна використати також для доведення теореми 1.3.

У випадку, $f''(x^*)=0$, у точці x^* може бути екстремум, а може і не бути. Узагальненням теореми 1.5 є наступна теорема.

Теорема 1.6 (третя достатня умова екстремуму). Нехай:

- 1) функція $f(x)$ $(n-1)$ разів диференційовна у деякому околі точки x^* і n разів диференційовна у самій точці x^* ;

2) у точці x^* виконуються умови:

$$f'(x^*) = f''(x^*) = \dots = f^{(n-1)}(x^*) = 0, \quad f^{(n)}(x^*) \neq 0.$$

Тоді, якщо n – непарне, то точка x^* не є точкою екстремуму, а якщо n – парне, то функція $f(x)$ має у точці x^* мінімум, якщо $f^{(n)}(x^*) > 0$, і максимум, якщо $f^{(n)}(x^*) < 0$.

◀ Нехай n – парне і $f^{(n)}(x^*) > 0$. Тоді функція $f^{(n-1)}(x)$ зростає у точці x^* . Оскільки $f^{(n-1)}(x^*) = 0$, то існує окіл точки x^* , у межах якого зліва від точки x^* функція $f^{(n-1)}(x) < 0$, а справа від точки x^* – $f^{(n-1)}(x) > 0$. Згідно із формулою Тейлора із залишковим членом у формі Лагранжа для функції $f'(x)$ маємо

$$\begin{aligned} f'(x) = & f'(x^*) + f''(x^*)(x - x^*) + \dots + \frac{f^{(n-2)}(x^*)}{(n-3)!}(x - x^*)^{n-3} + \\ & + \frac{f^{(n-1)}(x')}{(n-2)!}(x - x^*)^{n-2}. \end{aligned} \quad (1.5)$$

Із (1.5) випливає, що

$$f'(x) = \frac{f^{(n-1)}(x')}{(n-2)!}(x - x^*)^{n-2},$$

де точка x' лежить між x^* та x . Із останнього співвідношення випливає (n – парне), що знак функції $f'(x)$ визначається знаком $f^{(n-1)}(x')$. Оскільки $f^{(n-1)}(x') < 0$ при $x' < x^*$ і $f^{(n-1)}(x') > 0$ при $x' > x^*$, то при переході через точку x^* перша похідна $f'(x)$ змінює свій знак з «–» на «+», тобто x^* – точка мінімуму. Випадок $f^{(n)}(x^*) < 0$ доводиться аналогічно.

Якщо n – непарне, то при переході через точку x^* похідна $f'(x)$ не змінює свого знаку. ►

Зауважимо, що коли йдеться про те, що функція $f(x)$ n

($n \geq 1$) разів диференційовна у точці, то це означає, що у деякому околі цієї точки вона має всі похідні до $(n-1)$ -го порядку включно, що власне забезпечує існування n -ої похідної (для визначення похідної у точці потрібно, щоб функція у деякому околі цієї точки була визначеною). Крім того, всі ці похідні неперервні у точці, бо із диференційовності випливає неперервність. Отже, у вищенаведених теоремах у зв'язку з цим можливі різні еквівалентні формулювання.

Доповнюючи теореми 1.5, 1.6, зазначимо, що у випадку, коли у стаціонарній точці для парного n $f^n(x^*) = 0$ ($n \geq 2$), для з'ясування характеру цієї точки потрібні додаткові дослідження, наприклад для функції $f(x) = \sin x + \cos x + e^{-x}$ стаціонарна точка $x^* = 0$ ($f'(0) = 0$), але $f''(0) = 0$. Дослідження показують, що ця точка є точкою локального максимуму.

Підсумовуючи викладений вище матеріал, зацентруємося на кількох важливих висновках.

1. У теоремах 1.2–1.6 не підкреслювалось, яким є екстремум. Звичайно, у загальному випадку йшлося про локальні екстремуми, оскільки глобальний екстремум завжди локальний.

2. При використанні класичного методу мінімізації (чи максимізації) функції однієї змінної, щоб знайти глобальний екстремум функції $f(x)$ на X , потрібно перебрати всі точки локального екстремуму і вибрати серед них точку з найменшим (чи найбільшим) значенням функції, якщо таке існує. Для множини $X = \mathbb{R}^1$, $X = (-\infty, b)$, $X = (-\infty, b]$, $X = (a, +\infty)$, $X = [a, +\infty)$ поряд з вище-зазначеними дослідженнями потрібно також вивчати поведінку функції при $x \rightarrow -\infty$ чи $x \rightarrow +\infty$.

3. На початку даного пункту йшлося про відкриту

область визначення функції X ($X = \mathbb{R}^1$ або $X = (a, b)$), тобто область, що складається лише із внутрішніх точок, тому точки екстремуму також були внутрішніми, що і пов'язано зі специфікою поняття похідної та її застосуванням для визначення точок екстремуму диференційовних функцій. У загальному випадку допустима множина X задачі (1.1) може бути довільним проміжком із $\mathbb{R}^1$ або збігатись з $\mathbb{R}^1$. Залежно від того X відкрита чи замкнена (або напівзамкнена) множина, можна запропонувати відповідну схему розв'язування класичним методом задачі (1.1), у якій $f(x)$ – диференційовна достатню кількість разів у всіх внутрішніх точках X .

Якщо X – відкрита множина, то для знаходження розв'язку задачі (1.1), тобто точки $x^* \in X$, у якій функція досягає свого мінімуму (вважається, що точка існує), потрібно:

- 1) знайти всі стаціонарні точки функції $f(x)$ на X (тобто точки, які є коренями рівняння (1.3));
- 2) дослідити стаціонарні точки за допомогою достатніх умов оптимальності і вибрати точки локального мінімуму, а з них – точку глобального мінімуму.

Якщо X – замкнена чи напівзамкнена множина ($X = [a, b]$, $X = (a, b]$ та ін.), то для знаходження розв'язку задачі (1.1) потрібно:

- 1) знайти всі стаціонарні точки функції $f(x)$ на множині усіх внутрішніх точок X ;
- 2) за допомогою достатніх умов оптимальності вибрати серед стаціонарних точок ті, які є точками локального мінімуму;
- 3) значення $f(x)$ у точках мінімуму і у точках замикання множини X (кінцях проміжку або одному з них) порівняти між собою і вибрати найменше з них, якому відповідає точка

глобального мінімуму або розв'язок задачі (1.1).

Зауважимо, що на практиці при реалізації даних схем простіше порівняти значення функції $f(x)$ у всіх стаціонарних точках, не досліджуючи їх характер.

Приклад 1.2. Знайти найбільше та найменше значення функції $f(x) = x^3 - 3x^2 + 5$ на множині $X = [-2, 4]$.

◀ Згідно із зазначеною вище схемою розв'язування задачі (1.1) маємо: 1) $f'(x) = 3x^2 - 6x$, тобто стаціонарними точками на $(-2, 4)$ є точки $x_1 = 0, x_2 = 2$, які є коренями рівняння $3x^2 - 6x = 0$; 2) оскільки $f''(x) = 6x - 6$, то $f''(0) = -6 < 0$, $f''(2) = 6 > 0$, тобто $x_1 = 0$ – точка локального максимуму, а $x_2 = 2$ – точка локального мінімуму; 3) враховуючи, що $f(-2) = -15$, $f(0) = 5$, $f(2) = 1$, $f(4) = 21$, то найменше значення функції досягається у точці $x = -2$, а найбільше – у точці $x = 4$. ▶

4. Класичний метод мінімізації функції однієї змінної доцільно використовувати у тих випадках, коли достатньо просто виявити всі стаціонарні точки та реалізувати описану вище схему розв'язування задачі (1.1). При розв'язуванні практичних задач класичний метод, на жаль, має обмежене застосування. Це пояснюється тим, що, по-перше, у багатьох випадках значення цільової функції $f(x)$ є результатом певних вимірювань або експериментів, а обчислення похідної складне чи неможливе, а, по-друге, навіть якщо похідна $f'(x)$ задана аналітично, далеко не завжди просто розв'язати рівняння (1.3). Тому потрібно мати також інші методи розв'язування задачі (1.1), зокрема такі, які більш зручні для сучасних обчислювальних систем.

5. Зазначимо також, що задача (1.1) називається задачею безумовної мінімізації, якщо $X = \mathbb{R}^1$, тобто ніяких

обмежень на змінну x не накладається. Якщо X – деяка власна підмножина із $\mathbb{R}^1$, то задача (1.1) називається задачею умовної мінімізації (умовною задачею).

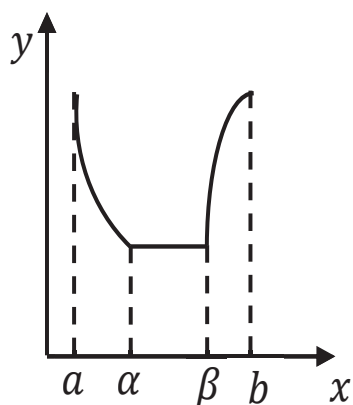
1.3. Унімодальні та опуклі функції

Розв'язування задачі (1.1) значно ускладнюється, якщо функція $f(x)$ має, крім глобального, також, відмінні від нього, локальні мінімуми і, навпаки, спрощується, коли локальний і глобальний мінімуми збігаються. Нижче розглянемо два класи функцій, у яких кожний локальний мінімум є одночасно і глобальним. Це унімодальні та опуклі функції.

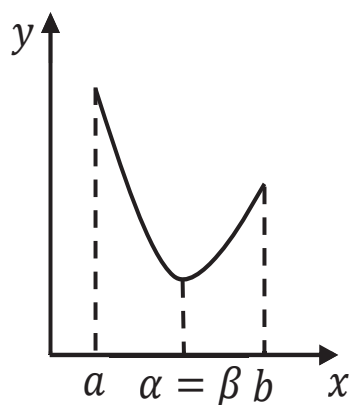
Означення 1.5. Неперервна на відрізку $[a,b]$ функція $f(x)$ називається унімодальною, якщо існують числа α та β ($a \leq \alpha \leq \beta \leq b$) такі, що:

- 1) при $a < \alpha$ функція $f(x)$ на відрізку $[a, \alpha]$ монотонно спадає;
- 2) при $\beta < b$ функція $f(x)$ на відрізку $[\beta, b]$ монотонно зростає;
- 3) при $x \in [\alpha, \beta]$ функція $f(x)$ набуває мінімальних значень, тобто $f(x) = f^* = \min_{[a,b]} f(x)$, $X^* = [\alpha, \beta]$.

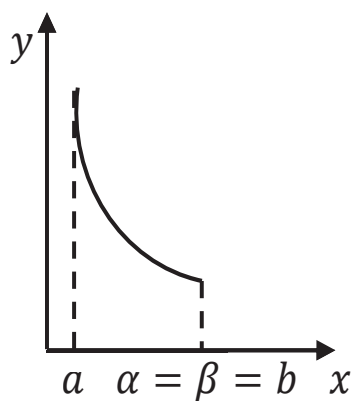
Зауважимо, що означення 1.5 охоплює також випадки, коли один або два відрізки із $[a, \alpha]$, $[\alpha, \beta]$, $[\beta, b]$ вироджуються у точку (рис. 1.2), зокрема якщо $\alpha = \beta$, то функція $f(x)$ називається строго унімодальною (рис. 1.2, 2)).



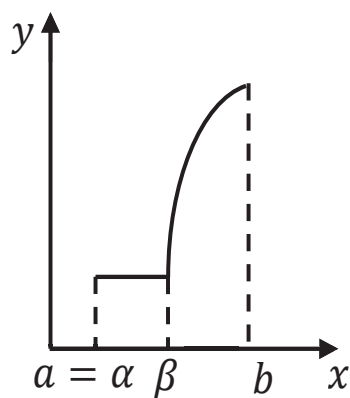
1)



2)



3)



4)

Рис. 1.2. Графічна ілюстрація деяких неперервних унімодальних функцій $y = f(x)$ (1), 2), 3), 4))

Унімодальні функції мають такі властивості:

1) будь-яка точка локального мінімуму унімодальної функції є одночасно і точкою її глобального мінімуму на відрізку $[a, b]$;

2) унімодальна на відрізку $[a, b]$ функція є унімодальною на будь-якому меншому відрізку $[c, d] \subset [a, b]$;

3) у випадку, коли функція $f(x)$ унімодальна та $a \leq x_1 < x_2 \leq b$, то $x^* \in [a, x_2]$, якщо $f(x_1) \leq f(x_2)$ і $x^* \in [x_1, b]$,

якщо $f(x_1) > f(x_2)$, де x^* – одна з точок мінімуму $f(x)$ на $[a, b]$.

Перелічені вище властивості впливають з означення 1.5.

Поняття унімодальності узагальнюється також на всі визначені на відрізку $X = [a, b]$ функції, які набувають скінченні значення, але не обов'язково є неперервними. У зв'язку з цим у більш загальному випадку функцію $f(x)$ називають строго унімодальною на X , якщо існує така точка $x^* \in X$, що

$$f(x_1) > f(x_2), \text{ якщо } x_1 < x_2 < x^*, \quad x_1, x_2 \in X,$$

$$f(x_1) < f(x_2), \text{ якщо } x^* < x_1 < x_2, \quad x_1, x_2 \in X.$$

У випадку, коли існує точка $x^* \in X$ (або й множина точок), що $f(x_1) \geq f(x_2)$ при $x_1 < x_2 < x^*$ та $f(x_1) \leq f(x_2)$ при $x^* < x_1 < x_2$ ($x_1, x_2 \in X$), функція $f(x)$ буде просто унімодальною. Очевидно, що дані означення строгої унімодальності та унімодальності допускають ситуації, коли $f(x)$ має усувні точки або точки розривів першого роду. Зауважимо, що найширше використовуються неперервні унімодальні функції.

Аналогічно визначаються й унімодальні функції у випадку задачі максимізації.

Ще один клас функцій, для яких поняття локального та глобального екстремуму еквівалентні, – це опуклі функції.

Означення 1.6. *Задана на відрізку $[a, b]$ функція $f(x)$ називається опуклою (опуклою вниз), якщо для $x_1, x_2 \in [a, b]$ і довільного числа $\alpha \in [0, 1]$ виконується нерівність*

$$f(\alpha x_1 + (1 - \alpha)x_2) \leq \alpha f(x_1) + (1 - \alpha)f(x_2). \quad (1.6)$$

Для опуклої вгору (угнутої) функції справедлива нерівність

$$f(\alpha x_1 + (1-\alpha)x_2) \geq \alpha f(x_1) + (1-\alpha)f(x_2). \quad (1.7)$$

Якщо $f(x)$ – опукла, то $(-f(x))$ – угнута і навпаки, що і впливає з нерівностей (1.6), (1.7). Якщо нерівності (1.6) і (1.7) перетворюються у рівності лише при $\alpha = 0$ та $\alpha = 1$, то функція $f(x)$ буде відповідно строго опуклою та строго угнутою. Такі, наприклад, функції $f(x) = x^2$ і $f(x) = -x^2$. Графіки опуклих і угнутих функцій продемонстровані на рис. 1.3.

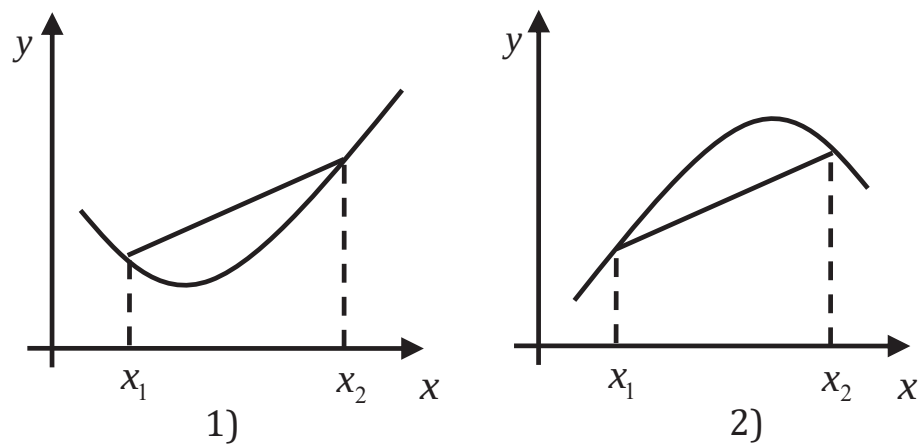


Рис. 1.3. Графіки опуклої (1) та угнутої (2) функції $y = f(x)$

Враховуючи зв'язок між опуклими та угнутими функціями, достатньо обмежитись вивченням властивостей тільки опуклих функцій. Перелічимо основні властивості опуклих функцій.

1. На довільному відрізку $[x_1, x_2] \subset [a, b]$ графік опуклої на $[a, b]$ функції $f(x)$ розміщений під хордою, проведеною через точки графіка з абсцисами x_1 і x_2 (рис. 1.3, 1)).

2. Опукла на відрізку $[a, b]$ функція $f(x)$ неперервна у кожній внутрішній точці x цього відрізка ($x \in (a, b)$) і має скінченні праву і ліву похідні.

Зазначимо, що на кінцях відрізка $[a, b]$ опукла функція може не тільки не мати відповідних односторонніх похідних, але й мати розрив. Наприклад, функція

$$f(x) = \begin{cases} 2, & x = 0, \\ x, & 0 < x < 1, \\ 2, & x = 1 \end{cases}$$

опукла на $[0, 1]$, але на кінцях відрізка має розриви. Функція $f(x) = -\sqrt{1-x^2}$ опукла і неперервна на $[-1, 1]$, але на кінцях відрізка не має скінченних похідних $f'(-1+0)$, $f'(1-0)$.

3. Довільна точка локального мінімуму опуклої та неперервної на $[a, b]$ функції $f(x)$ є точкою її глобального мінімуму на $[a, b]$.

У цьому твердженні підкреслюється, що функція $f(x)$ обов'язково має бути неперервною на кінцях відрізка, тобто у точках $x=a$ і $x=b$ мають існувати відповідно права та ліва границі, причому $f(a+0) = f(a)$ і $f(b-0) = f(b)$, оскільки у внутрішніх точках опукла функція завжди неперервна.

4. На практиці користуватися означенням (1.6) для дослідження опуклості функції незручно. Для цього використовують більш зручні, особливо у випадку диференційовних функцій, критерії опуклості, які сформулюємо у вигляді теорем для $X = [a, b]$.

Теорема 1.7. Для того, щоб диференційовна на відрізку $[a, b]$ функція $f(x)$ була опуклою на цьому відрізку, необхідно й достатньо, щоб її похідна $f'(x)$ була неспадною на $[a, b]$.

Зростання $f'(x)$ на (a, b) є необхідною й достатньою умовою строгої опуклості $f(x)$.

Теорема 1.8. Для того, щоб двічі диференційовна на відрізку $[a, b]$ функція $f(x)$ була опуклою на цьому відрізку,

необхідно й достатньо, щоб при всіх $x \in [a, b]$ виконувалася нерівність $f''(x) \geq 0$.

Достатньою умовою опуклості (угнутості) функції $f(x)$ на $[a, b]$ є умова: $f''(x) > 0$ ($f''(x) < 0$), $x \in [a, b]$.

Можна також довести, що для того, щоб $f(x)$ була строго опуклою на $[a, b]$, необхідно й достатньо, щоб при всіх $x \in [a, b]$ виконувалася нерівність $f''(x) \geq 0$, але, крім того, не існувало інтервалу $(c, d) \subset [a, b]$, в якому $f''(x) = 0$.

Зауважимо, що у теоремах 1.7, 1.8 диференційовність функцій на кінцях відрізка розуміється як існування відповідної скінченної односторонньої похідної.

Приклад 1.3. Використовуючи критерії опуклості (теорема 1.7, 1.8), дослідити на опуклість в області свого визначення функції $f_1(x) = 2^x$, $f_2(x) = -\ln x$.

◀ Функція $f_1(x) = 2^x$ визначена на всій числовій осі $\mathbb{R}^1$, а функція $f_2(x) = -\ln x$ – на $(0, +\infty)$. Оскільки $f_1''(x) = (\ln 2)^2 2^x > 0$, $f_2''(x) = \frac{1}{x^2} > 0$, то згідно з теоремами 1.7 і 1.8 функції $f_1(x)$ і $f_2(x)$ опуклі. ▶

Умова опуклості на відрізку $[a, b]$ диференційовної функції $f(x)$ означає, що на цьому відрізку довільна дотична до графіка $f(x)$ розміщена не вище цього графіка. Наслідком критеріїв опуклості диференційовних функцій є наступне твердження.

Теорема 1.9. Якщо $f(x)$ – опукла диференційовна на відрізку $[a, b]$ функція і у точці $x^* \in [a, b]$ виконується рівність $f'(x^*) = 0$, то x^* – точка глобального мінімуму $f(x)$ на $[a, b]$.

Доведення теореми 1.9 випливає з того, що рівняння

дотичної до графіка функції $f(x)$ у точці x^* (тобто рівняння $y(x) = f(x^*) + f'(x^*)(x - x^*)$) набуде вигляду $y(x) = f(x^*)$, але для опуклої функції $f(x) \geq f(x^*)$ (графік розміщений не нижче дотичної) для всіх $x \in [a, b]$, тобто x^* – точка глобального мінімуму $f(x)$. Зауважимо також, що x^* – внутрішня точка $[a, b]$, хоча умови теореми будуть виконуватися і для крайніх точок відрізка, якщо функція буде диференційовною в деяких околах цих точок.

Отже, рівність $f'(x^*) = 0$ для опуклої диференційовної функції є не тільки необхідною умовою глобального екстремуму (як і необхідною умовою екстремуму для будь-якої диференційовної функції), але і його достатньою умовою.

5. Довільна опукла неперервна на відрізку $[a, b]$ функція є унімодальною на цьому відрізку. Обернене твердження у загальному випадку хибне. Наприклад, на рис. 1.4 зображений графік унімодальної, але не опуклої функції, оскільки на відрізку $[x', x''] \subset [a, b]$ графік $f(x)$ розміщений не під хордою, що проходить через точки $(x', f(x'))$, $(x'', f(x''))$, а над нею.

Оскільки опукла функція неперервна у всіх внутрішніх точках $[a, b]$, то її неперервність на $[a, b]$ означає, що $\lim_{x \rightarrow a+0} f(x) = f(a)$, $\lim_{x \rightarrow b-0} f(x) = f(b)$. Зауважимо також, що остання властивість дозволяє на практиці здійснювати перевірку унімодальності достатньо гладких функцій, користуючись критеріями опуклості. Якщо функція опукла, то вона також унімодальна. Якщо опуклість не встановлена, то висновок про унімодальність у такий спосіб зробити неможливо.

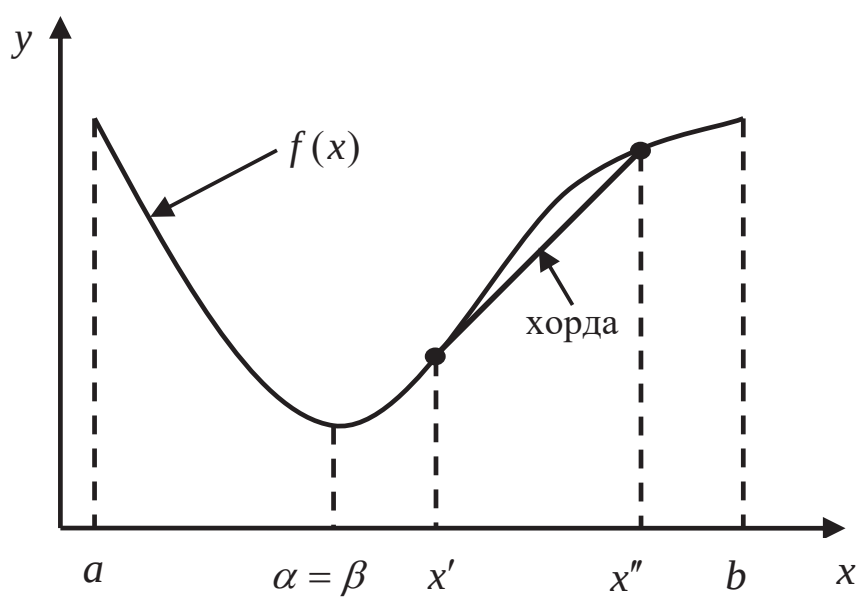


Рис. 1.4. Графік унімодальної, але не опуклої
функції $y = f(x)$

Розділ 2. Прямі числові методи

У вступі (пункт В.4) вже йшлося про те, що більшість задач оптимізації (мінімізації, максимізації) може бути розв'язана не точними, а наближеними методами, зокрема наближеними числовими методами [12, 17]. Це стосується також задач мінімізації функції однієї змінної.

Числові методи, які використовують тільки значення функції та не потребують обчислення її похідних, називаються *прямими методами мінімізації, або методами нульового порядку*. Особливість прямих методів ще й у тому, що цільова функція може бути не задана аналітично. Важливо лише, щоб можна було знаходити її значення у заданих точках. Нижче розглянемо деякі найбільш поширені на практиці прямі методи пошуку точки мінімуму унімодальної функції $f(x)$ на відрізку $[a, b]$.

2.1. Метод перебору (рівномірного пошуку)

Це один з найпростіших прямих методів мінімізації. Його суть полягає у тому, що відрізок $[a, b]$ розбивається на n рівних частин точками $x_i = a + i(b - a)/n$ ($i = \overline{0, n}$), а потім обчислюються значення $f(x)$ у точках x_i ($i = \overline{0, n}$), порівняння яких дозволяє визначити точку x_k ($0 \leq k \leq n$), у якій

$$f(x_k) = \min_{i=\overline{0, n}} f(x_i). \quad (2.1)$$

За наближений розв'язок вибирається $x^* \approx x_k$, $f^* \approx f(x_k)$. Похибка (точність) визначення точки мінімуму x^* методом перебору (рівномірного пошуку) не перевищує величини $\varepsilon(n) = (b - a)/n$.

Справді, якщо x_k – внутрішня точка відрізка $[a, b]$, тобто

$1 \leq k \leq n-1$, то із (2.1) з урахуванням властивості 3) унімодальної функції випливає, що якщо $f(x_{k-1}) \geq f(x_k)$, то $x^* \in [x_{k-1}, b]$, а якщо $f(x_k) \leq f(x_{k+1})$, то $x^* \in [a, x_{k+1}]$. Отже, $x^* \in [x_{k-1}, b] \cap [a, x_{k+1}] = [x_{k-1}, x_{k+1}]$. Довжина цього відрізка дорівнює $2(b-a)/n$. Оскільки точка x^* є його середньою точкою, то $|x_k - x^*| \leq (b-a)/n = \varepsilon(n)$, що й потрібно було встановити. Випадки $k=0$, $k=n$ доводяться аналогічно. Число відрізків n знаходиться з умови $\varepsilon(n) \leq \varepsilon$, де ε – задана точність пошуку точки мінімуму.

Підсумовуючи сказане вище, підкреслимо, що алгоритм методу перебору (рівномірного пошуку) зводиться до реалізації таких кроків.

Крок 1. Задання функції $f(x)$, відрізка $[a, b]$, на якому дана функція унімодальна та похибки $\varepsilon \geq \varepsilon(n)$;

Крок 2. Визначення числа n відрізків розбиття відрізка $[a, b]$ ($n \geq (b-a)/\varepsilon(n)$) та точок розбиття x_i ($i = \overline{0, n}$).

Крок 3. Знаходження у точках x_i ($i = \overline{0, n}$) значень функції $f(x)$ та уточнення x^* і $f(x^*)$.

Приклад 2.1. Методом перебору розв'язати задачу мінімізації функції $f(x) = \frac{x^2}{2} + e^{-x}$ на відрізку $[0, 1]$ з точністю $\varepsilon = 0,1$.

◀ Легко переконатись, що функція $f(x)$ унімодальна на $[0, 1]$. Число відрізків розбиття $n \geq \frac{1-0}{0,1} = 10$, тобто можна взяти $n = 10$. Точки $x_i = 0 + i(1-0)/10 = 0,1i$ ($i = \overline{0, 10}$) та значення $f(x_i)$ у цих точках зведені у табл. 2.1.

Таблиця 2.1

x_i	$f(x_i)$
0,0	1,00
0,1	0,910
0,2	0,839
0,3	0,786
0,4	0,750
0,5	0,732
0,6	0,729
0,7	0,742
0,8	0,769
0,9	0,812
1,0	0,868

З табл. 2.1 випливає, що $x^* \approx 0,6$, $f^* \approx 0,729$. ►

На закінчення даного підпункту зауважимо таке.

1. Метод перебору належить до класу так званих *пасивних* методів (алгоритмів), оскільки всі точки (значення аргумента функції) вибираються одночасно до початку обчислень. Якщо кожне наступне значення аргумента вибирається тоді, коли уже вибрані попередні значення, то такий метод належить до класу *послідовних* методів. Метод перебору має ряд модифікацій, спрямованих на зменшення кількості значень $f(x)$, які потрібно знаходити у процесі мінімізації.

2. У методі перебору точки x_i , у яких визначаються значення функції $f(x)$, вибираються наперед. Пошук точки мінімуму може бути більш ефективним, якщо скоротити число значень функції $f(x)$, використовуючи інформацію, яка міститься в уже знайдених значеннях функції. Таку інформацію дозволяє отримати властивість 3) унімодальної

функції. Справді, якщо $a < x_1 < x_2 < b$ (x_1, x_2 – пробні точки), то порівнявши значення $f(x_1)$ і $f(x_2)$, можна зменшити відрізок пошуку точки x^* , перейшовши до відрізка $[a, x_2]$, якщо $f(x_1) \leq f(x_2)$, або до відрізка $[x_1, b]$, якщо $f(x_1) > f(x_2)$. Цю процедуру можна повторювати до тих пір, коли довжина відрізка, що містить точку x^* , буде достатньо малою (такою, що не перевищує заданої величини похибки обчислень). Тоді можна вважати, що $x^* \approx \tilde{x}$, де $\tilde{x}$ – одна з точок, наприклад середина, «мінімального» відрізка. Методи мінімізації, які базуються на цьому принципі, називаються *методами виключення відрізків*.

Для того, щоб відносне зменшення довжини відрізка на кожній ітерації не залежало від того, яка з його частин виключається з подальшого розгляду, *пробні точки вибирають симетричними відносно середини відрізка*. Залежно від способу вибору пробних точок отримують різні методи виключення відрізків. Деякі з них розглянемо нижче.

2.2. Методи поділу відрізка пополам

Перший метод – метод дихотомії

Пробні точки x_1, x_2 ($a < x_1 < x_2 < b$) вибирають близькими до середини відрізка, тобто

$$x_1 = \frac{a+b}{2} - \frac{\delta}{2}, \quad x_2 = \frac{a+b}{2} + \frac{\delta}{2}, \quad (2.2)$$

де $\delta > 0$ – деяке мале число. Із (2.2) випливає, що

$$\tau = \frac{b-x_1}{b-a} = \frac{x_2-a}{b-a} \approx \frac{1}{2},$$

що й оправдовує назву методу. Звичайно, $\tau > \frac{1}{2}$, але вибір пробних точок забезпечує максимально можливе відносне

зменшення відрізка на кожній ітерації. За наближений розв'язок x^* береться середина останнього із знайдених відрізків $[a, b]$, для якого $(b - a)/2 \leq \varepsilon$ (ε – похибка пошуку точки мінімуму).

Надалі через $\varepsilon_n = (b - a)/2$ позначимо точність визначення точки мінімуму x^* на n -ій ітерації (після $(n - 1)$ -ої ітерації), яка власне починається з визначення ε_n (a, b – крайні точки відрізка – на n -ій ітерації).

Уточнимо алгоритм даного методу.

Крок 1. Визначення $\varepsilon_n = (b - a)/2$. Якщо $\varepsilon_n \leq \varepsilon$, то переходимо до кроку 4, інакше ($\varepsilon_n > \varepsilon$) – до кроку 2.

Крок 2. Знаходження (за формулами (2.2)) точок x_1, x_2 та значень функцій $f(x_1), f(x_2)$.

Крок 3. Порівняння $f(x_1)$ і $f(x_2)$. Якщо $f(x_1) \leq f(x_2)$, то вибирається відрізок $[a, x_2]$ (при цьому $b := x_2$, де « $:=$ » – операція присвоєння (перепозначення)), а якщо $f(x_1) > f(x_2)$, то вибирається відрізок $[x_1, b]$ ($a := x_1$). Перехід до кроку 1.

Крок 4. Визначення $x^* \approx \frac{a + b}{2}$, $f^* \approx f(x^*)$.

Зробимо кілька зауважень.

1. Якщо точність (похибка) пошуку точки мінімуму ε задана, то число δ вибирають на інтервалі $(0, 2\varepsilon)$, виходячи з таких міркувань:

а) при зменшенні δ досягається вища швидкість збіжності методу дихотомії;

б) при дуже малому δ порівняння значень $f(x)$ у точках x_1 і x_2 , які відрізняються на величину δ , може виявитись практично неможливим, тому вибір δ повинен

бути узгодженим з точністю (похибкою) визначення $f(x)$ та з кількістю правильних десяткових знаків при заданні значення аргумента x .

2. Номер ітерації, на якій визначається точка x^* з точністю ε , задовольняє нерівність

$$n \geq \log_2 \left[\frac{2(b-a-\delta)}{2\varepsilon-\delta} \right]. \quad (2.3)$$

◀ Доведемо нерівність (2.3). Нехай $[a, b]$ – початковий відрізок, а $\Delta_1 = (b-a)$ – довжина відрізка першої ітерації.

Тоді довжина відрізка другої ітерації $\Delta_2 = \left(\frac{b-a}{2} + \frac{\delta}{2} \right)$,

третьої ітерації – $\Delta_3 = \frac{\Delta_2}{2} + \frac{\delta}{2} = \frac{b-a}{4} + \delta \left(\frac{1}{4} + \frac{1}{2} \right)$, n -ої

ітерації – $\Delta_n = \frac{b-a}{2^{n-1}} + \delta \left(\frac{1}{2^{n-1}} + \dots + \frac{1}{2} \right) = \frac{b-a}{2^{n-1}} + \left(1 - \frac{1}{2^{n-1}} \right) \delta = 2\varepsilon_n$.

Із нерівності

$$\varepsilon_n = \frac{\Delta_n}{2} = \frac{b-a}{2^n} + \left(1 - \frac{1}{2^{n-1}} \right) \frac{\delta}{2} \leq \varepsilon \quad (2.4)$$

отримуємо нерівність (2.3). ▶

3. У випадку, коли величина δ достатньо мала, нею можна знехтувати у нерівностях (2.3), (2.4) і отримати співвідношення $\varepsilon \geq (b-a)/2^n \approx \varepsilon_n$. Оскільки кожна наступна ітерація починається з визначення $\varepsilon_n = (b-a)/2$, а точки a, b визначаються на попередній ітерації, то за перших $(n-1)$ ітерацій було зроблено $N = 2(n-1)$ обчислень значень функції $f(x)$ (на кожній ітерації обчислюється два значення). Це означає, що після $N = 2(n-1)$ обчислень $((n-1)$ ітерацій)

$$\varepsilon \geq \frac{b-a}{2^n} = \frac{b-a}{2^{\frac{N}{2}+1}} \approx \varepsilon_n = \varepsilon_{\frac{N}{2}+1} = \varepsilon(N). \quad (2.5)$$

Співвідношення (2.5) пов'язує задану точність обчислень ε з номером останньої ітерації n і кількістю N обчислень значень функції у методі дихотомії.

Приклад 2.2. Першим методом поділу відрізка пополам (методом дихотомії) розв'язати задачу мінімізації з прикладу 2.1.

◀ Виберемо $\delta = 0,02$ і переходимо до першої ітерації.

Крок 1. Визначаємо $\varepsilon_n = (b-a)/2$. Оскільки у даному випадку $\varepsilon_n = (1-0)/2 = 0,5 > 0,1 = \varepsilon$, то переходимо до кроку 2.

Крок 2. За формулами (2.2) знаходимо $x_1 = \frac{1}{2} - \frac{\delta}{2} = \frac{1}{2} - \frac{0,02}{2} = 0,49$; $x_2 = \frac{1}{2} + \frac{\delta}{2} = \frac{1}{2} + \frac{0,02}{2} = 0,51$, а потім значення $f(0,49) = 0,733$; $f(0,51) = 0,731$.

Крок 3. Оскільки $f(x_1) > f(x_2)$, то вибираємо відрізок $[x_1, b] = [x_1, 1] = [0,49; 1]$ і розпочинаємо другу ітерацію, повертаючись до кроку 1, де визначаємо $\varepsilon_n = (b-a)/2 = (1-0,49)/2 = 0,255 > 0,1 = \varepsilon$.

Очевидно, обчислення на цій ітерації треба продовжувати.

Результати повної реалізації алгоритму відображені у табл. 2.2 (k – номер ітерації, p – параметр методу).

Таблиця 2.2

$\begin{matrix} k \\ p \end{matrix}$	1	2	3	4
a	0	0,49	0,49	0,49
b	1	1	0,755	0,633
$\varepsilon_n = \frac{b-a}{2}$	0,5	0,255	0,133	0,07
x_1	0,49	0,735	0,623	Достигнута точність 0,07 < 0,1 -
x_2	0,51	0,755	0,633	
$f(x_1)$	0,733	0,750	0,730	
$f(x_2)$	0,731	0,755	0,731	
	$f(x_1) > f(x_2)$	$f(x_1) < f(x_2)$	$f(x_1) < f(x_2)$	

Аналіз табл. 2.2 показує, що на четвертій ітерації $x^* \approx \frac{a+b}{2} = \frac{0,49+0,633}{2} \approx 0,56$, $f^* \approx f(0,56) = 0,728$. ►

Другий метод поділу відрізка пополам

У цьому методі на кожній ітерації використовуються *три пробні точки* і забезпечується послідовне зменшення довжини відрізка, що містить x^* , рівно вдвічі. Суть методу полягає у тому, що відрізок $[a, b]$ ділиться пробними точками $x_i = a + \frac{b-a}{4}i$, $i=1,2,3$ на чотири рівні частини, а потім порівнюються значення $f(x_1)$ та $f(x_2)$. Якщо $f(x_1) \leq f(x_2)$, то відрізком пошуку точки x^* буде $[a, x_2]$. Якщо $f(x_1) > f(x_2)$, то при $f(x_2) \leq f(x_3)$ шуканим відрізком буде $[x_1, x_3]$, а в протилежному випадку – $[x_2, b]$. Очевидно, яким би не був новий відрізок, одна із трьох пробних точок попереднього відрізка буде його середньою точкою.

Для реалізації наступної ітерації на отриманому у

результаті попередньої ітерації відрізка потрібно уже обчислити значення функції $f(x)$ не більше, ніж у двох точках (або лише у точці x_1 , або у точках x_1 і x_3).

Алгоритм другого методу поділу відрізка пополам складається з перелічених нижче кроків.

Крок 1. Знаходження $x_2 = \frac{a+b}{2}$ та $f(x_2)$. Перехід до кроку 2.

Крок 2. Визначення $\varepsilon_n = (b-a)/2$. Якщо $\varepsilon_n > \varepsilon$, то здійснюється перехід до кроку 3. При $\varepsilon_n \leq \varepsilon$ пошук точки мінімуму завершується, причому $x^* \approx \frac{a+b}{2}$, $f^* \approx f(x^*)$.

Крок 3. Знаходження $x_1 = \frac{a+x_2}{2}$ та $f(x_1)$. Перехід до кроку 4.

Крок 4. Порівняння $f(x_1)$ та $f(x_2)$. При $f(x_1) \leq f(x_2)$ вибирається відрізок $[a, x_2]$ (при цьому $b := x_2$, $x_2 := x_1$, $f(x_2) := f(x_1)$) і здійснюється перехід до кроку 2. При $f(x_1) > f(x_2)$ знаходяться $x_3 = \frac{x_2+b}{2}$, $f(x_3)$ і здійснюється перехід до кроку 5.

Крок 5. Порівняння $f(x_2)$ і $f(x_3)$. При $f(x_2) \leq f(x_3)$ вибирається відрізок $[x_1, x_3]$ (при цьому $a := x_1$, $b := x_3$), а при $f(x_2) > f(x_3)$ – відрізок $[x_2, b]$ (при цьому $a := x_2$, $x_2 := x_3$, $f(x_2) := f(x_3)$) і здійснюється перехід до кроку 2.

Зауважимо, що у другому методі поділу відрізка пополам довжина відрізка, який містить точку x^* , на кожній ітерації дорівнює половині довжини всього відрізка, що відповідає даній ітерації. Інакше кажучи, якщо $[a, b]$ –

початковий відрізок, то довжина відрізка першої ітерації $\Delta_1 = b - a$, другої ітерації – $\Delta_2 = (b - a)/2$, ..., n -ої ітерації – $\Delta_n = (b - a)/2^{n-1}$. Це означає, що на n -ій ітерації (після $(n - 1)$ ітерацій) точність визначення точки мінімуму дорівнює $\left(\varepsilon_n = \frac{\Delta_n}{2} = \frac{b - a}{2^n} \right)$ (за наближений розв'язок

береться серединна точка відрізка, для якого половина довжини не перевищує ε). З умови $\varepsilon_n \leq \varepsilon$ одержимо умову на число ітерацій, що забезпечує задану точність знаходження точки мінімуму x^* :

$$n \geq \log_2 \left(\frac{b - a}{\varepsilon} \right). \quad (2.6)$$

Якщо врахувати, що відрізок n -ої ітерації є результатом перших $(n - 1)$ ітерацій, після яких може бути зроблено не більше $N = 3 + 2(n - 2) = 2n - 1$ обчислень значень функції $f(x)$ (на першій ітерації не більше трьох, а на всіх інших не більше двох), то матимемо рівності

$$\varepsilon_n = \frac{b - a}{2^n} = \frac{b - a}{2^{\frac{N+1}{2}}} = \varepsilon_{\frac{N+1}{2}} = \varepsilon(N). \quad (2.7)$$

На відміну від (2.6), співвідношення (2.7) зв'язують точність пошуку мінімуму як з числом ітерацій, так і з числом обчислень значень функції.

Приклад 2.3. Розв'язати задачу мінімізації з прикладу 2.1 другим методом поділу відрізка пополам.

◀ Перша ітерація починається з кроку 1.

Крок 1. Знаходимо $x_2 = (0 + 1)/2 = 0,5$; $f(x_2) = 0,732$.

Крок 2. Визначаємо $\varepsilon_n = (b - a)/2$. На першій ітерації $\varepsilon_n = 0,5 > 0,1 = \varepsilon$, тому переходимо до кроку 3.

Крок 3. Знаходимо $x_1 = (a + x_2)/2 = 0,25$ та $f(x_1) = 0,810$.

Переходимо до кроку 4.

Крок 4. Оскільки $f(x_1) > f(x_2)$, то знаходимо $x_3 = (x_2 + b)/2 = 0,75$ і $f(x_3) = 0,754$. Переходимо до кроку 5.

Крок 5. Враховуючи, що $f(x_2) < f(x_3)$, вибираємо відрізок $[x_1, x_3] = [0,25; 0,75]$ і розпочинаємо другу ітерацію із кроку 2.

Результати обчислень на першій та на інших ітераціях продемонстровані у табл. 2.3 (k – номер ітерації, p – параметр методу).

Таблиця 2.3

$p \backslash k$	1	2	3	4
a	0	0,25	0,5	0,5
b	1	0,75	0,75	0,625
$\varepsilon_n = \frac{b-a}{2}$	0,5	0,25	0,125	0,06
x_1	0,25	0,375	0,563	0,06 < 0,1 – ТОЧНІСТЬ ДОСЯГНУТА
x_2	0,5	0,5	0,625	
x_3	0,75	0,625	0,688	
$f(x_1)$	0,810	0,758	0,728	
$f(x_2)$	0,732	0,732	0,731	
$f(x_3)$	0,754	0,731	0,739	
	$f(x_1) > f(x_2),$ $f(x_2) < f(x_3)$	$f(x_1) > f(x_2),$ $f(x_2) > f(x_3)$	$f(x_1) < f(x_2),$ $f(x_2) < f(x_3)$	

Висновок, який впливає з табл. 2.3, такий: з точністю $\varepsilon = 0,1$ розв'язок знайдений на четвертій ітерації, причому $x^* \approx \frac{a+b}{2} = \frac{0,5+0,625}{2} \approx 0,56$, $f^* \approx f(0,56) = 0,728$.

Зауважимо, що обидва методи поділу відрізка пополам у даному випадку дали однаковий результат. ►

2.3. Метод золотого перерізу

Як і методи поділу відрізка пополам, *метод золотого перерізу* належить до методів виключення відрізків, однак спосіб вибору у ньому пробних точок x_1 і x_2 на відрізку $[a, b]$ такий, що одна з двох, симетричних відносно середини, точок стає пробною і на новому відрізку, одержаному після виключення частини попереднього відрізка. Це дозволяє на кожній ітерації, крім першої, визначати лише одне значення $f(x)$, оскільки друге значення уже знайдено на одній з попередніх ітерацій.

З'ясуємо, які пробні точки x_1 і x_2 мають таку властивість, скориставшись рис. 2.1.

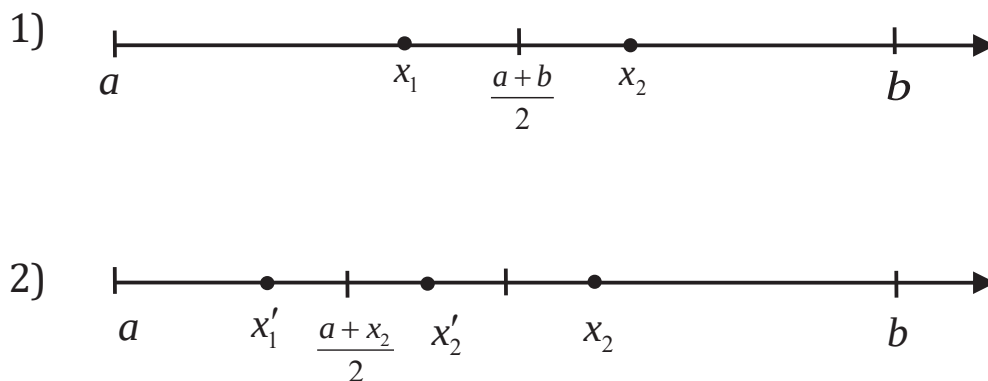


Рис. 2.1. Ілюстрація пробних точок у методі золотого перерізу

Для визначеності будемо вважати, що при зменшенні відрізка $[a, b]$ виключається права частина цього відрізка $[x_2, b]$. Точки x_1 і x_2 симетричні відносно серединної точки $(a + b)/2$ (рис. 2.1, 1)). Очевидно,

$$\begin{cases} x_2 = a + \frac{x_2 - a}{b - a}(b - a), \\ x_1 = a + \frac{x_1 - a}{b - a}(b - a) = a + \frac{b - x_2}{b - a}(b - a). \end{cases} \quad (2.8)$$

Після виключення відрізка $[x_2, b]$ пробна точка $x_1 \in [a, b]$ перейде у пробну точку $x'_2 \in [a, x_2]$. Симетричною до x'_2 відносно серединної точки $(a + x_2)/2$ буде точка x'_1 (рис. 2.1, 2)). Для того, щоб точки x_2 та x'_2 ділили відрізки $[a, b]$ та $[a, x_2]$ в одному і тому ж відношенні, має виконуватись рівність $(x'_2 - a = x_1 - a)$

$$\frac{x_2 - a}{b - a} = \frac{x_1 - a}{x_2 - a}. \quad (2.9)$$

Якщо $(x_2 - a)/(b - a) = \alpha$, то $(b - x_2)/(b - a) = 1 - \alpha$, тому із (2.8) випливає, що $x_2 - a = \alpha(b - a)$, $x_1 - a = (1 - \alpha)(b - a)$. Це означає, що рівність (2.9) набуде вигляду $\alpha = (1 - \alpha)/\alpha$ або $\alpha^2 = 1 - \alpha$. Додатний корінь цього рівняння $\alpha = \frac{\sqrt{5} - 1}{2} = 0,61803$. Оскільки $1 - \alpha = (3 - \sqrt{5})/2$, то співвідношення (2.8) конкретизуються так:

$$x_1 = a + \frac{3 - \sqrt{5}}{2}(b - a), \quad x_2 = a + \frac{\sqrt{5} - 1}{2}(b - a). \quad (2.10)$$

Відзначимо деякі особливості даного методу.

1. Точки x_1 і x_2 , які задаються формулами (2.10), ділять відрізок $[a, b]$ на дві нерівні частини у такий спосіб, що відношення довжини всього відрізка до довжини його більшої частини дорівнює відношенню довжини більшої частини до довжини меншої частини. Точки, які мають таку властивість, називаються *точками золотого перерізу* відрізка $[a, b]$. Саме з цим і пов'язана назва методу золотого перерізу.

2. Виключення відрізків із пробними точками (2.10) гарантує перехід однієї з цих точок на наступний відрізок, тому значення $f(x)$ у цій точці обчислювати уже не потрібно.

3. Оскільки $x_1 = a + b - x_2$ і $x_2 = a + b - x_1$ (це легко перевірити, скориставшись (2.10)), то на кожній ітерації методу золотого перерізу другу пробну точку можна знайти за допомогою пробної точки, яка перейшла з попередньої ітерації, не використовуючи формул (2.10).

4. За наближений розв'язок задачі мінімізації, розв'язаної методом золотого перерізу, можна взяти серединну точку останнього з отриманих відрізків $[a, b]$, тобто точку $x^* \approx (a + b)/2$, якщо $(b - a)/2 \leq \varepsilon$ (ε – точність (похибка) визначення точки мінімуму).

Оскільки довжина відрізка, що містить точку мінімуму x^* після $(n - 1)$ ітерацій, тобто на n -ій ітерації, визначається величиною $\Delta_n = \alpha^{n-1}(b - a)$, то точність визначення точки x^* на n -ій ітерації

$$\varepsilon_n = \frac{\Delta_n}{2} = \frac{1}{2} \left(\frac{\sqrt{5} - 1}{2} \right)^{n-1} (b - a), \quad (2.11)$$

а умовою завершення пошуку точки x^* з точністю ε є умова $\varepsilon_n \leq \varepsilon$.

Уточнимо алгоритм методу золотого перерізу.

Крок 1. Знаходження (за формулами (2.10)) пробних точок x_1, x_2 та значень $f(x_1), f(x_2)$. Перехід до кроку 2.

Крок 2. Визначення $\varepsilon_n = (b - a)/2$, де a, b – крайні точки відрізка поточної ітерації (при застосуванні початкових a, b використовується формула $\varepsilon_n = \alpha^{n-1}(b - a)/2$, де n – номер ітерації, а $\alpha = (\sqrt{5} - 1)/2$). Якщо $\varepsilon_n \leq \varepsilon$, то здійснюється перехід

до кроку 4, а якщо $\varepsilon_n > \varepsilon$ – до кроку 3.

Крок 3. Перехід до нового відрізка і нових пробних точок. Якщо $f(x_1) \leq f(x_2)$, то вибирається відрізок $[a, x_2]$ (при цьому $b := x_2$, $x_2 := x_1$, $f(x_2) := f(x_1)$) та обчислюються нові $x_1 = b - \alpha(b - a)$, $f(x_1)$. Якщо $f(x_1) > f(x_2)$, то вибирається відрізок $[x_1, b]$ (при цьому $a := x_1$, $x_1 := x_2$, $f(x_1) := f(x_2)$) та обчислюються нові $x_2 = a + \alpha(b - a)$, $f(x_2)$. Перехід до кроку 2.

Крок 4. Завершення пошуку мінімуму: $x^* \approx (a + b)/2$, $f^* \approx f(x^*)$.

Враховуючи (2.11), можна конкретизувати число ітерацій n , необхідне для забезпечення заданої точності ε , розв'язавши відносно n нерівність $\varepsilon_n \leq \varepsilon$:

$$n \geq 1 + \ln\left(\frac{2\varepsilon}{b-a}\right) / \ln \alpha \approx 1 - 2,1 \ln\left(\frac{2\varepsilon}{b-a}\right). \quad (2.12)$$

Після $(n-1)$ ітерацій методу золотого перерізу здійснюється $N = n$ обчислень значень функції $f(x)$, тому

$$\begin{aligned} \varepsilon_n &= \frac{1}{2} \left(\frac{\sqrt{5}-1}{2} \right)^{n-1} (b-a) = \\ &= \frac{1}{2} \left(\frac{\sqrt{5}-1}{2} \right)^{N-1} (b-a) = \varepsilon_N = \varepsilon(N). \end{aligned} \quad (2.13)$$

Отже, співвідношення (2.11), (2.12) пов'язані тільки з числом ітерацій, а співвідношення (2.13) – як з числом ітерацій, так і з числом обчислень значень функції.

Приклад 2.4. Розв'язати наведену у прикладі 2.1 задачу мінімізації методом золотого перерізу.

◀ Опишемо, як і раніше, тільки першу ітерацію.

Крок 1. Знаходимо $x_1 = 0 + \frac{3-\sqrt{5}}{2} \cdot 1 = 0,38$; $x_2 = 0 +$

$$+\frac{\sqrt{5}-1}{2} \cdot 1 = 0,62; f(x_1) = 0,756; f(x_2) = 0,730.$$

Крок 2. Оскільки $\varepsilon_n = (b-a)/2$, то $\varepsilon_1 = (1-0)/2 = 0,5 > 0,1 = \varepsilon$. Отже, переходимо до кроку 3.

Крок 3. Очевидно, $f(x_1) > f(x_2)$, тому вибираємо відрізок $[x_1, b] = [0,38; 1]$. На цьому відрізку нам уже відомі $b = 1$; $a = 0,38$ (попереднє значення x_1), $x_1 = 0,62$ (попереднє значення x_2), $f(x_1) = 0,730$ (попереднє значення $f(x_2)$). Обчислимо для

$$[a, b] = [0,38; 1] \quad \text{значення} \quad x_2 = a + \alpha(b-a) = 0,38 + \frac{\sqrt{5}-1}{2} \times (1-0,38) = 0,76 \quad \text{та} \quad f(x_2) = f(0,76) = 0,756.$$

Перейшовши до кроку 2, починаємо другу ітерацію. Оскільки $\varepsilon_2 = (1-0,38)/2 = \alpha(1-0)/2 = \frac{\alpha}{2} = 0,31 > 0,1 = \varepsilon$, то переходимо

до кроку 3 і продовжуємо обчислення на другій ітерації.

Остаточні результати обчислень зведені у табл. 2.4 (k – номер ітерації, p – параметр методу).

Таблиця 2.4

$p \backslash k$	1	2	3	4	5
a	0	0,38	0,38	0,38	0,47
b	1	1	0,76	0,62	0,62
$\varepsilon_n = \frac{b-a}{2}$	0,5	0,31	0,19	0,12	0,08
x_1	0,38	0,62	0,52	0,47	0,08 < 0,1 – точність досягнута
x_2	0,62	0,76	0,62	0,52	
$f(x_1)$	0,756	0,730	0,730	0,735	
$f(x_2)$	0,730	0,756	0,730	0,730	
	$f(x_1) > f(x_2)$	$f(x_1) < f(x_2)$	$f(x_1) \leq f(x_2)$	$f(x_1) > f(x_2)$	

Отже, із табл. 2.4 випливає, що розв'язок знайдений на п'ятій ітерації. При цьому $x^* \approx \frac{a+b}{2} = \frac{0,47+0,62}{2} = 0,55$, $f(x^*) \approx 0,728$. ►

2.4. Метод Фібоначчі

Метод Фібоначчі отримав свою назву від так званих чисел Фібоначчі, тобто чисел, які визначаються співвідношеннями

$$F_1 = F_2 = 1, \quad F_{n+2} = F_{n+1} + F_n, \quad n = 1, 2, \dots$$

До чисел Фібоначчі належать, наприклад числа $F_3 = 2$, $F_4 = 3$, $F_5 = 5$, $F_6 = 8$, $F_7 = 13$, $F_{10} = 55$, $F_{15} = 610$, $F_{20} = 6765$, $F_{25} = 75025$.

Застосовуючи індукцію, легко встановити, що n – є число Фібоначчі знаходиться за формулою

$$F_n = \frac{1}{\sqrt{5}} \left[\left(\frac{1+\sqrt{5}}{2} \right)^n - \left(\frac{1-\sqrt{5}}{2} \right)^n \right], \quad n = 1, 2, \dots \quad (2.14)$$

Із формули (2.14) випливає, що $F_n \approx \frac{1}{\sqrt{5}} \left(\frac{1+\sqrt{5}}{2} \right)^n$ при $n \rightarrow \infty$, тобто зі збільшенням n числа Фібоначчі дуже швидко зростають.

У методі Фібоначчі числа Фібоначчі використовуються для задання пробних точок на відрізку мінімізації (метод також належить до методів виключення відрізків). Припустимо, що для знаходження розв'язку задачі мінімізації із заданою точністю ε потрібно n ітерацій. Тоді на першій ітерації здійснюють два обчислення значень функції $f(x)$ у пробних точках $x_1^{(1)} = a + \frac{F_n}{F_{n+2}}(b-a)$,

$$x_2^{(1)} = a + b - x_1 = a + \frac{F_{n+1}}{F_{n+2}}(b - a), \quad \text{які розміщені симетрично}$$

відносно середини відрізка. Якщо $f(x_1^{(1)}) \leq f(x_2^{(1)})$, то відрізком локалізації мінімуму буде відрізок $[a, x_2^{(1)}]$, а якщо $f(x_1^{(1)}) > f(x_2^{(1)})$ – відрізок $[x_1^{(1)}, b]$. Починаючи з другої ітерації, на кожній з ітерацій вибирається лише одна точка обчислень, яка є симетричною до тієї з точок, що попала у локалізований на попередній ітерації відрізок. Якщо $[a_k, b_k]$ – відрізок мінімізації функції на k -ій ітерації ($a_1 = a, b_1 = b$), то пробні точки цього відрізка знаходяться за формулами ($k = \overline{1, n}$)

$$x_1^{(k)} = a_k + \frac{F_{n+1-k}}{F_{n+3-k}}(b_k - a_k) = a_k + \frac{F_{n+1-k}}{F_{n+2}}(b - a), \quad (2.15)$$

$$x_2^{(k)} = a_k + \frac{F_{n+2-k}}{F_{n+3-k}}(b_k - a_k) = a_k + \frac{F_{n+2-k}}{F_{n+2}}(b - a). \quad (2.16)$$

Довжина відрізка локалізації мінімуму на k -тій ітерації $\Delta_k = b_k - a_k$ залежить від точок (2.15), (2.16), вибраних на попередній ітерації ($k-1$). Якщо $[a_k, b_k] = [a_{k-1}, x_2^{(k-1)}]$ (його довжина дорівнює довжині відрізка $[a_k, b_k] = [x_1^{(k-1)}, b_{k-1}]$), то із (2.16) випливає, що

$$\begin{aligned} \Delta_k &= x_2^{(k-1)} - a_{k-1} = \frac{F_{n+2-(k-1)}}{F_{n+2}}(b - a) = \\ &= \frac{F_{n+3-k}}{F_{n+2}}(b - a), \quad (k = \overline{1, n}). \end{aligned} \quad (2.17)$$

При $k = n$ (тобто на n -ій ітерації) процес пошуку мінімуму закінчується. Із (2.15), (2.16) випливає, що

$$x_1^{(n)} = a_n + \frac{F_1}{F_{n+2}} = x_2^{(n)} = a_n + \frac{F_2}{F_{n+2}}, \quad \text{тобто точки збігаються і ділять}$$

відрізок $[a_n, b_n]$ пополам. При цьому $x^* \approx x_1^{(n)} = x_2^{(n)} = \frac{a_n + b_n}{2}$,

$f^* \approx f(x_1^{(n)}) = f(x_2^{(n)})$. Довжина відрізка $[a_n, b_n]$ знаходиться

за формулою (2.17), тобто $\Delta_n = b_n - a_n = \frac{F_3}{F_{n+2}}(b - a) = \frac{2(b - a)}{F_{n+2}}$.

Це означає, що гарантована точність

$$\varepsilon_n = \frac{\Delta_n}{2} = \frac{b - a}{F_{n+2}} \leq \varepsilon. \quad (2.18)$$

Число ітерації n (а також число обчислень значень функції $N = n + 1$), що забезпечує точність ε , визначається із (2.18).

Опишемо алгоритм методу Фібоначчі.

Крок 1. Знаходження $\varepsilon_1 = (b_1 - a_1)/2 = (b - a)/2$. При $\varepsilon_1 \leq \varepsilon$ здійснюється перехід до кроку 5 ($n = 1$), а при $\varepsilon_1 > \varepsilon$ – до кроку 2.

Крок 2. Визначення числа ітерацій n , тобто послідовності чисел Фібоначчі до числа F_{n+2} включно, для якого $F_{n+2} \geq (b - a)/\varepsilon$.

Крок 3. Знаходження при $k = 1$ перших пробних точок $x_1^{(1)}, x_2^{(1)}$ (за формулами (2.15), (2.16)) і значень $f(x_1^{(1)}), f(x_2^{(1)})$.

Крок 4. Побудова послідовності пробних точок $\{x_1^{(k)}, x_2^{(k)}\}$ при $k = \overline{2, n}$. При цьому, якщо $f(x_1^{(k-1)}) \leq f(x_2^{(k-1)})$, то відрізком локалізації мінімуму вибирається відрізок $[a_k, b_k] = [a_{k-1}, x_2^{(k-1)}]$, для якого $x_2^{(k)} := x_1^{(k-1)}$, $f(x_2^{(k)}) := f(x_1^{(k-1)})$, за формулою (2.15) знаходиться точка $x_1^{(k)}$ і обчислюється значення $f(x_1^{(k)})$. Якщо $f(x_1^{(k-1)}) > f(x_2^{(k-1)})$, то відрізком

локалізації мінімуму буде відрізок $[a_k, b_k] = [x_1^{(k-1)}, b_{k-1}]$, для якого $x_1^{(k)} := x_2^{(k-1)}$, $f(x_1^{(k)}) := f(x_2^{(k-1)})$, за формулою (2.16) знаходиться $x_2^{(k)}$ і обчислюється значення $f(x_2^{(k)})$.

Крок 5. Закінчення пошуку мінімуму: $x^* \approx \frac{a_n + b_n}{2}$ (при

$$n > 1 \quad \frac{a_n + b_n}{2} = x_1^{(n)} = x_2^{(n)}), \quad f^* \approx f\left(\frac{a_n + b_n}{2}\right).$$

Приклад 2.5. Розв'язати задачу із прикладу 2.1 методом Фібоначчі.

◀ *Крок 1.* $\varepsilon_1 = (1 - 0)/2 = 0,5 > 0,1 = \varepsilon$. Перехід до кроку 2.

Крок 2. Оскільки $(b - a)/\varepsilon = 1/0,1 = 10$, то $F_{n+2} \geq 10$ ($F_1 = F_2 = 1, F_3 = 2, F_4 = 3, F_5 = 5, F_6 = 8, F_7 = 13$) і $(n + 2) = 7$ або $n = 5$.

Крок 3. $x_1^{(1)} = F_5/F_7 = 0,38$; $x_2^{(1)} = F_6/F_7 = 0,62$; $f(x_1^{(1)}) = 0,755$; $f(x_2^{(1)}) = 0,730$.

Усі результати обчислень відображені у табл. 2.5 (k – номер ітерації, p – параметр методу).

Дані табл. 2.5 підтвердили гарантовану точність методу ($0,08 < 0,1 = \varepsilon$) на п'ятій ітерації ($n = 5$). При цьому $x^* \approx x_1^{(5)} = x_2^{(5)} = (a_5 + b_5)/2 = 0,539$, а $f^* \approx 0,729$. ►

Таблиця 2.5

$\begin{matrix} p \backslash k \\ \end{matrix}$	1	2	3	4	5
a_k	0	0,385	0,385	0,385	0,462
b_k	1	1	0,770	0,615	0,615
$\varepsilon_k = \frac{b_k - a_k}{2}$	0,5	0,31	0,19	0,12	0,08
$x_1^{(k)}$	0,385	0,615	0,539	0,462	0,539
$x_2^{(k)}$	0,615	0,770	0,615	0,539	0,539
$f_1^{(k)} = f(x_1^{(k)})$	0,755	0,730	0,729	0,737	0,729
$f_2^{(k)} = f(x_2^{(k)})$	0,730	0,759	0,730	0,729	0,729
	$f_1^{(k)} > f_2^{(k)}$	$f_1^{(k)} < f_2^{(k)}$	$f_1^{(k)} < f_2^{(k)}$	$f_1^{(k)} > f_2^{(k)}$	$f_1^{(k)} = f_2^{(k)}$

Порівняння методів виключення відрізків і перебору. При порівнянні прямих методів мінімізації здебільшого враховують число N обчислень значень функції $f(x)$, яке гарантує задану точність ε пошуку точки мінімуму x^* (у багатьох практичних ситуаціях обчислення значень цільової функції потребує чималих витрат). Найбільш ефективним вважається той метод, у якому число N найменше (методи золотого перерізу та Фібоначчі).

З іншого боку, ефективність методу [12] можна також оцінити за допомогою точності ε_N знаходження точки мінімуму x^* , оскільки дана точність залежить від числа N . Аналізуючи величини ε_N для методів перебору і виключення відрізків (для методу перебору $\varepsilon_n = \frac{b-a}{n} = \frac{b-a}{N-1} = \varepsilon(N)$, а для

першого та другого методів поділу відрізка пополам і методу золотого перерізу для визначення ε_N використовуються формули (2.5), (2.7), (2.13)), приходимо до висновку, що найбільш ефективними з цих методів є методи золотого перерізу, а найменш ефективним – метод перебору. Встановлено також, що при великих n (а отже, і N) метод Фібоначчі більш ефективний від методу золотого перерізу, хоча обидва методи практично починаються в одній точці.

2.5. Метод квадратичної інтерполяції

Специфіка методів виключення відрізків (першого та другого методів поділу відрізка пополам, методу золотого перерізу) полягає у тому, що всі вони базуються на порівнянні значень функції у двох точках з метою встановлення знаку їх різниці. Однак є група методів, які враховують інформацію про відносні зміни значень функції $f(x)$ у пробних точках. Це *методи поліноміальної апроксимації*. Основна ідея цих методів полягає у тому, що для функції $f(x)$ будується апроксимуючий многочлен, точка мінімуму якого слугує наближенням до x^* . При цьому, крім унімодальності, передбачається також, що функція $f(x)$ достатньо гладка, принаймні неперервна. Ця додаткова умова потрібна для того, щоб виконувалась відома з математичного аналізу теорема Вейєштраса, згідно з якою *неперервна на відрізку функція може бути з будь-якою точністю апроксимована на цьому відрізку деяким поліномом*.

Точність апроксимації звичайно залежить від порядку (степеня) полінома. Оскільки зростання порядку полінома приводить до ускладнення обчислювальних процедур, то на практиці використовуються апроксимуючі поліноми не вище третього порядку. До найбільш простих і поширених методів

поліноміальної апроксимації належить метод квадратичної інтерполяції (метод парабол), на кожній ітерації якого будується поліном другого порядку (квадратичний тричлен), графік якого (парабола) проходить через три вибрані точки графіка функції $f(x)$ (рис. 2.2).

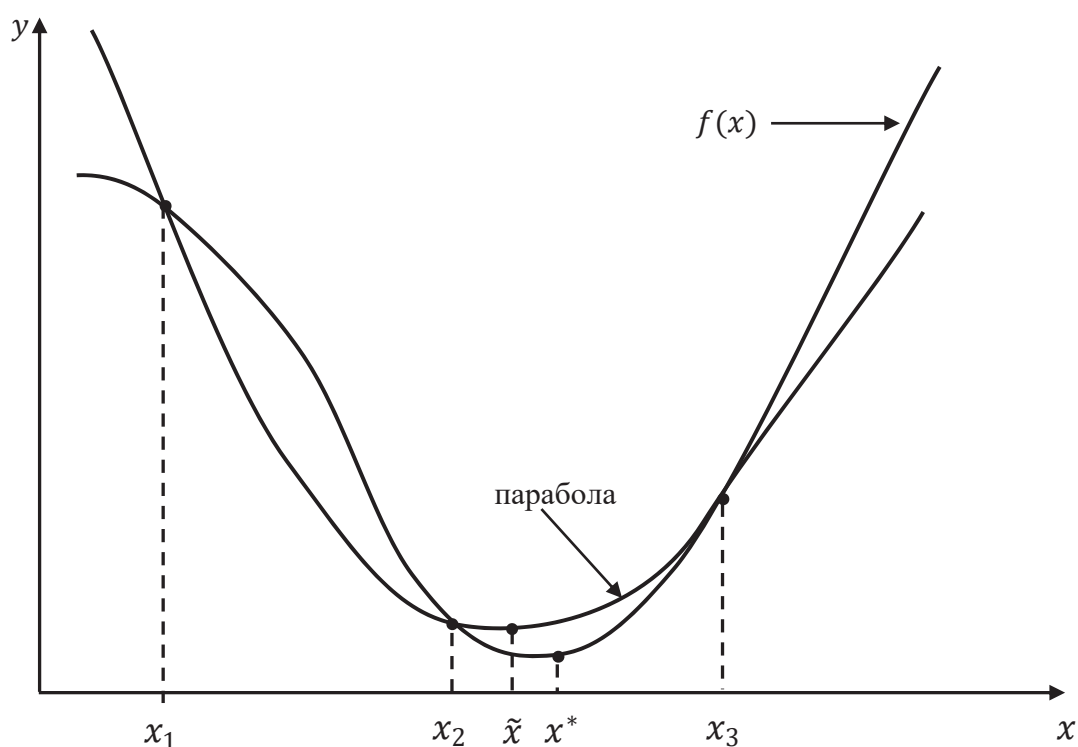


Рис. 2.2. Ілюстрація до методу квадратичної інтерполяції

Обґрунтуємо метод квадратичної інтерполяції. Нехай $f(x)$ унімодальна на $[a, b]$ функція (рис. 2.2), яка досягає свого мінімального значення у внутрішній точці $x^* \in [a, b]$, а x_1, x_2, x_3 , – три точки відрізка $[a, b]$, для яких

$$x_1 < x_2 < x_3, \quad f(x_1) \geq f(x_2) \leq f(x_3). \quad (2.19)$$

Із (2.19) випливає, що $x^* \in [x_1, x_3]$. Зауважимо, що на практиці для визначення точок, які задовольняють нерівність

(2.19), як правило, потрібно здійснити кілька проб. Будемо також вважати, що хоча б одна з нерівностей (2.19) для значень функції $f(x)$ є строгою, інакше при $f(x_1) = f(x_2) = f(x_3)$ пошук точки x^* завершується, оскільки з унімодальності функції $f(x)$ випливає, що вона досягає мінімального значення у кожній точці відрізка $[x_1, x_3]$. При зроблених припущеннях побудуємо квадратний тричлен $q(x) = a_0 + a_1(x - x_1) + a_2(x - x_1)(x - x_2)$, графік якого проходить через точки $(x_1, f(x_1))$, $(x_2, f(x_2))$, $(x_3, f(x_3))$ графіка функції $f(x)$. Із (2.19) можна дійти висновку, що парабола буде спрямована вітками вгору, а точка її мінімуму $\tilde{x} \in [x_1, x_3]$. Розв'язавши систему рівнянь $q(x_1) = f(x_1) = f_1$, $q(x_2) = f(x_2) = f_2$, $q(x_3) = f(x_3) = f_3$, знайдемо коефіцієнти a_0, a_1, a_2 :

$$a_0 = f_1, \quad a_1 = \frac{f_2 - f_1}{x_2 - x_1}, \quad a_2 = \frac{1}{x_3 - x_2} \left(\frac{f_3 - f_1}{x_3 - x_1} - \frac{f_2 - f_1}{x_2 - x_1} \right).$$

З умови $q'(x) = 0$ знайдемо точку мінімуму $\tilde{x}$ квадратного тричлена $q(x)$:

$$\begin{aligned} \tilde{x} &= \frac{1}{2} \left(x_1 + x_2 - \frac{a_1}{a_2} \right) = \\ &= \frac{1}{2} \left[x_1 + x_2 - \frac{(f_2 - f_1)(x_3 - x_2)}{x_2 - x_1} \right] / \left(\frac{f_3 - f_1}{x_3 - x_1} - \frac{f_2 - f_1}{x_2 - x_1} \right) = \\ &= \frac{1}{2} \cdot \frac{(x_2^2 - x_3^2)f_1 + (x_3^2 - x_1^2)f_2 + (x_1^2 - x_2^2)f_3}{(x_2 - x_3)f_1 + (x_3 - x_1)f_2 + (x_1 - x_2)f_3}. \end{aligned} \quad (2.20)$$

Точка (2.20) є одним з наближень до точки x^* . Наступне наближення – результат застосування методу парабол до нових трьох точок, які задовольняють умови (2.19). Їх можна вибирати за допомогою методу виключення відрізків з

попередніх точок $x_1, x_2, x_3, \tilde{x}$, перейшовши до нового відрізка $[x_1, x_3]$, який містить точку x^* , і використавши пробні точки x_2 і $\tilde{x}$ (якщо виявилось, що $x_2 = \tilde{x}$, то, крім x_2 , другою пробною точкою може бути довільна точка інтервала (x_1, x_3) , наприклад $(x_1 + x_2)/2$ чи $(x_2 + x_3)/2$). Початкова і кінцева точки нового відрізка, а також пробна точка, що попала на нього, утворюють нову трійку точок, які задовольняють умови (2.19).

Зауважимо також, що на кожній ітерації, крім першої, визначається тільки одне нове значення $f(x)$. Процес знаходження точки мінімуму x^* завершується, якщо різниця Δ значень $\tilde{x}$, знайдених на даній та попередній ітераціях, буде близькою до нуля, тобто якщо $|\Delta| \leq \varepsilon$, де ε – задана точність пошуку x^* .

Алгоритм методу квадратичної інтерполяції (методу парабол) конкретизується такими кроками.

Крок 1. Вибір точок x_1, x_2, x_3 , які задовольняють умови (2.19). Якщо $f(x_1) = f(x_2) = f(x_3)$, то пошук точки мінімуму x^* завершений (нею буде будь-яка точка відрізка $[x_1, x_3]$). У протилежному випадку переходять до кроку 2.

Крок 2. Знаходження $\tilde{x}$ за формулою (2.20). На першій ітерації переходять до кроку 4, на інших – до кроку 3.

Крок 3. Перевірка на завершення процесу пошуку точки мінімуму. Знаходиться різниця Δ значень $\tilde{x}$, знайдених на поточній та попередній ітераціях. Якщо $|\Delta| \leq \varepsilon$, то пошук завершений, причому $x^* \approx \tilde{x}$, $f^* \approx f(\tilde{x})$, а якщо $|\Delta| > \varepsilon$ – переходять до кроку 4.

Крок 4. Обчислення значення $f(\tilde{x})$ і перехід до кроку 5.

Крок 5. Визначення нової трійки чисел x_1, x_2, x_3 і

присвоєння $f(x_1), f(x_2), f(x_3)$ відповідних значень функції $f(x)$, знайдених раніше. Перехід до кроку 2.

Приклад 2.6. Розв'язати методом квадратичної інтерполяції (методом парабол) задачу мінімізації функції

$$f(x) = \frac{x^2}{2} + e^{-x} \text{ на відрізку } [0,1] \text{ з точністю } |\Delta| \leq \varepsilon = 0,01.$$

◀ *Ітерація 1*

Крок 1. Виберемо точки $x_1 = 0,25$; $x_2 = 0,5$; $x_3 = 0,75$. Тоді $f(x_1) = 0,8101$; $f(x_2) = 0,7315$; $f(x_3) = 0,7536$, тобто умови (2.19) виконуються. Переходимо до кроку 2.

Крок 2. За формулою (2.20) знаходимо $\tilde{x} = 0,5701$. Переходимо до кроку 4.

Крок 4. Обчислюємо $f(\tilde{x}) = 0,7280$ і переходимо до кроку 5.

Крок 5. Оскільки $x_1 < x_2 < \tilde{x} < x_3$, $f(x_3) > f(\tilde{x})$, $f(x_2) > f(\tilde{x})$, тобто $x^* \in [x_2, x_3] = [0,5; 0,75]$, то за нову трійку x_1, x_2, x_3 для другої ітерації візьмемо $x_1 = 0,5$ (дорівнює значенню x_2 на першій ітерації); $x_2 = \tilde{x} = 0,5701$; $x_3 = 0,75$. Значення функції у цих точках уже відомі ($f(x_1) = 0,7315$; $f(x_2) = 0,7280$; $f(x_3) = 0,7536$). Переходимо до другої ітерації (до кроку 2).

Ітерація 2

Крок 2. Визначаємо за формулою (2.20) $\tilde{x} = 0,5675$. Переходимо до кроку 3.

Крок 3. Оскільки

$\Delta = |0,5675 - 0,5701| = 0,0026 < 0,01 = \varepsilon$, то процес пошуку точки мінімуму завершений і $x^* \approx \tilde{x} \approx 0,568$, $f^* \approx f(0,568) \approx 0,728$. У результаті п'яти обчислень значень функції $f(x)$ точка x^* була знайдена з досить високою точністю. ►

Розділ 3. Методи мінімізації першого та другого порядків

До методів мінімізації першого та другого порядку належать методи, які використовують не тільки інформацію про значення самої функції, але й інформацію про значення її першої та другої похідної. Основна ідея цих методів ґрунтується на тому, що для опуклої диференційовної функції $f(x)$ (а значить, і унімодальної) умова $f'(x^*) = 0$ є необхідною й достатньою умовою глобального мінімуму у внутрішній точці $x^* \in [a, b]$, процес пошуку якої завершується, якщо $|f'(x)| \leq \varepsilon$ (ε – задана точність знаходження точки мінімуму).

3.1. Метод середньої точки

Нехай $f(x)$ – опукла диференційовна на $[a, b]$ функція, а $\bar{x} = (a + b)/2$. Тоді можливі ситуації:

- 1) $f'(\bar{x}) < 0$, тобто у точці $\bar{x}$ функція монотонно спадає, тому шукана точка мінімуму $x^* > \bar{x}$ і $x^* \in [\bar{x}, b]$;
- 2) $f'(\bar{x}) = 0$, тобто $x^* = \bar{x}$;
- 3) $f'(\bar{x}) > 0$, тобто у точці $\bar{x}$ функція монотонно зростає, тому $x^* < \bar{x}$ і $x^* \in [a, \bar{x}]$.

Отже, у випадку мінімізації опуклої диференційовної функції для виключення відрізка, що не містить точку мінімуму, на кожній ітерації потрібно тільки одне обчислення $f'(x)$. При цьому довжина відрізка пошуку точки x^* зменшується удвічі.

Вищезазначена інформація складає основу алгоритму методу середньої точки.

Крок 1. Знаходження $\bar{x} = (a + b)/2$ і $f'(\bar{x})$.

Крок 2. Перевірка на завершення пошуку точки мінімуму. Якщо $|f'(\bar{x})| \leq \varepsilon$, то задача мінімізації функції розв'язана і $x^* \approx \bar{x}$, $f^* \approx f(\bar{x})$. Якщо $|f'(\bar{x})| > \varepsilon$, то здійснюється перехід до кроку 3.

Крок 3. Вибір наступного відрізка мінімізації. При $f'(\bar{x}) < 0$ таким відрізком буде $[\bar{x}, b]$ ($a := \bar{x}$), а при $f'(\bar{x}) > 0$ – $[a, \bar{x}]$ ($b := \bar{x}$). Перехід до кроку 1.

Приклад 3.1. Методом середньої точки знайти точку мінімуму функції $f(x) = \frac{x^2}{2} + e^{-x}$ на відрізку $[0, 1]$ з точністю $|f'(\bar{x})| \leq 0,02 = \varepsilon$.

◀ *Ітерація 1*

Крок 1. Знаходимо $\bar{x} = 0,5$, $f'(x) = x - \frac{1}{e^x}$, $f'(\bar{x}) = -0,107$.

Крок 2. $|f'(\bar{x})| = 0,107 > 0,02$, тому переходимо до кроку 3.

Крок 3. Оскільки $f'(\bar{x}) < 0$, то новим відрізком буде $[\bar{x}, 1] = [0,5; 1] \equiv [a, b]$. Переходимо до нової ітерації, починаючи з кроку 1.

Результати обчислень на всіх ітераціях ілюструє табл. 3.1 (k – номер ітерації, p – параметр методу).

Таблиця 3.1

$k \backslash p$	a	b	$\bar{x}$	$f'(\bar{x})$	Знак $f'(\bar{x})$
1	0	1	0,5	-0,107	–
2	0,5	1	0,75	0,278	+
3	0,5	0,75	0,625	0,090	+
4	0,5	0,625	0,563	-0,006	–

На четвертій ітерації (табл. 3.1) точність пошуку точки мінімуму досягнута, оскільки $|f'(\bar{x})| < 0,02 = \varepsilon$.

Отже, $x^* \approx 0,563$, $f^* \approx 0,728$. ►

3.2. Метод хорд

Як і раніше, будемо вважати функцію $f(x)$ опуклою і диференційовною на $[a,b]$, причому такою, що $f'(a)f'(b) < 0$ (на кінцях відрізка похідна має різні знаки). Це означає, що на інтервалі (a,b) існує точка x^* , у якій $f'(x^*) = 0$, тому пошук точки мінімуму у даному разі зводиться до розв'язування рівняння

$$f'(x) = 0, \quad x \in (a,b). \quad (3.1)$$

Будь-який наближений метод розв'язування рівняння (3.1) у цій ситуації одночасно є методом мінімізації функції $f(x)$ на відрізку $[a,b]$. Один із таких – метод хорд, у якому на кожному відрізку визначається абсциса $\tilde{x}$ точки перетину хорди графіка функції $f'(x)$, що проходять через точки $(a, f'(a))$, $(b, f'(b))$, з віссю Ox (рис. 3.1).

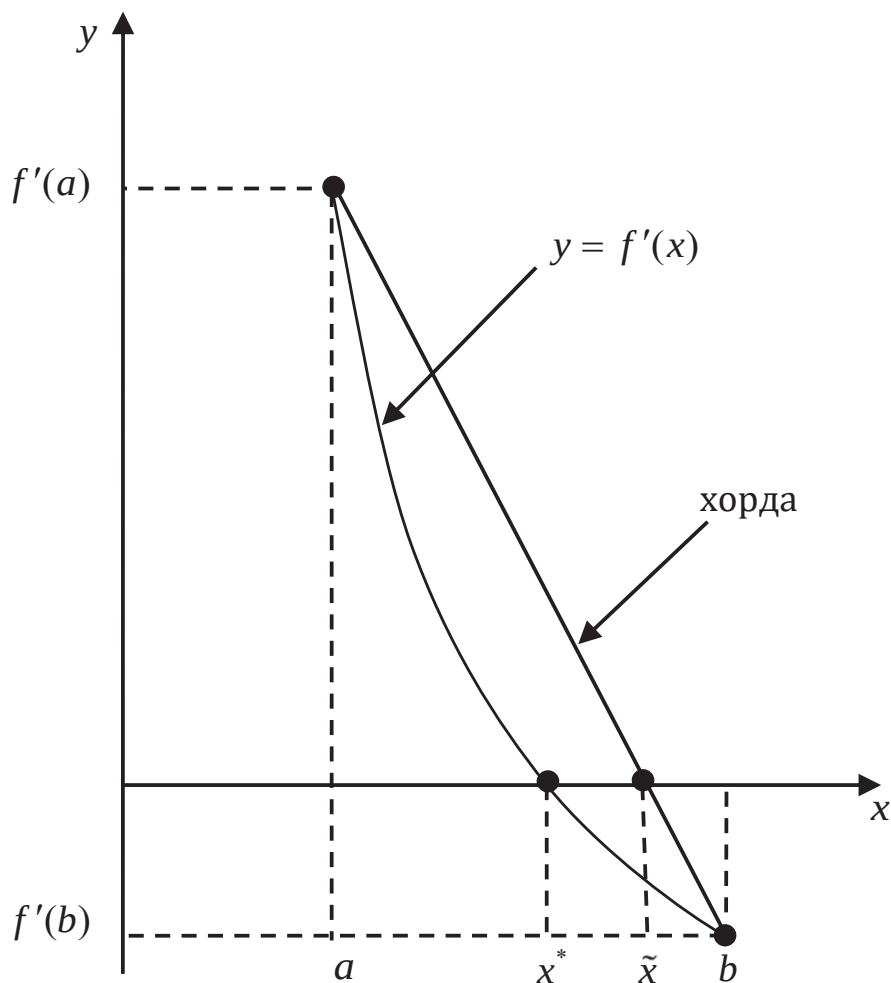


Рис. 3.1. Виключення відрізків методом хорд

Формула для знаходження точки $\tilde{x}$ має вигляд

$$\tilde{x} = a - \frac{f'(a)}{f'(b) - f'(a)}(b - a). \quad (3.2)$$

Знайшовши за формулою (3.2) $\tilde{x}$, конкретизуємо відрізок подальшого пошуку точки x^* , аналізуючи знак $f'(\tilde{x})$ (як і в методі середньої точки). Зауважимо також, що на кожній ітерації, крім першої, обчислюється одне значення $f'(x)$.

Опишемо алгоритм методу хорд.

Крок 1. Знаходження $\tilde{x}$ (за формулою (3.2)) і значення $f'(\tilde{x})$. Перехід до кроку 2.

Крок 2. Перевірка на завершення пошуку точки мінімуму.

Якщо $|f'(\tilde{x})| \leq \varepsilon$, то пошук завершений, причому $x^* \approx \tilde{x}$, $f^* \approx f(\tilde{x})$. При $|f'(\tilde{x})| > \varepsilon$ здійснюється перехід до кроку 3.

Крок 3. Вибір наступного відрізка мінімізації. Якщо $f'(\tilde{x}) < 0$, то це буде відрізок $[\tilde{x}, b]$, інакше (при $f'(\tilde{x}) > 0$) – відрізок $[a, \tilde{x}]$ (у першому випадку $a := \tilde{x}$, у другому – $b := \tilde{x}$). Після цього повертаються до кроку 1.

Приклад 3.2. Методом хорд розв'язати задачу мінімізації функції $f(x) = \frac{x^2}{2} + e^{-x}$ на відрізку $[0, 1]$ з точністю $|f'(\tilde{x})| \leq 0,001 = \varepsilon$.

◀ Ітерація 1

Крок 1. Користуючись формулою (3.2), знаходимо $\tilde{x} = 0,613$, а потім $f'(\tilde{x}) = 0,071$, $\left(f'(x) = x - \frac{1}{e^x} \right)$.

Крок 2. $|f'(\tilde{x})| = 0,071 > 0,001$, тому переходимо до кроку 3.

Крок 3. Оскільки $f'(\tilde{x}) > 0$, то відрізком мінімізації на другій ітерації буде відрізок $[0, \tilde{x}] = [0; 0,613] \equiv [a, b]$. Повертаємось до кроку 1 і продовжуємо обчислення на другій ітерації.

Процес подальшого розв'язування даного прикладу відображений у табл. 3.2 (k – номер ітерації, p – параметр методу).

Таблиця 3.2

$\begin{matrix} p \\ k \end{matrix}$	a	b	$\tilde{x}$	$f'(\tilde{x})$	Знак $f'(\tilde{x})$
1	0	1	0,613	0,071	+
2	0	0,613	0,572	0,008	+
3	0	0,572	0,567	-0,0002	–

З табл. 3.2 випливає, що на третій ітерації $|f'(\tilde{x})| < 0,001 = \varepsilon$, тому шуканий розв'язок такий: $x^* \approx 0,567$; $f^* \approx f(0,567) = 0,728$. ►

3.3. Метод Ньютона

У випадку, коли функція $f(x)$ двічі диференційовна і $f''(x) > 0$ (це гарантує опуклість функції), рівняння (3.1) можна наближено розв'язати, використовуючи *метод дотичних*. Цей метод дозволяє будувати послідовність точок x_k ($k = 0, 1, \dots$), у кожній з яких будується дотична до графіка функції $f'(x)$ (лінійна апроксимація функції $f'(x)$), точка перетину якої з віссю Ox є точкою x_{k+1} (рис. 3.2).

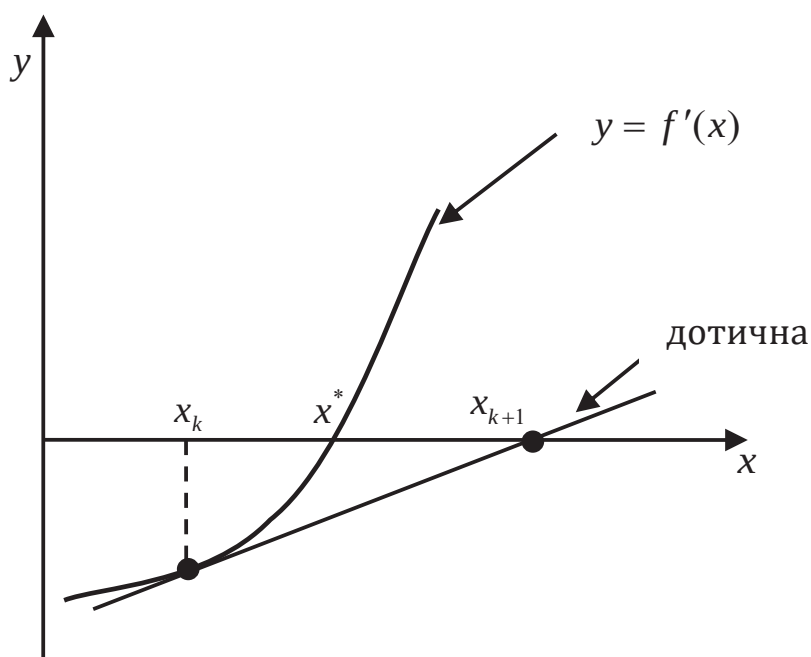


Рис. 3.2. Ілюстрація до методу дотичних розв'язування нелінійних рівнянь

Рівняння дотичної до графіка $f'(x)$ у точці $x = x_k$ має вигляд $y = f'(x_k) + f''(x_k)(x - x_k)$. Оскільки у точці x_{k+1} дотична

перетинає вісь Ox , тобто $y = 0$, то

$$x_{k+1} = x_k - \frac{f'(x_k)}{f''(x_k)}, k = 0, 1, \dots, \quad (3.3)$$

де x_0 – точка початкового наближення.

На практиці обчислення за формулою (3.3) закінчуються, якщо $|f'(x_k)| \leq \varepsilon$ (ε – задана точність пошуку точки мінімуму). Розв'язком задачі вважається точка $x^* \approx x_k$. Описана процедура пошуку точки мінімуму функції $f(x)$ називається *методом Ньютона* (він є методом мінімізації другого порядку).

Деталізуємо алгоритм методу Ньютона.

Крок 1. Вибір початкової точки x_0 , обчислення $f'(x_0)$ і перевірка умови закінчення процесу обчислень. При $|f'(x_0)| \leq \varepsilon$ процедура пошуку припиняється ($x^* \approx x_0$, $f^* \approx f(x_0)$), а при $|f'(x_0)| > \varepsilon$ здійснюється перехід до кроку 2.

Крок 2. Знаходження x_{k+1} (за формулою (3.3)) і $f'(x_{k+1})$. Перехід до кроку 3.

Крок 3. Перевірка умови закінчення пошуку точки мінімуму. Якщо $|f'(x_{k+1})| \leq \varepsilon$, то пошук припиняється і $x^* \approx x_{k+1}$, $f^* \approx f(x_{k+1})$. При $|f'(x_{k+1})| > \varepsilon$ знаходиться наступна точка x_{k+1} , тобто здійснюється перехід до кроку 2.

Приклад 3.3. Методом Ньютона розв'язати задачу мінімізації функції $f(x) = \frac{x^2}{2} + e^{-x}$ на відрізку $[0, 1]$ з точністю $|f'(x_k)| \leq 0,001 = \varepsilon$.

◀ Зазначимо, що дана функція двічі диференційовна, причому $f'(x) = x - e^{-x}$, $f''(x) = 1 + e^{-x} > 0$, тобто вона опукла (і унімодальна).

Ітерація 1

Крок 1. Виберемо $x_0 = 0,4$. Тоді $f'(x_0) = -0,270$ і $|f'(x_0)| > \varepsilon$, тому переходимо до кроку 2.

Крок 2. За формулою (3.3) знаходимо $x_1 = 0,562$ і $f'(x_1) = -0,008$. Переходимо до кроку 3.

Крок 3. Оскільки $|f'(x_1)| > \varepsilon = 0,001$, переходимо до кроку 2 і наступної ітерації.

Результати обчислень подані у табл. 3.3 (k – номер ітерації, p – параметр методу).

Таблиця 3.3

$k \backslash p$	x_k	$f'(x_k)$
0	0,4	-0,270
1	0,562	-0,008
2	0,567	-0,0002

На другій ітерації $|f'(x_2)| = 0,0002 < 0,001 = \varepsilon$. Отже, наближений розв'язок знайдений: $x^* \approx 0,567$; $f^* \approx 0,728$. ►

Зробимо деякі зауваження стосовно методу Ньютона.

1. Якщо початкова точка x_0 вибрана не достатньо близькою до x^* , то послідовність точок (3.3) може також бути розбіжною. У такому випадку потрібно вибрати іншу початкову точку x_0 , скориставшись, наприклад, кількома ітераціями методу золотого перерізу.

2. Дослідження методу Ньютона [12] підтвердили, що для функції $f(x)$, яка задовольняє певні вимоги, і достатньо близької до x^* початкової точки x_0 гарантується висока швидкість збіжності послідовності (3.3).

3.4. Метод кубічної апроксимації

Метод кубічної апроксимації разом з методами середньої точки та хорд належить до методів мінімізації першого порядку. Якщо $f(x)$ – опукла диференційовна функція, а точка мінімуму $x^* \in (x_1, x_2)$, тобто $f'(x_1) < 0$, $f'(x_2) > 0$, то, використовуючи дві пробні точки x_1, x_2 можна побудувати апроксимуючий кубічний поліном

$$g(x) = a_0 + a_1(x - x_1) + a_2(x - x_1)(x - x_2) + a_3(x - x_1)^2(x - x_2),$$

для якого

$$\begin{cases} g(x_1) = a_0 = f(x_1), \\ g(x_2) = a_0 + a_1(x_2 - x_1) = f(x_2), \\ g'(x_1) = a_1 + a_2(x_1 - x_2) = f'(x_1), \\ g'(x_2) = a_1 + a_2(x_2 - x_1) + a_3(x_2 - x_1)^2 = f'(x_2). \end{cases} \quad (3.4)$$

Інакше кажучи, невідомі коефіцієнти функції $g(x)$ знаходимо, розв'язавши систему (3.4). У ролі наближеної до x^* точки $\tilde{x}$ виберемо корінь квадратного рівняння

$$g'(x) = a_1 + a_2(2x - x_1 - x_2) + a_3(x - x_1)^2 + 2a_3(x - x_1)(x - x_2) = 0, \quad (3.5)$$

який належить інтервалу (x_1, x_2) (такий корінь існує, оскільки з того, що $g'(x_1) < 0, g'(x_2) > 0$, випливає існування внутрішньої точки $\tilde{x}$, для якої $g'(\tilde{x}) = 0$). Знак похідної $f'(\tilde{x})$ дозволяє вибрати для наступної ітерації один з відрізків $[a, \tilde{x}]$ або $[\tilde{x}, b]$, який містить точку x^* , а умова $|f'(\tilde{x})| \leq \varepsilon$ – зробити висновок про закінчення або продовження пошуку (ε – задана точність знаходження точки мінімуму).

Опишемо алгоритм методу кубічної апроксимації.

Крок 1. Знаходження точок $x_1, x_2 \in [a, b]$, для яких

$x^* \in (x_1, x_2)$ і відповідних значень $f(x_1), f(x_2), f'(x_1), f'(x_2)$.

Крок 2. Розв'язування системи (3.4) і знаходження коефіцієнтів a_0, a_1, a_2, a_3 . Розв'язування рівняння (3.5) і визначення кореня $\tilde{x} \in (x_1, x_2)$ та значення $f'(\tilde{x})$.

Крок 3. Перевірка умови закінчення пошуку точки мінімуму. Якщо $|f'(\tilde{x})| \leq \varepsilon$, то пошук закінчений і $x^* \approx \tilde{x}$, $f^* \approx f(\tilde{x})$. При $|f'(\tilde{x})| > \varepsilon$ здійснюється перехід до кроку 4.

Крок 4. Обчислення значення $f(\tilde{x})$. При $f'(\tilde{x}) < 0$ для наступної ітерації вибирається інтервал $(\tilde{x}, x_2)$ (при цьому $x_1 := \tilde{x}$, $f(x_1) := f(\tilde{x})$, $f'(x_1) := f'(\tilde{x})$), а при $f'(\tilde{x}) > 0$ – інтервал $(x_1, \tilde{x})$ (при цьому $x_2 := \tilde{x}$, $f(x_2) := f(\tilde{x})$, $f'(x_2) := f'(\tilde{x})$).

Перехід до кроку 2.

Зауважимо, що на кожній ітерації, крім першої, знаходяться два нові значення $f(x)$ і $f'(x)$.

Приклад 3.4. Виконати одну ітерацію методу кубічної апроксимації у випадку мінімізації функції $f(x) = \frac{x^2}{2} + e^{-x}$ на відріжку $[0, 1]$.

◀ **Крок 1.** Візьмемо $x_1 = 0,25$, $x_2 = 0,75$. Тоді $f'(0,25) = -0,5288 < 0$, $f'(0,75) = 0,2776 > 0$, тобто $x^* \in (0,25; 0,75)$. Крім того, $f(0,25) = 0,8101$, $f(0,75) = 0,7536$.

Крок 2. Розв'язавши систему (3.4), отримаємо $a_0 = 0,8101$, $a_1 = -0,113$, $a_2 = 0,8316$, $a_3 = -0,1008$. Рівняння (3.5) набуде вигляду $0,3024x^2 - 1,9152x + 0,9887 = 0$. Його коренем із інтервалу $(0,25; 0,75)$ буде $\tilde{x} = 0,567$. Отже, $x^* \approx \tilde{x} = 0,567$, тобто після однієї ітерації точка x^* знайдена з високою точністю. ►

Розділ 4. Методи мінімізації багатомодальних функцій

4.1. Умова Ліпшиця

У попередніх пунктах ішлося про мінімізацію унімодальних функцій. Якщо функція $f(x)$ багатомодальна (багатоекстремальна), то у цьому випадку задача знаходження глобального екстремуму суттєво ускладнюється. Нижче розглянемо два методи мінімізації багатомодальних функцій, припускаючи, що функція $f(x)$ задовольняє на відрізку $X = [a, b]$ умову Ліпшиця.

Означення 4.1. Функція $f(x)$ задовольняє на відрізку $[a, b]$ умову Ліпшиця, якщо існує таке число $L > 0$, що

$$|f(x') - f(x'')| \leq L|x' - x''| \quad (4.1)$$

для всіх $x', x'' \in [a, b]$.

Із означення 4.1 випливає кілька важливих фактів.

1. Якщо нерівність (4.1) виконується з константою L , то вона також буде виконуватися з будь-якою константою $L' > L$. Оптимальні результати в алгоритмах мінімізації досягаються у випадку мінімальної константи Ліпшиця.

2. Якщо функція $f(x)$ задовольняє на відрізку $[a, b]$ умову Ліпшиця, то вона неперервна на цьому відрізку і, згідно з теоремою Вейєштраса, досягає свого мінімального значення.

3. Модуль кутового коефіцієнта будь-якої хорди графіка функції $f(x)$ не перевищує L .

Цей факт відображає геометричний зміст нерівності (4.1).

Якщо перейти у (4.1) до границі при $|x' - x''| \rightarrow 0$, то, припускаючи існування дотичної у відповідній точці, переконаємося, що модуль кутового коефіцієнта дотичної

також не може перевищувати L . У протилежному випадку функція не задовольняє умову Ліпшиця. Наприклад, $f(x) = \sqrt{x}$ на відрізку $[0,1]$ умову Ліпшиця не задовольняє, тому що при $x \rightarrow 0+0$ кутовий коефіцієнт дотичної до її графіка необмежено зростає.

4. Якщо функція $f(x)$ має на відрізку $[a,b]$ неперервну похідну, то для неї виконується на цьому відрізку умова Ліпшиця з константою $L = \max_{[a,b]} |f'(x)|$.

Цей факт впливає з формули скінченних приростів $f(x') - f(x'') = f'(\xi)(x' - x'')$, яка виконується для довільних $x', x'' \in [a,b]$ (ξ – деяка внутрішня точка із (x', x'')). Враховуючи, що $|f'(\xi)| \leq \max_{[a,b]} |f'(x)| = L$, отримуємо формулу (4.1).

5. Якщо функція $f(x)$ неперервна на відрізку $[a,b]$ і для неї виконується умова (4.1) на кожному з відрізків $[x_i, x_{i+1}]$ з константою L_i ($i = \overline{0, n-1}$, $x_0 = a$, $x_n = b$), то вона задовольняє умову Ліпшиця на всьому відрізку $[a,b]$ з константою $L = \max_{i=0, n-1} L_i$.

Отже, як впливає із вищесказаного, задача мінімізації на відрізку $[a,b]$ функції $f(x)$, для якої виконується умова (4.1), має розв'язок, хоча точка глобального мінімуму може бути не єдиною.

4.2. Метод перебору

Для унімодальної на відрізку $[a,b]$ функції $f(x)$ цей метод строго обґрунтований раніше. Якщо функція $f(x)$ не є унімодальною, але вона задовольняє на відрізку $[a,b]$ умову

Ліпшиця (4.1), то метод перебору гарантує знаходження мінімального значення f^* з будь-якою наперед заданою точністю (похибкою) δ (відстань точки x^* , для якої $f(x^*) = f^*$, до точки (чи точок) глобального мінімуму тут до уваги не береться). Обґрунтуємо це твердження детально.

Нехай відрізок $[a, b]$ розбитий на n рівних частин точками $x_i = a + i(b-a)/n$, $i = \overline{0, n}$, у яких знайдені значення $f(x_i)$, $i = \overline{0, n}$ і вибрані $x^* \approx x_k$, $f^* \approx f(x_k) = \min_{i=\overline{0, n}} f(x_i)$. Якщо припустити, що x^* – точка глобального мінімуму функції $f(x)$ на $[a, b]$, то, очевидно, знайдеться така точка x_i , що $|x_i - x^*| \leq \frac{b-a}{2n}$. З урахуванням умови (4.1) матимемо

$$0 \leq f(x_k) - f^* \leq f(x_i) - f^* \leq L|x_i - x^*| \leq L \frac{b-a}{2n} = \delta_n, \quad (4.2)$$

де δ_n – точність визначення мінімального значення f^* у результаті розбиття відрізка $[a, b]$ на n рівних частин.

Із (4.2) випливає, що:

1) для забезпечення заданої точності δ знаходження мінімального значення f^* число n повинно бути виbrane з умови $\delta_n \leq \delta$, тобто $n \geq L \frac{b-a}{2\delta}$;

2) після $N = n + 1$ обчислень значень функції метод перебору дозволяє знайти значення f^* з точністю

$$\delta(N) = \delta_{N-1} = L \frac{b-a}{2(N-1)}.$$

Ще раз підкреслимо, що метод перебору для багатомодальних функцій дає можливість знайти із заданою точністю δ лише мінімальне значення функції $f(x)$. Похибка

точки мінімуму може бути значною, однак у багатьох практичних задачах вимагається знайти тільки мінімальне значення функції, що й оправдовує застосування методу перебору для мінімізації багатомодальних функцій.

Алгоритм методу перебору у випадку мінімізації багатомодальних функцій, для яких виконується умова (4.1), передбачає такі кроки.

Крок 1. Задання похибки $\delta \geq \delta_n = L \frac{b-a}{2n}$.

Крок 2. Визначення числа n відрізків розбиття відрізка $[a, b]$ $\left(n \geq L \frac{b-a}{2\delta} \right)$ і точок розбиття $x_i = a + i(b-a)/n$, $i = \overline{0, n}$.

Крок 3. Знаходження значень $f(x_i)$, $i = \overline{0, n}$, та уточнення $f^* \approx \min_{i=\overline{0, n}} f(x_i)$.

Приклад 4.1. Методом перебору розв'язати задачу знаходження мінімального значення функції $f(x) = 0,2(x^3 - x) + 2$ на відрізку $[-1, 1]$ з точністю $\delta = 0,025$.

◀ Очевидно, що функція $f(x)$ багатомодальна на $[-1, 1]$, у чому легко переконатися, побудувавши її графік. Оскільки її похідна $f'(x) = 0,2(3x^2 - 1)$ неперервна на $[-1, 1]$ і $\max_{[-1, 1]} |f'(x)| = 0,4$, то функція $f(x)$ задовольняє умову Ліпшиця з константою $L = 0,4$.

Крок 1. $\delta = 0,025 \geq 0,4 \cdot 2/2n = 0,4/n$.

Крок 2. $n \geq 0,4/\delta = 0,4/0,025 = 16$, тому можна вибрати $n = 16$.

При цьому $x_i = -1 + i/8$, $i = \overline{0, 16}$.

Крок 3. Знаходимо значення $f_i = f(x_i)$ ($i = \overline{0, 16}$) і вибираємо мінімальне з них значення f^* (результати обчислень продемонстровані у табл. 4.1).

Таблиця 4.1

x_i	$f(x_i)$
-1,000	2,000
-0,875	2,041
-0,750	2,066
-0,625	2,076
-0,500	2,075
-0,375	2,064
-0,250	2,047
-0,125	2,025
0,000	2,000
0,125	1,975
0,250	1,953
0,375	1,936
0,500	1,925
0,625	1,924
0,750	1,934
0,875	1,959
1,000	2,000

Із табл. 4.1 випливає, що мінімальне значення $f^* \approx 1,924$ (воно відповідає точці $x^* \approx x_{13} = 0,625$). ►

4.3. Метод ламаних

Як і раніше, будемо вважати, що функція $f(x)$ задовольняє на відрізку $[a, b]$ умову Ліпшиця (4.1). Зафіксуємо яку-небудь точку $\tilde{x} \in [a, b]$ і визначимо допоміжну функцію $g(x, \tilde{x}) = f(\tilde{x}) - L|x - \tilde{x}|$ змінної x . Ця функція є кусково-лінійною, а її графік – це ламана лінія, що складається із двох відрізків прямих (з кутовими коефіцієнтами $(+L)$ і $(-L)$), які перетинаються у точці $(\tilde{x}, f(\tilde{x}))$ (рис. 4.1).

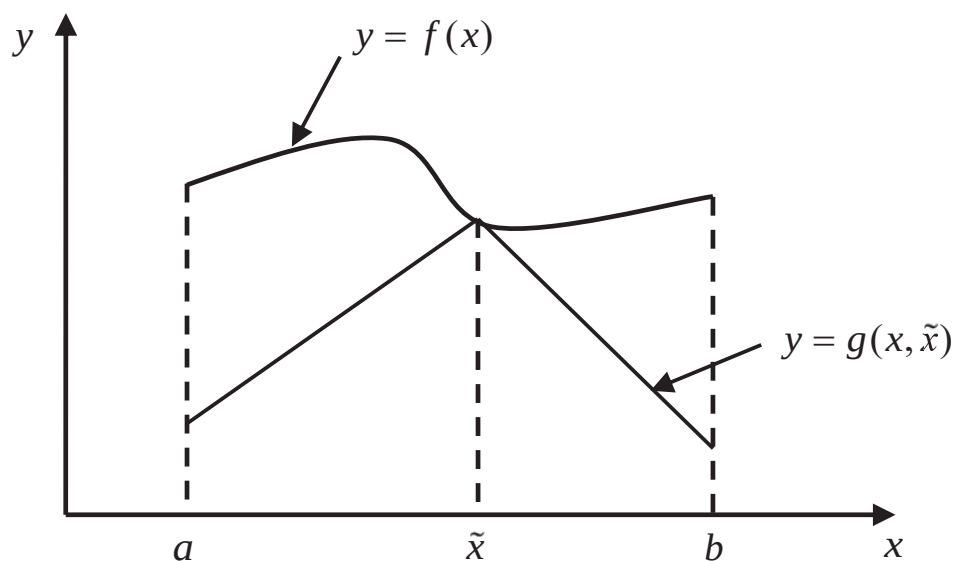


Рис. 4.1. Графік функції $g(x, \tilde{x})$

Цей висновок прямо випливає з умови (4.1), оскільки

$$g(x, \tilde{x}) = f(\tilde{x}) - L|x - \tilde{x}| \leq f(x) \quad (4.3)$$

при всіх $x \in [a, b]$ і $g(\tilde{x}, \tilde{x}) = f(\tilde{x})$, тобто графік функції $f(x)$ розміщений вище ламаної $g(x, \tilde{x})$ і має з нею єдину спільну точку $(\tilde{x}, f(\tilde{x}))$.

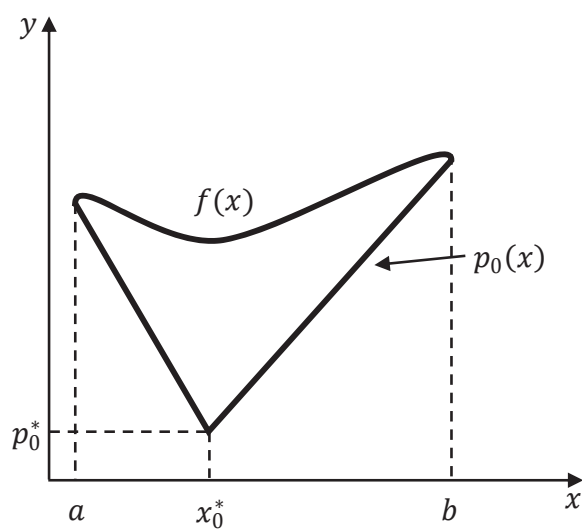
Властивість (4.3) функції $g(x, \tilde{x})$ використовується для побудови так званого *методу ламаних*, який полягає у побудові збіжної до множини точок мінімуму функції $f(x)$ на $[a, b]$ послідовності точок x_k^* ($k=0, 1, \dots$), які є точками мінімуму апроксимуючих кусково-лінійних функцій $p_k(x) = g(x, x_k)$ ($k=0, 1, \dots$).

Процес побудови даної послідовності розпочнемо з вибору точки $x_0^* = x_0$ та значення $p_0^* = p_0(x_0) = y_0$, де x_0 і y_0 – відповідно абсциса й ордината точки перетину прямих $y = f(a) - L(x - a)$ та $y = f(b) + L(x - b)$. Очевидно,

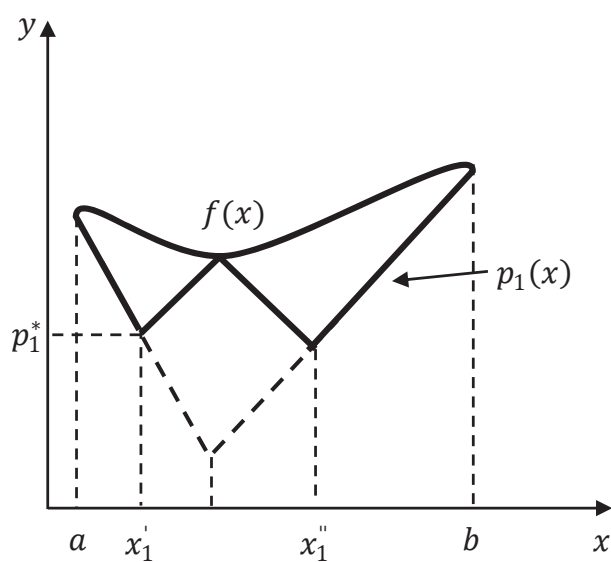
$$\begin{cases} x_0 = \frac{1}{2L} [f(a) - f(b) + L(a+b)], \\ y_0 = \frac{1}{2} [f(a) + f(b) + L(a-b)], \end{cases} \quad (4.4)$$

$$p_0(x) = \begin{cases} f(a) - L(x-a), & x \in [a, x_0], \\ f(b) + L(x-b), & x \in [x_0, b]. \end{cases} \quad (4.5)$$

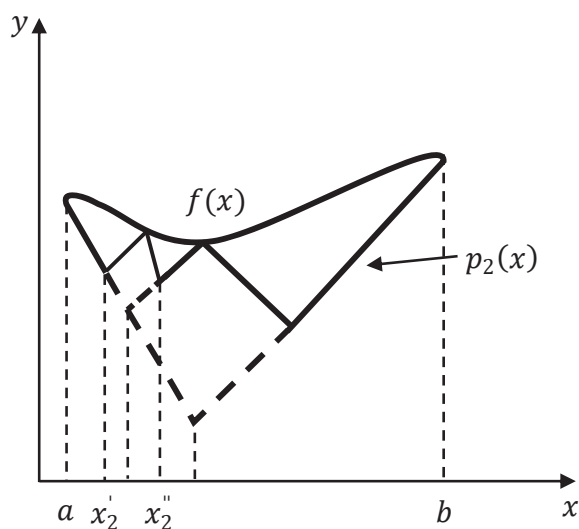
Графік функції (4.5) проілюстрований на рис. 4.2, 1).



1)



2)



3)

Рис. 4.2. Ілюстрація процесу побудови ламаних $p_0(x)$, $p_1(x)$, $p_2(x)$ (1), 2), 3))

Наступна апроксимуюча функція

$$p_1(x) = \max \{ g(x, x_0^*), p_0(x) \}. \quad (4.6)$$

Функція (4.6) має дві точки локального мінімуму x_1' та x_1'' (рис. 4.2, 2)). Очевидно, x_1' є абсцисою точки перетину прямої, що проходить через точку $(x_0^*, f(x_0^*))$, і має кутовий коефіцієнт $(+L)$, з прямою, яка проходить через точку (x_0^*, p_0^*) і має кутовий коефіцієнт $(-L)$, а x_1'' – абсцисою точки перетину прямої, яка проходить через точку $(x_0^*, f(x_0^*))$ і має кутовий коефіцієнт $(-L)$, з прямою, яка проходить через точку (x_0^*, p_0^*) і має кутовий коефіцієнт $(+L)$. Провівши відповідні обчислення, отримаємо

$$x_1' = x_0^* - \Delta_1, \quad x_1'' = x_0^* + \Delta_1, \quad (4.7)$$

$$\text{де } \Delta_1 = [f(x_0^*) - p_0^*]/2L, \quad p_1(x_1') = p_1(x_1'') = p_1 = \frac{1}{2} [f(x_0^*) + p_0^*].$$

Оскільки значення апроксимуючої функції $p_1(x)$ однакові у точках x'_1 і x''_1 , то за точку глобального мінімуму можна вибрати довільну з них, наприклад $x_1^* = x'_1$. Далі, використовуючи допоміжну функцію $g(x, x_1^*)$, будуємо апроксимуючу функцію

$$p_2(x) = \max \{ g(x, x_1^*), p_1(x) \}. \quad (4.8)$$

Порівняно з функцією (4.6) у функції (4.8) замість x_1^* появилися дві нові точки локального мінімуму x'_2 і x''_2 (рис. 4.2, 3)), які визначаються за формулами

$$x'_2 = x_1^* - \Delta_2, \quad x''_2 = x_1^* + \Delta_2, \quad (4.9)$$

де $\Delta_2 = [f(x_1^*) - p_1^*]/2L$,

$$p_2(x'_2) = p_2(x''_2) = p_2 = \frac{1}{2} [f(x_1^*) + p_1^*].$$

Нехай у описаний вище спосіб побудована функція $p_{k-1}(x)$ і вибрана точка (одна з двох) її глобального мінімуму x_{k-1}^* . Тоді функція $p_k(x)$ і точки x'_k , x''_k визначаються так:

$$p_k(x) = \max \{ g(x, x_{k-1}^*), p_{k-1}(x) \}, \quad (4.10)$$

$$x'_k = x_{k-1}^* - \Delta_k, \quad x''_k = x_{k-1}^* + \Delta_k, \quad (4.11)$$

де $\Delta_k = [f(x_{k-1}^*) - p_{k-1}^*]/2L$,

$$p_k(x'_k) = p_k(x''_k) = p_k = \frac{1}{2} [f(x_{k-1}^*) + p_{k-1}^*].$$

Формули (4.10) і (4.11) є узагальненням формул (4.6), (4.8) і (4.7), (4.9).

Відзначимо деякі важливі факти, що стосуються описаного вище методу ламаних.

1. Графіком кусково-лінійної функції $p_k(x)$ є неперервна ламана лінія, що складається із відрізків прямих

з кутовими коефіцієнтами $(+L)$ і $(-L)$. Функція $p_k(x)$ ($k \geq 0$) задовольняє умову (4.1) з тією ж константою L , що і функція $f(x)$, і має $(k+1)$ локальних мінімумів, точками яких є абсциси вершин ламаної.

2. Для всіх $x \in [a, b]$ і $k \geq 1$ справджуються нерівності

$$p_{k-1}(x) \leq p_k(x) \leq f(x). \quad (4.12)$$

Нерівності (4.12) випливають із (4.3) і означають, що на кожному кроці методу ламаних задача мінімізації функції $f(x)$ замінюється більш простою задачею мінімізації функції $p_k(x)$ ($k \geq 0$), яка наближає $f(x)$ знизу. Послідовність функцій $p_k(x)$ ($k \geq 0$) монотонно зростаюча.

3. При необмеженому зростанні k метод ламаних збігається. Для довільної функції $f(x)$, яка задовольняє на відрізку $[a, b]$ умову (4.1), доведено [12] такі твердження:

а) $\lim_{k \rightarrow \infty} f(x_k^*) = f^*$;

б) якщо точка x^* глобального мінімуму функції $f(x)$ на відрізку $[a, b]$ єдина, то $\lim_{k \rightarrow \infty} x_k^* = x^*$, а якщо множина точок глобального мінімуму X^* містить більше, ніж одну точку, то $\lim_{k \rightarrow \infty} \rho_k = 0$, де ρ_k – відстань від точки x_k^* до найближчої точки множини X^* ;

в) для точності δ наближення $f(x_{k-1}^*)$ до мінімального значення f^* справедлива оцінка:

$$0 \leq f(x_{k-1}^*) - f^* \leq f(x_{k-1}^*) - p_{k-1}^* = 2L\Delta_k = \delta_k \leq \delta. \quad (4.13)$$

Співвідношення (4.13) конкретизують умову завершення процесу пошуку мінімуму методом ламаних.

4. Реалізація методу ламаних передбачає знання константи Ліпшиця L . У загальному випадку на практиці

константу L визначають так: знаходять кутові коефіцієнти деякого числа хорд графіка $f(x)$ і вибирають максимальний за модулем, який і є наближеною оцінкою L знизу (якість оцінки поліпшується зі зменшенням довжин хорд). Зауважимо також, що використання надто завищеної оцінки для мінімальної константи Ліпшиця знижує швидкість збіжності методу ламаних, а використання заниженої оцінки практично приводить до неправильних результатів.

Деталізуємо алгоритм методу ламаних.

Крок 1. Обчислення за формулами (4.4) значень x_0, y_0 і визначення пари $(x, p) = (x_0, y_0)$.

Крок 2. Порівняння наявних на даній ітерації пар (x, p) і визначення серед них пари $(\tilde{x}, \tilde{p})$, у якій значення $\tilde{p}$ є мінімальним. Обчислення $f(\tilde{x})$.

Крок 3. Перевірка умови (4.13) (умови завершення процесу пошуку мінімуму), у якій $\Delta_k = \frac{1}{2L}[f(\tilde{x}) - \tilde{p}] = \Delta$, $\delta_k = 2L\Delta$. Якщо $\delta_k \leq \delta$, то пошук завершений і $x^* \approx \tilde{x}$, $f^* \approx f(\tilde{x})$. У протилежному випадку здійснюється перехід до кроку 4.

Крок 4. Визначення нових пар $(x', p), (x'', p)$, де $x' = \tilde{x} - \Delta$, $x'' = \tilde{x} + \Delta$, $p = \frac{1}{2}[f(\tilde{x}) + \tilde{p}]$, і введення їх у розгляд замість пари $(\tilde{x}, \tilde{p})$. Перехід до кроку 2.

Приклад 4.2. Методом ламаних розв'язати задачу мінімізації функції $f(x) = 0,2(x^3 - x) + 2$ на відрізку $[-1, 1]$ з точністю $\delta = 0,03$.

◀ Раніше (приклад 4.1) вже зазначалося, що функція $f(x)$ не є унімодальною на $[-1, 1]$ і задовольняє умову

Ліпшиця з константою $L=0,4$. Звичайно, за значення L можна взяти довільне число, яке більше від $0,4$, але тоді збіжність методу уповільниться.

Ітерація 1

Крок 1. Із (4.4) при $a=-1$, $b=1$, $L=0,4$ випливає, що $(x, p) = (x_0, y_0) = (0; 1,6)$.

Крок 2. Пара (x, p) єдина, тому $(\tilde{x}, \tilde{p}) = (0; 1,6)$. При цьому $f(\tilde{x}) = f(0) = 2$.

Крок 3. $\Delta = (2 - 1,6)/0,8 = 0,5$; $\delta_k = 2 \cdot 0,4 \cdot 0,5 = 0,4 > 0,03 = \delta$, тому переходимо до кроку 4.

Крок 4. Оскільки $x' = 0 - 0,5 = -0,5$; $x'' = 0 + 0,5 = 0,5$; $p = \frac{1}{2}(2 + 1,6) = 1,8$, то виключивши із розгляду пару $(0; 1,6)$, маємо дві нові пари $(-0,5; 1,8)$, $(0,5; 1,8)$. Переходимо до ітерації 2, починаючи з кроку 2 (результати обчислень на наступних ітераціях відображені у такому порядку: спочатку записані пари, з яких починається ітерація (вибрана пара підкреслена), потім записані оцінка δ_k і знову утворені пари (x', p) , (x'', p) , які переходять на наступну ітерацію замість $(\tilde{x}, \tilde{p})$).

Ітерація 2

$(-0,5; 1,8)$, $(0,5; 1,8)$, $\delta_2 = 0,125 > 0,03 = \delta$, $(0,344; 1,863)$, $(0,656; 1,863)$.

Ітерація 3

$(-0,5; 1,8)$, $(0,344; 1,863)$, $(0,656; 1,863)$, $\delta_3 = 0,275 > 0,03 = \delta$, $(-0,844; 1,938)$, $(-0,156; 1,938)$.

Ітерація 4

$(0,344;1,863), \underline{(0,656;1,863)}, (-0,844;1,938), (-0,156;1,938),$

$\delta_4 = 0,062 > 0,03 = \delta, (0,578;1,894), (0,734;1,894).$

Ітерація 5

$\underline{(0,344;1,863)}, (-0,844;1,938), (-0,156;1,938), (0,578;1,894),$

$(0,734;1,894), \delta_5 = 0,076 > 0,03 = \delta, (0,249;1,901), (0,439;1,901).$

Ітерація 6

$(-0,844;1,938), (-0,156;1,938), \underline{(0,578;1,894)}, (0,734;1,894),$

$(0,249;1,901), (0,439;1,901), \delta_6 = 0,029 < 0,03 = \delta.$

Отже, точність досягнута, тому $x^* \approx 0,578,$
 $f^* \approx f(0,578) = 1,923. \blacktriangleright$

ЧАСТИНА II. БАГАТОВИМІРНА ОПТИМІЗАЦІЯ

Розділ 5. Теоретичні основи та методи безумовної мінімізації функції багатьох змінних

5.1. Постановка задачі безумовної мінімізації функції багатьох змінних

Позначимо через $\mathbb{R}^n$ ($n > 1$) n -вимірний дійснозначний лінійний простір точок (векторів) $x = (x_1, \dots, x_n)^T$, тобто простір, у якому введені операції додавання та множення на дійсне число $(x + y = (x_1 + y_1, \dots, x_n + y_n), \quad \lambda x = (\lambda x_1, \dots, \lambda x_n)^T, \quad \lambda \in \mathbb{R})$. Якщо у $\mathbb{R}^n$ введена операція скалярного добутку $\left(\langle x^{(1)}, x^{(2)} \rangle = \sum_{i=1}^n x_i^{(1)} x_i^{(2)}, \quad x^{(1)}, x^{(2)} \in \mathbb{R}^n \right)$, то простір $\mathbb{R}^n$ перетворюється у n -вимірний евклідовий простір. Отже, надалі $\mathbb{R}^n$ – це евклідовий простір. В евклідовому просторі для вектора x визначаються довжина (норма)

$$\|x\| = \sqrt{\langle x, x \rangle} = \left(\sum_{i=1}^n x_i^2 \right)^{1/2}$$

та відстань між векторами (точками) $x^{(1)}$ і $x^{(2)}$:

$$\begin{aligned} \rho(x^{(1)}, x^{(2)}) &= \|x^{(1)} - x^{(2)}\| = \sqrt{\langle x^{(1)} - x^{(2)}, x^{(1)} - x^{(2)} \rangle} = \\ &= \left(\sum_{i=1}^n (x_i^{(1)} - x_i^{(2)})^2 \right)^{1/2}. \end{aligned}$$

Для норми і скалярного добутку відповідно виконуються нерівність трикутника

$$\|x^{(1)} + x^{(2)}\| \leq \|x^{(1)}\| + \|x^{(2)}\|$$

і нерівність Коші-Буняковського

$$|\langle x^{(1)}, x^{(2)} \rangle| \leq \|x^{(1)}\| \|x^{(2)}\|.$$

Нехай функція n змінних $f(x) \equiv f(x_1, \dots, x_n)$ (як і у випадку функції однієї змінної для позначення функції багатьох змінних використаємо символ « f ») визначена у всіх точках простору $\mathbb{R}^n$.

Означення 5.1. Точка $x^* \in \mathbb{R}^n$ називається точкою глобального мінімуму функції $f(x)$ на $\mathbb{R}^n$, якщо $f(x^*) \leq f(x)$ для всіх $x \in \mathbb{R}^n$. Значення $f(x^*) = \min_{\mathbb{R}^n} f(x) = f^*$ називається мінімальним значенням (мінімумом) функції $f(x)$ на $\mathbb{R}^n$.

Множину всіх точок глобального мінімуму функції $f(x)$ на $\mathbb{R}^n$, як і раніше, будемо позначати через X^* .

Означення 5.2. Точка $x^* \in \mathbb{R}^n$ називається точкою локального мінімуму функції $f(x)$, якщо $f(x^*) \leq f(x)$ у деякому ε -околі точки x^* , тобто для всіх $x \in O(x^*, \varepsilon) = \{x \in \mathbb{R}^n / \rho(x, x^*) < \varepsilon, \varepsilon > 0\}$.

Задача мінімізації функції n змінних на допустимій множині X , яка збігається із простором $\mathbb{R}^n$, тобто задача

$$\begin{cases} f(x) \rightarrow \min, \\ x \in \mathbb{R}^n, \end{cases} \quad (5.1)$$

називається задачею безумовної мінімізації функції n змінних (безумовної тому, що ніяких допоміжних умов (обмежень) на x не накладається). Очевидно, що задача (5.1) є частинним випадком задачі (В.1) при $X = \mathbb{R}^n$. Задача безумовної мінімізації має принципове значення у теорії оптимізації функцій, оскільки більшість методів

розв'язування задач математичного програмування (задач умовної оптимізації) зводиться останні до розв'язування послідовності задач безумовної оптимізації.

Що ж до задач безумовної оптимізації, то зауважимо таке.

1. Якщо $X^* = \emptyset$, то задача не має змісту, тому замість мінімуму функції $f(x)$ інколи шукають її точку нижню межу $f_* = \inf_{\mathbb{R}^n} f(x)$. Поняття обмеженості функції багатьох змінних та її точних меж (нижньої, верхньої) визначаються аналогічно до цих понять у випадку функції однієї змінної.

2. Задача максимізації функції $f(x)$ ($x \in \mathbb{R}^n$) зводиться до задачі мінімізації функції $(-f(x))$ на $\mathbb{R}^n$, тому надалі обмежимося вивченням лише задач мінімізації.

5.2. Класичний метод мінімізації функції багатьох змінних

Як і у випадку мінімізації функції однієї змінної, під класичним методом мінімізації функції багатьох змінних будемо розуміти метод пошуку точок екстремуму (точок локального мінімуму і максимуму), що використовує диференціальне числення. У зв'язку з цим уточнимо деякі поняття і факти із теорії диференціального числення функції багатьох змінних $f(x)$, вважаючи що вона визначена, як мінімум, у деякому околі точки $x^* \in \mathbb{R}^n$, тобто x^* є внутрішньою точкою.

Означення 5.3. Функція $f(x)$ називається диференційовною у точці x^* , якщо її повний приріст $\Delta f(x)$ у цій точці можна записати у вигляді

$$\Delta f(x^*) = \sum_{j=1}^n A_j \Delta x_j + \sum_{j=1}^n \alpha_j \Delta x_j = \sum_{j=1}^n A_j \Delta x_j + o(\rho), \quad (5.2)$$

де $A_j \left(j = \overline{1, n} \right)$ – деякі, не залежні від приростів аргументів $\Delta x_j \left(j = \overline{1, n} \right)$ числа; $\alpha_j \left(j = \overline{1, n} \right)$ – нескінченно малі функції при $\Delta x_j \rightarrow 0 \left(j = \overline{1, n} \right)$, які приймають нульові значення при $\Delta x_j = 0 \left(j = \overline{1, n} \right)$; $\rho = \|\Delta x\| = \sqrt{(\Delta x_1)^2 + \dots + (\Delta x_n)^2}$ – відстань між точками $x^* = (x_1^*, \dots, x_n^*)^T$ та $(x^* + \Delta x) = (x_1^* + \Delta x_1, \dots, x_n^* + \Delta x_n)^T$.

Умова (5.2) називається умовою диференційовності функції $f(x)$ у точці x^* . Необхідні та достатні умови диференційовності функції багатьох змінних конкретизуються у наступних теоремах.

Теорема 5.1. Якщо функція $f(x)$ диференційовна у точці x^* , то у цій точці вона має частинні похідні $\frac{\partial f(x^*)}{\partial x_j} = A_j$, де $A_j \left(j = \overline{1, n} \right)$ визначені із умови (5.2), і є неперервною.

Теорема 5.2. Якщо функція $f(x)$ має всі частинні похідні у деякому околі точки x^* , причому ці похідні неперервні у самій точці x^* , то дана функція диференційовна у точці x^* .

З теореми 5.2 робимо висновок, що неперервність функції впливає із неперервності її частинних похідних.

Означення 5.4. Функція $f(x)$ називається k ($k > 1$) разів диференційовною у точці x^* , якщо всі її частинні похідні $(k-1)$ -го порядку є диференційовними функціями у цій точці.

Теорема 5.3. Для того, щоб функція $f(x)$ була k ($k > 1$) разів диференційовною у точці x^* , достатньо, щоб усі її частинні похідні k -го порядку були неперервними функціями у точці x^* .

Теорема 5.3 впливає із визначення диференційовності

функції багатьох змінних і теореми 5.2.

Зазначимо також, що для функції $f(x)$ ($x \in \mathbb{R}^n$) значення довільної змішаної частинної похідної k -го порядку у точці x^* не залежить від порядку послідовних диференціювань даної функції, наприклад, якщо функція двох змінних $f(x_1, x_2)$ двічі диференційовна у точці (x_1^*, x_2^*) , то значення її змішаних похідних рівні.

Означення 5.5. Диференціалом (першим диференціалом) $df(x^*)$ функції $f(x)$ у точці x^* називається головна лінійна відносно приростів Δx_j ($j = \overline{1, n}$) частина (5.2).

Назвавши при $j = \overline{1, n}$ диференціалами dx_j незалежних змінних x_j відповідні прирости Δx_j і врахувавши теорему 5.1, отримаємо для першого диференціала у точці x^* таку формулу:

$$df(x^*) = \sum_{j=1}^n \frac{\partial f(x^*)}{\partial x_j} \Delta x_j \equiv \sum_{j=1}^n \frac{\partial f(x^*)}{\partial x_j} dx_j. \quad (5.3)$$

Означення 5.6. Вектор $f'(x^*) = \left(\frac{\partial f(x^*)}{\partial x_1}, \dots, \frac{\partial f(x^*)}{\partial x_n} \right)^T$

називається градієнтом функції $f(x)$ у точці x^* .

Для позначення градієнта інколи використовують символ $grad f$. У малому околі точки x^* градієнт показує напрямок найшвидшого зростання функції $f(x)$, а його норма характеризує швидкість цього зростання (антиградієнт $(-f'(x^*))$ показує напрямок найшвидшого спадання функції $f(x)$). Градієнт у точці x^* перпендикулярний до лінії (поверхні) рівня $f(x) = const$, що проходить через цю точку.

Якщо функція $f(x)$ k разів ($k > 1$) диференційовна, то для неї вводиться поняття диференціала k -го порядку $d^k f(x) = d(d^{k-1} f(x))$ як диференціала від функції $d^{k-1} f(x)$ (вважається, що диференціал $(k-1)$ порядку є диференційовною функцією від x , а Δx_j ($j = \overline{1, n}$) – сталі величини). Формула для обчислення $d^k f(x)$ у точці $x^* \in \mathbb{R}^n$ така:

$$\begin{aligned} d^k f(x^*) &= \left(\frac{\partial}{\partial x_1} \Delta x_1 + \dots + \frac{\partial}{\partial x_n} \Delta x_n \right)^k f(x^*) \equiv \\ &\equiv \left(\frac{\partial}{\partial x_1} dx_1 + \dots + \frac{\partial}{\partial x_n} dx_n \right)^k f(x^*). \end{aligned} \quad (5.4)$$

У (5.4) піднесення виразу у дужках до степеня k належить розуміти як звичайне піднесення до степеня k многочлена. Очевидно, при $k = 1$ із (5.4) отримуємо (5.3).

Із (5.4) випливає формула для другого диференціала:

$$d^2 f(x^*) = \sum_{i=1}^n \sum_{j=1}^n \frac{\partial^2 f(x^*)}{\partial x_i \partial x_j} dx_i dx_j = \langle f''(x^*) dx, dx \rangle, \quad (5.5)$$

де $f''(x^*) = \left(\frac{\partial^2 f(x^*)}{\partial x_i \partial x_j} \right)_{i,j=1}^n$ – матриця Гессе (матриця

других похідних), яку позначають також через $H(x^*)$.

Принципово важливе значення при дослідженні функцій багатьох змінних має наступна теорема.

Теорема 5.4. Нехай функція $f(x) \equiv f(x_1, \dots, x_n)$ диференційовна $(k-1)$ разів у деякому околі точки $x^* = (x_1^*, \dots, x_n^*)^T$ і k разів диференційовна у самій точці x^* . Тоді для довільної точки $x = x^* + \Delta x = (x_1^* + \Delta x_1, \dots, x_n^* + \Delta x_n)^T$ із цього

околу $(x^* + \Delta x \rightarrow x^*)$ справджується рівність

$$f(x) = f(x^*) + df(x^*) + \frac{1}{2!} d^2 f(x^*) + \dots + \frac{1}{k!} d^k f(x^*) + o(\rho) = P_k^*(x, x^*) + R_{k+1}^*(x, x^*), \quad (5.6)$$

де $P_k^*(x, x^*)$ – многочлен Тейлора порядку k функції $f(x)$ у точці x^* , $R_{k+1}^*(x, x^*)$ – залишковий член формули Тейлора у формі Пеано.

Формулу (5.6) називають формулою Тейлора для функції $f(x)$ з центром розкладу у точці x^* . Більш жорсткі вимоги до функції $f(x)$ дозволяють сформулювати та довести [12, 14, 20] такий аналог теореми 5.4:

якщо функція $f(x) \equiv f(x_1, \dots, x_n)$ диференційовна k разів у деякому околі точки $x^* = (x_1^*, \dots, x_n^*)^T$, то для кожної точки $x = x^* + \Delta x = (x_1^* + \Delta x_1, \dots, x_n^* + \Delta x_n)^T$ із цього околу

$$f(x) = P_{k-1}^*(x, x^*) + \tilde{R}_k^*(x, x^*),$$

де $P_{k-1}^*(x, x^*)$ – многочлен Тейлора порядку $(k-1)$ функції $f(x)$ у точці x^* , а $\tilde{R}_k^*(x, x^*) = \frac{1}{k!} d^k f(x')$ – залишковий член у формі Лагранжа ($x' = (x_1', \dots, x_n')^T$ – деяка точка відрізка, що з'єднує точки x^* та $(x^* + \Delta x)$).

Вищевведені поняття та факти дозволяють дійти таких висновків.

1. Якщо функція $f(x)$ диференційовна у деякому околі точки x^* , то її приріст

$$\Delta f(x^*) = df(x^*) + o(\|\Delta x\|) = \langle f'(x^*), \Delta x \rangle + o(\|\Delta x\|). \quad (5.7)$$

Формула (5.7) випливає із (5.6) і (5.3).

2. Якщо функція $f(x)$ двічі диференційовна у деякому

околі точки x^* , то

$$\begin{aligned}\Delta f(x^*) &= df(x^*) + \frac{1}{2} d^2 f(x^*) + o(\|\Delta x\|^2) = \\ &= \langle f'(x^*), \Delta x \rangle + \frac{1}{2} \langle f''(x^*) \Delta x, \Delta x \rangle + o(\|\Delta x\|^2).\end{aligned}\quad (5.8)$$

Формула (5.8) випливає із (5.6), (5.3) і (5.5).

3. Із (5.7) і (5.8) для малих значень $\|\Delta x\|$ маємо наближену рівність

$$f(x) \approx f(x^*) + \langle f'(x^*), \Delta x \rangle \quad (5.9)$$

або

$$f(x) \approx f(x^*) + \langle f'(x^*), \Delta x \rangle + \frac{1}{2} \langle f''(x^*) \Delta x, \Delta x \rangle. \quad (5.10)$$

Отже, у малому околі точки x^* поведінка диференційовної функції $f(x)$ наближено описується формулою (5.9), а двічі диференційовної – формулою (5.10), яка більш точна.

Тепер сформулюємо необхідні та достатні умови екстремуму для диференційовних функцій багатьох змінних.

Теорема 5.5 (необхідні умови екстремуму першого порядку). Нехай:

- 1) x^* – точка локального екстремуму функції $f(x)$;
- 2) функція $f(x)$ диференційовна у точці x^* .

Тоді градієнт функції $f(x)$ у точці x^* є нульовим вектором, тобто

$$f'(x^*) = 0 \quad \text{або} \quad \frac{\partial f(x^*)}{\partial x_j} = 0, \quad j = \overline{1, n}. \quad (5.11)$$

◀ Зафіксувавши значення всіх змінних, за винятком, наприклад x_1 , будемо мати функцію однієї змінної x_1 , яка має у

точці x_1^* екстремум. Із теореми 1.2 випливає, що перша похідна цієї функції у точці x_1^* дорівнює нулеві, тобто $\frac{\partial f(x^*)}{\partial x_1} = 0$.

Аналогічно встановлюються інші рівності (5.11). ►

Точки, у яких виконуються умови (5.11), називаються стаціонарними точками диференційовної функції $f(x)$. Ці точки можуть бути точками екстремуму, а можуть і не бути ними, тобто умови (5.11) не достатні. Наприклад, функція двох змінних $f(x_1, x_2) = 2(x_1^2 - x_2^2)$ має стаціонарну точку $(x_1^*, x_2^*) = (0, 0)$, яка не є точкою локального екстремуму, оскільки у довільному околі цієї точки наявні точки $(0, x_2)$, у яких $f(0, x_2) < 0 = f(0, 0)$ (тобто $(0, 0)$ не є точкою локального мінімуму) і точки $(x_1, 0)$, у яких $f(x_1, 0) > 0 = f(0, 0)$ (тобто $(0, 0)$ не є точкою локального максимуму).

Для виявлення «зайвих» стаціонарних точок (як і у випадку функцій однієї змінної) використовуються необхідні умови екстремуму другого порядку.

Теорема 5.6 (необхідні умови екстремуму другого порядку). Нехай:

1) x^* – точка локального мінімуму (максимуму) функції $f(x)$;

2) функція $f(x)$ диференційовна у деякому околі точки x^* і двічі диференційовна у самій точці x^* .

Тоді другий диференціал $d^2 f(x^*)$ функції $f(x)$ у точці x^* невід'ємний (недодатний), тобто матриця Гессе $H(x^*) \equiv f''(x^*)$ невід'ємно (недодатно) визначена.

◀ Доведення теореми 5.6 аналогічне до доведення теореми 1.3. Потрібно скористатися формулою (5.8) (формулою Тейлора із залишковим членом у формі Пеано). ►

Нагадаємо, що функція

$$Q(h) = \sum_{i=1}^n \sum_{j=1}^n a_{ij} h_i h_j = \langle Ah, h \rangle, \quad (5.12)$$

у якій матриця $A = (a_{ij})_{i,j=1}^n$ – симетрична, називається квадратичною формою (квадратичною формою від змінних $\Delta x_i = dx_i$ ($i = \overline{1, n}$) є другий диференціал (5.5)). При цьому функція (5.12) називається *додатно визначеною (від’ємно визначеною)*, якщо $Q(h) > 0$ ($Q(h) < 0$) при всіх $h \neq 0$ ($Q(0) = 0$). Обидва випадки об’єднані під назвою *знаковизначених квадратичних форм*.

Квадратична форма (5.12) називається *квазізнаковизначеною*, якщо вона набуває або лише невід’ємних, або лише недодатних значень, але нульові значення має не лише при $h_1 = \dots = h_n = 0$. Наприклад, $Q(h_1, h_2) = h_1^2 + 2h_1 h_2 + h_2^2$ – квазізнаковизначена форма, оскільки $Q(h_1, h_2) \geq 0$ у всіх точках (h_1, h_2) , але $Q(h_1, h_2) = 0$ не лише у точці $(0, 0)$, але й у точках $(-1, 1)$, $(1, -1)$. Отже, квазізнаковизначені форми – це невід’ємно та недодатно визначені форми.

Якщо форма (5.12) має як додатні, так і від’ємні значення, то вона називається *знакозмінною, або невизначеною*.

Для матриці A визначники

$$\Delta_1 = a_{11}, \quad \Delta_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}, \dots, \Delta_n = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}$$

називаються *головними (або кутовими) мінорами*.

Теорема 5.7 (критерій Сильвестра). Нехай A – матриця квадратичної форми (5.12), а Δ_i ($i = \overline{1, n}$) – її головні мінори. Тоді форма (5.12) додатно визначена (від’ємно визначена) тоді і лише тоді, коли $\Delta_1 > 0, \dots, \Delta_n > 0$ ($\Delta_1 < 0, \Delta_2 > 0, \dots, (-1)^n \Delta_n > 0$).

Приклад 5.1. Установити, якими є квадратичні форми $Q(h_1, h_2) = 2h_1^2 + 3h_2^2$ та $Q(h_1, h_2) = 2h_1^2 - 2h_2^2$.

◀ Для першої форми $A = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}$, тобто $\Delta_1 = 2 > 0$ і $\Delta_2 = 6 > 0$. Згідно з критерієм Сильвестра, дана форма є додатно визначеною.

Матриця другої форми $A = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix}$, тому $\Delta_1 = 2 > 0$, $\Delta_2 = -4 < 0$ і за допомогою критерію Сильвестра не можна зробити висновок про характер квадратичної форми, але цей висновок можна зробити, проаналізувавши функцію. Оскільки $Q(0, h_2) = -2h_2^2 < 0$ ($h_2 \neq 0$), а $Q(h_1, 0) = 2h_1^2 > 0$ ($h_1 \neq 0$), то дана форма є знакозмінною (невизначеною). ▶

Теорема 5.8 (достатні умови екстремуму). Нехай:

- 1) x^* – стаціонарна точка функції $f(x)$, тобто розв’язок системи (5.11);
- 2) функція $f(x)$ диференційовна у деякому околі точки x^* і двічі диференційовна у самій точці x^* .

Якщо другий диференціал $d^2 f(x^*)$ функції $f(x)$ у точці x^* додатний (від’ємний), тобто матриця Гессе $H(x^*) \equiv f''(x^*)$ додатно (від’ємно) визначена, то x^* – точка локального мінімуму (максимуму) функції $f(x)$.

Якщо $d^2 f(x^*)$ є знакозмінною (невизначеною) квадратичною формою, то у точці x^* функція $f(x)$ не має екстремуму.

◀ Доведення цієї теореми також здійснюється з використанням формули Тейлора. ▶

Зауважимо, що у випадку, коли $d^2 f(x^*)$ є квазізнаковизначеною квадратичною формою, то $f(x)$ може мати у точці x^* локальний екстремум, а може і не мати його. Наприклад, для кожної з функцій $f(x_1, x_2) = x_1^4 + x_2^4$ і $f(x_1, x_2) = x_1^3 x_2^3$ у стаціонарній точці $(0, 0)$ $d^2 f(0, 0) = 0$, тобто кожна з них є квазізнаковизначеною формою, але перша функція має у точці $(0, 0)$ локальний мінімум, а друга у цій точці не має екстремуму.

Знаковизначеність другого диференціала (як квадратичної форми) та знаковизначеність матриці Гессе $f''(x^*) = H(x^*)$ – поняття рівносильні, тому на практиці, згідно з теоремою 5.8, перевіряють знаковизначеність матриці Гессе. Крім того, для перевірки достатніх умов оптимальності можна скористатися власними значеннями матриці Гессе (вони будуть дійсними числами, якщо всі елементи матриці є дійсними числами). Якщо власні значення $\lambda_i > 0$ ($i = \overline{1, n}$), то x^* – точка локального мінімуму, якщо $\lambda_i < 0$ ($i = \overline{1, n}$), то x^* – точка локального максимуму. У випадку, коли λ_i ($i = \overline{1, n}$) мають різні знаки, екстремуму немає, а у інших випадках (крім різних за знаком, зустрічаються також нульові власні значення) потрібні додаткові дослідження.

Для функції двох змінних достатні умови екстремуму

конкретизуються у наступному твердженні (наслідок теореми 5.8).

Теорема 5.9 (достатні умови екстремуму у випадку функції двох змінних). Нехай:

1) $x^* = (x_1^*, x_2^*)^T$ – стаціонарна точка функції $f(x_1, x_2)$, тобто розв’язок системи (5.11) при $n = 2$;

2) функція $f(x_1, x_2)$ диференційовна у деякому околі точки x^* і двічі диференційовна у самій точці x^* ;

$$3) \quad a_{11} = \frac{\partial^2 f(x_1^*, x_2^*)}{\partial x_1^2}, \quad a_{12} = \frac{\partial^2 f(x_1^*, x_2^*)}{\partial x_1 \partial x_2}, \quad a_{22} = \frac{\partial^2 f(x_1^*, x_2^*)}{\partial x_2^2},$$

$$\Delta(x_1^*, x_2^*) = a_{11}a_{22} - a_{12}^2.$$

Тоді:

1) якщо $\Delta(x_1^*, x_2^*) > 0$, то у точці $(x_1^*, x_2^*)^T$ функція $f(x_1, x_2)$ має локальний мінімум (максимум), якщо $a_{11} > 0$ ($a_{11} < 0$);

2) якщо $\Delta(x_1^*, x_2^*) < 0$, то у точці $(x_1^*, x_2^*)^T$ функція не має екстремуму;

3) якщо $\Delta(x_1^*, x_2^*) = 0$, то у точці $(x_1^*, x_2^*)^T$ функція може мати локальний екстремум, а може і не мати його.

Зауважимо, що як і у випадку одновимірної оптимізації, теореми про необхідні та достатні умови екстремуму функції багатьох змінних $f(x)$ часто формулюють, припускаючи, що $f(x)$ є неперервно диференційовною та двічі неперервно диференційовною функцією у деякому околі точки x^* . Такі вимоги до функції $f(x)$ сильніші (жорсткіші).

Підсумовуючи попередні твердження та факти, можна зробити такі висновки:

1) $d^2 f(x^*) > 0$ (матриця $H(x^*)$ додатно визначена) – x^* є точкою локального мінімуму;

2) $d^2 f(x^*) < 0$ (матриця $H(x^*)$ від'ємно визначена) – x^* є точкою локального максимуму;

3) $d^2 f(x^*) \geq 0$ (матриця $H(x^*)$ невід'ємно визначена, тобто всі її головні мінори невід'ємні) – x^* може бути точкою локального мінімуму, потрібні додаткові дослідження;

4) $d^2 f(x^*) \leq 0$ (матриця $H(x^*)$ недодатно визначена, тобто всі її мінори парного порядку невід'ємні, а непарного – недодатні) – x^* може бути точкою локального максимуму, потрібні додаткові дослідження;

5) $d^2 f(x^*) = 0$ (матриця $H(x^*)$ складається із нульових елементів) – для з'ясування характеру точки x^* потрібні додаткові дослідження;

6) $d^2 f(x^*)$ – знаковмінна квадратична форма (жодна з умов 1)-5) не виконується), тому у точці x^* немає екстремуму.

Отже, теореми 5.5 і 5.8 (чи 5.5 і 5.9) є теоретичною основою класичного методу мінімізації функції багатьох змінних, алгоритм якого складається з двох основних кроків.

Крок 1. Розв'язування системи (5.11) і знаходження стаціонарних точок функції $f(x)$.

Крок 2. Дослідження стаціонарних точок на екстремум за допомогою достатніх умов і визначення точок глобального мінімуму та відповідного значення функції або встановлення відсутності таких точок.

Приклад 5.2. Класичним методом розв'язати задачу мінімізації функції $f(x_1, x_2) = -x_1^3 + x_2^3 + 3x_1x_2$ на $\mathbb{R}^2$.

◀ *Крок 1.* $\frac{\partial f(x_1, x_2)}{\partial x_1} = -3x_1^2 + 3x_2$, $\frac{\partial f(x_1, x_2)}{\partial x_2} = 3x_2^2 + 3x_1$, тому,

розв'язавши систему (5.11), тобто систему рівнянь $-3x_1^2 + 3x_2 = 0$, $3x_2^2 + 3x_1 = 0$, отримаємо дві стаціонарні точки

$(0,0)^T, (-1,1)^T$.

Крок 2. Знайдемо матрицю Гессе у стаціонарних точках. Оскільки $\partial^2 f(x_1, x_2)/\partial x_1^2 = -6x_1$, $\partial^2 f(x_1, x_2)/\partial x_1 \partial x_2 = 3$, $\partial^2 f(x_1, x_2)/\partial x_2^2 = 6x_2$, то гессіан у точках $(0,0)^T, (-1,1)^T$ визначається відповідно матрицями $\begin{pmatrix} 0 & 3 \\ 3 & 0 \end{pmatrix}, \begin{pmatrix} 6 & 3 \\ 3 & 6 \end{pmatrix}$.

Отже, $\Delta(0,0) = -9 < 0$, а $\Delta(-1,1) = 27 > 0$ і $a_{11} = 6 > 0$, тому у точці $(0,0)^T$ немає локального екстремуму, а точка $(-1,1)^T$ є точкою мінімуму. При цьому $f^* = f(-1,1) = -1$. ►

Приклад 5.3. Нехай задані точки площини (x_i, y_i) $i = \overline{1,5}$ (табл. 5.1), де x_i – значення незалежної змінної x , а y_i – відповідне йому значення (як правило, наближене) залежної змінної y . Побудувати для цих точок оптимальну апроксимуючу функцію $g(x) = a + bx$ (a, b – параметри), використавши критерій мінімізації суми квадратів відхилень значень $g(x_i)$ від заданих y_i ($i = \overline{1,5}$).

Таблиця 5.1

x_i	1	2	3	4	5
y_i	1	3	2	4	4,5

◄ Очевидно, критерій мінімізації формалізується залежною від параметрів a та b функцією

$$F(a, b) = \sum_{i=1}^5 (a + bx_i - y_i)^2,$$

а задача побудови функції $g(x)$ або задача її параметризації (ідентифікації) – задачею безумовної мінімізації

$$F(a, b) \rightarrow \min, (a, b)^T \in \mathbb{R}^2.$$

Остання задача відображає сутність так званого *методу найменших квадратів*, який широко використовується при розв'язуванні задач апроксимації.

Отриману задачу безумовної мінімізації розв'яжемо класичним методом.

Крок 1. Система (5.11) має вигляд

$$\begin{cases} \frac{\partial F}{\partial a} = 2 \sum_{i=1}^5 (a + bx_i - y_i) = 0, \\ \frac{\partial F}{\partial b} = 2 \sum_{i=1}^5 (a + bx_i - y_i)x_i = 0 \end{cases}$$

або

$$\begin{cases} 5a + \left(\sum_{i=1}^5 x_i \right) b = \sum_{i=1}^5 y_i, \\ \left(\sum_{i=1}^5 x_i \right) a + \left(\sum_{i=1}^5 x_i^2 \right) b = \sum_{i=1}^5 x_i y_i. \end{cases}$$

Її розв'язок такий: $a = 0,5$; $b = 0,8$.

Крок 2. Матриця других похідних у точці $(a,b)^T = (0,5;0,8)^T$

$$f''(a,b) = \begin{pmatrix} 10 & 2 \sum_{i=1}^5 x_i \\ 2 \sum_{i=1}^5 x_i & 2 \sum_{i=1}^5 x_i^2 \end{pmatrix} = \begin{pmatrix} 10 & 30 \\ 30 & 110 \end{pmatrix}$$

додатно визначена $(\Delta_1 = 10 > 0, \Delta_2 = 200 > 0)$, тому $(0,5;0,8)^T$ – точка безумовного мінімуму функції $F(a,b)$ і $g(x) = 0,5 + 0,8b$. ►

5.3. Поняття опуклої задачі оптимізації

Для того, щоб з'ясувати зміст опуклої задачі оптимізації функції багатьох змінних, нагадаємо деякі базові поняття з теорії опуклих множин і опуклих функцій (теорії опуклого аналізу). При цьому надалі будемо вважати, що X – це множина точок евклідового простору $\mathbb{R}^n$ ($X \subset \mathbb{R}^n$ або $X = \mathbb{R}^n$), а $f(x)$ – визначена на X функція ($x \in X$).

Означення 5.7. Множина X називається опуклою, якщо для довільних $x^{(1)}, x^{(2)} \in X$ і числа $\alpha \in [0, 1]$ точка $x = \alpha x^{(1)} + (1 - \alpha)x^{(2)} \in X$. Інакше кажучи, множина X опукла, якщо разом із довільними двома точками $x^{(1)}, x^{(2)} \in X$ вона повністю містить відрізок, що з'єднує ці точки.

Очевидно, простір $\mathbb{R}^n$ – опукла множина. Прийнято також вважати опуклими порожню множину і множину, що складається з однієї точки. Тоді з означення 5.7 випливає, що перетин довільного числа опуклих множин також буде опуклою множиною.

Означення 5.8. 1) Визначена на опуклій множині X функція $f(x)$, називається опуклою, якщо для довільних $x^{(1)}, x^{(2)} \in X$ і числа $\alpha \in [0, 1]$ виконується нерівність

$$f(\alpha x^{(1)} + (1 - \alpha)x^{(2)}) \leq \alpha f(x^{(1)}) + (1 - \alpha)f(x^{(2)}). \quad (5.13)$$

2) Якщо при всіх $\alpha \in (0, 1)$ нерівність (5.13) строга, то $f(x)$ називається строго опуклою.

3) Функція $f(x)$ називається угнутою (строго угнутою), якщо функція $(-f(x))$ опукла (строго опукла).

У плані геометричної інтерпретації опуклості функції $f(x)$ означає, що будь-яка точка довільної хорди графіка $f(x)$ розміщена не нижче відповідної точки самого графіка (для

угнутої функції навпаки). Із означення 5.8 випливає також, що лінійна функція одночасно опукла й угнута, але не строго.

Довільна опукла на $\mathbb{R}^n$ функція $f(x)$ неперервна, а довільна опукла на $X \subset \mathbb{R}^n$ функція $f(x)$ ($X \neq \mathbb{R}^n$) неперервна у кожній внутрішній точці цієї множини [12].

Наведемо також інші факти з опуклого аналізу, які відіграють важливу роль у теорії оптимізації.

Означення 5.9. *Задача мінімізації (максимізації) функції $f(x)$ на опуклій множині X називається опуклою задачею мінімізації (максимізації), якщо $f(x)$ опукла (угнута) на X функція.*

Теорема 5.10. *Лінійна комбінація опуклих на опуклій множині X функцій $f_i(x)$ ($i = \overline{1, k}$), тобто $f(x) = \sum_{i=1}^k \beta_i f_i(x)$, є опуклою функцією на X , якщо $\beta_i \geq 0$ ($i = \overline{1, k}$).*

Теорема 5.11. *Якщо задана на $\mathbb{R}^n$ функція $g(x)$ опукла, то опуклою буде множина $\{x \in \mathbb{R}^n / g(x) \leq b\}$ ($b \in \mathbb{R}$).*

Теорема 5.12. *Якщо задача мінімізації опукла, то будь-який її локальний розв'язок є також глобальним розв'язком, тобто локальний мінімум цільової функції збігається з її глобальним мінімумом.*

Теорема 5.13. *Якщо функція $f(x)$ опукла на $\mathbb{R}^n$ і диференційовна у точці $x^* \in \mathbb{R}^n$, у якій виконуються необхідні умови екстремуму першого порядку, то x^* – розв'язок задачі безумовної мінімізації (5.1).*

Теорема 5.14. *Множина X^* розв'язків опуклої задачі оптимізації опукла, причому у випадку строго опуклої функції $f(x)$ вона складається з однієї точки.*

Ілюстрація теореми 5.14 подана на рис. 5.1.

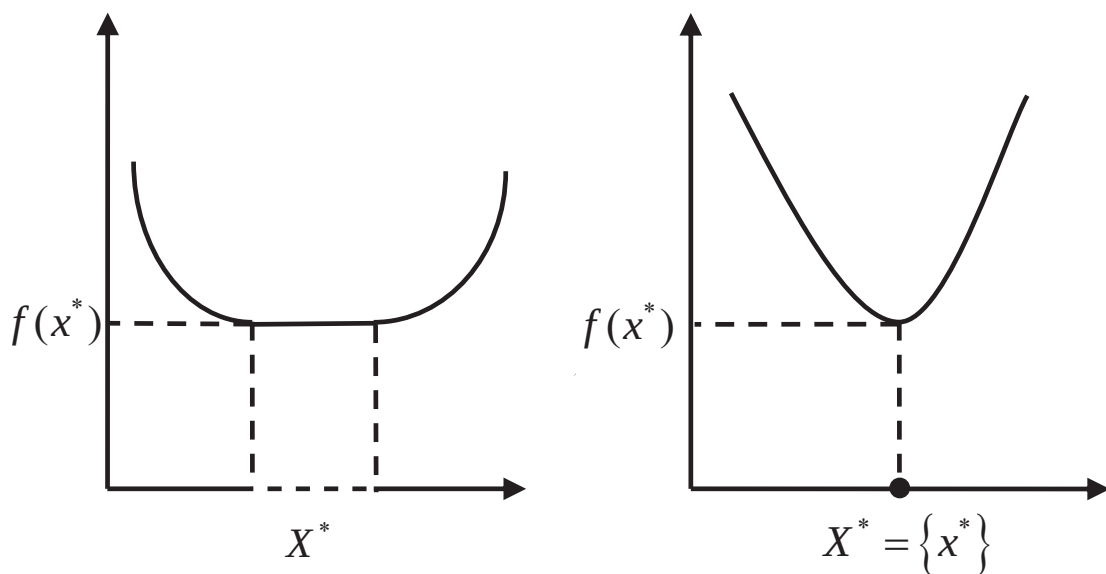


Рис. 5.1. Ілюстрація опуклості множини X^*

У тих випадках, коли опукла задача оптимізації не має розв'язку, X^* є порожньою множиною ($X^* = \emptyset$), яку також зручно вважати опуклою (наприклад, для функції однієї змінної $f(x) = x$ при $x \in (0,1)$ і $f(0) = f(1) = 2$ $X^* = \emptyset$).

Сформулюємо диференціальні критерії опуклості для гладких функцій багатьох змінних [12, 14], а саме: для неперервно диференційовних та двічі неперервно диференційовних функцій. Ці критерії є логічними узагальненнями відповідних критеріїв для функції однієї змінної (теореми 1.5 і 1.6).

Терема 5.15. Нехай $f(x)$ неперервно диференційовна функція на опуклій множині X . Для того, щоб $f(x)$ була опуклою на X , необхідно й достатньо, щоб для всіх $x, x^* \in X$ виконувалася нерівність

$$f(x) \geq f(x^*) + \langle f'(x^*), x - x^* \rangle. \quad (5.14)$$

Опираючись на теорему 5.15 і нерівність (5.14), можна встановити критерій опуклості у термінах перших похідних.

Теорема 5.16. Нехай $f(x)$ неперервно диференційовна функція на опуклій множині X . Для того, щоб функція $f(x)$ була опуклою на X , необхідно й достатньо, щоб для всіх $x, x^* \in X$ виконувалася нерівність

$$\langle f'(x) - f'(x^*), x - x^* \rangle \geq 0. \quad (5.15)$$

Критерій опуклості $f(x)$, що базується на використанні других похідних, конкретизує наступна теорема.

Теорема 5.17. Нехай $f(x)$ двічі неперервно диференційовна функція на опуклій множині X , для якої множина внутрішніх точок $\text{int } X \neq \emptyset$. Для того, щоб $f(x)$ була опуклою на X , необхідно й достатньо, щоб для всіх $x^* \in X$ і $h \in \mathbb{R}^n$ виконувалася нерівність

$$\langle f''(x^*)h, h \rangle \geq 0. \quad (5.16)$$

Зміст нерівностей (5.15) і (5.16) легко зрозуміти при $n=1$. Для функції однієї змінної умова (5.15) означає, що похідна $f'(x)$ монотонно не спадає на X , а умова (5.16) – невід’ємність другої похідної на X . Для функції багатьох змінних (5.16) – це умова невід’ємної визначеності матриці других похідних $f''(x^*)$, яку широко використовують для перевірки опуклості функцій.

Зауважимо також, що у теоремі 5.17 умова $\text{int } X \neq \emptyset$ суттєва, у чому можна переконатися, аналізуючи функцію $f(x_1, x_2) = x_1^2 - x_2^2$ на $X = \{(x_1, x_2) \in \mathbb{R}^2 / x_2 = 0\}$. Дана функція опукла на X , але умова $\langle f''(x^*)h, h \rangle = 2h_1^2 - 2h_2^2 \geq 0$ не виконується, наприклад для $h = (0, 1)^T$. Це пов’язано з тим, що тут $\text{int } X = \emptyset$, тобто немає жодної точки $x \in X$, яка б належала множині X разом з деяким її оточенням.

Ще одне важливе зауваження стосовно теорем 5.15–5.17. Якщо в умовах (5.14) і (5.15) замінити знак

нестрогої нерівності « $\geq$ » на знак строгої нерівності « $>$ », то при $x, x^* \in X$ ($x \neq x^*$) отримаємо необхідні й достатні умови строгої опуклості функції $f(x)$ на X . Теорема 5.17 переноситься лише частково: умова $\langle f''(x^*)h, h \rangle > 0$ при всіх $x^* \in X, h \in \mathbb{R}^n$ ($h \neq 0$) є лише достатньою для строгої опуклості $f(x)$ на X . На прикладі функції $f(x) = x^4$ ($x \in \mathbb{R}$) легко переконатися, що обернене твердження хибне (функція опукла, але у точці $x=0$ її друга похідна $f''(x) = 12x^2$ є нульовою, а не додатною).

Приклад 5.4. Визначити, при яких a, b, c функція $f(x) = x_1^2 + 2ax_1x_2 + bx_2^2 + cx_3^2$ буде опуклою на $\mathbb{R}^3$.

◀ Матриця других похідних

$$f''(x) = \begin{pmatrix} 2 & 2a & 0 \\ 2a & 2b & 0 \\ 0 & 0 & 2c \end{pmatrix},$$

тому $\Delta_1 = 2$, $\Delta_2 = 4(b - a^2)$, $\Delta_3 = 8c(b - a^2)$.

Отже, згідно з критерієм [1] (для невід'ємної визначеності квадратичної форми або відповідної їй симетричної матриці необхідно й достатньо, щоб всі головні мінори цієї матриці були невід'ємними), матриця $f''(x)$ невід'ємно визначена при $b \geq a^2$ і $c \geq 0$. Ці умови забезпечують опуклість функції $f(x)$ на $\mathbb{R}^3$. ►

У теорії оптимізації важливу роль відіграє ще один підклас опуклих функцій – так звані *сильно опуклі функції*, для яких мінімум у $\mathbb{R}^n$ обов'язково існує (як зазначалося раніше не для всіх опуклих функцій мінімум існує, наприклад $f(x) = e^x$ опукла у просторі $\mathbb{R}^1$, але мінімум для неї не досягається).

Означення 5.10. Визначена на опуклій множині X функція $f(x)$ називається сильно опуклою з константою $\theta > 0$ на X , якщо для довільних $x^{(1)}, x^{(2)} \in X$ і числа $\alpha \in [0, 1]$ виконується нерівність

$$f(\alpha x^{(1)} + (1 - \alpha)x^{(2)}) \leq \alpha f(x^{(1)}) + (1 - \alpha)f(x^{(2)}) - \alpha(1 - \alpha)\theta \|x^{(1)} - x^{(2)}\|^2. \quad (5.17)$$

Очевидно, сильно опукла функція є строго опуклою (а значить, і опуклою), але не навпаки. Можна довести, що сильно опукла функція завжди має єдину точку глобального мінімуму.

Необхідні та достатні умови (критерії) сильної опуклості для гладких функцій з константою $\theta > 0$ формулюються у теоремах, аналогічних до теорем 5.15 – 5.17.

Теорема 5.18. Нехай $f(x)$ неперервно диференційовна функція на опуклій множині X . Для того, щоб функція $f(x)$ була сильно опуклою на X , необхідне й достатнє існування такої константи $\theta > 0$, щоб для всіх $x, x^* \in X$ виконувалася нерівність

$$f(x) \geq f(x^*) + \langle f'(x^*), x - x^* \rangle + \theta \|x - x^*\|^2. \quad (5.18)$$

Теорема 5.19. Нехай $f(x)$ неперервно диференційовна функція на опуклій множині X . Для того, щоб функція $f(x)$ була сильно опуклою на X , необхідне й достатнє існування такої константи $\theta > 0$, щоб для всіх $x, x^* \in X$ виконувалася нерівність

$$\langle f'(x) - f'(x^*), x - x^* \rangle \geq 2\theta \|x - x^*\|^2. \quad (5.19)$$

Теорема 5.20. Нехай $f(x)$ двічі неперервно диференційовна функція на опуклій множині X , для якої множина точок $\text{int } X \neq \emptyset$. Для того, щоб функція $f(x)$ була

сильно опуклою на X , необхідне й достатнє існування такої константи $\theta > 0$, щоб для всіх $x^* \in X$ і $h \in \mathbb{R}^n$ виконувалась нерівність

$$\langle f''(x^*)h, h \rangle \geq 2\theta \|h\|^2. \quad (5.20)$$

Як у означенні 5.10, так і у нерівностях (5.18)–(5.20), константа θ не залежить від точок множини X . Із (5.20) випливає, що критерій сильної опуклості функції $f(x)$ еквівалентний додатній визначеності матриці $\tilde{H} = (f''(x^*) - 2\theta I)$, де I – діагональна одинична матриця розмірності $(n \times n)$.

Приклад 5.5. Показати, що функція $f(x) = \|x\|^2 = x_1^2 + \dots + x_n^2$ сильно опукла.

◀ Матрица

$$\tilde{H} = (f''(x^*) - 2\theta I) = \begin{pmatrix} 2(1-\theta) & 0 & 0 \dots 0 \\ 0 & 2(1-\theta) & 0 \dots 0 \\ \dots & \dots & \dots \\ 0 & 0 & 0 \dots 2(1-\theta) \end{pmatrix},$$

тому існує константа θ ($\theta \in (0,1)$), при якій всі головні мінори матриці $\tilde{H}$ будуть додатними, а матриця $\tilde{H}$ – додатно визначеною. ►

У теорії n -вимірної оптимізації важливу роль відіграють опуклі квадратичні функції вигляду

$$f(\mathbf{x}) = \sum_{i=1}^n \sum_{j=1}^n \alpha_{ij} x_i x_j + \sum_{j=1}^n b_j x_j + c, \quad (5.21)$$

де $\alpha_{ij} (i, j = \overline{1, n}), b_j (j = \overline{1, n}), c$ - деякі дійсні числа.

Увівши симетричну матрицю $A = (a_{ij})_{i,j=1}^n$, де $a_{ij} = \alpha_{ij} + \alpha_{ji}$, функцію (5.21) перепишемо у вигляді

$$f(x) = \frac{1}{2} \langle Ax, x \rangle + \langle b, x \rangle + c, \quad (5.22)$$

де $b = (b_1, \dots, b_n)^T \in \mathbb{R}^n$.

Уточнимо основні властивості функції (5.22):

- 1) градієнт $f'(x) = Ax + b$;
- 2) матриця Гессе збігається з матрицею A : $f''(x) = A$;
- 3) якщо A додатно визначена, то функція (5.22) сильно опукла, тому вона має єдину точку глобального мінімуму.

5.4. Деякі прямі методи безумовної мінімізації

Як і у випадку мінімізації функції однієї змінної прямі методи (методи нульового порядку) мінімізації функцій багатьох змінних (задача (5.1)) використовують тільки значення функцій, причому для їх застосування не вимагається диференційовність функції чи її аналітичне задання. Достатньо лише мати можливість обчислювати (вимірювати) значення функції $f(x)$ у будь-якій точці області визначення. Нижче розглянемо деякі з прямих методів безумовної багатовимірної мінімізації, які базуються на обчислювальних процедурах $x^{(k+1)} = x^{(k)} + \alpha_k p^{(k)}$ ($k = 0, 1, 2, \dots$), які залежно від способу вибору напрямку $p^{(k)}$ та кроку α_k приводять до різних алгоритмів мінімізації.

5.4.1. Метод циклічного покоординатного спуску

Ідейною платформою цього методу є послідовна мінімізація функції $f(x)$ спочатку вздовж першого базисного вектора $e^{(1)}$, потім другого – $e^{(2)}$ і т.д. ($e^{(j)}$ –

n -вимірний вектор, у якого j -та координата дорівнює одиниці, а всі інші нулеві). Закінчивши мінімізацію вздовж останнього базисного вектора $e^{(n)}$, цикл знову повторюють. Алгоритм реалізації методу циклічного покоординатного спуску конкретизується такими кроками.

Крок 1. Вибір початкової точки $x \in \mathbb{R}^n$, точності наближення ε та критерію її досягнення, знаходження $f(x)$, присвоєння $j := 1$.

Крок 2. Розв'язування одновимірної задачі мінімізації $\Phi(\alpha) = f(x + \alpha e^{(j)}) \rightarrow \min_{\alpha} (\alpha \in \mathbb{R}^1)$, знаходження відповідних α^* , $\tilde{x} = x + \alpha^* e^{(j)}$, $f(\tilde{x})$.

Крок 3. Порівняння j з n . Якщо $j < n$, то здійснюється переприсвоєння $x := \tilde{x}$, $j := j + 1$ та перехід до кроку 2. Якщо $j \geq n$, то здійснюється перехід до кроку 4.

Крок 4. Перевірка умови досягнення точності $\|x - \tilde{x}\| < \varepsilon$ або $|f(x) - f(\tilde{x})| < \varepsilon$. Якщо вона виконується, то $x^* \approx \tilde{x}$, $f^* \approx f(\tilde{x})$ і пошук завершується. У протилежному випадку здійснюється переприсвоєння $x := \tilde{x}$, $f(x) := f(\tilde{x})$, $j := 1$ та перехід до кроку 2.

Зауважимо, що допоміжну одновимірну задачу мінімізації (крок 2) можна розв'язати одним із методів одновимірної оптимізації, зокрема методом перебору чи його модифікаціями. Зазначимо також, що ефективність методу циклічного покоординатного спуску суттєво залежить від властивостей функції $f(x)$.

Приклад 5.6. Методом циклічного покоординатного спуску розв'язати задачу безумовної мінімізації функції $f(x) = x_1^2 + x_2^2$.

◀ Очевидно, лініями рівня функції $f(x)$ є кола з центром у початку координат $(0,0)^T$. Виберемо довільну початкову точку, наприклад $x^{(0)} = (4,4)^T$. Тоді для досягнення мети потрібно зробити два спуски – спочатку вздовж напрямку $e^{(1)}$, а потім – $e^{(2)}$, які власне кажучи і приведуть у точку мінімуму $x^* = x^{(2)} = (0,0)^T$ (рис. 5.2). ▶

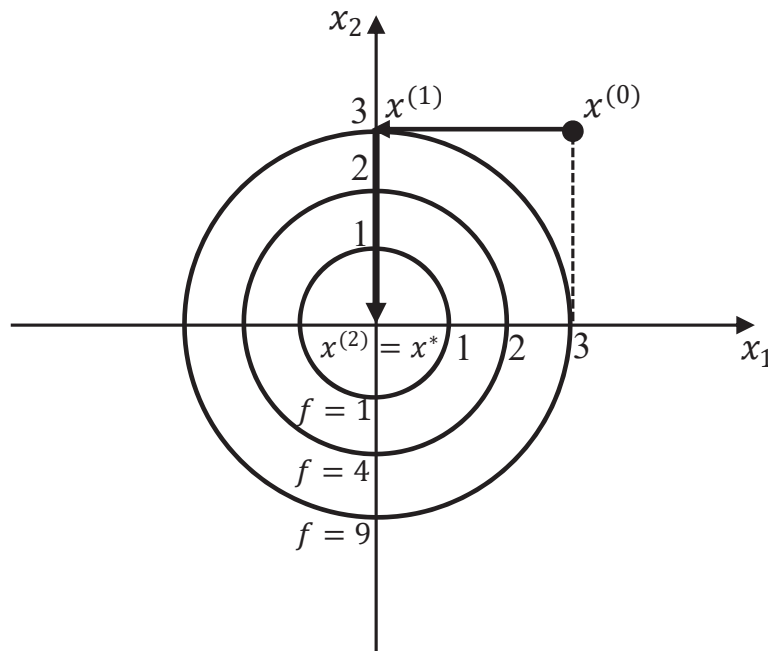


Рис. 5.2. Ілюстрація до прикладу 5.6

Ситуація, коли точку мінімуму функції вдається знайти точно за скінченне число кроків (як у прикладі 5.6), є винятком. У більшості випадків це не так.

Приклад 5.7. Використовуючи метод циклічного покоординатного спуску, розв'язати задачу (5.1) у випадку, коли $n = 2$, $f(x) = 5x_1^2 + 5x_2^2 + 8x_1x_2$.

◀ Оскільки

$$\begin{aligned} 5x_1^2 + 5x_2^2 + 8x_1x_2 &= (10x_1^2 + 10x_2^2 + 16x_1x_2)/2 = \\ &= [(x_1^2 + x_2^2 - 2x_1x_2) + (9x_1^2 + 9x_2^2 + 18x_1x_2)]/2 = \end{aligned}$$

$$= \frac{(x_1 - x_2)^2}{2} + \frac{9(x_1 + x_2)^2}{2},$$

то зробивши заміну змінних $y_1 = \frac{x_1 - x_2}{\sqrt{2}}$, $y_2 = \frac{x_1 + x_2}{\sqrt{2}}$ або

$$x_1 = \frac{y_1 + y_2}{\sqrt{2}}, \quad x_2 = \frac{y_2 - y_1}{\sqrt{2}} \quad (\text{ця заміна рівносильна повороту}$$

системи координат на кут (-45°) , тобто на 45° за годинниковою стрілкою, оскільки старі та нові координати зв'язані формулами $x_1 = y_1 \cos(-45^\circ) - y_2 \sin(-45^\circ)$,

$x_2 = y_1 \sin(-45^\circ) + y_2 \cos(-45^\circ)$), від функції $f(x)$ прийдемо до функції $\varphi(y) = y_1^2 + 9y_2^2$. Лінії рівня функції $\varphi(y)$ – еліпси $y_1^2/9 + y_2^2 = c^2$ або $y_1^2/9c^2 + y_2^2/c^2 = 1$ (рис. 5.3).

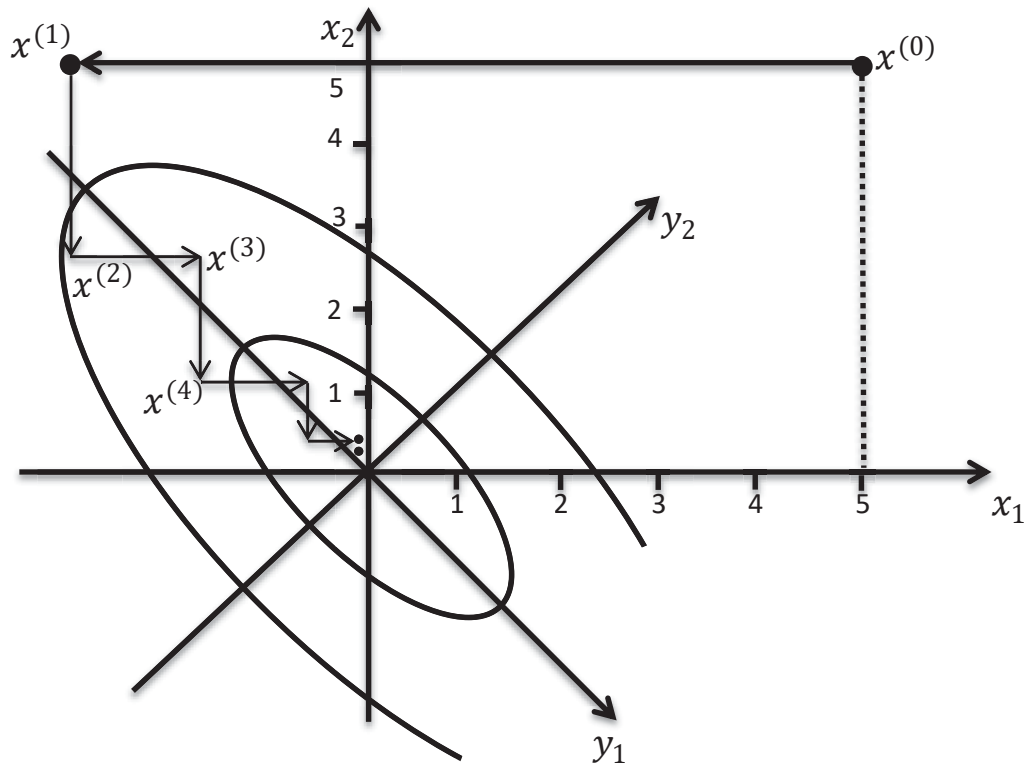


Рис. 5.3. Ілюстрація до прикладу 5.7

Процес знаходження розв'язку задачі здійснюється згідно з описаним вище алгоритмом.

Крок 1. Виберемо початкову точку $x^{(0)} = (5; 5)^T$, ε , критерій близькості точок, $j = 1$. Тоді $f(5; 5) = 450$.

Крок 2. Оскільки $e^{(j)} = e^{(1)}$, то $x + \alpha e^{(1)} = (5; 5)^T + \alpha(1; 0)^T = (5 + \alpha; 5)^T$ і $f(x + \alpha e^{(1)}) = \alpha^2/2 + 9(10 + \alpha)^2/2$. Мінімальне значення цієї функції досягається при $\alpha^* = -9$, тому $\tilde{x} = x - 9e^{(1)} = (5; 5)^T - 9(1; 0)^T = (-4; 5)^T$, $f(\tilde{x}) = 45$.

Крок 3. $j < n = 2$, тому $x := \tilde{x} = (-4; 5)^T$, $j := 2$ та повертаємося до кроку 2. Тоді $e^{(j)} = e^{(2)}$, $x + \alpha e^{(2)} = (-4; 5)^T + \alpha(0; 1)^T = (-4; 5 + \alpha)^T$, $f(x + \alpha e^{(2)}) = (-9 - \alpha)^2/2 + 9(1 + \alpha)^2/2$. Шукане значення $\alpha^* = -1,8$, тому $\tilde{x} = x - e^{(2)} = (-4; 5)^T - 1,8(0; 1)^T = (-4; 3,2)^T$, $f(\tilde{x}) = 28,8$. При цьому $j = 2 \geq 2$, тому переходимо до кроку 4.

Крок 4. Оскільки $\|x - \tilde{x}\| = \sqrt{(-4 + 4)^2 + (3,2 - 5)^2} = 1,8 > \varepsilon$, то, здійснивши переприсвоєння $x := \tilde{x}$, $f(x) := f(\tilde{x})$, $j := 1$, переходимо до кроку 2 і наступної ітерації.

Результати знаходження точки мінімуму відображені у табл. 5.2.

Таблиця 5.2

Номер Ітерації	$(x_1; x_2)$	$f(x)$
0	$(-5; 5)$	450
1	$(-4; 5)$	45
2	$(-4; 3,2)$	28,8
3	$(-2,56; 3,2)$	18,43
4	$(-2,56; 2,05)$	11,8
5	$(-1,64; 2,05)$	7,55
6	$(-1,64; 1,31)$	4,83
7	$(-1,05; 1,31)$	3,09
8	$(-1,05; 0,84)$	1,98
9	$(-0,67; 0,84)$	1,27
10	$(-0,67; 0,54)$	0,81

Очевидно, що мінімізуюча послідовність $\{x^{(k)}\}_1^\infty$ збігається до точки $x^* = (0; 0)^T$, хоча тут немає гарантії, що ця точка буде досягнута за скінченне число кроків. ►

Аналізуючи приклад 5.7, нескладно переконатися, що ефективність методу циклічного покоординатного спуску можна підвищити, періодично повторюючи пошук точки мінімуму у напрямку $p^{(i)} = x^{(i)} - x^{(i-2)}$ з точки $x^{(i)}$. При цьому здійснюється вичерпуючий або оптимальний спуск, тобто спуск, при якому величина кроку α_i в ітераційному процесі $x^{(i+1)} = x^{(i)} + \alpha_i p^{(i)}$, $i = 0, 1, \dots$ є розв'язком одновимірної задачі мінімізації $f(x^{(i)} + \alpha_i p^{(i)}) \rightarrow \min_\alpha$. Наприклад, якщо у прикладі 5.7 з точки $x^{(4)}$ здійснити вичерпуючий спуск у напрямку $p^{(4)} = x^{(4)} - x^{(2)}$, то отримаємо точку, яка розміщена до точки мінімуму $(0, 0)$ значно ближче, ніж точки $x^{(5)}$, $x^{(6)}$, $x^{(7)}$. Ідея послідовного знаходження напрямків спадання функції та її

мінімізація вздовж цих напрямків основоположна для багатьох методів мінімізації, зокрема у методі Хука-Джівса.

5.4.2. Метод Хука-Джівса

В основу цього методу закладено дві процедури:

1) покоординатний пошук в околі даної точки з метою визначення напрямку спадання $f(x)$;

2) переміщення у напрямку спадання.

Алгоритм використання покоординатного пошуку із точки x і приростами Δ_j ($j = \overline{1, n}$) полягає у виконанні таких кроків.

Крок 1. Присвоєння $\bar{x} := x$, $j := 1$.

Крок 2. Аналіз пробного кроку $y = \bar{x} - \Delta_j e^{(j)}$, де $e^{(j)}$ – j -тий базисний вектор. Якщо $f(\bar{x}) \leq f(y)$, то здійснюється перехід до кроку 3, у протилежному випадку – до кроку 4.

Крок 3. Аналіз пробного кроку $y = \bar{x} + \Delta_j e^{(j)}$. Якщо $f(\bar{x}) \leq f(y)$, то здійснюється перехід до кроку 5, у протилежному випадку – до кроку 4.

Крок 4. Присвоєння $\bar{x} := y$.

Крок 5. Присвоєння $j := j + 1$. Якщо $j \leq n$, то здійснюється перехід до кроку 2, у протилежному випадку точка $\bar{x}$, для якої $f(\bar{x}) < f(x)$ знайдена ($\bar{x} \neq x$).

Стосовно цього алгоритму зазначимо, що його результатом може також бути точка $\bar{x} = x$. У цій ситуації пошук вважається невдалим. Проте, якщо $\|\Delta\| < \varepsilon$, де ε – задана точність, $\Delta = (\Delta_1, \dots, \Delta_n)^T$, то вважають, що $x^* = x$. Якщо задана точність не досягнута, то прирости Δ_j зменшують до

Δ_j/γ ($j=\overline{1,n}$; $\gamma > 1$ – сталий коефіцієнт) і повторюють пошук, повертаючись до початку даного алгоритму.

Реалізація обох процедур приводить до повного алгоритму (методу) Хука-Джівса.

Крок 1. Вибір початкової точки $x^{(0)} \in \mathbb{R}^n$, вектора приростів $\Delta = (\Delta_1, \dots, \Delta_n)^T$, коефіцієнта зменшення кроку $\gamma > 1$, параметра завершення пошуку $\varepsilon > 0$.

Крок 2. Проведення покоординатного пошуку із точки $x^{(0)}$ та знаходження точки $\bar{x}^{(0)}$. Якщо $\bar{x}^{(0)} \neq x^{(0)}$, то здійснюється перехід до кроку 4, у протилежному випадку – до кроку 3.

Крок 3. Перевірка умови $\|\Delta\| < \varepsilon$. При її виконанні пошук завершується і $x^* = x^{(0)}$, а при її невиконанні зменшуються прирости Δ_j до Δ_j/γ ($j=\overline{1,n}$) та здійснюється перехід до кроку 2.

Крок 4. Переміщення із точки $\bar{x}^{(0)}$ у напрямку спадання $(\bar{x}^{(0)} - x^{(0)})$: $x^{(1)} = \bar{x}^{(0)} + (\bar{x}^{(0)} - x^{(0)}) = 2\bar{x}^{(0)} - x^{(0)}$.

Крок 5. Проведення покоординатного пошуку в точці $x^{(1)}$ і знаходження точки $\bar{x}^{(1)}$. Якщо $f(\bar{x}^{(1)}) < f(\bar{x}^{(0)})$, то здійснюють переприсвоєння $x^{(0)} := \bar{x}^{(0)}$, $\bar{x}^{(0)} := \bar{x}^{(1)}$ і перехід до кроку 4, інакше – $x^{(0)} := \bar{x}^{(1)}$ і переходять до кроку 2.

Приклад 5.8. Проілюструвати процес розв’язування задачі $f(x) = (x_1 + 1)^2 + x_2^2 \rightarrow \min$ методом Хука-Джівса із початкової точки $x^{(0)} = (2, 3)^T$ і вектором переміщень $\Delta = (1/2, 1)^T$.

◀ Очевидно, функція $f(x)$ є квадратом відстані між точками $x = (x_1, x_2)^T$ і $x^* = (-1, 0)^T$ (у зв’язку з цим значення

функції $f(x)$ у проміжних точках алгоритму легко порівнювати, оцінюючи їх відстані від x^*). Лініями рівня є кола з центром у точці x^* . На рис. 5.4 пунктиром виділені траєкторії покоординатного пошуку напрямків спадання функції, а суцільні лінії – це безпосереднє переміщення у напрямку спадання. Нескладно переконатися, що графічна ілюстрація (рис. 5.4) адекватно відображає процес розв'язування задачі методом Хука-Джівса. ►

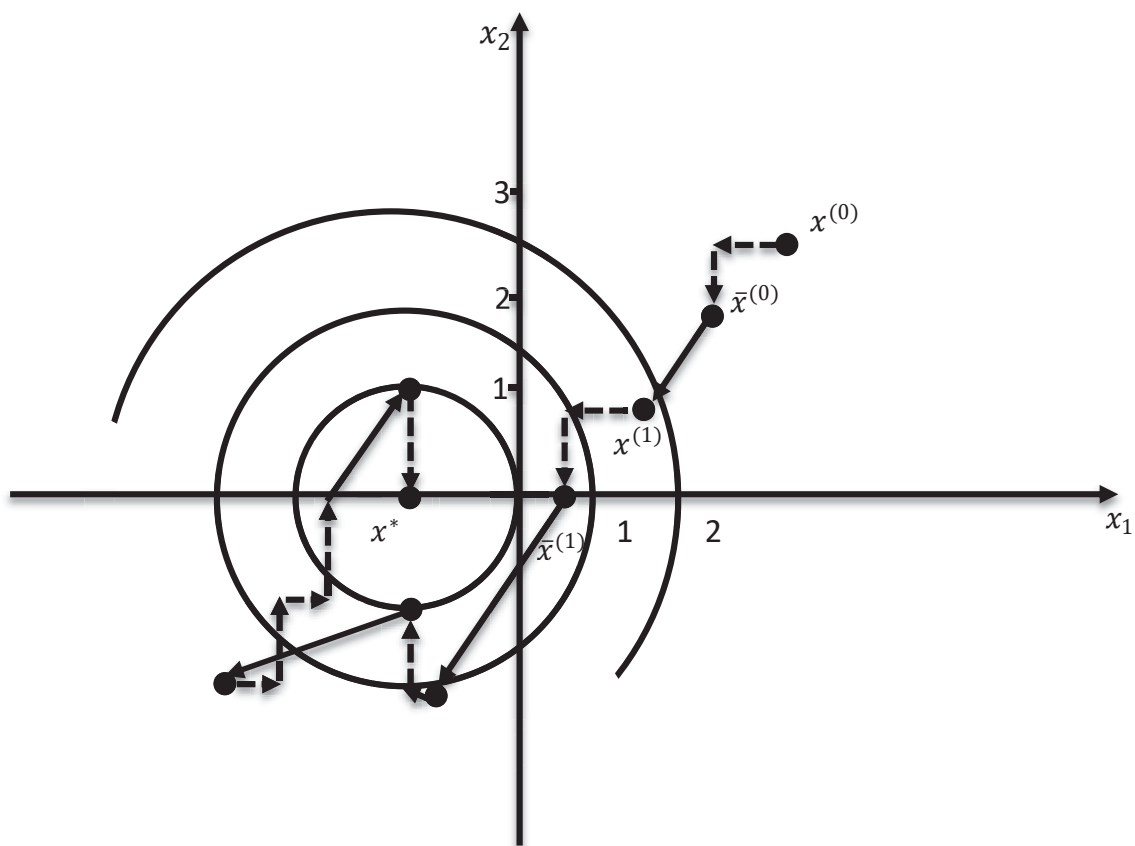


Рис. 5.4. Ілюстрація до прикладу 5.8

5.4.3. Методи випадкового пошуку

Крім методів безумовної мінімізації функції багатьох змінних, у яких процедури пошуку точок з меншим значенням функції детерміновані, використовуються також

методи з процедурами випадкового пошуку. Тоді ітераційний процес формалізується співвідношеннями

$$x^{(k+1)} = x^{(k)} + \alpha_k \frac{\xi}{\|\xi\|}, \quad k = 0, 1, \dots, \quad (5.23)$$

де, як і раніше, α_k – величина кроку, а $\xi = (\xi_1, \dots, \xi_n)^T$ – деяка реалізація n – вимірною випадкового вектора ξ , $\|x^{(k+1)} - x^{(k)}\| = \|\alpha_k\|$. Зазвичай вважають, що координати вектора ξ є незалежними випадковими величинами, рівномірно розподіленими на $[-1, 1]$. Нижче наведено два алгоритми, які можуть бути використані і як самостійні мінімізуючі процедури, і як складові інших алгоритмів.

Алгоритм 1 (з поверненням у випадку невдалого кроку)

Крок 1. Вибір початкової точки x , початкового кроку $\alpha > 0$, коефіцієнта зменшення кроку $\gamma > 1$, граничного числа невдалих спроб N , точності $\varepsilon > 0$ та обчислення $f(x)$.

Крок 2. Фіксація лічильника числа невдалих спроб $j = 1$.

Крок 3. Знаходження реалізації випадкового вектора ξ .

Крок 4. Знаходження пробної точки $y = x + \alpha \xi / \|\xi\|$ і обчислення $f(y)$.

Крок 5. Порівняння значень $f(y)$ і $f(x)$. Якщо $f(y) < f(x)$, то здійснюється переприсвоєння $x := y$, $f(x) := f(y)$ і перехід до кроку 4, у протилежному випадку – перехід до кроку 6.

Крок 6. Присвоєння $j := j + 1$. Якщо $j \leq N$, то здійснюється перехід до кроку 3, інакше – до кроку 7.

Крок 7. Перевірка умови досягнення точності. Якщо $\alpha < \varepsilon$, то пошук завершується і $x^* \approx x$, $f^* \approx f(x)$. В іншому випадку $\alpha := \alpha / \gamma$ і здійснюється перехід до кроку 2.

Стосовно максимального числа невдалих спроб N зауважимо, що на практиці $N = 3n$, де n – число змінних функції.

На рис. 5.5 наведена можлива ілюстрація процесу побудови послідовності (5.23) за допомогою цього алгоритму для функції двох змінних (пунктиром виділені невдалі спроби визначення точки $x^{(k+1)}$, тобто точки, що приводить до зменшення $f(x)$).

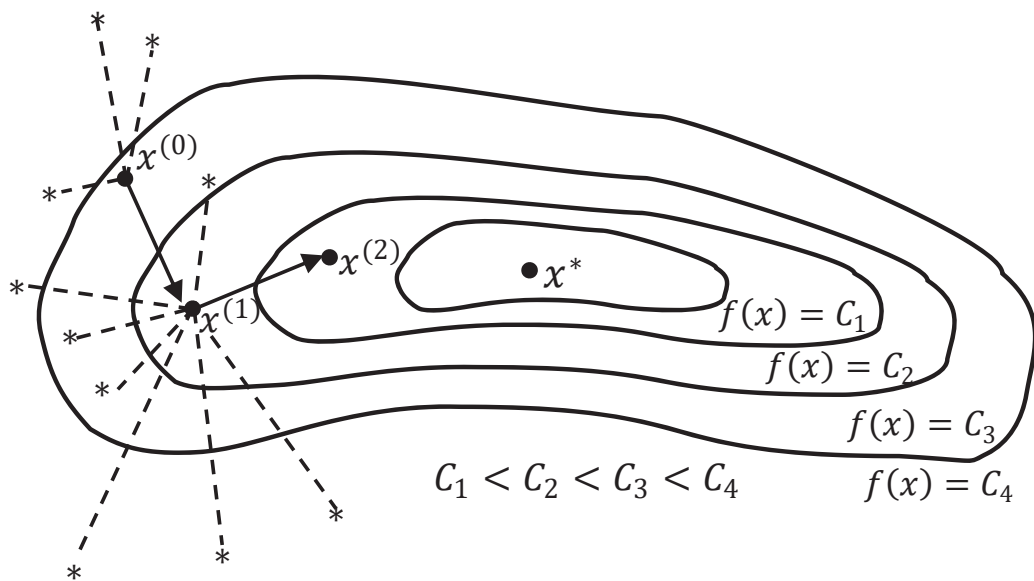


Рис. 5.5. Ілюстрація роботи алгоритму 1 у просторі $\mathbb{R}^2$

Алгоритм 2 (найкращої спроби)

Він відрізняється від алгоритму 1 тільки кроками 3 і 4.

Крок 3. Знаходження m реалізацій випадкового вектора ξ , тобто $\xi^{(1)}, \dots, \xi^{(m)}$.

Крок 4. Знаходження пробних точок $y^{(i)} = x + \alpha \xi^{(i)} / \|\xi^{(i)}\|$ ($i = \overline{1, m}$) та обчислення значень функцій $f(y^{(i)})$ ($i = \overline{1, m}$). Вибір $y^{(k)}$ із умови $f(y^{(k)}) = \min_{i=1, m} f(y^{(i)})$ та визначення $y = y^{(k)}$.

5.4.4. Метод спряжених напрямків

Описані вище методи у загальному випадку потребують для точного знаходження точки мінімуму нескінченного числа ітерацій, навіть якщо функція $f(x)$ є сильно опуклою квадратичною функцією (квадратична функція сильно опукла, якщо матриця її других похідних $A = f''(x)$ додатно визначена). Однак існують і такі прямі ітераційні методи, які дозволяють знайти точку мінімуму сильно опуклої квадратичної функції за скінченне число кроків. Оскільки широкий клас цільових функцій в околі точки мінімуму може бути апроксимований квадратичними функціями, то ці методи мають також високу ефективність і у випадку опуклої неквадратичної функції. Нижче розглянемо один із них – так званий метод спряжених напрямків.

Означення 5.11. Ненульові вектори $p^{(1)}, \dots, p^{(k)}$ називаються спряженими відносно матриці $A = (a_{ij})_{i,j=1}^n$ (A – ортогональними), якщо

$$\langle Ap^{(i)}, p^{(j)} \rangle = 0, \quad i \neq j; \quad j = \overline{1, k}. \quad (5.24)$$

Розглянемо задачу мінімізації у $\mathbb{R}^n$ квадратичної функції (5.22) з додатно визначеною матрицею A за допомогою ітераційного процесу

$$x^{(k)} = x^{(k-1)} + \alpha_k p^{(k)}, \quad k = 1, 2, \dots, \quad (5.25)$$

у якому вектори $p^{(k)}$ A – ортогональні.

Мають місце такі твердження:

1) система із n векторів $p^{(1)}, \dots, p^{(n)}$, спряжених відносно додатно визначеної матриці A , лінійно незалежна, тобто утворює базис в $\mathbb{R}^n$;

2) якщо в ітераційному процесі (5.25) на кожному кроці використовується вичерпуючий спуск, то величина кроку α_k визначається за формулою

$$\alpha_k = -\frac{\langle f'(x^{(0)}), p^{(k)} \rangle}{\langle Ap^{(k)}, p^{(k)} \rangle}, k=1,2,\dots; \quad (5.26)$$

3) послідовний вичерпуючий спуск по A – ортогональних напрямках (5.25) приводить до точки мінімуму квадратичної функції за $l \leq n$ кроків.

Стосовно ітераційних процедур (5.25), (5.26) зауважимо таке. Знаходження базису в $\mathbb{R}^n$ із A – ортогональних векторів реалізується неоднозначно. Таким базисом може бути ортогональний базис із власних векторів матриці A , хоча їх пошук при $n > 2$ є складною задачею. Крім того, послідовна одновимірна мінімізація по спряжених напрямках $p^{(k)}$ може бути здійснена і без попередньої побудови векторів $p^{(1)}, \dots, p^{(n)}$, послідовно знаходячи їх у процесі мінімізації.

Тепер сформулюємо алгоритм методу спряжених напрямків для мінімізації функції n змінних.

Крок 1. Вибір початкової точки $x^{(0)} \in \mathbb{R}^n$.

Крок 2. Присвоєння $p^{(1)} := e^{(1)}$ і знаходження точки $x^{(1)}$ з умови $f(x^{(1)}) = \min_{\alpha \in \mathbb{R}} f(x^{(0)} + \alpha p^{(1)})$.

Крок 3. а) $y^{(0)} := x^{(1)} + e^{(2)}$; б) знаходження точки $y^{(1)}$ з умови вичерпуючого спуску із точки $y^{(0)}$ у напрямку $p^{(1)}$: $f(y^{(1)}) = \min_{\alpha \in \mathbb{R}} f(y^{(0)} + \alpha p^{(1)})$; в) $p^{(2)} := x^{(1)} - y^{(1)}$, знаходження точки $x^{(2)}$ з умови $f(x^{(2)}) = \min_{\alpha \in \mathbb{R}} f(x^{(1)} + \alpha p^{(2)})$.

Крок 4. а) $y^{(1)} := x^{(2)} + e^{(3)}$; б) знаходження точки $y^{(2)}$, мінімізуючи $f(x)$ послідовно по напрямках $p^{(1)}$ і $p^{(2)}$, починаючи з точки $y^{(1)}$; в) $p^{(3)} := x^{(2)} - y^{(2)}$, знаходження точки $x^{(3)}$ з умови $f(x^{(3)}) = \min_{\alpha \in \mathbb{R}} f(x^{(2)} + \alpha p^{(3)})$.

.....
Крок $n+1$. а) $y^{(n-1)} := x^{(n-1)} + e^{(n)}$; б) знаходження точки $y^{(n)}$, мінімізуючи $f(x)$ послідовно по напрямках $p^{(1)}, \dots, p^{(n-1)}$, починаючи з точки $y^{(n-1)}$; в) $p^{(n)} := x^{(n-1)} - y^{(n-1)}$, знаходження точки $x^{(n)}$ з умови $f(x^{(n)}) = \min_{\alpha \in \mathbb{R}} f(x^{(n-1)} + \alpha p^{(n)})$.

Зауважимо, що побудовані у такий спосіб вектори $p^{(1)}, \dots, p^{(n)}$ будуть А – ортогональними, тому, якщо $f(x)$ є квадратичною функцією з додатно визначеною матрицею А і всі задачі одновимірної мінімізації розв'язуються точно, то $x^* = x^{(n)}$ і обчислення завершуються. Якщо $f(x)$ не є квадратичною функцією або допоміжні задачі одновимірної мінімізації розв'язуються наближено, то потрібно перейти до наступного кроку.

Крок $n+2$. Перевірка умови закінчення обчислень. Якщо $\|x^{(n)} - x^{(n+1)}\| \leq \varepsilon$, де ε – задана точність, то пошук точки мінімуму завершується і $x^* \approx x^{(n)}$, інакше $x^{(0)} := x^{(n)}$ і здійснюється перехід до кроку 2.

Описаний у цьому пункті метод спряжених напрямків один з найбільш ефективних прямих методів, хоча при його застосуванні доводиться розв'язувати велику кількість задач одновимірної мінімізації.

Приклад 5.9. Методом спряжених напрямків знайти точку безумовного мінімуму функції $f(x) = 3x_1^2 - 2x_1x_2 + 3x_2^2 + x_1 + 2x_2$.

◀ Крок 1. Виберемо початкову точку $x^{(0)} = (0, 0)^T$.

Крок 2. $p^{(1)} = e^{(1)} = (1, 0)^T$, $f(x^{(1)}) = \min_{\alpha} f(x^{(0)} + \alpha p^{(1)}) = \min_{\alpha} f(\alpha, 0)$, тому $\alpha = -1/6$ і $x^{(1)} = (-1/6, 0)^T$.

Крок 3. а) $y^{(0)} = x^{(1)} + e^{(2)} = (-1/6, 0)^T + (0, 1)^T = (-1/6, 1)^T$;

б) $f(y^{(1)}) = \min_{\alpha} f(y^{(0)} + \alpha p^{(1)}) = \min_{\alpha} f\left(-\frac{1}{6} + \alpha, 1\right)$, тому

$\alpha = 1/3$ і $y^{(1)} = \left(-\frac{1}{6}, 1\right)^T + \frac{1}{3}(1, 0)^T = \left(\frac{1}{6}, 1\right)^T$;

в) $p^{(2)} = x^{(1)} - y^{(1)} = (-1/6, 0)^T - (1/6, 1)^T = (-1/3, -1)^T$,

$f(x^{(2)}) = \min_{\alpha} f(x^{(1)} + \alpha p^{(2)}) = \min_{\alpha} f\left(-\frac{1}{6} - \frac{1}{3}\alpha, -\alpha\right)$, тому

$\alpha = \frac{7}{16}$, $x^{(2)} = x^{(1)} + \alpha p^{(2)} = \left(-\frac{5}{16}, -\frac{7}{16}\right)^T$.

Оскільки матриця $A = f''(x) = \begin{pmatrix} 6 & -2 \\ -2 & 6 \end{pmatrix}$ додатно визначена,

вектори $p^{(1)}$, $p^{(2)}$ A – ортогональні (це легко перевірити) і всі задачі одновимірної мінімізації розв'язані точно, то шуканою точкою мінімуму є $x^* = x^{(2)} = \left(-\frac{5}{16}, -\frac{7}{16}\right)^T$. ►

5.5. Методи безумовної мінімізації першого та другого порядків

Знову розглянемо задачу (5.1), тобто задачу безумовної мінімізації функції $f(x)$ на допустимій множині $\mathbb{R}^n$, але від функції $f(x)$ будемо вимагати, щоб вона була

диференційовною в $\mathbb{R}^n$ певну кількість разів, оскільки напрямок спадання $p^{(k)}$ в ітераційних процедурах мінімізації

$$x^{(k+1)} = x^{(k)} + \alpha_k p^{(k)}, \quad k = 0, 1, 2, \dots \quad (5.27)$$

буде визначатися з урахуванням частинних похідних функції $f(x)$. Критерієм завершення розрахунків у випадку диференційовних цільових функцій та ітераційного процесу (5.27) зазвичай вибирається умова $\|f'(x^{(k)})\| < \varepsilon$, де $f'(x^{(k)})$ – градієнт функції $f(x)$ у точці $x^{(k)}$, а ε – заданий параметр точності пошуку мінімуму.

5.5.1. Метод градієнтного спуску

Як відомо, градієнт функції у точці показує напрямок зростання цієї функції, а антиградієнт – напрямок спадання. У зв'язку з цим, якщо функція $f(x)$ диференційовна, то у (5.27) на k – му кроці можна вибрати $p^{(k)} = -f'(x^{(k)})$, якщо $f'(x^{(k)}) \neq 0$, тобто градієнт не є нульовим вектором. При цьому у деякому ε – околі точки $x^{(k)}$ цей напрямок $p^{(k)}$ забезпечує найшвидше спадання функції. Величину кроку $\alpha_k > 0$ потрібно вибирати такою, щоб

$$f(x^{(k)}) < f(x^{(k-1)}), \quad k = 1, 2, \dots \quad (5.28)$$

Нижче наведено алгоритм одного з варіантів методу градієнтного спуску з постійним кроком $\alpha_k \equiv \alpha$, тобто методу, що формалізується ітераційним процесом (5.27) при $\alpha_k \equiv \alpha$.

Крок 1. Вибір початкової точки $x \in \mathbb{R}^n$ і кроку $\alpha > 0$, параметра точності $\varepsilon > 0$ та знаходження $f(x)$.

Крок 2. Знаходження $f'(x)$ і перевірка умови $\|f'(x)\| < \varepsilon$.

При її виконанні обчислення завершуються і вважають, що $x^* = x$, $f^* = f(x)$, а при невиконанні – здійснюється перехід до кроку 3.

Крок 3. Знаходження $y = x - \alpha f'(x)$ і $f(y)$. Якщо $f(y) < f(x)$, то $x := y$, $f(x) := f(y)$ і переходять до кроку 2, у протилежному випадку – переходять до кроку 4.

Крок 4. Переприсвоєння $\alpha := \alpha/2$ і перехід до кроку 3.

Зазначимо, що в околі стаціонарної точки функції $f(x)$ величина $\|f'(x)\|$ стає малою, що уповільнює збіжність послідовності $\{x^{(k)}\}$. У зв'язку з цим у (5.27) вектор $p^{(k)}$ інколи нормують, тобто приводять до вектора одиничної довжини $\left(p^{(k)} = -f'(x^{(k)}) / \|f'(x^{(k)})\|\right)$.

Враховуючи ту обставину, що часто в околі точки мінімуму цільова функція $f(x)$ може бути апроксимована опуклою квадратичною функцією (5.22), важливо мати обґрунтування збіжності градієнтного методу з кроком $\alpha_k \equiv \alpha$ для опуклої функції (5.22). Нехай A – додатно визначена матриця, λ_1 і λ_2 – її найменше і найбільше власні значення ($0 < \lambda_1 \leq \lambda_2$). Тоді встановлено, що при будь-яких

$\alpha \in \left(0, \frac{2}{\lambda_2}\right)$ і $x^{(0)} \in \mathbb{R}^n$ ітераційна послідовність (5.27) (при $p^{(k)} = -f'(x^{(k)})$) збігається до єдиної точки x^* глобального мінімуму функції (5.22) лінійно (зі швидкістю геометричної прогресії), тобто $\|x^{(k)} - x^*\| \leq q^k \|x^{(0)} - x^*\|$, де $q = \max(|1 - \alpha\lambda_1|, |1 - \alpha\lambda_2|)$. Величина q набуває мінімального

значення $q^* = (\lambda_2 - \lambda_1)/(\lambda_2 + \lambda_1)$ при $\alpha = \alpha^* = 2/(\lambda_2 + \lambda_1)$. Число ітерацій суттєво залежить від співвідношення між λ_1 і λ_2 , зокрема при $\lambda_2 = \lambda_1 > 0$ точка мінімуму знаходиться за один крок (прикл. 5.10).

Приклад 5.10. За допомогою методу градієнтного спуску знайти точку безумовного мінімуму функції $f(x) = x_1^2 + x_2^2$.

◀ Функція є сильно опуклою квадратичною функцією (має єдину точку глобального мінімуму), причому $A = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$, $\lambda_1 = \lambda_2 = 2$ (власні значення є коренями рівняння $(2 - \lambda)^2 = 0$).

Крок 1. Виберемо початкову точку $x^{(0)} = (2, 2)^T$, $\alpha = \alpha^* = 2/(2 + 2) = 1/2$, $\varepsilon > 0$ ($\varepsilon \in (0, 1)$). При цьому $f(x^{(0)}) = 8$.

Крок 2. Оскільки $f'(x^{(0)}) = (2 \cdot 2, 2 \cdot 2)^T = (4, 4)^T$, то $\|f'(x)\| > \varepsilon$ і переходимо до кроку 3.

Крок 3. $y = x^{(0)} - \alpha f'(x^{(0)}) = (2, 2)^T - \frac{1}{2}(4, 4)^T = (0, 0)^T$, $f(y) = 0$. Враховуючи, що $f(y) < f(x^{(0)})$, змінюємо початкову точку $x^{(0)}$ на y , а $f(x^{(0)})$ на $f(y)$ і переходимо до кроку 2, припиняючи обчислення, бо $\|f'(x^{(0)})\| = 0 < \varepsilon$. Отже, шукана точка мінімуму $x^* = (0, 0)^T$. ►

Зауважимо також, що число $\mu = \lambda_2/\lambda_1 \geq 1$ для симетричної додатно визначеної матриці A називається *числом обумовленості*. У задачах мінімізації сильно опуклих функцій (не обов'язково квадратичних) число μ (для

матриці Гессе) характеризує ступінь витягнутості (видовженості) ліній (поверхонь) рівня $f(x)=c$. Якщо μ велике, то лінії рівня сильно видовжені, тобто мають «пагорбний» характер (сильно змінюються по одних напрямках і слабо – по інших). Такі задачі мінімізації називають *погано обумовленими*. Задачі мінімізації з близьким до одиниці числом μ називають *добре обумовленими*.

5.5.2. Метод найшвидшого градієнтного спуску

Цей варіант градієнтного методу відрізняється від попереднього тим, що на кожній ітерації вздовж антиградієнта $(-f'(x^{(k)}))$ здійснюють вичерпуючий спуск, тобто величина кроку α_k знаходиться як розв'язок задачі

$$f(x^{(k)} - \alpha f'(x^{(k)})) \rightarrow \min_{\alpha > 0}. \quad (5.29)$$

Уточнимо алгоритм цього методу.

Крок 1. Вибір початкової точки $x \in \mathbb{R}^n$ та параметра точності $\varepsilon > 0$.

Крок 2. Знаходження $f'(x)$ і перевірка умови $\|f'(x)\| < \varepsilon$.

При її виконанні вважають, що $x^* = x$, $f^* = f(x)$ і обчислення завершується, а при її невиконанні – здійснюється перехід до кроку 3.

Крок 3. Розв'язування одновимірної задачі мінімізації (5.29) при $x^{(k)} = x$ і знаходження $\alpha = \alpha^*$, переприсвоєння $x := x - \alpha^* f'(x)$ і перехід до кроку 2.

Стосовно задачі (5.29) зазначимо, що вона може бути розв'язана одним із методів одновимірної мінімізації, а у випадку, коли $f(x)$ є квадратичною (тобто функцією (5.22)), для визначення α^* можна скористатися формулою

$$\alpha_k = -\frac{\langle Ax^{(k)} + b, p^{(k)} \rangle}{\langle Ap^{(k)}, p^{(k)} \rangle} \quad (5.30)$$

при $p^{(k)} = -f'(x^{(k)})$.

Приклад 5.11. Методом найшвидшого градієнтного спуску розв'язати задачу безумовної мінімізації функції $f(x) = x_1^2 + 100x_2^2$.

◀ **Крок 1.** Візьмемо початкову точку $x^{(0)} = (1, 1)^T$ та $\varepsilon = 10^{-3}$.

$$\text{Крок 2.} \quad \frac{\partial f}{\partial x_1} = 2x_1, \quad \frac{\partial f}{\partial x_2} = 200x_2, \quad A = \begin{pmatrix} 2 & 0 \\ 0 & 200 \end{pmatrix},$$

$f'(x^{(0)}) = (2, 200)^T$. Оскільки $\|f'(x^{(0)})\| > \varepsilon$, то переходимо до кроку 3.

Крок 3. Для знаходження розв'язку задачі мінімізації (5.29) при $x^{(k)} = x^{(0)} = (1, 1)^T$, $p^{(0)} = -f'(x^{(0)}) = (-2, -200)^T$, скористаємося формулою (5.30), згідно з якою отримаємо $\alpha^* = \alpha^{(0)} \approx 0,005$ (точне значення $\alpha^* = 0,005000495$). Тоді $x^{(1)} = x^{(0)} - \alpha^* f'(x^{(0)}) = (9,9 \cdot 10^{-1}; -9,9 \cdot 10^{-5})^T$ і переходимо до кроку 2 та наступної ітерації.

Отримані після перших трьох ітерацій результати наведено у табл. 5.3. ►

Таблиця 5.3

Номер ітерації	$(x_1^{(k)}, x_2^{(k)})^T$	$f(x^{(k)})$	$\ f'(x^{(k)})\ $
0	$(1; 1)^T$	101	200
1	$(9,9 \cdot 10^{-1}; -9,9 \cdot 10^{-5})^T$	$9,8 \cdot 10^{-1}$	1,98
2	$(9,7 \cdot 10^{-3}; 9,7 \cdot 10^{-3})^T$	$9,5 \cdot 10^{-3}$	1,94
3	$(9,6 \cdot 10^{-3}; -9,6 \cdot 10^{-7})^T$	$9,2 \cdot 10^{-5}$	$1,92 \cdot 10^{-2}$

5.5.3. Метод Ньютона

Якщо функція $f(x)$ двічі диференційовна в $\mathbb{R}^n$, то її можна розкласти в околі точки $x^{(k)}$ за формулою Тейлора (5.8), тобто записати її у вигляді

$$f(x) = f(x^{(k)}) + \langle f'(x^{(k)}), x - x^{(k)} \rangle + \frac{1}{2} \langle f''(x^{(k)})(x - x^{(k)}), (x - x^{(k)}) \rangle + o(\|x - x^{(k)}\|^2),$$

звідки і випливає, що з точністю до величини $o(\|x - x^{(k)}\|^2)$ в околі точки $x^{(k)}$ функція $f(x)$ може бути наближена квадратичною функцією

$$\varphi_k(x) = \frac{1}{2} \langle f''(x^{(k)})(x - x^{(k)}), (x - x^{(k)}) \rangle + \langle f'(x^{(k)}), x - x^{(k)} \rangle + f(x^{(k)}). \quad (5.31)$$

Мінімізуючи функцію (5.31) замість функції $f(x)$, точку її мінімуму $x^{(k+1)}$ знайдемо з умови

$$\varphi'_k(x) = f''(x^{(k)})(x - x^{(k)}) + f'(x^{(k)}) = 0. \quad (5.32)$$

Вважаючи матрицю $f''(x^{(k)})$ додатно визначеною, при всіх $x \in \mathbb{R}^n$ (тоді (5.32) матиме єдиний розв'язок, оскільки

функція (5.31) буде сильно опуклою, тобто точка глобального мінімуму єдина), а значить, невиродженою ($\det f''(x^{(k)}) > 0$), можна знайти обернену матрицю $[f''(x^{(k)})]^{-1}$, помножити її зліва на обидві частини (5.32) і знайти точку $x^{(k+1)}$:

$$x^{(k+1)} = x^{(k)} - [f''(x^{(k)})]^{-1} \cdot f'(x^{(k)}), \quad k = 0, 1, \dots \quad (5.33)$$

При довільній початковій точці $x^{(0)} \in \mathbb{R}^n$ співвідношення (5.33) формалізують ітераційний процес, який прийнято називати методом *Ньютона мінімізації функції багатьох змінних*.

Якщо функція $f(x)$ квадратична з додатно визначеною матрицею Гессе (тоді вона буде сильно опуклою), то точку глобального мінімуму цієї функції методом Ньютона знаходять за один крок із будь-якої початкової точки $x^{(0)} \in \mathbb{R}^n$.

Якщо функція $f(x)$ не є квадратичною, то у більшості випадків метод Ньютона забезпечує швидку збіжність і є більш ефективним, ніж градієнтні методи. При цьому слід зазначити, що при вдалому виборі початкової точки $x^{(0)} \in \mathbb{R}^n$ для сильно опуклої двічі диференційовної функції $f(x)$ мінімізуюча послідовність $\{x^{(k)}\}$ збігається до точки мінімуму з квадратичною швидкістю (відстань між точками $x^{(k)}$ і x^* $\rho(x^{(k)}, x^*) \leq cq^{2^k}$, $q \in (0, 1)$), а у випадку, коли точка $x^{(0)}$ недостатньо близька до x^* , послідовність (5.33) може навіть розбігатися, що підкреслює принципову важливість початкової точки у цій ситуації.

Недоліком методу Ньютона є те, що навіть збіжна послідовність не завжди гарантує монотонне спадання

значень функції $f(x)$ (може порушуватись умова $f(x^{(k+1)}) < f(x^{(k)})$ для деяких $k = 0, 1, \dots$), однак цього недоліку позбавлений так званий *узагальнений метод Ньютона*, який описується ітераційним процесом

$$x^{(k+1)} = x^{(k)} - \alpha_k \left[f''(x^{(k)}) \right]^{-1} \cdot f'(x^{(k)}), \quad k = 0, 1, \dots \quad (5.34)$$

У (5.34) величина кроку $\alpha_k > 0$ на кожному кроці вибирається такою, щоб був реалізований вичерпуючий спуск вздовж напрямку $p^{(k)} = - \left[f''(x^{(k)}) \right]^{-1} \cdot f'(x^{(k)})$.

Побудова послідовності $\{x^{(k)}\}$ завершується у точці, у якій $\|f'(x^{(k)})\| < \varepsilon$, де ε – заданий параметр точності (можливі й інші умови зупинки обчислень (або навіть їх поєднання)), зокрема $k \geq N$, де N – граничне число ітерацій, або двократне одночасне виконання нерівностей $\|x^{(k+1)} - x^{(k)}\| < \tilde{\varepsilon}$, $|f(x^{(k+1)}) - f(x^{(k)})| < \tilde{\varepsilon}$, де $\tilde{\varepsilon}$ – параметр точності).

Алгоритм методу Ньютона, який поєднує у собі процедури (5.33) і (5.34), конкретизується такими кроками.

Крок 1. Вибір початкової точки x та параметра точності ε .

Крок 2. Знаходження $f'(x)$ і перевірка умови $\|f'(x)\| < \varepsilon$. Якщо ця умова виконується, то обчислення завершуються і приймається $x^* = x$, $f^* = f(x)$, у протилежному випадку здійснюється перехід до кроку 3.

Крок 3. Обчислення $f''(x)$, $[f''(x)]^{-1}$, $y = x - [f''(x)]^{-1} f'(x)$, $f(y)$.

Крок 4. Перевірка умови $f(y) < f(x)$. У випадку її виконання здійснюють переприсвоєння $x := y$, $f(x) := f(y)$ і перехід до кроку 2, а у випадку невиконання – перехід до кроку 5.

Крок 5. Знаходження $y = x - \alpha^* [f''(x)]^{-1} f'(x)$, $f(y)$, де α^* – значення величини кроку, отриманого у результаті вичерпуючого спуску вздовж вектора $p = -[f''(x)]^{-1} f'(x)$ (тобто α^* – розв'язок задачі $f(y) \rightarrow \min_{\alpha > 0}$), переприсвоєння $x := y$, $f(x) := f(y)$ та перехід до кроку 2.

Як уже було сказано, збіжність методу Ньютона доведена тільки для сильно опуклих функцій і достатньо доброго початкового наближення, вибір якого при розв'язуванні практичних задач проблематичний. У зв'язку з цим на практиці результати обчислень обов'язково аналізують на предмет адекватності розв'язку.

Завершуючи даний підпункт, зауважимо, що недоліком методу Ньютона є потреба в обчисленні на кожній з ітерацій матриці Гессе та оберненої до неї матриці. Цього недоліку позбавлені так звані квазіньютонівські методи.

Приклад 5.12. Знайти точку безумовного мінімуму функції $f(x) = 2x_1^2 + x_1x_2 + x_2^2$, використовуючи метод Ньютона.

◀ *Крок 1.* Виберемо початкову точку $x = \left(\frac{1}{2}, 1\right)^T$ та $\varepsilon = 0,1$.

Крок 2. $f'(x) = (4x_1 + x_2, x_1 + 2x_2)^T$, тобто $f'\left(\frac{1}{2}, 1\right) = \left(3, \frac{5}{2}\right)^T$.

Оскільки $\|f'(x)\| = \sqrt{61}/2 > \varepsilon$, то переходимо до кроку 3.

$$\text{Крок 3.} \quad f''(x) = \begin{pmatrix} 4 & 1 \\ 1 & 2 \end{pmatrix}, \quad [f''(x)]^{-1} = \begin{pmatrix} \frac{2}{7} & -\frac{1}{7} \\ -\frac{1}{7} & \frac{4}{7} \end{pmatrix},$$

$$\begin{aligned} y &= x - [f''(x)]^{-1} f'(x) = \begin{pmatrix} 1/2 \\ 1 \end{pmatrix} - \begin{pmatrix} 2/7 & -1/7 \\ -1/7 & 4/7 \end{pmatrix} \begin{pmatrix} 3 \\ 5/2 \end{pmatrix} = \\ &= \begin{pmatrix} 1/2 \\ 1 \end{pmatrix} - \begin{pmatrix} 1/2 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad f(y) = 0. \end{aligned}$$

Крок 4. $f(y) = 0 < f(x) = 2$, тому переходимо до кроку 2 і наступної ітерації, вибравши за початкову точку x уже знайдену точку y .

Враховуючи, що для цієї початкової точки $\|f'(x)\| = 0 < \varepsilon$, завершуємо обчислення, отримавши $x^* = (0, 0)^T$, $f^* = 0$ (у точці x^* виконується умова стаціонарності). ►

Розділ 6. Основи теорії та методи умовної мінімізації функції багатьох змінних

6.1. Постановка та класифікація задач умовної мінімізації функції багатьох змінних

Задача (В.1) називається задачею умовної оптимізації (або умовною задачею), якщо X – власна підмножина простору $\mathbb{R}^n$. Як і раніше, більш детально будемо розглядати лише задачу мінімізації, тобто задачу

$$\begin{cases} f(x) \rightarrow \min, \\ x \in X, \end{cases} \quad (6.1)$$

оскільки задача максимізації функції $f(x)$ на X еквівалентна задачі мінімізації функції $(-f(x))$. Порівнюючи задачу (6.1) із задачею безумовної мінімізації (5.10, легко зрозуміти, що всі твердження відносно необхідних і достатніх умов оптимальності розв'язку безумовної задачі мають місце і для умовної задачі, якщо точка її локального екстремуму x^* є внутрішньою точкою допустимої множини X ($x^* \in \text{int } X$). Однак принципова різниця між умовною та безумовною задачами оптимізації полягає у тому, що розв'язок умовної задачі може бути граничною точкою множини X , у якій класичні методи дослідження екстремуму (методи безумовної оптимізації) застосувати неможливо.

Означення точок глобального та локального мінімумів функції $f(x)$ відрізняються від означень 5.1 та 5.2 лише уточненням допустимої множини оптимізаційної задачі. Конкретизуємо їх.

Означення 6.1. Точка $x^* \in X$ ($X \subset \mathbb{R}^n$) називається точкою глобального мінімуму функції $f(x)$ на X , якщо

$f(x^*) \leq f(x)$ для всіх $x \in X$.

Означення 6.2. Точка $x^* \in X$ називається точкою локального мінімуму функції $f(x)$ у задачі (6.1), якщо $f(x^*) \leq f(x)$ у деякому ε -околі точки x^* , тобто для всіх $x \in O(x^*, \varepsilon) = \{x \in X / \rho(x, x^*) < \varepsilon, \varepsilon > 0\}$.

Задача (6.1) є задачею на знаходження точки (чи множини точок) глобального мінімуму. Відповісти на питання про існування розв'язку задачі (6.1) дозволяють відомі з математичного аналізу теореми Вейєштраса.

Теорема 6.1 (Вейєштраса). Нехай X – компакт у $\mathbb{R}^n$, тобто замкнена обмежена множина, а $f(x)$ – неперервна функція на X . Тоді задача (6.1) має розв'язок (існує точка глобального мінімуму функції $f(x)$ на X).

Часто використовується інша форма даної теореми.

Теорема 6.2. Нехай X – замкнена множина у $\mathbb{R}^n$, $f(x)$ – неперервна функція на X і для деякого $\bar{x} \in X$ є обмеженою множина

$$M(\bar{x}) = \{x \in X / f(x) \leq f(\bar{x})\}.$$

Тоді існує точка глобального мінімуму функції $f(x)$ на X .

◀ Множина $M(\bar{x})$ компакт, оскільки її замкненість впливає із замкненості X і неперервності $f(x)$ на X . Згідно з теоремою 6.1, точка глобального мінімуму $f(x)$ на $M(\bar{x})$ існує. ►

Задача (6.1), у якій допустима множина

$$X = \{x \in X_0 / g_i(x) \leq 0, i = \overline{1, k}; g_i(x) = 0, i = \overline{k+1, m}\},$$

тобто задається системою скінченного числа нерівностей і рівнянь на деякій множині $X_0 \subset \mathbb{R}^n$, називається загальною задачею математичного

програмування на мінімум. Зазвичай ця задача формалізується співвідношеннями:

$$f(x) \rightarrow \min, \quad (6.2)$$

$$g_i(x) \leq 0, i = \overline{1, k}, \quad (6.3)$$

$$g_i(x) = 0, i = \overline{k+1, m}, \quad (6.4)$$

$$x \in X_0. \quad (6.5)$$

У задачі (6.2)–(6.5) обмеження типу (6.3) називаються *обмеженнями-нерівностями*, обмеження типу (6.4) – *обмеженнями-рівностями*, а обидва ці типи обмежень – *функціональними обмеженнями*. Умова $x \in X_0$, тобто обмеження (6.5), називається *прямим обмеженням*.

Постановка загальної задачі математичного програмування (6.2)–(6.5) охоплює всі можливі її частинні випадки, зокрема при $k=0$ у задачі не буде обмежень-нерівностей, при $m=0$ – функціональних обмежень, а при $X_0 = \mathbb{R}^n$ – прямого обмеження.

Залежно від типу включених обмежень частинними випадками задачі (6.2)–(6.5) є задача (6.2), (6.4) з обмеженнями-рівностями, задача (6.2), (6.3) – з обмеженнями-нерівностями, задача (6.2)–(6.4) – зі змішаними функціональними обмеженнями, задача (6.2), (6.5) – з прямим обмеженням. Якщо перелічені вище задачі, крім (6.2), (6.5), доповнити прямим обмеженням (6.5), то отримаємо ще групу задач, які є частинними випадками загальної задачі математичного програмування (6.2)–(6.5).

Зазначимо також, що поділ обмежень задачі (6.2)–(6.5) на функціональні та прямі умовний, адже будь-які функціональні обмеження (обмеження-нерівності чи обмеження-рівності) можна ввести у пряме обмеження, наприклад, обмеження-нерівності, і перейти від (6.2)–(6.5)

до задачі

$$\begin{cases} f(x) \rightarrow \min, \\ g_i(x) = 0, \quad i = \overline{k+1, m}, \\ x \in X'_0, \end{cases}$$

$$\text{де } X'_0 = \{x \in X_0 / g_i(x) \leq 0, i = \overline{1, k}\}.$$

Виключивши усі функціональні обмеження, ми повернемось до задачі (6.1). У більшості випадків немає сенсу повертатись до задачі (6.1), оскільки наявність функціональних обмежень дозволяє детальніше вивчити властивості та специфіку задачі.

Важливе значення, особливо у прикладних задачах, має множина «простої» структури, на якій власне задані функціональні обмеження множини X . У багатьох задачах математичного програмування множина

$X_0 = \{x \in \mathbb{R}^n / a_j \leq x_j \leq b_j, j = \overline{1, n}\}$, тобто є n -вимірним паралелепіпедом (випадки $a_j = -\infty$ і $b_j = +\infty$ для деяких j тут не виключаються). Часто $X_0 = \mathbb{R}^n$ або $X_0 = \mathbb{R}_+^n = \{x \in \mathbb{R}^n / x_j \geq 0, j = \overline{1, n}\}$.

Залежно від типу функцій $f(x), g_i(x) (i = \overline{1, m})$ із (6.2)–(6.5) можна виділити спеціальні класи задач математичного програмування, зокрема задачі *лінійного* (всі функції лінійні) та *нелінійного програмування* (принаймні одна із функцій нелінійна), а також інші класи задач, які будуть вивчені пізніше.

6.2. Класична задача на умовний екстремум

Під класичною задачею на умовний мінімум розуміють задачу математичного програмування з функціональними обмеженнями-рівностями, тобто задачу (6.2), (6.4) при $k = 0$. У цій задачі допустима множина X задається системою скінченного числа рівнянь:

$$X = \{x \in \mathbb{R}^n / g_i(x) = 0, i = \overline{1, m}\}. \quad (6.6)$$

Як правило, класичну задачу на умовний мінімум записують у вигляді

$$\begin{cases} f(x) \rightarrow \min, \\ g_i(x) = 0, i = \overline{1, m}. \end{cases} \quad (6.7)$$

Запис (6.7) дозволяє явно вказати не саму допустиму множину (6.6), а систему, що її визначає.

Не порушуючи загальності міркувань, надалі при формулюванні необхідних і достатніх умов оптимальності розв'язку задачі (6.7) будемо вважати функції $f(x)$ та $g_i(x)$ ($i = \overline{1, m}$) достатньо гладкими в околі точки мінімуму, тобто неперервно диференційовними або двічі неперервно диференційовними [1, 12, 14].

Зауважимо, що розв'язок задачі (6.7) обов'язково лежить на границі області (допустимої множини) X , оскільки X не має жодної внутрішньої точки. У зв'язку з цим поняття точки мінімуму для задачі (6.7) потрібно конкретизувати.

Означення 6.3. Точка $x^* \in X$ називається точкою локального мінімуму у задачі (6.7), якщо $f(x^*) \leq f(x)$ для всіх допустимих $x \in X$, достатньо близьких до x^* ($\rho(x, x^*) = \|x - x^*\| < \varepsilon, \varepsilon > 0$).

При дослідженні задачі (6.7) особливу роль відіграє її

узагальнена функція Лагранжа, тобто функція

$$\tilde{L}(x, \lambda_0, \lambda) = \lambda_0 f(x) + \sum_{i=1}^m \lambda_i g_i(x), \quad (6.8)$$

де $x \in \mathbb{R}^n$, $\lambda_0 \in \mathbb{R}$, $\lambda \in \mathbb{R}^m$ (числа $\lambda_0, \lambda_i, i = \overline{1, m}$, називають множниками Лагранжа).

Очевидно, частинні похідні

$$\frac{\partial \tilde{L}(x, \lambda_0, \lambda)}{\partial x_j} = \lambda_0 \frac{\partial f(x)}{\partial x_j} + \sum_{i=1}^m \lambda_i \frac{\partial g_i(x)}{\partial x_j}, \quad j = \overline{1, n}. \quad (6.9)$$

Часто рівності (6.9) записують у векторній формі (6.10):

$$\begin{aligned} \tilde{L}'_x(x, \lambda_0, \lambda) &\equiv \frac{\partial \tilde{L}(x, \lambda_0, \lambda)}{\partial x} = \lambda_0 \frac{\partial f(x)}{\partial x} + \sum_{i=1}^m \lambda_i \frac{\partial g_i(x)}{\partial x} = \\ &= \lambda_0 f'(x) + \sum_{i=1}^m \lambda_i g'_i(x). \end{aligned} \quad (6.10)$$

Має місце наступна теорема про необхідні умови локального умовного екстремуму у задачі (6.7) з функцією Лагранжа (6.8).

Теорема 6.1 (необхідні умови мінімуму першого порядку для задачі (6.7) у випадку узагальненої функції Лагранжа). Нехай:

- 1) x^* – точка локального мінімуму у задачі (6.7);
- 2) функції $f(x), g_i(x) \quad (i = \overline{1, m})$ неперервно диференційовні у деякому околі точки x^* .

Тоді існують числа $\lambda_0^*, \lambda_1^*, \dots, \lambda_m^*$, які одночасно не дорівнюють нулю і при яких виконуються умови:

- 1) стаціонарності узагальненої функції Лагранжа (6.8) по x :

$$\tilde{L}'_x(x^*, \lambda_0^*, \lambda^*) = 0 \text{ або } \frac{\partial \tilde{L}(x^*, \lambda_0^*, \lambda^*)}{\partial x_j} = 0, \quad j = \overline{1, n}; \quad (6.11)$$

2) допустимості розв'язку:

$$g_i(x^*) = 0, \quad i = \overline{1, m}. \quad (6.12)$$

Якщо при цьому градієнти $g'_1(x^*), \dots, g'_m(x^*)$ лінійно незалежні, то виконується умова регулярності функції Лагранжа і $\lambda_0^* \neq 0$.

Умови (6.11), (6.12) мають місце і у випадку задачі максимізації.

У векторній рівності (6.11) $0 \in \mathbb{R}^n$. Умова (6.11) означає що градієнти $f'(x^*), g'_1(x^*), \dots, g'_m(x^*)$ лінійно залежні, зокрема при $m=1$ $f'(x^*)$ і $g'_1(x^*)$ колінеарні. Довільна точка x^* , що задовольняє при деяких $\lambda_0^*, \lambda^* \in \mathbb{R}^m$ $\left(\sum_{i=1}^m (\lambda_i^*)^2 \neq 0 \right)$ умови (6.11) та (6.12), називається *стаціонарною (критичною) точкою* задачі (6.7). Усі стаціонарні точки визначаються системою (6.11), (6.12), яка має $(n+m)$ рівнянь відносно $(n+m+1)$ невідомих. Якщо $\lambda_0^* \neq 0$, то, поділивши (6.11) на λ_0^* , фактично отримаємо нове $\lambda_0^* = 1$. Це означає, що при розв'язуванні системи (6.11), (6.12) достатньо припускати лише два випадки: $\lambda_0^* = 0$ або $\lambda_0^* = 1$. Отже, число невідомих у цій системі також дорівнює $(n+m)$. У випадку, коли умова регулярності виконується $\lambda_0^* = 1$, розглядають *регулярну або класичну функцію Лагранжа*

$$L(x, \lambda) = \tilde{L}(x, 1, \lambda) = f(x) + \sum_{i=1}^m \lambda_i g_i(x). \quad (6.13)$$

Щоб функція Лагранжа була регулярною, тобто мала вигляд (6.13), потрібно щоб вектори $g'_1(x^*), \dots, g'_m(x^*)$ були лінійно незалежними.

Вектори $g'_1(x^*), \dots, g'_m(x^*)$ називаються лінійно

залежними, якщо існують числа $\lambda_i^* (i = \overline{1, m})$, одночасно не рівні нулю, для яких виконується векторна рівність

$$\lambda_1^* g'_1(x^*) + \dots + \lambda_m^* g'_m(x^*) = 0.$$

Якщо дана рівність виконується лише при всіх $\lambda_i^* = 0 (i = \overline{1, m})$, то $g'_1(x^*), \dots, g'_m(x^*)$ лінійно незалежні. Якщо вектори лінійно залежні, то один з них є лінійною комбінацією інших. Залежною буде система векторів, що містить нульовий вектор, бо тоді лінійна комбінація векторів буде нульовим вектором, якщо всі коефіцієнти дорівнюють нулю, крім одного, що відноситься до вказаного нульового вектора. Один вектор також є лінійно залежною системою (візьмемо для прикладу вектор $g'_1(x^*)$), якщо $g'_1(x^*) = 0$, і лінійно незалежною системою, якщо $g'_1(x^*) \neq 0$.

Згідно з відомою теоремою із математичного аналізу рівняння зв'язку можна розв'язати у деякому околі точки x^* відносно m залежних змінних, якщо ранг матриці Якобі

$$\left(\frac{\partial g_i(x)}{\partial x_j} \right)_{i,j=1}^{m,n} \text{ для функцій } g_i(x), i = \overline{1, m}, \text{ у точці } x^*, \text{ дорівнює}$$

m , тобто відмінний від нуля деякий визначник порядку m , наприклад

$$\begin{vmatrix} \frac{\partial g_1(x^*)}{\partial x_1} & \dots & \frac{\partial g_1(x^*)}{\partial x_m} \\ \dots & \dots & \dots \\ \frac{\partial g_m(x^*)}{\partial x_1} & \dots & \frac{\partial g_m(x^*)}{\partial x_m} \end{vmatrix} \neq 0.$$

Саме ця умова означає, що градієнти $g'_1(x^*), \dots, g'_m(x^*)$ лінійно незалежні, тобто виконується умова регулярності.

Із лінійної незалежності градієнтів випливає, що вони одночасно не перетворюються у нуль у точці x^* , оскільки $\sum_{i=1}^m \lambda_i^* g'_i(x^*) \neq 0$ при жодних одночасно не рівних нулю $\lambda_i^* (i = \overline{1, m})$.

Нехай рівняння зв'язку $g_i(x) = 0, i = \overline{1, m}$ можуть бути розв'язані відносно деяких m ($m < n$) змінних x_j . Тоді ці змінні будуть залежними, а інші $(n - m)$ змінні – незалежними. Нехай для визначеності залежними змінними є $x_1, \dots, x_m$. Тоді $x_j = \varphi_j(x_{m+1}, \dots, x_n), j = \overline{1, m}$. Підставивши їх у функцію $f(x)$, перейдемо від умовної задачі (6.7) до безумовної задачі

$$f(\varphi_1, \dots, \varphi_m, x_{m+1}, \dots, x_n) = F(x_{m+1}, \dots, x_n) \rightarrow \min.$$

Однак на практиці розв'язати рівняння зв'язку відносно залежних змінних у більшості випадків складно, тому для розв'язування задачі (6.7) розроблені інші методи, зокрема *метод множників Лагранжа*, про який ідеться у цьому пункті.

Якщо у околі точки x^* функції $f(x)$ та $g_i(x)$ неперервно диференційовні, а градієнти $g'_i(x^*) (i = \overline{1, m})$ лінійно незалежні, то точку x^* ще називають *регулярною точкою локального мінімуму*.

У цьому випадку варіантом теореми 6.1 є така теорема.

Теорема 6.2 (необхідні умови мінімуму першого порядку для задачі (6.7) у випадку регулярної (класичної) функції Лагранжа). Нехай:

1) x^* – регулярна точка локального мінімуму у задачі (6.7);

2) функції $f(x)$, $g_i(x)$ ($i = \overline{1, m}$) неперервно диференційовні у деякому околі точки x^* .

Тоді існують числа $\lambda_1^*, \dots, \lambda_m^*$ ($\|\lambda^*\| \neq 0$), які одночасно не рівні нулю і при яких виконуються умови:

1) стаціонарності регулярної функції Лагранжа (6.13) по x :

$$L'_x(x^*, \lambda^*) = 0 \quad \text{або} \quad \frac{\partial L(x^*, \lambda^*)}{\partial x_j} = 0, \quad j = \overline{1, n}; \quad (6.14)$$

2) допустимості розв'язку (6.12): $g_i(x^*) = 0$, $i = \overline{1, m}$.

Умови (6.14), (6.12) необхідні і для відповідної задачі на умовний максимум.

Теорему 6.2 часто називають *правилом множників Лагранжа*. Зазначимо також, що теорема 6.2 впливає з теореми 6.1.

Зауважимо, що умова регулярності принципова. Наприклад, у задачі $x \rightarrow \min$, $x^2 = 0$ ($x \in \mathbb{R}^1$) точка $x^* = 0$ є точкою мінімуму (але не регулярною) і рівність (6.14) не виконується при жодному λ^* , оскільки $f'(x^*) = 1$, а $g'(x^*) = 0$ (тобто не виконується умова $\lambda^* g'(x^*) \neq 0$ при жодному $\lambda^* \neq 0$). У випадку, коли $\lambda_0 = 0$, узагальнена функція Лагранжа (6.8) не має доданка $\lambda_0 f(x)$, тобто цільова функція $f(x)$ надалі у дослідженні не використовується. Точка екстремуму у даній ситуації називається *нерегулярною*. Перевірка умови регулярності на практиці це не проста задача, тому найчастіше здійснюють безпосередню перевірку умови $\lambda_0 \neq 0$.

Як у задачах на безумовний мінімум, так і у задачі (6.7) необхідні умови не визначають характер стаціонарної

точки. Для цього потрібні достатні умови, які пов'язані з вимогою знаковизначеності у точці x^* другого диференціала функції Лагранжа при $\lambda = \lambda^*$, тобто $d^2L(x^*, \lambda^*)$. Із конструкції функції Лагранжа (6.13) очевидно, що при наявності умов зв'язку $g_i(x) = 0$ ($i = \overline{1, m}$) екстремуми (у даному разі мінімуми) функцій $f(x)$ і $L(x, \lambda)$ на множині (6.6) збігаються.

Нехай x^* – точка локального мінімуму у задачі (6.7), а вектор $\Delta x = (\Delta x_1, \dots, \Delta x_n)^T = (dx_1, \dots, dx_n)^T = dx$ такий, що $(x^* + dx) \in X$ (тобто $g_i(x^* + dx) = 0, i = \overline{1, m}$). Тоді

$$\Delta g_i(x^*) = \langle g'_i(x^*), dx \rangle + o(\|dx\|) = 0, i = \overline{1, m}.$$

З точністю до нескінченно малих більш високого порядку, ніж $\|\Delta x\| = \|dx\|$, маємо

$$dg_i(x^*) = \langle g'_i(x^*), dx \rangle = 0, i = \overline{1, m}. \quad (6.15)$$

Очевидно, із (6.15) при кожному i можна визначити прирости m залежних змінних через прирости $(n - m)$ незалежних змінних. Для цих приростів (тобто тих, що задовольняють (6.15)) $\Delta f(x^*) = f(x^* + dx) - f(x^*) \geq 0$, оскільки x^* – точка мінімуму, але у цій точці (як уже в точці безумовного мінімуму складної функції) перший диференціал дорівнює нулю (форма першого диференціала інваріантна), тобто

$$\langle f'(x^*), dx \rangle = \sum_{j=1}^n \frac{\partial f(x^*)}{\partial x_j} dx_j = 0. \quad (6.16)$$

Зауважимо, що із (6.15) і (6.16) і випливає при $\lambda_0 = 1$ необхідна умова (6.11) чи (6.14). Із урахуванням (6.15) і (6.16) приріст функції Лагранжа у точці x^* при фіксованому

λ^* і малих приростах Δx обчислюється за формулою

$$\begin{aligned}\Delta L(x^*, \lambda^*) &= dL(x^*, \lambda^*) + \frac{1}{2} d^2 L(x^*, \lambda^*) + o(\|dx\|^2) = \\ &= \frac{1}{2} d^2 L(x^*, \lambda^*) + o(\|dx\|^2),\end{aligned}$$

тобто знак приросту $\Delta L(x^*, \lambda^*)$ визначається другим диференціалом

$$d^2 L(x^*, \lambda^*) = \sum_{j=1}^n \sum_{k=1}^n \frac{\partial^2 L(x^*, \lambda^*)}{\partial x_j \partial x_k} dx_j dx_k + \sum_{j=1}^n \frac{\partial L(x^*, \lambda^*)}{\partial x_j} d^2 x_j.$$

У цій рівності $d^2 x_j \neq 0$, оскільки диференціали dx_j залежні, але із (6.15) і (6.16) (або із (6.14)) випливає, що другий доданок дорівнює нулю, тобто

$$d^2 L(x^*, \lambda^*) = \langle L''_{xx}(x^*, \lambda^*) dx, dx \rangle,$$

де через

$$L''_{xx}(x^*, \lambda^*) = f''(x^*) + \sum_{i=1}^m \lambda_i^* g_i''(x^*)$$

позначено матрицю других похідних функції Лагранжа (6.13) по змінним x_j ($j = \overline{1, n}$). Отже, другий диференціал функції Лагранжа у точці x^* має таку ж форму запису, як і випадку, коли всі змінні x_j ($j = \overline{1, n}$) незалежні.

Теорема 6.3 (необхідні умови мінімуму другого порядку для задачі (6.7) у випадку регулярної функції Лагранжа). Нехай:

1) x^* – регулярна точка локального мінімуму у задачі (6.7), а $(x^*, \lambda^*)^T$ – розв'язок системи (6.14), (6.12) ($\|\lambda^*\| \neq 0$);

2) функції $f(x)$, $g_i(x)$ ($i = \overline{1, m}$) двічі неперервно диференційовні у деякому околі x^* .

Тоді у точці $(x^*, \lambda^*)^T$ другий диференціал

$$d^2L(x^*, \lambda^*) = \langle L''_{xx}(x^*, \lambda^*) dx, dx \rangle \geq 0 \quad (6.17)$$

при всіх $dx \in \mathbb{R}^n$, що задовольняють умови (6.15):

$$dg_i(x^*) = 0 \quad \text{або} \quad \sum_{j=1}^n \frac{\partial g_i(x^*)}{\partial x_j} dx_j = 0, \quad i = \overline{1, m}.$$

Очевидно, необхідна умова мінімуму (6.17) у теоремі 6.3 означає, що другий диференціал класичної функції Лагранжа у точці $(x^*, \lambda^*)^T$ невід'ємний, тобто матриця $L''_{xx}(x^*, \lambda^*)$ невід'ємно визначена на дотичному підпросторі (6.15). Для класичної задачі на умовний максимум у точці $(x^*, \lambda^*)^T$ другий диференціал недодатний, тобто матриця $L''_{xx}(x^*, \lambda^*)$ недодатно визначена (у (6.17) потрібно поміняти знак нерівності).

Теорема 6.4 (достатні умови мінімуму другого порядку для задачі (6.7) у випадку регулярної функції Лагранжа). Нехай:

1) $(x^*, \lambda^*)^T$ – розв'язок системи (6.14), (6.12) ($\|\lambda^*\| \neq 0$);

2) функції $f(x)$, $g_i(x)$ ($i = \overline{1, m}$) двічі неперервно диференційовні у деякому околі точки x^* .

Якщо у точці $(x^*, \lambda^*)^T$ другий диференціал

$$d^2L(x^*, \lambda^*) = \langle L''_{xx}(x^*, \lambda^*) dx, dx \rangle > 0 \quad (6.18)$$

при всіх ненульових $dx \in \mathbb{R}^n$, що задовольняють умови (6.15):

$$dg_i(x^*) = \sum_{j=1}^n \frac{\partial g_i(x^*)}{\partial x_j} dx_j = 0, \quad i = \overline{1, m},$$

то x^* – точка локального мінімуму у задачі (6.7).

Інакше кажучи, якщо у точці x^* виконана необхідна

умова першого порядку, а $d^2L(x^*, \lambda^*) > 0$, тобто $L''_{xx}(x^*, \lambda^*)$ додатно визначена на дотичному підпросторі (6.15), то x^* – точка мінімуму.

Із (6.18) випливає, що для класичної задачі на умовний максимум достатньою умовою є умова $d^2L(x^*, \lambda^*) < 0$, тобто матриця $L''_{xx}(x^*, \lambda^*)$ повинна бути від'ємно визначеною.

Якщо матриця $L''_{xx}(x^*, \lambda^*)$ додатно визначена, тобто нерівність (6.18) виконується при всіх ненульових векторах dx , то перевірка достатньої умови суттєво спрощується. У протилежному випадку для такої перевірки потрібно із системи (6.15) виражати залежні dx_j та підставляти їх у квадратичну форму для дослідження її додатної визначеності.

Усі можливі результати дослідження другого диференціала $d^2L(x^*, \lambda^*)$ $((x^*, \lambda^*)^T$ – розв'язок системи (6.14), (6.12)) такі:

- 1) $d^2L(x^*, \lambda^*) > 0$ ($dx \neq 0$, виконані умови (6.15)) – x^* є точкою локального мінімуму у задачі (6.7);
- 2) $d^2L(x^*, \lambda^*) < 0$ ($dx \neq 0$, виконані умови (6.15)) – x^* є точкою локального максимуму у задачі максимізації функції $f(x)$ на допустимій множині (6.6);
- 3) $d^2L(x^*, \lambda^*) \geq 0$ (виконані умови (6.15)) – x^* може бути точкою локального мінімуму у задачі (6.7), однак для з'ясування цього факту потрібні додаткові дослідження;
- 4) $d^2L(x^*, \lambda^*) \leq 0$ (виконані умови (6.15)) – x^* може бути точкою локального максимуму у задачі максимізації функції $f(x)$ на допустимій множині (6.6), однак для з'ясування цього факту потрібні

додаткові дослідження;

- 5) $d^2L(x^*, \lambda^*) = 0$ (виконані умови (6.15)) – для з'ясування характеру точки x^* потрібні додаткові дослідження;
- 6) $d^2L(x^*, \lambda^*)$ (виконані умови (6.15)) – знакозмінна квадратична форма (може набувати додатних, від'ємних та нульових значень при $dx \neq 0$), тому у точці x^* немає екстремуму.

Звичайно знаковизначеність $d^2L(x^*, \lambda^*)$ визначається визначеністю матриці $L''_{xx}(x^*, \lambda^*)$, яка у даному разі є певним аналогом матриці Гессе $H(x^*)$.

Алгоритм розв'язування задачі (6.7) методом множників Лагранжа передбачає такі кроки.

Крок 1. Перевірка умови регулярності (якщо це можливо) та побудова функції Лагранжа (6.13).

Крок 2. Знаходження розв'язків системи (6.11), (6.12) чи (6.14), (6.12), тобто знаходження стаціонарних точок функції Лагранжа.

Крок 3. Дослідження стаціонарних точок за допомогою достатніх умов і відбір точок локального умовного мінімуму.

Крок 4. Порівняння значень функції у точках локального умовного мінімуму і знаходження точки глобального мінімуму.

Приклад 6.1. Розв'язати задачу мінімізації функції

$$f(x) = (x_1 - 4)^2 + (x_2 - 4)^2, \text{ якщо } x_1^2 + x_2^2 = 8.$$

◀ *Крок 1.* Оскільки $g(x) = x_1^2 + x_2^2 - 8$, то $g'(x) = (2x_1, 2x_2)^T \neq 0$ для всіх x , що задовольняють умову зв'язку $g(x) = 0$. Один ненульовий вектор є лінійно незалежним, тому умова регулярності виконується. Отже,

використовуємо класичну функцію Лагранжа
 $L(x, \lambda) = (x_1 - 4)^2 + (x_2 - 4)^2 + \lambda(x_1^2 + x_2^2 - 8)$.

Крок 2.

$$\begin{cases} \frac{\partial L}{\partial x_1} = 2(x_1 - 4) + 2\lambda x_1 = 0, \\ \frac{\partial L}{\partial x_2} = 2(x_2 - 4) + 2\lambda x_2 = 0, \\ \frac{\partial L}{\partial \lambda} = g(x) = x_1^2 + x_2^2 - 8 = 0. \end{cases}$$

Розв'язавши дану систему, отримаємо два розв'язки:

- 1) $x_1^{(1)} = x_2^{(1)} = 2, \lambda^{(1)} = 1;$
- 2) $x_1^{(2)} = x_2^{(2)} = -2, \lambda^{(2)} = -3.$

Крок 3. Матриця других похідних

$$L''_{xx}(x, \lambda) = \begin{pmatrix} 2(1 + \lambda) & 0 \\ 0 & 2(1 + \lambda) \end{pmatrix}$$

для першого розв'язку ($\lambda^{(1)} = 1$) додатно визначена, тому $x^{(1)} = (2, 2)^T$ – точка локального умовного мінімуму. У другому випадку ($\lambda^{(2)} = -3$) матриця $L''_{xx}(x^{(2)}, \lambda^{(2)})$ від'ємно визначена, тобто $x^{(2)} = (-2, -2)^T$ – точка локального умовного максимуму. Цей висновок можна також зробити, дослідивши квадратичну форму (другий диференціал) $\langle L''_{xx}(x^{(2)}, \lambda^{(2)}) dx, dx \rangle = 2(1 + \lambda^{(2)}) dx_1^2 + 2(1 + \lambda^{(2)}) dx_2^2$ у точці $x^{(2)}$. Із (6.15), тобто із рівності $2x_1^{(2)} dx_1 + 2x_2^{(2)} dx_2 = 0$ або $(-4dx_1 - 4dx_2) = 0$ випливає, що $dx_2 = -dx_1$, тому $\langle L''_{xx}(x^{(2)}, \lambda^{(2)}) dx, dx \rangle = 2(1 + \lambda^{(2)})(dx_1^2 + (-dx_1)^2) = 2(1 - 3) \cdot 2dx_1^2 = -8dx_1^2 < 0$ при всіх $dx_1 \neq 0$, що й потрібно було встановити.

Крок 4. Розв'язком задачі є точка $x^* = x^{(1)} = (2, 2)^T$. При

цьому мінімальне значення функції $f^* = 8$. ►

Приклад 6.2. Знайти умовний екстремум функції $f(x) = 2x_1 + x_2$, якщо $x_1^2 + x_2^2 = 5$.

◄ Крок 1. Не перевіряючи умову регулярності, розглянемо для даної задачі узагальнену функцію Лагранжа

$$L(x, \lambda_0, \lambda) = \lambda_0(2x_1 + x_2) + \lambda(x_1^2 + x_2^2 - 4).$$

Крок 2. Система (6.11), (6.12) у даному випадку має вигляд

$$\begin{cases} \frac{\partial L}{\partial x_1} = 2\lambda_0 + 2\lambda x_1 = 0, \\ \frac{\partial L}{\partial x_2} = \lambda_0 + 2\lambda x_2 = 0, \\ \frac{\partial L}{\partial \lambda} = x_1^2 + x_2^2 - 5 = 0. \end{cases}$$

Проаналізуємо систему для двох випадків.

1) $\lambda_0 = 0$. Тоді $2\lambda x_1 = 0$, $2\lambda x_2 = 0$, $x_1^2 + x_2^2 = 5$. Оскільки всі множники Лагранжа не можуть бути одночасно нульовими, то $\lambda \neq 0$ і $x_1 = 0$, $x_2 = 0$, але точка $(0,0)^T$ не належить допустимій множині, яка описується рівнянням зв'язку. Отже, дана система несутісна.

2) $\lambda_0 = 1$ Тоді система рівнянь $2 + 2\lambda x_1 = 0$, $1 + 2\lambda x_2 = 0$, $x_1^2 + x_2^2 = 5$ має два розв'язки а) $x_1 = 2$, $x_2 = 1$, $\lambda = -\frac{1}{2}$;
б) $x_1 = -2$, $x_2 = -1$, $\lambda = \frac{1}{2}$.

Крок 3. Матриця других похідних

$$L''_{xx}(x, \lambda) = \begin{pmatrix} 2\lambda & 0 \\ 0 & 2\lambda \end{pmatrix}$$

у випадку а), тобто при $\lambda = -\frac{1}{2}$, від'ємно визначена, а у випадку б), тобто при $\lambda = \frac{1}{2}$ – додатно визначена. Це означає, що $x^{(1)} = (2, 1)^T$ – точка умовного максимуму, а $x^{(2)} = (-2, -1)^T$ – точка умовного мінімуму. При цьому $f_{\max}^* = 5$, а $f_{\min}^* = -5$. ►

Завершуючи даний пункт, зазначимо, що множники Лагранжа (або невизначені множники у функції Лагранжа) – це не просто довільні коефіцієнти, введені для певних математичних маніпуляцій, а *коефіцієнти чутливості* оптимального значення f^* цільової функції класичної задачі на умовний екстремум до змін правих частин умов зв'язку. У цьому легко переконатися на прикладі задачі $f(x_1, x_2) \rightarrow \min$, $g_1(x_1, x_2) = b_1$, для якої функція Лагранжа $L(x, \lambda) = f(x) + \lambda(g_1(x) - b_1)$. Нехай $(x^*, \lambda^*)^T$ – розв'язок цієї задачі. Тоді $g_1(x^*) = b_1$, тому $f^* = f(x^*) = L(x^*, \lambda^*)$. Очевидно, що x^*, λ^* і f^* функціонально залежать від b_1 , тобто $x^* = x^*(b_1)$, $\lambda^* = \lambda^*(b_1)$, $f^* = f^*(b_1)$. З одного боку,

$$\frac{\partial f^*}{\partial b_1} = \frac{\partial f^*}{\partial x_1} \cdot \frac{\partial x_1^*}{\partial b_1} + \frac{\partial f^*}{\partial x_2} \cdot \frac{\partial x_2^*}{\partial b_1}, \quad (6.19)$$

а з іншого – (враховуючи, що $g_1(x^*) = b_1$)

$$\frac{\partial g_1}{\partial x_1} \cdot \frac{\partial x_1^*}{\partial b_1} + \frac{\partial g_1}{\partial x_2} \cdot \frac{\partial x_2^*}{\partial b_1} - 1 = 0. \quad (6.20)$$

Помноживши (6.20) на λ^* і додавши (6.19), отримаємо

$$\frac{\partial f^*}{\partial b_1} = -\lambda^* + \sum_{j=1}^2 \left(\frac{\partial f^*}{\partial x_j} + \lambda^* \frac{\partial g_1}{\partial x_j} \right) \cdot \frac{\partial x_j^*}{\partial b_1}. \quad (6.21)$$

Оскільки $(x^*, \lambda^*)^T$ – стаціонарна точка функції Лагранжа, то вирази у круглих дужках правої частини (6.21) дорівнюють нулю і $\frac{\partial f^*}{\partial b_1} = -\lambda^*$, тобто множник Лагранжа визначає швидкість зміни оптимального значення f^* залежно від зміни b_1 .

У випадку класичної задачі на умовний екстремум з m обмеженнями-рівностями $g_i(x) = b_i$ ($i = \overline{1, m}$) аналогічно доводиться, що $\frac{\partial f^*}{\partial b_i} = -\lambda_i^*$ ($i = \overline{1, m}$).

6.3. Необхідні та достатні умови екстремуму у випадку задачі математичного програмування із функціональними обмеженнями-нерівностями

У цьому пункті з'ясуємо диференціальні умови екстремуму (оптимальності) для задачі (6.2), (6.3) при $k = m$, тобто задачі

$$\begin{cases} f(x) \rightarrow \min, \\ g_i(x) \leq 0, \quad i = \overline{1, m}. \end{cases} \quad (6.22)$$

Як і раніше, будемо вважати, що функції $f(x)$, $g_i(x)$ ($i = \overline{1, m}$) визначені та двічі неперервно диференційовні при всіх $x \in \mathbb{R}^n$. З метою знаходження розв'язку задачі (6.22) зведемо її до еквівалентної задачі на класичний мінімум, перейшовши від обмежень-нерівностей до обмежень-рівностей

$$g_i(x) + y_i^2 = 0, \quad i = \overline{1, m}, \quad (6.23)$$

де y_i ($i = \overline{1, m}$) – допоміжні змінні.

З урахуванням (6.23) задача (6.22) рівносильна задачі

$$\begin{cases} f(x) \rightarrow \min, \\ g_i(x) + y_i^2 = 0, \quad i = \overline{1, m}. \end{cases} \quad (6.24)$$

Рівносильність задач (6.22) і (6.24) полягає у тому, що, якщо x^* – точка локального мінімуму у задачі (6.22), то $(x^*, y^*)^T$, де

$$y^* = \left(y_1^* = (-g_1(x^*))^{1/2}, \dots, y_m^* = (-g_m(x^*))^{1/2} \right)^T,$$

буде точкою локального мінімуму у задачі (6.24), і навпаки, якщо $(x^*, y^*)^T$ – точка локального мінімуму у задачі (6.24), то x^* – точка локального мінімуму у задачі (6.22).

Отже, еквівалентом задачі (6.22) є задача на класичний мінімум (6.24). Функція Лагранжа (узагальнена) для задачі (6.24) має вигляд

$$L_1(x, \lambda_0, \lambda) = \lambda_0 f(x) + \sum_{i=1}^m \lambda_i [g_i(x) + y_i^2]. \quad (6.25)$$

Перш ніж сформулювати для задачі (6.22) на основі задачі (6.24) із функцією Лагранжа (6.25) теореми про необхідні та достатні умови екстремуму, зауважимо, що для функції (6.25) похідні $\frac{\partial L_1(x, \lambda_0, \lambda)}{\partial x_j} \quad (j = \overline{1, n})$ знаходяться за

формулами (6.9), а

$$\frac{\partial L_1(x, \lambda_0, \lambda)}{\partial \lambda_i} = g_i(x) + y_i^2 = 0, \quad i = \overline{1, m}, \quad (6.26)$$

$$\frac{\partial L_1(x, \lambda_0, \lambda)}{\partial y_i} = 2\lambda_i y_i = 0, \quad i = \overline{1, m}. \quad (6.27)$$

Очевидно, умови (6.26) рівносильні умовам (6.3) при $k = m$. Помноживши (6.27) на $y_i/2$, отримаємо $\lambda_i y_i^2 = 0$

$(i = \overline{1, m})$ або

$$\lambda_i g_i(x) = 0, \quad i = \overline{1, m}. \quad (6.28)$$

Перед тим, як сформулювати теорему про необхідні умови екстремуму першого порядку, зауважимо, що на відміну від класичної задачі на екстремум (задачі з обмеженнями-рівностями), ці умови формулюються окремо для мінімуму і максимуму.

Стосовно задачі математичного програмування на максимум зазначимо таке. Оскільки довільна задача максимізації функції $f(x)$ на довільній множині X рівносильна задачі мінімізації функції $(-f(x))$ на X (для значення функції у екстремальній точці виконується рівність: $\max_{x \in X} f(x) = -\min_{x \in X} (-f(x))$), тобто точка максимуму функції $f(x)$ є точкою мінімуму функції $(-f(x))$, то у випадку задачі

$$\begin{cases} f(x) \rightarrow \max, \\ g_i(x) \leq 0, \quad i = \overline{1, m} \end{cases} \quad (6.29)$$

розглядаємо рівносильну задачу $(-f(x)) \rightarrow \min$, $x \in X = \{x \in \mathbb{R}^n / g_i(x) \leq 0, \quad i = \overline{1, m}\}$. Однак для цієї задачі функція Лагранжа (6.8) має вигляд

$$\begin{aligned} \tilde{L}(x, \lambda_0, \lambda) &= \lambda_0(-f(x)) + \sum_{i=1}^m \lambda_i g_i(x) = (-\lambda_0) f(x) + \sum_{i=1}^m \lambda_i g_i(x) = \\ &= -\left[\lambda_0 f(x) + \sum_{i=1}^m (-\lambda_i) g_i(x) \right]. \end{aligned} \quad (6.30)$$

Урахувавши (6.25)-(6.30), сформулюємо наступну теорему.

Теорема 6.5 (необхідні умови мінімуму (максимуму) першого порядку). Нехай:

1) x^* – точка локального мінімуму (максимуму) у задачі

(6.22) (у задачі (6.29));

2) функції $f(x)$, $g_i(x)$ ($i = \overline{1, m}$) неперервно диференційовні у деякому околі точки x^* .

Тоді існують одночасно не рівні нулю числа $\lambda_0^* \geq 0$ і $\lambda_1^*, \dots, \lambda_m^*$, що виконуються умови:

1) стаціонарності (6.11) узагальненої функції Лагранжа (6.8) по x :

$$\frac{\partial \tilde{L}(x^*, \lambda_0^*, \lambda^*)}{\partial x_j} = 0, \quad j = \overline{1, n};$$

2) допустимості розв'язку:

$$g_i(x^*) \leq 0, \quad i = \overline{1, m}; \quad (6.31)$$

3) доповнюючої нежорсткості:

$$\lambda_i^* g_i(x^*) = 0, \quad i = \overline{1, m}; \quad (6.32)$$

4) невід'ємності (недодатності) числових множників Лагранжа:

$$\lambda_i^* \geq 0 \quad (\lambda_i^* \leq 0), \quad i = \overline{1, m}. \quad (6.33)$$

Якщо при цьому градієнти активних у точці x^* обмежень лінійно незалежні, то виконується умова регулярності і $\lambda_0^* \neq 0$.

Обґрунтуємо детальніше теорему 6.5, яка узагальнює теорему 6.1 на випадок обмежень-нерівностей. Умови стаціонарності впливають з того, що похідні $\partial L_1(x, \lambda_0, \lambda) / \partial x_j$ та $\partial \tilde{L}(x, \lambda_0, \lambda) / \partial x_j$ ($j = \overline{1, n}$) збігаються. Умови допустимості (6.31) – очевидний наслідок (6.26), оскільки $y_i^2 \geq 0$ ($i = \overline{1, m}$). Аналогічно із (6.28) впливають умови доповнюючої нежорсткості (6.32), які дозволяють зробити такі висновки:

1) якщо $\lambda_i^* \neq 0$, то $g_i(x^*) = 0$ (i – те обмеження

називається активним);

2) якщо $g_i(x^*) < 0$ (i – те обмеження пасивне), то $\lambda_i^* = 0$.

Умови невід’ємності (недодатності) числових множників Лагранжа (6.33) гарантують збереження знака функцій $g_i(x)$ ($i = \overline{1, m}$) у функції Лагранжа.

Усі множники Лагранжа визначаються з точністю до додатної константи, зокрема, якщо $\lambda_0^* > 0$, то при $\alpha = 1/\lambda_0^*$ величини $\alpha\lambda_i^*$ ($i = \overline{1, m}$) також будуть множниками Лагранжа. Це дає підстави розглядати лише два випадки: $\lambda_0^* = 0$ і $\lambda_0^* = 1$. Будь-яка допоміжна умова, що гарантує випадок $\lambda_0^* = 1$, як ми уже знаємо, називається *умовою регулярності*. Тоді для задачі (6.22) (чи (6.29)) використовується *регулярна, або класична* функція Лагранжа (6.13). Як стверджує теорема 6.5, для того, щоб функція Лагранжа (а значить, і задача у цілому) була регулярною, потрібно, щоб градієнти активних у точці x^* обмежень були лінійно незалежні.

Конкретизуємо кілька важливих зауважень.

1. Як випливає з теореми 6.5, необхідні умови мінімуму та максимуму для задач (6.22) і (6.29) відрізняються лише умовами знаковизначеності множників Лагранжа λ_i^* ($i = \overline{1, m}$). Зауважимо також, що із (6.30) випливає, що для задачі (6.29) умови $\lambda_0^* \geq 0$, $\lambda_i^* \leq 0$ ($i = \overline{1, m}$) можна замінити еквівалентними умовами $\lambda_0^* \leq 0$, $\lambda_i^* \geq 0$ ($i = \overline{1, m}$).

2. Крім того, з теореми (6.5) випливає, що:

1) якщо функції $f(x)$, $g_i(x)$ ($i = \overline{1, m}$) опуклі, то умови теореми є необхідними й достатніми умовами глобального

мінімуму;

2) якщо функції $(-f(x)), g_i(x)$ ($i = \overline{1, m}$) опуклі, то умови теореми є необхідними й достатніми умовами глобального максимуму.

3. Якщо виконується умова регулярності для задач (6.22), (6.29), то теорема 6.5 переформулюється без умови на λ_0^* та останнього речення про умову регулярності.

Нехай I_a – множина індексів обмежень, активних у точці x^* , $L(x, \lambda)$ – регулярна (класична) функція Лагранжа (6.13).

Теорема 6.6 (необхідні умови мінімуму (максимуму) другого порядку). Нехай:

1) x^* – регулярна точка локального мінімуму (максимуму) у задачі (6.22) (у задачі (6.29)), а $(x^*, \lambda^*)^T$ – розв’язок системи (6.11), (6.31)–(6.33) при $\lambda_0^* = 1$ ($\|\lambda^*\| \neq 0, \lambda^* = (\lambda_1^*, \dots, \lambda_m^*)^T$);

2) функції $f(x), g_i(x)$ ($i = \overline{1, m}$) двічі неперервно диференційовні у деякому околі x^* .

Тоді у точці $(x^*, \lambda^*)^T$ другий диференціал

$$d^2 L(x^*, \lambda^*) \geq 0 \quad (d^2 L(x^*, \lambda^*) \leq 0) \quad (6.34)$$

при всіх $dx \in \mathbb{R}^n$, що задовольняють умови:

$$1) \quad \begin{cases} dg_i(x^*) = 0 \text{ для всіх } i \in I_a, \\ \text{для яких } \lambda_i^* > 0 \text{ } (\lambda_i^* < 0); \end{cases} \quad (6.35)$$

$$2) \quad dg_i(x^*) \leq 0 \text{ для всіх } i \in I_a, \text{ для яких } \lambda_i^* = 0. \quad (6.36)$$

Для задачі з обмеженнями-нерівностями для з’ясування характеру стаціонарних точок використовують достатні умови екстремуму першого та другого порядків.

Теорема 6.7 (достатні умови мінімуму (максимуму))

першого порядку). Нехай:

1) $(x^*, \lambda^*)^T$ – розв’язок системи (6.11), (6.31)–(6.33) при $\lambda_0^* = 1$ ($\|\lambda^*\| \neq 0$);

2) функції $f(x)$, $g_i(x)$ ($i = \overline{1, m}$) неперервно диференційовні у деякому околі x^* ;

3) число l активних обмежень у точці x^* збігається з числом n ($l = n$) змінних (при цьому виконується умова регулярності).

Якщо $\lambda_i^* > 0$ для всіх $i \in I_a$, то x^* – точка локального мінімуму у задачі (6.22), а якщо $\lambda_i^* < 0$ для всіх $i \in I_a$, то x^* – точка локального максимуму у задачі (6.29).

Зазначимо, що теорема 6.7 дозволяє у деяких випадках визначити характер стаціонарної точки x^* , користуючись лише першими похідними і не вимагаючи опуклості оптимізаційної задачі. Теорема 6.7 не має аналогів ні в задачах безумовної оптимізації, ні в задачах з обмеженнями-рівностями (тобто у класичних задачах оптимізації). У цьому її унікальність.

Теорема 6.8 (достатні умови мінімуму (максимуму) другого порядку). Нехай:

1) $(x^*, \lambda^*)^T$ – розв’язок системи (6.11), (6.31)–(6.33) при $\lambda_0^* = 1$ ($\|\lambda^*\| \neq 0$);

2) функції $f(x)$, $g_i(x)$ ($i = \overline{1, m}$) двічі неперервно диференційовні у деякому околі x^* .

Якщо у точці $(x^*, \lambda^*)^T$ другий диференціал

$$d^2L(x^*, \lambda^*) > 0 \quad (d^2L(x^*, \lambda^*) < 0) \quad (6.37)$$

при всіх ненульових $dx \in \mathbb{R}^n$, що задовольняють умови (6.35)

для мінімуму (максимуму) та (6.36), то x^* – точка локального мінімуму (максимуму) у задачі (6.22) (у задачі (6.29)).

При дослідженні знайдених стаціонарних точок спочатку перевіряють достатні умови першого порядку. Якщо число l активних у точці x^* обмежень менше від числа невідомих n ($l < n$) або відповідні множники Лагранжа не задовольняють достатні умови першого порядку, то перевіряють достатні умови другого порядку. Якщо достатні умови першого та другого порядку не виконуються, то треба перевірити необхідні умови другого порядку. Якщо останні умови виконуються, то потрібні додаткові дослідження. У протилежному випадку точка x^* не є точкою умовного екстремуму.

Результати дослідження стаціонарних точок за допомогою достатніх умов екстремуму другого порядку приводять до таких висновків:

- 1) $d^2L(x^*, \lambda^*) > 0$ ($dx \neq 0$, виконані умови (6.35) для мінімуму та умови (6.36)) – x^* є точкою локального мінімуму;
- 2) $d^2L(x^*, \lambda^*) < 0$ ($dx \neq 0$, виконані умови (6.35) для максимуму та умови (6.36)) – x^* є точкою локального максимуму;
- 3) $d^2L(x^*, \lambda^*) \geq 0$ (виконані умови (6.35) для мінімуму та умови (6.36)) – x^* може бути точкою локального мінімуму, потрібні додаткові дослідження;
- 4) $d^2L(x^*, \lambda^*) \leq 0$ (виконані умови (6.35) для максимуму та умови (6.36)) – x^* може бути точкою локального максимуму, потрібні додаткові дослідження;

- 5) $d^2L(x^*, \lambda^*) = 0$ (виконані умови (6.35) для мінімуму та умови (6.36)) – потрібні додаткові дослідження;
- 6) $d^2L(x^*, \lambda^*) = 0$ (виконані умови (6.35) для максимуму та умови (6.36)) – потрібні додаткові дослідження;
- 7) $d^2L(x^*, \lambda^*)$ – знакозмінна квадратична форма (виконані умови (6.35) для мінімуму та умови (6.36)) – x^* не є точкою екстремуму;
- 8) $d^2L(x^*, \lambda^*)$ – знакозмінна квадратична форма (виконані умови (6.35) для максимуму та умови (6.36)) – x^* не є точкою екстремуму.

Алгоритм розв'язування задач (6.22), (6.29) конкретизується такими кроками.

Крок 1. Побудова узагальненої функції Лагранжа (6.8) або класичної функції Лагранжа (6.13), якщо попередньо вдається встановити регулярність задачі.

Крок 2. Формування системи співвідношень (6.11), (6.31)–(6.33) та знаходження її розв'язків (при цьому систему необхідних умов екстремуму потрібно розв'язати при $\lambda_0^* = 0$ (якщо регулярність не встановлена) і $\lambda_0^* = 1$).

Крок 3. Дослідження знайдених регулярних стаціонарних розв'язків за допомогою достатніх умов екстремуму першого та другого порядку.

Крок 4. Обчислення значень функції у знайдених точках умовного екстремуму і вибір точки (чи точок) глобального екстремуму.

Практична реалізація цього алгоритму навіть у найпростіших випадках не проста, що і буде продемонстровано нижче.

На першому кроці задачу математичного програмування

потрібно привести до вигляду (6.22) чи (6.29). Зазначимо, що для кожної задачі можна виписати багато функцій Лагранжа залежно від того, які з обмежень будемо вважати функціональними, а які – прямими, тобто включеними у множину X_0 .

Якщо вдається на першому кроці перевірити яку-небудь умову регулярності, то звичайно будується регулярна (класична) функція Лагранжа.

На другому кроці для задачі будується система співвідношень (6.11), (6.31)–(6.33), що є необхідними умовами оптимальності.

Якщо задача регулярна, то у (6.11) $\lambda_0^* = 1$.

Другий крок найскладніший. Далеко не завжди вдається розв'язати систему необхідних умов оптимальності. Це один з проблемних етапів процесу розв'язування задачі.

Третій крок також складний. Якщо задача є регулярною задачею опуклого програмування, то знайдена стаціонарна точка x^* – розв'язок (глобальний) задачі. Інколи вдається використати не зовсім зручні та громіздкі умови оптимальності другого порядку (достатні), а інколи безпосередньо досліджують поведінку цільової функції у стаціонарній точці, наприклад, якщо відомо, що задача має розв'язок (глобальний), то цим розв'язком буде та стаціонарна точка, у якій цільова функція має найменше значення. Є й інші підходи, які пов'язані зі специфікою конкретної задачі. Деякі з них будуть продемонстровані у наступних прикладах (для зручності викладення матеріалу у прикладах символ «*» біля λ_0 і λ ставити не будемо).

Крок четвертий – найпростіший і зводиться до виконання елементарних обчислювальних операцій.

Приклад 6.3. Знайти точки локального екстремуму

функції $f(x) = \frac{1}{2}x_1^2 + x_2^2$ на множині $X = \{x \in \mathbb{R}^2 / x_1^2 + x_2^2 \geq 1, x_1^3 + x_2^3 \leq 1\}$.

◀ Крок 1. Функція Лагранжа (узагальнена) має вигляд

$$\tilde{L}(x, \lambda_0, \lambda) = \lambda_0 \left(\frac{1}{2}x_1^2 + x_2^2 \right) + \lambda_1 (1 - x_1^2 - x_2^2) + \lambda_2 (x_1^3 + x_2^3 - 1).$$

Крок 2. Система необхідних умов оптимальності така (умови $\lambda_0 \geq 0$, $\lambda_i \leq 0$ ($i = \overline{1, m}$) для задачі на максимум у даному прикладі замінимо еквівалентними умовами $\lambda_0 \leq 0$, $\lambda_i \geq 0$ ($i = \overline{1, m}$), тобто надалі $\lambda_0 \geq 0$ у задачі на мінімум і $\lambda_0 \leq 0$ у задачі на максимум):

$$\frac{\partial L}{\partial x_1} = x_1(\lambda_0 - 2\lambda_1 + 3\lambda_2 x_1) = 0, \quad (6.38)$$

$$\frac{\partial L}{\partial x_2} = x_2(2\lambda_0 - 2\lambda_1 + 3\lambda_2 x_2) = 0, \quad (6.39)$$

$$\lambda_1(1 - x_1^2 - x_2^2) = 0, \quad x_1^2 + x_2^2 \geq 1, \quad \lambda_1 \geq 0, \quad (6.40)$$

$$\lambda_2(x_1^3 + x_2^3 - 1) = 0, \quad x_1^3 + x_2^3 \leq 1, \quad \lambda_2 \geq 0, \quad (6.41)$$

$$\lambda_0^2 + \lambda_1^2 + \lambda_2^2 \neq 0. \quad (6.42)$$

Проаналізуємо можливі розв'язки системи (6.38)–(6.42) у загальному випадку (без будь-яких обмежень на λ_0).

Припустивши, що $x_1 = 0$, із (6.40) і (6.41) отримаємо, що $x_2^2 \geq 1$ і $x_2^3 \leq 1$ або $x_2 \in (-\infty, -1] \cup [1, +\infty)$ і $x_2 \leq 1$. Отже, або $x_2 = 1$ або $x_2 \leq -1$. При $x_2 \leq -1$ із (6.41) випливає, що $\lambda_2 = 0$. Якщо $x_2 < -1$, то згідно з (6.40) маємо $\lambda_1 = 0$, а із (6.39) – $\lambda_0 = 0$, що суперечить (6.42). Тепер, взявши $x_2 = \pm 1$, із системи (6.38)–(6.42) отримаємо дві групи розв'язків:

$$1) \quad x_1 = 0, \quad x_2 = 1, \quad 2\lambda_0 - 2\lambda_1 + 3\lambda_2 = 0, \quad \lambda_1 \geq 0, \quad \lambda_2 \geq 0, \\ \lambda_0^2 + \lambda_1^2 + \lambda_2^2 \neq 0;$$

$$2) \quad x_1 = 0, \quad x_2 = -1, \quad -2\lambda_0 + 2\lambda_1 = 0, \quad \lambda_1 \geq 0, \quad \lambda_2 = 0, \quad \lambda_0^2 + \lambda_1^2 \neq 0.$$

Припустивши, що $x_2 = 0$, із (6.41) і (6.40) відповідно матимемо $x_1^3 \leq 1$ і $x_1^2 \geq 1$, тобто $x_1 \leq 1$ і $x_1 \in (-\infty, -1] \cup [1, +\infty)$. Це означає, що або $x_1 = 1$ або $x_1 \leq -1$. При $x_1 \leq -1$ із (6.41) випливає, що $\lambda_2 = 0$, а при $x_1 < -1$ із (6.40) маємо $\lambda_1 = 0$, а із (6.38) – $\lambda_0 = 0$, що неможливо. Отже, взявши $x_1 = \pm 1$, із (6.38)–(6.42) одержимо ще дві групи розв'язків:

$$3) \quad x_1 = 1, \quad x_2 = 0, \quad \lambda_0 - 2\lambda_1 + 3\lambda_2 = 0, \quad \lambda_1 \geq 0, \quad \lambda_2 \geq 0, \\ \lambda_0^2 + \lambda_1^2 + \lambda_2^2 \neq 0;$$

$$4) \quad x_1 = -1, \quad x_2 = 0, \quad -\lambda_0 + 2\lambda_1 = 0, \quad \lambda_1 \geq 0, \quad \lambda_2 = 0, \quad \lambda_0^2 + \lambda_1^2 \neq 0.$$

Якщо припустити, що $x_1 \neq 0$, $x_2 \neq 0$, то рівняння (6.38), (6.39) набудуть вигляду

$$\lambda_0 - 2\lambda_1 + 3\lambda_2 x_1 = 0, \quad 2\lambda_0 - 2\lambda_1 + 3\lambda_2 x_2 = 0. \quad (6.43)$$

При $\lambda_2 = 0$ з цих рівнянь маємо $\lambda_0 = \lambda_1 = 0$, що суперечить (6.42), тому нехай $\lambda_2 > 0$. Тоді із (6.41) випливає, що $x_1^3 + x_2^3 = 1$, звідки робимо висновок про неможливість ситуації $x_1^2 + x_2^2 = 1$, оскільки для всіх точок кола з центром у початку координат і одиничним радіусом $x_1^3 + x_2^3 < 1$. Отже, $x_1^2 + x_2^2 > 1$ і з (6.40) маємо $\lambda_1 = 0$. Тепер з рівнянь (6.38)–(6.42) при $\lambda_1 = 0$ (тобто з рівнянь (6.43) при $\lambda_1 = 0$ і (6.42)) і умови $x_1^3 + x_2^3 = 1$ отримаємо п'яту групу розв'язків:

$$5) \quad x_1 = \frac{1}{\sqrt[3]{9}}, \quad x_2 = \frac{2}{\sqrt[3]{9}}, \quad \lambda_0 < 0, \quad \lambda_1 = 0, \quad \lambda_2 = -\frac{\lambda_0 \sqrt[3]{9}}{3} > 0.$$

Після проведеного аналізу стало очевидним, що

множник λ_0 у випадках 1) і 3) може бути як від'ємним, так і додатним, у випадках 2) і 4) – тільки додатним, а у випадку 5) – від'ємним. Це свідчить про те, що точки $(0,1)^T$ і $(1,0)^T$ можуть бути як точками мінімуму, так і точками максимуму, точки $(0,-1)^T$ і $(-1,0)^T$ – тільки точками мінімуму, а точка $\left(\frac{1}{\sqrt[3]{9}}, \frac{2}{\sqrt[3]{9}}\right)^T$ – тільки точкою максимуму.

Крок 3. Знайшовши стаціонарні точки, дослідимо їх на оптимальність. Очевидно, цільова функція $f(x) = x_1^2/2 + x_2^2$ сильно опукла, оскільки матриця

$$\tilde{H} = f''(x) - 2\theta I = \begin{pmatrix} 1-2\theta & 0 \\ 0 & 2 \end{pmatrix}$$

додатно визначена при $\theta \in (0, 1/2)$ (тобто таке θ існує, при якому $\tilde{H}$ додатно визначена). Отже, вона досягає глобального мінімуму на довільній замкненій множині у $\mathbb{R}^2$, зокрема на множині X . Обчисливши значення функції в усіх точках-претендентах $(0,-1)^T$, $(-1,0)^T$, і $(0,1)^T$, $(1,0)^T$, матимемо

$$f(0,-1) = f(0,1) = 1, \quad f(-1,0) = f(1,0) = \frac{1}{2}.$$

Отже, точки $(-1,0)^T$ і $(1,0)^T$ – точки глобального мінімуму функції $f(x)$ на X .

Додатково дослідимо стаціонарні точки $(-1,0)^T$, $(1,0)^T$ також за допомогою достатніх умов оптимальності (теорема 6.8). Матриця других похідних функції Лагранжа по змінних x_1, x_2 має вигляд

$$L''_{xx} = \begin{pmatrix} \lambda_0 - 2\lambda_1 + 6\lambda_2 x_1 & 0 \\ 0 & 2\lambda_0 - 2\lambda_1 + 6\lambda_2 x_2 \end{pmatrix}.$$

При $x_1 = -1$, $x_2 = 0$ (тобто для точки $x^* = (-1, 0)^T$) множники Лагранжа $\lambda_2 = 0$, $\lambda_1 = \frac{\lambda_0}{2}$, причому $\lambda_0 > 0$ і $\lambda_1 > 0$, тому можна взяти $\lambda_0 = 1$, $\lambda_1 = 1/2$. Тоді

$$d^2L(x^*, \lambda^*) = \langle L''_{xx} dx, dx \rangle = dx_2^2.$$

Активним у точці $(-1, 0)^T$ є тільки перше обмеження ($\lambda_1 > 0$), тому, згідно із (6.35), маємо $dg_1(x^*) = -2 \cdot (-1) \cdot dx_1 - 2 \cdot 0 \cdot dx_2 = 0$, тобто $dx_1 = 0$. Для всіх ненульових векторів $dx = (0, dx_2)^T$, очевидно $d^2L(x^*, \lambda^*) = dx_2^2 > 0$, тобто $(-1, 0)^T$ – точка мінімуму.

Для точки $(1, 0)^T$ $\lambda_2 = \frac{2\lambda_1 - \lambda_0}{3} \geq 0$, тобто $\lambda_1 \geq \lambda_0/2 \geq 0$, звідки випливає, що $\lambda_0 \geq 0$. При $\lambda_0 > 0$ (випадок $\lambda_0 = 0$ немає сенсу розглядати, оскільки він не передбачений теоремою 6.8) можна взяти $\lambda_0 = 1$. Тоді $\lambda_1 > 0$, $\lambda_2 = \frac{2\lambda_1 - 1}{3} \geq 0$, $d^2L(x^*, \lambda^*) = (2\lambda_1 - 1)dx_1^2 + (2 - 2\lambda_1)dx_2^2$.

У точці $(1, 0)^T$ обидва обмеження активні, причому $dg_2(x^*) = 3 \cdot 1 \cdot dx_1 + 3 \cdot 0 \cdot dx_2 = 3dx_1$. Для першого ($\lambda_1 > 0$) $dx_1 = 0$, а для другого або $dx_1 = 0$ ($\lambda_2 > 0$) або $dx_1 \leq 0$ ($\lambda_2 = 0$), тобто умова $dx_1 = 0$ задовольняє як (6.35), так і (6.36). При $\lambda_0 = 1$ для всіх ненульових векторів $dx = (0, dx_2)^T$ згідно з теоремою 6.8 $d^2L(x^*, \lambda^*) > 0$ при $\frac{1}{2} \leq \lambda_1 < 1$. Отже, $(1, 0)^T$ – точка мінімуму.

Оскільки $f(x) = (x_1^2 + x_2^2)/2 + x_2^2/2$, а на колі $x_1^2 + x_2^2 = 1$ відповідно $f(x) = x_2^2/2 + 1/2$, то будь-яке зміщення із точок $(0, -1)^T$, $(0, 1)^T$ приводить до зменшення значення функції,

тому ці точки не можуть бути точками локального мінімуму. Крім того, при довільному $\varepsilon > 0$ точка $(-\varepsilon, 1)^T$ допустима ($\varepsilon^2 + 1 > 1, -\varepsilon^3 + 1 < 1$) і $f(0, 1) = 1 < f(-\varepsilon, 1) = \varepsilon^2/2 + 1$, тобто точка $(0, 1)^T$ не може бути і точкою локального максимуму $f(x)$ на X . Це стосується і точки $(0, -1)^T$ ($f(0, 1) = f(0, -1) = 1$). Отже, стаціонарні точки $(0, 1)^T$ і $(0, -1)^T$ виявились зайвими.

Залишилось дослідити останню стаціонарну точку $\left(\frac{1}{\sqrt[3]{9}}, \frac{2}{\sqrt[3]{9}}\right)^T$. Оскільки у цій точці матриця других похідних

$$L''_{xx} = \begin{pmatrix} -\lambda_0 & 0 \\ 0 & -2\lambda_0 \end{pmatrix}$$

додатно визначена ($\lambda_0 < 0$), то виконується достатня умова оптимальності (мінімуму функції $(-f(x)) = -x_1^2/2 - x_2^2$) (6.37), тобто $\left(\frac{1}{\sqrt[3]{9}}, \frac{2}{\sqrt[3]{9}}\right)^T$ – точка строгого локального максимуму функції $f(x)$ на X .

До такого ж висновку прийдемо, скориставшись теоремою 6.8 для функції $(-f(x)) = -x_1^2/2 - x_2^2$. Справді для точки $\left(\frac{1}{\sqrt[3]{9}}, \frac{2}{\sqrt[3]{9}}\right)^T$ $d^2L(x^*, \lambda^*) = -\lambda_0(dx_1^2 + 2dx_2^2)$. Активним у цій точці є тільки друге обмеження ($\lambda_2 > 0$), тому $dg_2(x^*) = (3/\sqrt[3]{81})dx_1 + (12/\sqrt[3]{81})dx_2 = 0$ або $dx_1 = -4dx_2$ і $d^2L(x^*, \lambda^*) = (-\lambda_0) \cdot 18dx_2^2 > 0$ при $(-\lambda_0) > 0$, тобто дана стаціонарна точка є точкою мінімуму функції $(-f(x))$ та максимуму функції $f(x)$ на X , що і потрібно було підтвердити.

Крок 4. На попередньому кроці точки глобального екстремуму уже встановлені. ►

Приклад 6.4. Розв'язати задачу мінімізації функції $f(x) = 4x_1^2 + 5x_1x_2 + 6x_2^2$ на множині $X = \{x \in \mathbb{R}^2 / x_1 + x_2 \geq 5, x_1 \geq 0, x_2 \geq 0\}$.

◄ **Крок 1.** Функція Лагранжа (узагальнена) для даної задачі

$$\tilde{L}(x, \lambda_0, \lambda) = \lambda_0(4x_1^2 + 5x_1x_2 + 6x_2^2) + \lambda_1(5 - x_1 - x_2) - \lambda_2x_1 - \lambda_3x_2.$$

Крок 2. Система необхідних умов оптимальності (6.11), (6.31)–(6.33) для задачі на мінімум у даному випадку така:

$$\frac{\partial L}{\partial x_1} = \lambda_0(8x_1 + 5x_2) - \lambda_1 - \lambda_2 = 0, \quad (6.44)$$

$$\frac{\partial L}{\partial x_2} = \lambda_0(5x_1 + 12x_2) - \lambda_1 - \lambda_3 = 0, \quad (6.45)$$

$$\lambda_1(5 - x_1 - x_2) = 0, \quad x_1 + x_2 \geq 5, \quad (6.46)$$

$$\lambda_2x_1 = 0, \quad x_1 \geq 0, \quad (6.47)$$

$$\lambda_3x_2 = 0, \quad x_2 \geq 0, \quad (6.48)$$

$$\lambda_0 \geq 0, \quad \lambda_i \geq 0, \quad (i = \overline{1,3}), \quad \sum_{i=0}^3 \lambda_i^2 \neq 0. \quad (6.49)$$

При $\lambda_0 = 0$ із (6.44), (6.45) випливає, що $\lambda_2 = -\lambda_1$, $\lambda_3 = -\lambda_1$, а із (6.46) (з урахуванням (6.47), (6.48)) – $5\lambda_1 = 0$, тобто $\lambda_1 = 0$. Однак така ситуація неможлива, оскільки тоді всі множники Лагранжа нульові, що суперечить умовам (6.49).

Отже, $\lambda_0 = 1$. Розглянемо два випадки: $\lambda_1 > 0$, $\lambda_1 = 0$.

При $\lambda_1 > 0$ із (6.46) маємо $x_1 + x_2 = 5$, тому проаналізуємо три можливі ситуації: 1) $x_1 > 0$, $x_2 > 0$; 2) $x_1 = 0$, $x_2 = 5$; 3) $x_1 = 5$, $x_2 = 0$.

1) Якщо $x_1 > 0$, $x_2 > 0$, то, згідно із (6.47), (6.48), $\lambda_2 = 0$,

$\lambda_3 = 0$. Тоді із (6.44), (6.45) випливає, що $8x_1 + 5x_2 = 5x_1 + 12x_2$, тобто $x_2 = \frac{3}{7}x_1$. Оскільки $x_1 + x_2 = \frac{10}{7}x_1 = 5$, то $x_1 = \frac{7}{2}$, $x_2 = \frac{3}{2}$ і точка $(x_1, x_2)^T = \left(\frac{7}{2}, \frac{3}{2}\right)^T$ стаціонарна.

2) Якщо $x_1 = 0$, $x_2 = 5$, то, згідно із (6.47)–(6.49), $\lambda_2 \geq 0$, $\lambda_3 = 0$. Система (6.44), (6.45) набуде вигляду

$$\begin{cases} 8x_1 + 5x_2 = \lambda_1 + \lambda_2, \\ 5x_1 + 12x_2 = \lambda_1. \end{cases} \quad (6.50)$$

Із (6.50) випливає, що $\lambda_2 = 3x_1 - 7x_2 = 3 \cdot 0 - 7 \cdot 5 = -35$, що суперечить умовам (6.49). Отже, ситуація 2) неможлива.

3) Якщо $x_1 = 5$, $x_2 = 0$, то $\lambda_2 = 0$, $\lambda_3 \geq 0$. Із (6.44), (6.45) отримаємо $\lambda_3 = -3x_1 + 7x_2 = -15 < 0$, тому дана ситуація також неможлива.

При $\lambda_1 = 0$ із (6.44), (6.45) матимемо $\lambda_2 = 8x_1 + 5x_2$, $\lambda_3 = 5x_1 + 12x_2$. Підставивши вирази для λ_2, λ_3 відповідно у (6.47), (6.48), одержимо $(8x_1 + 5x_2)x_1 = 0$, $(5x_1 + 12x_2)x_2 = 0$. Аналіз цих рівностей показує, що $x_1 \neq 0$ і $x_2 \neq 0$ (припустивши, що значення однієї зі змінних нульове, дійдемо висновку, що і значення іншої змінної нульове, але це неможливо, бо $x_1 + x_2 \geq 5$). А це означає, що $8x_1 + 5x_2 = 0$ і $5x_1 + 12x_2 = 0$, тобто $x_2 = -\frac{8}{5}x_1$ і $x_2 = -\frac{5}{12}x_1$, що також суперечить умовам допустимості.

Отже, знайдена вище стаціонарна точка єдина.

Крок 3. У стаціонарній точці $\left(\frac{7}{2}, \frac{3}{2}\right)^T$ число активних

обмежень $l = 1 \left(\frac{7}{2} + \frac{3}{2} = 5 \right)$. Оскільки $l = 1 \neq n = 2$, то достатні умови першого порядку не виконуються. Використаємо достатні умови мінімуму другого порядку (при $\lambda_0 = 1$). Враховуючи, що $\partial^2 L / \partial x_1^2 = 8$, $\partial^2 L / \partial x_1 \partial x_2 = \partial^2 L / \partial x_2 \partial x_1 = 5$, $\partial^2 L / \partial x_2^2 = 12$, то матриця других похідних

$$L''_{xx}(x^*, \lambda^*) = \begin{pmatrix} 8 & 5 \\ 5 & 12 \end{pmatrix}$$

додатно визначена, тобто $d^2 L(x^*, \lambda^*) > 0$ (виконується умова (6.37)) при всіх ненульових $dx \in \mathbb{R}^n$, зокрема і для тих, які задовольняють умову (6.35). Справді, активним є тільки перше обмеження ($\lambda_1 > 0$). Умова (6.35) для цього обмеження перетворюється у рівність $dx_2 = -dx_1$, враховуючи яку, отримаємо

$$d^2 L(x^*, \lambda^*) = \langle L''_{xx}(x^*, \lambda^*) dx, dx \rangle = 10 dx_1^2 > 0, \text{ тобто згідно з}$$

теоремою 6.8 $\left(\frac{7}{2}, \frac{3}{2} \right)^T$ – точка мінімуму досліджуваної функції $f(x)$ на X .

Крок 4. Точка мінімуму $\left(\frac{7}{2}, \frac{3}{2} \right)^T$ єдина, тому у даній

задачі вона є точкою глобального мінімуму. При цьому

$$f\left(\frac{7}{2}, \frac{3}{2}\right) = \frac{355}{4}. \blacktriangleright$$

Зауважимо, що розв'язана задача мінімізації опукла (цільова функція та допустима множина опуклі), тому її стаціонарна точка апіорі є точкою мінімуму.

6.4. Необхідні та достатні умови екстремуму у випадку задачі математичного програмування зі змішаними функціональними обмеженнями

Оскільки умови екстремуму для задач зі змішаними функціональними обмеженнями, як і для задач із функціональними обмеженнями-нерівностями, формуються окремо для мінімуму і максимуму, то одночасно розглянемо дві оптимізаційні задачі:

$$\begin{cases} f(x) \rightarrow \min, \\ g_i(x) \leq 0, \quad i = \overline{1, k}, \\ g_i(x) = 0, \quad i = \overline{k+1, m} \end{cases} \quad (6.51)$$

та

$$\begin{cases} f(x) \rightarrow \max, \\ g_i(x) \leq 0, \quad i = \overline{1, k}, \\ g_i(x) = 0, \quad i = \overline{k+1, m}. \end{cases} \quad (6.52)$$

Очевидно, задачі (6.51), (6.52) є частинними варіантами загальних задач математичного програмування на мінімум та максимум, зокрема (6.51) є задачею (6.2)–(6.4).

Допустиму множину задач (6.51), (6.52) формують змішані обмеження (обмеження-нерівності та обмеження-рівності), тому умови екстремуму обов'язково будуть містити умови доповнюючої нежорсткості та знаковизначеності множників Лагранжа, але тільки для обмежень-нерівностей. Ці умови легко встановити, замінивши задачі (6.51), (6.52) рівносильними задачами на класичний екстремум. Для цього потрібно обмеження-нерівності замінити обмеженнями (6.23) при $m = k$, перейти до задач на класичний екстремум, виписати відповідні функції Лагранжа й отримати аналоги співвідношень (6.26)–(6.28), з яких і випливають умови екстремуму.

Теорема 6.9 (необхідні умови мінімуму (максимуму) першого порядку). Нехай:

1) x^* – точка локального мінімуму (максимуму) у задачі (6.51) (у задачі (6.52));

2) функції $f(x)$, $g_i(x)$ ($i = \overline{1, m}$) неперервно диференційовні у деякому околі точки x^* .

Тоді існують числа $\lambda_0^* \geq 0$ і $\lambda_1^*, \dots, \lambda_m^*$, які одночасно не дорівнюють нулю і при яких виконуються умови:

1) стаціонарності (6.11) узагальненої функції Лагранжа (6.8) по x :

$$\frac{\partial \tilde{L}(x^*, \lambda_0^*, \lambda^*)}{\partial x_j} = 0, \quad j = \overline{1, n};$$

2) допустимості розв'язку:

$$g_i(x^*) \leq 0, \quad i = \overline{1, k}; \quad g_i(x^*) = 0, \quad i = \overline{k+1, m} \quad (6.53)$$

3) доповнюючої нежорсткості:

$$\lambda_i^* g_i(x^*) = 0, \quad i = \overline{1, k}; \quad (6.54)$$

4) невід'ємності (недодатності) числових множників Лагранжа:

$$\lambda_i^* \geq 0 \quad (\lambda_i^* \leq 0), \quad i = \overline{1, k}. \quad (6.55)$$

Якщо при цьому градієнти активних обмежень-нерівностей та усіх обмежень-рівностей у точці x^* лінійно незалежні, то виконується умова регулярності і $\lambda_0^* \neq 0$.

Зазначимо, що при $\lambda_0^* \neq 0$ справедливі твердження:

1) якщо функції $f(x)$, $g_i(x)$ ($i = \overline{1, k}$) опуклі, а функції $g_i(x)$ ($i = \overline{k+1, m}$) лінійні, то умови теореми 6.9 необхідні й достатні для існування глобального мінімуму;

2) якщо функції $(-f(x))$, $g_i(x)$ ($i = \overline{1, k}$) опуклі, а функції

$g_i(x)$ ($i = \overline{k+1, m}$) лінійні, то умови теореми 6.9 необхідні й достатні для існування глобального максимуму.

Ще одне важливе зауваження стосується умов доповнюючої нежорсткості (6.54), з яких випливає, що якщо i -те обмеження-нерівність пасивне у точці x^* , тобто $g_i(x^*) < 0$, то $\lambda_i^* = 0$, а якщо – активне, тобто $g_i(x^*) = 0$, то $\lambda_i^* \geq 0$ для мінімуму і $\lambda_i^* \leq 0$ – для максимуму.

Нехай I_a^* – множина активних у точці x^* обмежень-нерівностей, а $L(x, \lambda)$ – регулярна функція Лагранжа (6.13) для задачі (6.51) (задачі (6.52)).

Теорема 6.10 (необхідні умови мінімуму (максимуму) другого порядку). Нехай:

1) x^* – регулярна точка локального мінімуму (максимуму) у задачі (6.51) (у задачі (6.52)), а $(x^*, \lambda^*)^T$ – розв’язок системи (6.11), (6.53)–(6.55) при $\lambda_0^* = 1$ ($\|\lambda^*\| \neq 0$);

2) функції $f(x)$, $g_i(x)$ ($i = \overline{1, m}$) двічі неперервно диференційовні у деякому околі x^* .

Тоді у точці $(x^*, \lambda^*)^T$ другий диференціал задовольняє умову (6.34):

$$d^2L(x^*, \lambda^*) \geq 0 \quad (d^2L(x^*, \lambda^*) \leq 0)$$

при всіх $dx \in \mathbb{R}^n$, що задовольняють умови:

$$1) \quad \begin{cases} dg_i(x^*) = 0 \text{ для } i = \overline{k+1, m} \text{ та всіх } i \in I_a^*, \\ \text{для яких } \lambda_i^* > 0 \text{ } (\lambda_i^* < 0); \end{cases} \quad (6.56)$$

$$2) \quad dg_i(x^*) \leq 0 \text{ для всіх } i \in I_a^*, \text{ для яких } \lambda_i^* = 0. \quad (6.57)$$

Як і у випадку оптимізаційних задач з обмеженнями-нерівностями, для дослідження стаціонарних (умовно-

стаціонарних) точок використовують достатні умови екстремуму першого і другого порядків.

Теорема 6.11 (достатні умови мінімуму (максимуму) першого порядку). Нехай:

1) $(x^*, \lambda^*)^T$ – розв’язок системи (6.11), (6.53)–(6.55) при $\lambda_0^* = 1$ ($\|\lambda^*\| \neq 0$);

2) функції $f(x)$, $g_i(x)$ ($i = \overline{1, m}$) неперервно диференційовні у деякому околі x^* ;

3) сумарне число l активних обмежень-нерівностей у точці x^* та усіх обмежень-рівностей збігається з числом n змінних ($l = n$) (при цьому виконується умова регулярності).

Якщо $\lambda_i^* > 0$ для всіх $i \in I_a^*$, то x^* – точка локального мінімуму у задачі (6.51), а якщо $\lambda_i^* < 0$ для всіх $i \in I_a^*$, то x^* – точка локального максимуму у задачі (6.52).

Теорема 6.12 (достатні умови мінімуму (максимуму) другого порядку). Нехай:

1) $(x^*, \lambda^*)^T$ – розв’язок системи (6.11), (6.53)–(6.55) при $\lambda_0^* = 1$ ($\|\lambda^*\| \neq 0$);

2) функції $f(x)$, $g_i(x)$ ($i = \overline{1, m}$) двічі неперервно диференційовні у деякому околі x^* .

Якщо у точці $(x^*, \lambda^*)^T$ другий диференціал задовольняє умову (6.37):

$$d^2L(x^*, \lambda^*) > 0 \quad (d^2L(x^*, \lambda^*) < 0)$$

при всіх ненульових $dx \in \mathbb{R}^n$, що задовольняють умови (6.56) для мінімуму (максимуму) та (6.57), то x^* – точка локального мінімуму (максимуму) у задачі (6.51) (у задачі (6.52)).

Розв'язування задач (6.51), (6.52) за допомогою необхідних і достатніх умов екстремуму приводить до таких результатів:

- 1) $d^2L(x^*, \lambda^*) > 0$ ($dx \neq 0$, виконані умови (6.56) для мінімуму та умови (6.57)) – x^* є точкою локального мінімуму;
- 2) $d^2L(x^*, \lambda^*) < 0$ ($dx \neq 0$, виконані умови (6.56) для максимуму та умови (6.57)) – x^* є точкою локального максимуму;
- 3) $d^2L(x^*, \lambda^*) \geq 0$ (виконані умови (6.56) для мінімуму та умови (6.57)) – x^* може бути точкою локального мінімуму, потрібні додаткові дослідження;
- 4) $d^2L(x^*, \lambda^*) \leq 0$ (виконані умови (6.56) для максимуму та умови (6.57)) – x^* може бути точкою локального максимуму, потрібні додаткові дослідження;
- 5) $d^2L(x^*, \lambda^*) = 0$ (виконані умови (6.56) для мінімуму та умови (6.57)) – потрібні додаткові дослідження;
- 6) $d^2L(x^*, \lambda^*) = 0$ (виконані умови (6.56) для максимуму та умови (6.57)) – потрібні додаткові дослідження;
- 7) $d^2L(x^*, \lambda^*)$ – знакозмінна квадратична форма (виконані умови (6.56) для мінімуму та умови (6.57)) – x^* не є точкою екстремуму;
- 8) $d^2L(x^*, \lambda^*)$ – знакозмінна квадратична форма (виконані умови (6.56) для максимуму та умови (6.57)) – x^* не є точкою екстремуму.

Уточнимо основні кроки алгоритму розв'язування задач (6.51), (6.52) за допомогою необхідних і достатніх умов

екстремуму.

Крок 1. Побудова узагальненої функції Лагранжа (6.8) або класичної функції Лагранжа (6.13), якщо попередньо вдається встановити регулярність задачі.

Крок 2. Знаходження розв'язків системи (6.11), (6.53)–(6.55) (при цьому дану систему потрібно розв'язати при $\lambda_0^* = 0$ (якщо регулярність не встановлена) і $\lambda_0^* = 1$).

Крок 3. Дослідження знайдених регулярних стаціонарних (умовно-стаціонарних) розв'язків за допомогою достатніх умов екстремуму першого та другого порядку.

Крок 4. Знаходження значень функції у знайдених точках умовного екстремуму і вибір точки (чи точок) глобального екстремуму.

Приклад розв'язування оптимізаційної задачі зі змішаними обмеженнями розглянемо нижче.

Приклад 6.5. Дослідити на екстремум функцію $f(x) = x_1^2 + 2x_2^2$ на множині $X = \{x \in \mathbb{R}^2 / x_1^2 + x_2^2 \leq 5, x_1 - x_2 = 1/2\}$.

◀ *Крок 1.* Маємо задачу на екстремум зі змішаними обмеженнями. Узагальнена функція Лагранжа

$$\tilde{L}(x, \lambda_0, \lambda) = \lambda_0 (x_1^2 + 2x_2^2) + \lambda_1 (x_1^2 + x_2^2 - 5) + \lambda_2 \left(x_1 - x_2 - \frac{1}{2} \right).$$

Крок 2. Система необхідних умов оптимальності першого порядку (6.11), (6.53)–(6.55) для даної задачі формалізується співвідношеннями

$$\frac{\partial L}{\partial x_1} = 2\lambda_0 x_1 + 2\lambda_1 x_1 + \lambda_2 = 0, \quad (6.58)$$

$$\frac{\partial L}{\partial x_2} = 4\lambda_0 x_2 + 2\lambda_1 x_2 - \lambda_2 = 0, \quad (6.59)$$

$$x_1^2 + x_2^2 \leq 5, \quad x_1 - x_2 = \frac{1}{2}, \quad (6.60)$$

$$\lambda_1 (x_1^2 + x_2^2 - 5) = 0, \quad (6.61)$$

$$\lambda_0 \geq 0, \quad \lambda_1 \geq 0 \quad (\lambda_1 \leq 0), \quad \lambda_0^2 + \lambda_1^2 + \lambda_2^2 \neq 0. \quad (6.62)$$

Нагадаємо, що у дужках (6.62) указана умова недодатності множника Лагранжа λ_1 для задачі на максимум.

Нехай $\lambda_0 = 0$. Якщо припустити, що $\lambda_1 = 0$, то з умов (6.58), (6.59) отримаємо $\lambda_2 = 0$, але всі множники Лагранжа не можуть бути нульовими, оскільки це суперечить (6.62). Якщо $\lambda_1 \neq 0$, то із (6.58), (6.59) маємо рівність $2\lambda_1 x_2 = -2\lambda_1 x_1$ або $x_2 = -x_1$. Крім того, із (6.61) випливає, що $x_1^2 + x_2^2 = 5$.

Інакше кажучи, або $x_1 = -\sqrt{5}/\sqrt{2}$, $x_2 = \sqrt{5}/\sqrt{2}$, або $x_1 = \sqrt{5}/\sqrt{2}$,

$x_2 = -\sqrt{5}/\sqrt{2}$, але точки $\left(-\frac{\sqrt{5}}{\sqrt{2}}, \frac{\sqrt{5}}{\sqrt{2}}\right)^T$ і $\left(\frac{\sqrt{5}}{\sqrt{2}}, -\frac{\sqrt{5}}{\sqrt{2}}\right)^T$ не є

допустимими (умова $x_1 - x_2 = 1/2$ не виконується). Отже, $\lambda_0 \neq 0$.

Нехай $\lambda_0 = 1$. Припустивши, що $\lambda_1 = 0$, із (6.58), (6.59) отримаємо $x_2 = -\frac{x_1}{2}$, а із (6.60) – $x_2 = x_1 - \frac{1}{2}$. Прирівнявши

праві частини цих рівностей, знайдемо $x_1 = \frac{1}{3}$, $x_2 = -\frac{1}{6}$. Точка

$\left(\frac{1}{3}, -\frac{1}{6}\right)^T$ задовольняє всі співвідношення (6.58)–(6.62), тому

є стаціонарною. Якщо $\lambda_1 \neq 0$, то, згідно із (6.61), (6.60)

$x_1^2 + x_2^2 = 5$, $x_1 - x_2 = \frac{1}{2}$. Розв'язавши цю систему рівнянь,

одержимо дві пари розв'язків $\left(\frac{1-\sqrt{39}}{4}, -\frac{1+\sqrt{39}}{4}\right)^T$ і $\left(\frac{1+\sqrt{39}}{4}, \frac{\sqrt{39}-1}{4}\right)^T$, які будуть стаціонарними точками.

Із (6.58), (6.59) визначаємо $\lambda_1 = -(x_1 + 2x_2)/(x_1 + x_2)$, тому для першої із цих точок $\lambda_1 = -(1 + 3\sqrt{39})/2\sqrt{39} < 0$, а для другої – $\lambda_1 = -(3\sqrt{39} - 1)/2\sqrt{39} < 0$.

Крок 3. У точці $\left(\frac{1}{3}, -\frac{1}{6}\right)^T$ число активних обмежень $l = 1$

(виконується лише обмеження-рівність), тобто $l \neq n = 2$, тому для дослідження характеру даної стаціонарної точки використаємо достатні умови екстремуму другого порядку. Оскільки $\partial^2 L / \partial x_1^2 = 2 + 2\lambda_1 = 2(1 + \lambda_1)$, $\partial^2 L / \partial x_1 \partial x_2 = \partial^2 L / \partial x_2 \partial x_1 = 0$, $\partial^2 L / \partial x_2^2 = 4 + 2\lambda_1 = 2(2 + \lambda_1)$, то

$$L''_{xx}(x^*, \lambda^*) = \begin{pmatrix} 2(1 + \lambda_1) & 0 \\ 0 & 2(2 + \lambda_1) \end{pmatrix}.$$

При $\lambda_1 = 0$ матриця $L''_{xx}(x^*, \lambda^*)$ додатно визначена, причому у досліджуваній точці активним є тільки обмеження-рівність, для якого умова (6.56) (тобто умова $dg_2(x^*) = 0$) зводиться до рівності $dx_1 = dx_2$. Умова (6.57) відсутня, тому при $dx_1 = dx_2$ $d^2 L(x^*, \lambda^*) = 6dx_1^2 > 0$, тобто згідно із (6.37) (теорема 6.12) точка $\left(\frac{1}{3}, -\frac{1}{6}\right)^T$ є точкою мінімуму функції $f(x) = x_1^2 + 2x_2^2$ (глобального мінімуму, оскільки цільова функція опукла) на допустимій множині X .

Для точок $\left(\frac{1-\sqrt{39}}{4}, -\frac{1+\sqrt{39}}{4}\right)^T$ і $\left(\frac{1+\sqrt{39}}{4}, \frac{\sqrt{39}-1}{4}\right)^T$

виконуються достатні умови максимуму першого порядку (число усіх активних у цих точках обмежень $l=n=2$ і $\lambda_1 < 0$), тобто вони є точками локального максимуму функції $f(x)$ на X .

Із методичних міркувань перевіримо також достатні умови екстремуму другого порядку у цих точках.

У точці $\left(\frac{1-\sqrt{39}}{4}, -\frac{1+\sqrt{39}}{4}\right)^T$

$$dg_1(x^*) = \frac{1-\sqrt{39}}{2}dx_1 - \frac{1+\sqrt{39}}{2}dx_2 = 0,$$

$$dg_2(x^*) = dx_1 - dx_2 = 0.$$

З останніх двох рівностей випливає, що $dx_1 = dx_2 = 0$ і $d^2L(x^*, \lambda^*) = 0$, тому потрібні додаткові дослідження. Оскільки цільова функція $f(x) = x_1^2 + 2x_2^2$ не є опуклою вгору (угнутою), то зробити однозначний висновок на основі вигляду функції про дану стаціонарну точку неможливо. Однак геометрична інтерпретація задачі (графічний метод її розв'язування) підтверджує, що досліджувана стаціонарна точка є точкою максимуму.

Аналогічна ситуація із точкою $\left(\frac{1+\sqrt{39}}{4}, \frac{\sqrt{39}-1}{4}\right)^T$.

Крок 4. Точка мінімуму $\left(\frac{1}{3}, -\frac{1}{6}\right)^T$ єдина, тому вона буде

точкою глобального мінімуму функції $f(x)$ на X . Щодо інших двох точок (точок локального максимуму), то

$$f\left(\frac{1-\sqrt{39}}{4}, -\frac{1+\sqrt{39}}{4}\right) = f\left(\frac{1+\sqrt{39}}{4}, \frac{\sqrt{39}-1}{4}\right) = \frac{60-\sqrt{39}}{8},$$

тобто локальні максимуми збігаються з глобальним максимумом. ►

6.5. Задача умовної оптимізації з прямим обмеженням

Якщо у загальній задачі математичного програмування (6.2)–(6.4) зняти функціональні обмеження, то отримаємо задачу з прямим обмеженням (6.2), (6.5), тобто задачу

$$\begin{cases} f(x) \rightarrow \min, \\ x \in X_0, \end{cases} \quad (6.63)$$

Прийнято вважати, що у задачі з прямим обмеженням, у тому числі у задачі (6.63) $X_0 \subset \mathbb{R}^n$ – деяка «проста» множина або множина так званої «простої структури». У загальному випадку вимоги до множини X_0 найбільш загальні (опуклість, замкненість тощо), але змістовного характеру набувають лише оптимізаційні задачі, у яких множина X_0 відображає реальну природу (економічну, фізичну тощо) змінних задачі. До таких множин належать, наприклад паралелепіпед $X_0 = \{x \in \mathbb{R}^n / a_j \leq x_j \leq b_j, j = \overline{1, n}\}$ ($-\infty \leq a_j < b_j \leq +\infty, j = \overline{1, n}$ (якщо $a_j = -\infty$ або $b_j = +\infty$, то відповідний знак нерівності буде строгим)), лінійний многовид $X = \{x \in \mathbb{R}^n / Ax = b\}$ (A – матриця, x та b – узгоджені з нею вектори), куля $X_0 = \{x \in \mathbb{R}^n / \|x\| \leq \alpha\}$ (α – число) та ін. Усі ці множини можна описати функціональними обмеженнями, тому від задачі (6.63) з прямим обмеженням можна перейти до еквівалентної

оптимізаційної задачі із функціональними обмеженнями і на основі цієї задачі виписати відповідні умови екстремуму (оптимальності) розв'язку.

Для підтвердження вищесказаного розглянемо деякі, важливі для подальшого викладення матеріалу, приклади задач, для яких установимо необхідні умови екстремуму.

1. Нехай у задачі (6.63) множина X_0 є паралелепіпедом.

Спочатку розглянемо випадок, коли всі a_j, b_j ($j = \overline{1, n}$) є скінченними числовими значеннями. Тоді задачу (6.63) можна записати у формі типової задачі з обмеженнями-нерівностями, а саме задачі:

$$\begin{cases} f(x) \rightarrow \min, \\ g_j(x) = x_j - b_j \leq 0, \quad j = \overline{1, n}, \\ q_j(x) = -x_j + a_j \leq 0, \quad j = \overline{1, n}. \end{cases} \quad (6.64)$$

Узагальнена функція Лагранжа для задачі (6.64) має вигляд

$$L_2(x, \lambda_0, \lambda) = \lambda_0 f(x) + \sum_{j=1}^n \lambda_j (x_j - b_j) + \sum_{j=1}^n \mu_j (-x_j + a_j), \quad (6.65)$$

де $\lambda_0, \lambda_j, \mu_j$ ($j = \overline{1, n}$) – множники Лагранжа.

Згідно з теоремою 6.5, необхідні умови мінімуму першого порядку конкретизуються співвідношеннями (6.11), (6.31)–(6.33), тобто умовами ($\lambda_0^* \geq 0$)

$$\frac{\partial L_2(x^*, \lambda_0^*, \lambda^*)}{\partial x_j} = \lambda_0^* \frac{\partial f(x^*)}{\partial x_j} + \lambda_j^* - \mu_j^* = 0, \quad j = \overline{1, n}; \quad (6.66)$$

$$g_j(x^*) = x_j^* - b_j \leq 0, \quad q_j(x^*) = -x_j^* + a_j \leq 0, \quad j = \overline{1, n}; \quad (6.67)$$

$$\lambda_j^* (x_j^* - b_j) = 0, \quad \mu_j^* (-x_j^* + a_j) = 0, \quad j = \overline{1, n}; \quad (6.68)$$

$$\lambda_j^* \geq 0, \quad \mu_j^* \geq 0, \quad j = \overline{1, n}. \quad (6.69)$$

Якщо x_j^* – внутрішня точка (умови допустимості (6.67) для x_j^* є строгими нерівностями), то із умов доповнюючої нежорсткості (6.68) випливає, що $\lambda_j^* = 0$, $\mu_j^* = 0$, тобто, згідно із (6.66), при $\lambda_0^* \neq 0$ (у випадку регулярності функції (6.65)) умова регулярності для x_j^* збігається з умовою безумовного екстремуму функції $f(x)$ по змінній x_j : $\frac{\partial f(x^*)}{\partial x_j} = 0$.

Нехай x_j^* – гранична точка. Якщо $x_j^* = a_j$, то згідно із (6.68) і (6.69) $\lambda_j^* = 0$, $\mu_j^* \geq 0$, тому при $\lambda_0^* = 1$ із (6.66) отримаємо нерівність $\frac{\partial f(x^*)}{\partial x_j} = \mu_j \geq 0$. При $x_j^* = b_j$ згідно з (6.68), (6.69) $\lambda_j^* \geq 0$, $\mu_j^* = 0$, тому при $\lambda_0^* = 1$ із (6.66) отримаємо нерівність $\frac{\partial f(x^*)}{\partial x_j} = -\lambda_j^* \leq 0$.

Отже, якщо x^* – точка мінімуму у регулярній задачі (6.64), то

$$\frac{\partial f(x^*)}{\partial x_j} \begin{cases} = 0, & a_j < x_j^* < b_j, \\ \geq 0, & x_j^* = a_j \neq -\infty, \\ \leq 0, & x_j^* = b_j \neq +\infty. \end{cases} \quad (6.70)$$

Очевидно, умови (6.70) рівносильні умовам

$$\frac{\partial f(x^*)}{\partial x_j} (x_j - x_j^*) \geq 0, \quad x_j \in [a_j, b_j], \quad j = \overline{1, n}. \quad (6.71)$$

Умови (6.70) (чи (6.71)) є необхідними умовами локального мінімуму у задачі (6.64) (тобто у задачі (6.63), у якій X_0 – паралелепіпед).

Якщо маємо задачу з прямим обмеженням на максимум

$$\begin{cases} f(x) \rightarrow \max, \\ x \in X_0, \end{cases} \quad (6.72)$$

де X_0 – паралелепіпед, то із теореми 6.5 випливає, що необхідними умовами максимуму першого порядку у задачі (6.72) будуть умови (6.66)–(6.68) та умови недодатності числових множників Лагранжа

$$\lambda_j^* \leq 0, \quad \mu_j^* \leq 0, \quad j = \overline{1, n}. \quad (6.73)$$

Аналізуючи умови (6.66)–(6.68), (6.73) для регулярної задачі (6.72), приходимо до висновку, що із (6.68), (6.73), (6.66) при $x_j^* = a_j$ маємо $\lambda_j^* = 0, \quad \mu_j^* \leq 0, \quad \frac{\partial f(x^*)}{\partial x_j} = \mu_j \leq 0$, а при $x_j^* = b_j$ – $\lambda_j^* \leq 0, \quad \mu_j^* = 0, \quad \frac{\partial f(x^*)}{\partial x_j} = -\lambda_j^* \geq 0$. Інакше кажучи, якщо x^* – точка максимуму у регулярній задачі (6.72), то

$$\frac{\partial f(x^*)}{\partial x_j} \begin{cases} = 0, & a_j < x_j^* < b_j, \\ \leq 0, & x_j^* = a_j \neq -\infty, \\ \geq 0, & x_j^* = b_j \neq +\infty. \end{cases} \quad (6.74)$$

Очевидно, умови (6.74) рівносильні умовам (6.75):

$$\frac{\partial f(x^*)}{\partial x_j} (x_j - x_j^*) \leq 0, \quad x_j \in [a_j, b_j], \quad j = \overline{1, n}. \quad (6.75)$$

У випадку, коли $a_j = -\infty$ чи $b_j = +\infty$, у необхідних умовах екстремуму для задач (6.63), (6.72), тобто в умовах (6.70), (6.74), буде відсутня одна з нерівностей, що відповідає a_j чи b_j .

Частинним випадком розглянутих вище оптимізаційних задач з прямим обмеженням є задача з обмеженнями на

невід'ємність деяких або усіх змінних.

2. Нехай множина $X_0 = \{x \in \mathbb{R}^n / x_j \geq 0, j = \overline{1, m}\}$, де $0 \leq m \leq n$ (при $m = 0$ $X = \mathbb{R}^n$).

У цьому випадку задачі (6.63) і (6.72) можна формалізувати відповідно так:

$$\begin{cases} f(x) \rightarrow \min, \\ g_j(x) = -x_j \leq 0, j = \overline{1, m}; \end{cases} \quad (6.76)$$

$$\begin{cases} f(x) \rightarrow \max, \\ g_j(x) = -x_j \leq 0, j = \overline{1, m}. \end{cases} \quad (6.77)$$

Задачі (6.76) та (6.77) – це частинні варіанти задач (6.63), (6.72), у яких $a_j = 0$ ($j = \overline{1, m}$), $a_j = -\infty$ ($j = \overline{m+1, n}$), $b_j = +\infty$ ($j = \overline{1, n}$). З урахуванням умов (6.70) та (6.74) необхідні умови мінімуму у задачі (6.76) мають вигляд

$$\begin{cases} \frac{\partial f(x^*)}{\partial x_j} \geq 0, x_j^* \frac{\partial f(x^*)}{\partial x_j} = 0, j = \overline{1, m}, \\ \frac{\partial f(x^*)}{\partial x_j} = 0, j = \overline{m+1, n}, \end{cases} \quad (6.78)$$

а у задачі (6.77) –

$$\begin{cases} \frac{\partial f(x^*)}{\partial x_j} \leq 0, x_j^* \frac{\partial f(x^*)}{\partial x_j} = 0, j = \overline{1, m}, \\ \frac{\partial f(x^*)}{\partial x_j} = 0, j = \overline{m+1, n}. \end{cases} \quad (6.79)$$

Умови (6.78), (6.79) означають, що кожна з координат x_j^* ($j = \overline{1, n}$) може бути або внутрішньою, або граничною нульовою точкою.

Умови (6.78), (6.79), як і умови (6.70), (6.74), дозволяють у деяких найпростіших випадках явно розв'язати задачу оптимізації.

Приклад 6.6. Розв'язати задачу

$$\begin{cases} f(x_1, x_2) = 2x_1^2 + x_1x_2 + x_2^2 \rightarrow \min, \\ (x_1, x_2)^T \in X = \{(x_1, x_2)^T \in \mathbb{R}^2 / -1 \leq x_1 \leq 1, x_2 \geq 2\}. \end{cases} \quad (6.80)$$

◀ Із умов (6.70) випливає, що довільний розв'язок задачі (6.80) задовольняє умови:

$$\frac{\partial f(x_1, x_2)}{\partial x_1} = 4x_1 + x_2 \begin{cases} = 0, & -1 < x_1 < 1, \\ \geq 0, & x_1 = -1, \\ \leq 0, & x_1 = 1, \end{cases} \quad (6.81)$$

$$\frac{\partial f(x_1, x_2)}{\partial x_2} = x_1 + 2x_2 \begin{cases} = 0, & x_2 > 2, \\ \geq 0, & x_2 = 2. \end{cases} \quad (6.82)$$

Попарно комбінуючи співвідношення (6.81), (6.82), отримаємо шість систем, наприклад перші дві з них такі:

$$4x_1 + x_2 = 0, \quad x_1 + 2x_2 = 0, \quad x_1 \in (-1, 1), \quad x_2 > 2, \quad (6.83)$$

$$4x_1 + x_2 = 0, \quad x_1 + 2x_2 \geq 0, \quad x_1 \in (-1, 1), \quad x_2 = 2. \quad (6.84)$$

Далі потрібно знайти розв'язок кожної із систем (у тому числі (6.83), (6.84)) і дослідити його на оптимальність, хоча перед цим бажано здійснити якісний аналіз задачі.

Легко переконатись, що цільова функція задачі (6.80) є додатно визначеною квадратичною формою, вона сильно опукла і має єдиний розв'язок (що є одночасно локальним і глобальним). Отже, тільки одна із шести систем має єдиний розв'язок, який є розв'язком задачі (6.80). Цією системою є (6.84). Її розв'язком є $x^* = \left(-\frac{1}{2}, 2\right)^T$. Аналіз інших систем зайвий.

Зауважимо, що до такого ж самого висновку прийдемо,

якщо задачу (6.80) розв'яжемо методом множників Лагранжа як задачу з обмеженнями-нерівностями з урахуванням достатніх умов екстремуму. ►

6.6. Методи можливих напрямків

У загальному випадку методи можливих напрямків належать до прямих, які передбачають лише можливість обчислення значення функції $f(x)$ у будь-якій точці області визначення, а більш точно – допустимої множини задачі мінімізації (чи задачі максимізації), що достатньо для тих обчислювальних процедур, якими ці методи можуть формалізуватися, хоча вони можуть використовувати і похідні від функції, якщо останні диференційовні. Для презентації даної групи методів нижче зупинимось тільки на методах можливих напрямків, основу яких складає ітераційний процес

$$x^{(k+1)} = x^{(k)} + \alpha_k p^{(k)}, \quad k = 0, 1, 2, \dots, \quad x^{(0)} \in X, \quad (6.85)$$

де X – допустима множина задачі (6.1), а вектор $p^{(k)}$, вздовж якого здійснюється переміщення із точки $x^{(k)}$ у точку $x^{(k+1)}$, задовольняє умови:

1) для достатньо малих $\alpha_k > 0$ точка $x^{(k+1)} \in X$, тобто $p^{(k)}$ є вектором можливого напрямку;

2) для достатньо малих $\alpha_k > 0$ справджується нерівність $f(x^{(k+1)}) < f(x^{(k)})$, тобто $p^{(k)}$ є вектором спадання функції $f(x)$.

Величина кроку $\alpha_k > 0$ у (6.85) здебільшого вибирається з умови вичерпуючого спуску у напрямку $p^{(k)}$ (з урахуванням того, що $x^{(k+1)} \in X$).

6.6.1. Метод можливих напрямків у випадку лінійних обмежень

Нехай задача (6.1) має вигляд

$$f(x) \rightarrow \min, \quad (6.86)$$

$$\langle a^{(i)}, x \rangle \leq b_i, \quad i = \overline{1, m}, \quad (6.87)$$

$$x_j \geq 0, \quad j = \overline{1, n}, \quad (6.88)$$

де цільова функція $f(x) \equiv f(x_1, \dots, x_n)$ опукла та диференційовна на опуклій множині X , заданій обмеженнями (6.87), (6.88), а $\langle a^{(i)}, x \rangle = \sum_{j=1}^n a_{ij} x_j$. Розв'язок задачі (6.86)–(6.88)

будемо шукати, будуючи послідовність (6.85). При цьому вибір вектора $p^{(k)}$ залежить від двох випадків.

1) Точка $x^{(k)}$ є внутрішньою точкою множини X (всі нерівності (6.87), (6.88) у точці $x^{(k)}$ строгі). Тоді

$$p^{(k)} = -f'(x^{(k)}). \quad (6.89)$$

2) Точка $x^{(k)}$ є граничною точкою множини X (принаймні одна з нерівностей (6.87), (6.88) у точці $x^{(k)}$ перетворюється у рівність).

У цьому випадку вектор (6.89) (антиградієнт) може бути спрямований за межі області X , тобто при довільному $\alpha_k > 0$ точка $x^{(k+1)}$ із (6.85) може бути недопустимою ($x^{(k+1)} \notin X$). Це означає, що вибір вектора $p^{(k)}$ потрібно здійснювати в інший спосіб. З одного боку, цей вектор повинен бути спрямований у внутрішність множини X або вздовж її межі, а з іншого – логічно вимагати, щоб кут між векторами $p^{(k)}$ і $(-f'(x^{(k)}))$ був якомога меншим. Знаходження вектора $p^{(k)}$ з такими властивостями можна

звести до деякої задачі лінійного програмування (теорію задач лінійного програмування буде викладено пізніше). Обґрунтуємо зміст цієї задачі.

Нехай I_k та J_k – це множини індексів, що відповідають обмеженням (6.87) та (6.88), які у граничній точці $x^{(k)}$ є активними (перетворюються у рівності). Це означає, що $x^{(k)}$ належить гіперплощинам

$$\langle a^{(i)}, x \rangle = b_i \quad (i \in I_k), \quad \langle e^{(j)}, x \rangle = 0 \quad (j \in J_k), \quad (6.90)$$

де $e^{(j)} = (0, \dots, 0, 1, 0, \dots, 0)^T$ – j -й базисний вектор (j -та його компонента дорівнює 1). При цьому вектори $a^{(i)}$, $(-e^{(j)})$ є нормальними до гіперплощин (6.90) і спрямовані із допустимої множини X . Це випливає з того, що до прикладу гіперплощина $\langle a^{(i)}, x \rangle = b_i$ є поверхнею рівня функції $F_i(x) = \langle a^{(i)}, x \rangle$, градієнт якої $F'_i(x) = a^{(i)}$ є нормальним вектором до даної гіперплощини. Оскільки функція $F_i(x)$ уздовж градієнта зростає, то рівність $\langle a^{(i)}, x \rangle = b_i$ буде порушуватись, переходячи у нерівність зі знаком протилежним до (6.87), що підтверджує той факт, що нормальний вектор $a^{(i)} = F'_i(x)$ спрямований із множини X . Аналогічні міркування стосуються і вектора $(-e^{(j)})$, однак обмеження (6.88) потрібно записати у вигляді $-\langle e^{(j)}, x \rangle \leq 0$. Стосовно вектора $p^{(k)}$, який виходить із точки $x^{(k)}$, то по відношенню до множини X він буде можливим, якщо він буде утворювати з нормальними векторами $a^{(i)}$ та $(-e^{(j)})$ ($i \in I_k, j \in J_k$) кути, не менші, ніж $\pi/2$ (тобто кут між

векторами $p^{(k)}$ та $(-a^{(i)})$, $p^{(k)}$ та $e^{(j)}$ повинен бути не більшим, ніж $\pi/2$), тобто

$$\langle a^{(i)}, p^{(k)} \rangle \leq 0, i \in I_k, \quad (6.91)$$

$$-\langle e^{(j)}, p^{(k)} \rangle \leq 0, j \in J_k \text{ або } p_j^{(k)} \geq 0, j \in J_k. \quad (6.92)$$

Скалярний добуток двох векторів є монотонно спадною функцією від кута між ними (на $[0, \pi]$), тому мінімальному куту між векторами $p^{(k)}$ і $(-f'(x^{(k)}))$ буде відповідати максимальне значення скалярного добутку $\langle p^{(k)}, -f'(x^{(k)}) \rangle$ або мінімальне значення протилежної величини:

$$F(p^{(k)}) = \langle f'(x^{(k)}), p^{(k)} \rangle \rightarrow \min. \quad (6.93)$$

Функція (6.93) необмежена знизу, тому на норму вектора $p^{(k)}$ введемо обмеження

$$\sum_{j=1}^n |p_j^{(k)}| \leq 1. \quad (6.94)$$

Отже, для знаходження $p^{(k)}$ маємо задачу (6.91)–(6.94), яку можна звести до задачі лінійного програмування. Справді, якщо $J_k \neq \emptyset$, то, згідно із (6.92), $p_j \geq 0$ ($j \in J_k$). Стосовно інших змінних p_j (компонент шуканого вектора $p^{(k)}$) застосовуємо правило переходу до невід'ємних змінних, згідно з яким

$$p_j = p_j^+ - p_j^-, \text{ де } p_j^+ \geq 0, p_j^- \geq 0, p_j^+ p_j^- = 0.$$

Після цього для знаходження вектора $p^{(k)}$ матимемо задачу лінійного програмування

$$\langle f'(x^{(k)}), p \rangle = \sum_{j \in J_k} \frac{\partial f(x^{(k)})}{\partial x_j} p_j +$$

$$+ \sum_{j \notin J_k} \frac{\partial f(x^{(k)})}{\partial x_j} (p_j^+ - p_j^-) \rightarrow \min, \quad (6.95)$$

$$\langle a^{(i)}, p \rangle = \sum_{j \in J_k} a_{ij} p_j + \sum_{j \notin J_k} a_{ij} (p_j^+ - p_j^-) \leq 0, \quad i \in I_k, \quad (6.96)$$

$$\sum_{j=1}^n |p_j| = \sum_{j \in J_k} p_j + \sum_{j \notin J_k} (p_j^+ + p_j^-) \leq 1, \quad (6.97)$$

$$p_j \geq 0 \quad (j \in J_k), p_j^+ \geq 0, p_j^- \geq 0, p_j^+ p_j^- = 0 \quad (j \notin J_k). \quad (6.98)$$

При розв'язуванні задачі (6.95)–(6.98) симплексним методом для урахування умов $p_j^+ p_j^- = 0 \quad (j \notin J_k)$ змінні p_j^+ і p_j^- з однаковим номером $j \notin J_k$ не включають одночасно у число базисних змінних.

Розв'язавши задачу (6.95)–(6.98) і отримавши розв'язок $p^{(k)} \equiv p$, знаходять величину кроку α_k як розв'язок задачі мінімізації функції $f(x^{(k)} + \alpha p^{(k)})$ (як функції α) при $\alpha > 0$, $x^{(k)} + \alpha p^{(k)} \in X$, тобто

$$\alpha_k = \min \{ \bar{\alpha}_1, \dots, \bar{\alpha}_m, \hat{\alpha}_1, \dots, \hat{\alpha}_n, \alpha^* \},$$

де $\bar{\alpha}_i$ та $\hat{\alpha}_j$ – максимальні переміщення, при яких для точки $x^{(k+1)}$ із (6.85) виконується відповідно i -те обмеження (6.87) та j -те обмеження (6.88). Підставивши у (6.87) та (6.88) вираз (6.85), отримаємо

$$\bar{\alpha}_i = \begin{cases} +\infty, & \text{якщо } \langle a^{(i)}, p^{(k)} \rangle \leq 0 \\ \frac{b_i - \langle a^{(i)}, x^{(k)} \rangle}{\langle a^{(i)}, p^{(k)} \rangle}, & \text{якщо } \langle a^{(i)}, p^{(k)} \rangle > 0, \end{cases} \quad (6.99)$$

$$\hat{\alpha}_j = \begin{cases} +\infty, & \text{якщо } p_j^{(k)} \geq 0, \\ -\frac{x_j^{(k)}}{p_j^{(k)}}, & \text{якщо } p_j^{(k)} < 0. \end{cases} \quad (6.100)$$

Величина α^* знаходиться з умови безумовного вичерпуючого спуску вздовж вектора $p^{(k)}$ (без урахування (6.87), (6.88)), тобто α^* є розв'язком задачі

$$f(x^{(k)} + \alpha p^{(k)}) \rightarrow \min_{\alpha > 0}. \quad (6.101)$$

Нижче наведемо алгоритм $(k+1)$ -ої ітерації розв'язування задачі (6.86)–(6.88) (тобто побудови точки $x^{(k+1)}$ на основі наявної точки $x^{(k)}$) методом можливих напрямків.

Крок 1. Знаходження $f'(x^{(k)})$. У випадку, коли $\|f'(x^{(k)})\| \leq \varepsilon$, де ε – задана точність обчислень, процес пошуку розв'язку завершується і вважають, що $x^* = x^{(k)}$. У протилежному випадку здійснюється перехід до кроку 2.

Крок 2. Підстановка $x^{(k)}$ у нерівності (6.87), (6.88) та конкретизація на основі (6.90) множин активних обмежень I_k, J_k .

Крок 3. Аналіз множин I_k, J_k . Якщо $I_k = J_k = \emptyset$, то вектор $p^{(k)}$ визначається за формулою (6.89). У протилежному випадку вектор $p^{(k)}$ знаходять як розв'язок задачі (6.95)–(6.98).

Крок 4. Аналіз $p^{(k)}$. Якщо $p^{(k)} = 0$, то задача (6.86)–(6.88) розв'язана точно, тобто $x^* = x^{(k)}$. Якщо $p^{(k)} \neq 0$, то на основі (6.99)–(6.101) визначають α_k , а за допомогою (6.85) – наступне наближення $x^{(k+1)}$.

Проілюструємо цей алгоритм на відомому прикладі.

Приклад 6.7. За допомогою методу можливих напрямків знайти розв'язок задачі

$$\begin{cases} (x_1 - 4)^2 + (x_2 - 2)^2 \rightarrow \min, \\ x_1 + x_2 \leq 3, \\ x_1 + 2x_2 \leq 4, \\ x_1 \geq 0, x_2 \geq 0, \end{cases} \quad (6.102)$$

якщо $x^{(0)} = (1/2, 17/12)^T$, $\varepsilon = 0,01$.

◀ Ітерація 1

Крок 1. Оскільки $f'(x) = (2x_1 - 8, 2x_2 - 4)^T$, $f'(x^{(0)}) = (-7, -7/6)^T$, $\|f'(x^{(0)})\| > \varepsilon$, то переходимо до кроку 2.

Крок 2. У точці $x^{(0)}$ всі обмеження задачі (6.102) є строгими, тобто точка $x^{(0)}$ – внутрішня точка множини X і $I_0 = J_0 = \emptyset$.

Крок 3. Вектор можливих напрямків визначаємо за формулою (6.89), тобто $p^{(0)} = -f'(x^{(0)}) = (7, 7/6)^T$.

Крок 4. Вектор $p^{(0)} \neq 0$, тому із (6.99) та (6.100) знаходимо $\bar{\alpha}_1 = 13/98$, $\bar{\alpha}_2 = 1/14$, $\hat{\alpha}_1 = \hat{\alpha}_2 = +\infty$. Розв'язавши задачу (6.101), тобто задачу безумовної мінімізації

$$f(x^{(0)} + \alpha p^{(0)}) = \left(\frac{1}{2} + 7\alpha - 4\right)^2 + \left(\frac{17}{12} + \frac{7}{6}\alpha - 2\right)^2 \rightarrow \min_{\alpha},$$

отримаємо $\alpha^* = 1/2$.

Отже, $\alpha_0 = \min\{\bar{\alpha}_1, \bar{\alpha}_2, \hat{\alpha}_1, \hat{\alpha}_2, \alpha^*\} = 1/14$, тому $x^{(1)} = x^{(0)} + \alpha_0 p^{(0)} = (1/2, 17/12)^T + \frac{1}{14}(7, 7/6)^T = (1, 3/2)^T$.

Ітерація 2

Крок 1. $f'(x^{(1)}) = (-6, -1)^T$, $\|f'(x^{(1)})\| > \varepsilon$, тому переходимо до кроку 2.

Крок 2. У точці $x^{(1)}$ друге обмеження задачі (6.102) активне (перетворюється у рівність), тобто ця точка є

граничною точкою допустимої множини X , тому $I_1 = \{2\}$, $J_1 = \emptyset$.

Крок 3. Вектор можливих напрямків $p^{(1)}$ знаходимо як розв'язок задачі (6.95)–(6.98), тобто задачі

$$\begin{cases} \langle f'(x^{(1)}), p \rangle = -6p_1^+ + 6p_1^- - p_2^+ + p_2^- \rightarrow \min, \\ p_1^+ - p_1^- + 2p_2^+ - 2p_2^- \leq 0, \\ p_1^+ + p_1^- + p_2^+ + p_2^- \leq 1, \\ p_j^+ \geq 0, p_j^- \geq 0, p_j^+ p_j^- = 0, j = \overline{1,2}. \end{cases}$$

Розв'язавши цю задачу симплекс-методом, отримаємо $p_1^+ = 2/3$, $p_2^- = 1/3$, $p_1^- = p_2^+ = 0$, тому $p^{(1)} = (p_1^+ - p_1^-, p_2^+ - p_2^-)^T = (2/3, -1/3)^T$.

Крок 4. Оскільки $p^{(1)} \neq 0$, то за допомогою (6.99)–(6.101) знаходимо $\bar{\alpha}_1 = 3/2$, $\bar{\alpha}_2 = +\infty$, $\hat{\alpha}_1 = +\infty$, $\hat{\alpha}_2 = 9/2$, $\alpha^* = 156/85$, звідки $\alpha_1 = 3/2$, $x^{(2)} = x^{(1)} + \alpha_1 p^{(1)} = (1, 3/2)^T + 3/2(2/3, -1/3)^T = (2, 1)^T$.

Ітерація 3

Крок 1. $f'(x^{(2)}) = (-4, -2)^T$, тобто $\|f'(x^{(2)})\| > \varepsilon$, тому переходимо до кроку 2.

Крок 2. У точці $x^{(2)}$ перші два обмеження активні, тобто $I_2 = \{1, 2\}$, $J_2 = \emptyset$.

Крок 3. Для знаходження вектора $p^{(2)}$ формуємо задачу (6.95)–(6.98):

$$\begin{cases} -4p_1^+ + 4p_1^- - 2p_2^+ + 2p_2^- \rightarrow \min, \\ p_1^+ - p_1^- + p_2^+ - p_2^- \leq 0, \\ p_1^+ - p_1^- + 2p_2^+ - 2p_2^- \leq 0, \\ p_1^+ + p_1^- + p_2^+ + p_2^- \leq 1, \\ p_j^+ \geq 0, p_j^- \geq 0, p_j^+ p_j^- = 0, j = \overline{1,2}. \end{cases}$$

Розв'язком задачі є $p^{(2)} = (1/2, -1/2)^T$.

Крок 4. $p^{(2)} \neq 0$, тому за допомогою (6.99)–(6.101) знаходимо $\bar{\alpha}_1 = \bar{\alpha}_2 = +\infty$, $\hat{\alpha}_1 = +\infty$, $\hat{\alpha}_2 = 1$, $\alpha^* = 1$, звідки $\alpha_2 = 1$, $x^{(3)} = x^{(2)} + \alpha_2 p^{(2)} = (5/2, 1/2)^T$.

Ітерація 4

На даній ітерації (крок 3) отримаємо $p^{(3)} = 0$, що означає, що задача (6.102) розв'язана точно і $x^* = x^{(3)} = (5/2, 1/2)^T$, $f(x^*) = 9/2$. ►

6.6.2. Метод можливих напрямків у випадку нелінійних обмежень

Узагальненням задачі (6.86)–(6.88) є задача

$$f(x) \rightarrow \min, \quad (6.103)$$

$$g_i(x) \leq 0, i = \overline{1, m}, \quad (6.104)$$

де серед функцій $g_i(x)$ ($i = \overline{1, m}$) є нелінійні функції, причому $f(x)$ та $g_i(x)$ ($i = \overline{1, m}$) – опуклі та диференційовні в $\mathbb{R}^n$. Для побудови ітераційної послідовності (6.85) потрібно мати правило вибору вектора можливого напрямку $p^{(k)}$.

Якщо точка $x^{(k)}$ є внутрішньою точкою допустимої множини X (нерівності (6.104) виконуються строго), то вибір $p^{(k)}$ здійснюється згідно з (6.89).

Якщо точка $x^{(k)}$ є граничною точкою допустимої множини X , то множина індексів активних обмежень $I_k = \{i / g_i(x^{(k)}) = 0\} \neq \emptyset$. Враховуючи, що деякі з функцій $g_i(x)$ ($i = \overline{1, m}$) можуть бути лінійними, розіб'ємо множину I_k на дві підмножини $I_k^{(1)}$ та $I_k^{(2)}$ ($I_k^{(1)} \cap I_k^{(2)} = \emptyset, I_k^{(1)} \cup I_k^{(2)} = I_k$), які відносяться відповідно до нелінійних і лінійних функцій обмежень. Розглянемо у просторі $\mathbb{R}^n$ таку допоміжну задачу лінійного програмування:

$$\tilde{f}(p_1, \dots, p_n, u) = -u \rightarrow \min, \quad (6.105)$$

$$\langle f'(x^{(k)}), p \rangle + u \leq 0, \quad (6.106)$$

$$\langle g'_i(x^{(k)}), p \rangle + u \leq 0, i \in I_k^{(1)}, \quad (6.107)$$

$$\langle g'_i(x^{(k)}), p \rangle \leq 0, i \in I_k^{(2)}, \quad (6.108)$$

$$u \geq 0, p_j = p_j^+ - p_j^-, p_j^+ \geq 0, p_j^- \geq 0, p_j^+ p_j^- = 0, j = \overline{1, n}, \quad (6.109)$$

$$\sum_{j=1}^n |p_j| = \sum_{j=1}^n (p_j^+ + p_j^-) \leq 1, \quad (6.110)$$

де змінні p_j^+ та p_j^- мають той самий зміст, що і в задачі (6.95)–(6.98), а при розв'язуванні задачі (6.105)–(6.110) симплекс-методом p_j^+ та p_j^- одночасно не включають у число базисних змінних.

Сформулюємо наступну теорему.

Теорема 6.13. Якщо $x^{(k)}$ – гранична точка допустимої множини X задачі (6.103), (6.104), то розв'язок $(p^{(k)}, u_k)$

задачі лінійного програмування (6.105)–(6.110) існує. При цьому, якщо $u_k > 0$, то вектор $p^{(k)}$ є напрямком спадання функції $f(x)$ на множині X із точки $x^{(k)}$, а якщо $u_k = 0$, то $x^{(k)}$ – точка мінімуму функції $f(x)$ на X .

Знайшовши вектор можливого напрямку $p^{(k)}$ із точки $x^{(k)}$, визначають, як і у випадку задачі з лінійними обмеженнями, величину кроку α_k :

$$\alpha_k = \min \{ \alpha_{k1}, \dots, \alpha_{km}, \alpha_k^* \}, \quad (6.111)$$

де α_{ki} ($i = \overline{1, m}$) – максимально можливе переміщення із точки $x^{(k)}$ вздовж вектора $p^{(k)}$ з урахуванням i -го обмеження (6.104), тобто з урахуванням умови

$$g_i(x^{(k)} + \alpha_{ki} p^{(k)}) = 0, \quad (6.112)$$

а α_k^* – це розв'язок задачі

$$f(x^{(k)} + \alpha p^{(k)}) \rightarrow \min_{\alpha > 0}. \quad (6.113)$$

Критерієм закінчення обчислень і досягнення заданої точності $\varepsilon > 0$ при розв'язуванні задачі (6.103), (6.104) методом можливих напрямків є виконання принаймні однієї з умов $u_k \leq \varepsilon$, $\|f'(x^{(k)})\| \leq \varepsilon$. Тоді вважають, що розв'язок знайдений і $x^* = x^{(k)}$, $f^* = f(x^{(k)})$. Виконання умов $f'(x^{(k)}) = 0$, $u_k = 0$ свідчить про те, що точка мінімуму x^* знайдена точно.

Алгоритм $(k+1)$ -ої ітерації, результатом якої буде точка $x^{(k+1)}$, конкретизується такими кроками.

Крок 1. Знаходження $f'(x^{(k)})$ і перевірка умови $\|f'(x^{(k)})\| \leq \varepsilon$, де ε – задана точність розрахунків. При

виконанні цієї умови обчислення завершуються, тобто вважають, що $x^* = x^{(k)}$, $f^* = f(x^*) = f(x^{(k)})$, а при невиконанні – переходять до кроку 2.

Крок 2. Підстановка $x^{(k)}$ у нерівності (6.104) та уточнення множин індексів $I_k^{(1)}$, $I_k^{(2)}$.

Крок 3. Аналіз множин $I_k^{(1)}$, $I_k^{(2)}$. Якщо $I_k^{(1)} = I_k^{(2)} = \emptyset$, то вектор $p^{(k)}$ визначається за формулою (6.89). У протилежному випадку розв'язують задачу (6.105)–(6.110).

Крок 4. Аналіз $p^{(k)}$. Якщо знайдені $p^{(k)} = 0$ або $u_k \leq \varepsilon$, то здійснюють перехід до кроку 1, у протилежному випадку – до кроку 5.

Крок 5. Визначення максимально можливих переміщень α_{ki} ($i = \overline{1, m}$) вздовж вектора $p^{(k)}$ з урахуванням обмежень (6.104), тобто умов (6.112) (для $i \in I_k^{(2)}$, очевидно, $\alpha_{ki} = +\infty$), та α_k^* – розв'язку задачі (6.113).

Крок 6. Знаходження α_k за формулою (6.111) та чергового наближення $x^{(k+1)}$ до точки мінімуму x^* за формулою (6.85).

Приклад 6.8. Розв'язати задачу

$$\begin{cases} -x_1 - x_2 \rightarrow \min, \\ x_1^2 + x_2^2 - 25 \leq 0, \end{cases} \quad (6.114)$$

методом можливих напрямків, якщо $x^{(0)} = (0, 1)$, $\varepsilon = 0, 1$.

◀ Попередньо зауважимо, що всі вихідні положення щодо методу можливих напрямків у задачі (6.114) виконані, оскільки функції $f(x) = -x_1 - x_2$ та $g_1(x) = x_1^2 + x_2^2 - 25$ опуклі та диференційовні в $\mathbb{R}^2$.

Ітерація 1

Крок 1. Оскільки $f'(x^{(0)}) = (-1, -1)^T$ і $\|f'(x^{(0)})\| > \varepsilon$, то переходимо до кроку 2.

Крок 2. Точка $x^{(0)}$ є внутрішньою точкою допустимої множини задачі (6.114), тому $I_k = I_k^{(1)} = \emptyset$.

Крок 3. Вектор $p^{(0)}$ визначаємо за формулою (6.89), тобто $p^{(0)} = -f'(x^{(0)}) = (1, 1)^T$.

Крок 4. Оскільки $p^{(0)} \neq 0$, переходимо до кроку 5.

Крок 5. Знаходимо максимально можливе переміщення α_{01} із умови (6.112), тобто з умови $g_1(x^{(0)} + \alpha_{01}p^{(0)}) = g_1(\alpha_{01}, 1 + \alpha_{01}) = \alpha_{01}^2 + (1 + \alpha_{01})^2 - 25 = 0$. Додатним розв'язком отриманого рівняння є $\alpha_{01} = 3$. Величину α_0^* визначаємо як розв'язок задачі (6.113), тобто задачі

$$f(x^{(0)} + \alpha_0^* p^{(0)}) = f(\alpha_0^*, 1 + \alpha_0^*) = -\alpha_0^* - 1 - \alpha_0^* = -2\alpha_0^* - 1 \rightarrow \min_{\alpha_0^* > 0}.$$

Очевидно, $\alpha_0^* = +\infty$.

Крок 6. Із (6.111) випливає, що $\alpha_0 = 3$ і $x^{(1)} = x^{(0)} + \alpha_0 p^{(0)} = (0, 1)^T + 3(1, 1)^T = (3, 4)^T$.

Ітерація 2

Крок 1. $f'(x^{(1)}) = (-1, -1)^T$, $\|f'(x^{(1)})\| > \varepsilon$, тому переходимо до кроку 2.

Крок 2. Точка $x^{(1)} = (3, 4)^T$ є граничною точкою допустимої множини X задачі (6.114) ($g_1(x^{(1)}) = 0$), тому $I_k = I_k^{(1)} = \{1\}$.

Крок 3. Оскільки $I_k^{(1)} \neq \emptyset$, то для знаходження $p^{(1)}$ потрібно розв'язати задачу (6.105)–(6.110), яка у випадку задачі (6.114) має вигляд:

$$\begin{cases} \tilde{f}(p, u) = -u \rightarrow \min, \\ -p_1^+ + p_1^- - p_2^+ + p_2^- + u \leq 0, \\ 6p_1^+ - 6p_1^- + 8p_2^+ - 8p_2^- + u \leq 0, \\ p_1^+ + p_1^- + p_2^+ + p_2^- \leq 1, \\ u \geq 0, p_j^+ \geq 0, p_j^- \geq 0, p_j^+ p_j^- = 0, j = \overline{1, 2}. \end{cases}$$

Розв'язавши дану задачу симплекс-методом, отримаємо $u = 1/8$, $p_1^+ = 9/16$, $p_2^- = 7/16$, $p_1^- = p_2^+ = 0$, тобто $p^{(1)} = (9/16, -7/16)^T$.

Крок 4. $p^{(1)} \neq 0$, тому переходимо до кроку 5.

Крок 5. Із умови (6.112), тобто умови $g_1(x^{(1)} + \alpha_{11} p^{(1)}) =$

$$= g_1\left(3 + \frac{9}{16}\alpha_{11}, 4 - \frac{7}{16}\alpha_{11}\right) = \left(3 + \frac{9}{16}\alpha_{11}\right)^2 + \left(4 - \frac{7}{16}\alpha_{11}\right)^2 - 25 = 0,$$

знаходимо $\alpha_{11} = 16/65$. Величину α_1^* шукаємо як розв'язок задачі

$$f(x^{(1)} + \alpha_1^* p^{(1)}) = f\left(3 + \frac{9}{16}\alpha_1^*, 4 - \frac{7}{16}\alpha_1^*\right) = -\frac{1}{8}\alpha_1^* - 7 \rightarrow \min_{\alpha_1^* > 0}.$$

Очевидно, $\alpha_1^* = +\infty$.

Крок 6. Із (6.111) маємо $\alpha_1 = 16/65$ і

$$x^{(2)} = x^{(1)} + \alpha_1 p^{(1)} = (3, 4)^T + \frac{16}{65} \left(\frac{9}{16}, -\frac{7}{16}\right)^T = (204/65, 253/65)^T \approx (3, 138; 3, 892)^T.$$

Ітерація 3

Точка $x^{(2)}$ є граничною точкою ($g_1(x^{(2)}) = 0$). Реалізація алгоритму на цій ітерації приводить до таких результатів: $p^{(2)} = (0, 547; -0, 453)^T$; $u_2 = 0, 125$; $\alpha_2 = 0, 186$; $x^{(3)} = (3, 240; 3, 809)^T$.

Ітерація 4

Розв'язавши на цій ітерації задачу (6.105)–(6.110), отримаємо: $u_3 = 0,094 < 0,1 = \varepsilon$, тобто задана точність буде досягнута. Розв'язком задачі (6.114) буде $x^* = x^{(3)} = (3,240; 3,809)^T$, $f^* = f(x^{(3)}) = -7,049$.

Зауважимо, що задачу (6.114) можна розв'язати графічним методом, в результаті чого матимемо $x^* = (5/\sqrt{2}, 5\sqrt{2}) \approx (3,536; 3,536)^T$, $f^* = -5\sqrt{2} \approx -7,071$. ►

6.7. Градієнтні методи умовної мінімізації

Вище було викладено методи можливих напрямків для двох класів задач умовної мінімізації. Більш загальні варіанти цих методів, у яких похідні від функцій не використовуються, належать до прямих методів умовної мінімізації (методів нульового порядку). Ті методи умовної мінімізації, у яких використання похідних принципове, можна віднести до методів умовної мінімізації вищих порядків (першого, другого тощо). До цієї групи методів відносяться, зокрема, градієнтні методи умовної мінімізації, які розглянемо нижче також для опуклої задачі (задачі опуклого програмування) (6.103), (6.104), у якій функції $f(x)$ та $g_i(x)$ ($i = \overline{1, m}$) опуклі та диференційовні.

6.7.1. Метод проекції градієнта

Перш за все уточнимо поняття проекції точки на множину та пов'язані з цим поняттям факти з теорії.

Означення 6.4. Проекцією точки $x \in \mathbb{R}^n$ на замкнену множину $X \subset \mathbb{R}^n$ називається найближча до x точка $\bar{x}$ множини X , тобто точка, для якої

$$\rho(x, \bar{x}) = \|x - \bar{x}\| = \min_{y \in X} \|x - y\|,$$

якщо така точка існує.

Надалі проекцію $\bar{x}$ точки x на множину X будемо позначати через $P_X(x)$ (якщо $x \in X$, то $P_X(x) = x$).

Допустима множина X є опуклою замкненою множиною. Сформулюємо два важливих твердження, що стосуються опуклих задач:

1) якщо X – опукла замкнена множина із $\mathbb{R}^n$, то для будь-якої точки $x \in \mathbb{R}^n$ існує єдина проекція на X ;

2) щоб точка $x^* \in X$ була розв'язком задачі (6.103), (6.104), необхідно й достатньо, щоб для всіх $\alpha > 0$ виконувалась рівність

$$x^* = P_X(x^* - \alpha f'(x^*)). \quad (6.115)$$

Умова (6.115) в ідейному плані розкриває сутність методу проекції градієнта, у якому ітераційний процес пошуку чергового наближення $x^{(k+1)}$ до точки мінімуму x^* цільової функції $f(x)$ описується формулою

$$x^{(k+1)} = P_X(x^{(k)} - \alpha_k f'(x^{(k)})), \quad k = 0, 1, \dots, \quad x^{(0)} \in X, \quad (6.116)$$

де величина кроку α_k вибирається так само, як і у випадку градієнтних методів безумовної мінімізації. Зазначимо також, що згідно із (6.116) на кожній ітерації передбачена процедура повернення точки $(x^{(k)} - \alpha_k f'(x^{(k)}))$ на допустиму множину X , якщо ця точка не належить X .

Критерієм завершення обчислень за формулою (6.116) є виконання однієї з умов: $\|f'(x^{(k)})\| \leq \varepsilon$ або $\|x^{(k)} - x^{(k-1)}\| \leq \varepsilon$ ($\varepsilon > 0$ – задана точність знаходження розв'язку). При цьому вважають, що $x^* = x^{(k)}$, $f^* = f(x^{(k)})$.

Зауважимо, що у (6.116) задача проектування на множину X , тобто задача

$$\|y^{(k)} - x\|^2 \rightarrow \min, x \in X, y^{(k)} = x^{(k)} - \alpha_k f'(x^{(k)}) \quad (6.117)$$

є задачею нелінійного програмування, розв'язування якої може викликати певні ускладнення. У ряді випадків задача (6.117) може бути розв'язана у явному вигляді (це залежить від структури множини X) або є задачею квадратичного програмування, тому її розв'язок знаходиться за скінченне число кроків (множина X задається лише лінійними обмеженнями-нерівностями). Наприклад, якщо $X = \mathbb{R}_+^n$ (невід'ємний ортант простору $\mathbb{R}^n$), то $P_X(x) = \bar{x}$, де $\bar{x}_j = \max\{0, x_j\}$, $j = \overline{1, n}$. Якщо $X = \{x \in \mathbb{R}^n / a_j \leq x_j \leq b_j, a_j < b_j, j = \overline{1, n}\}$ – n -вимірний паралелепіпед, то $P_X(x) = \bar{x}$, де

$$\bar{x}_j = \begin{cases} a_j, & x_j < a_j, \\ b_j, & x_j > b_j \\ x_j, & a_j \leq x_j \leq b_j. \end{cases}$$

Для кулі радіуса $R > 0$ з центром у точці $x^{(0)}$ ($X = \{x \in \mathbb{R}^n / \|x - x^{(0)}\| \leq R\}$)

$$P_X(x) = \begin{cases} x, & x \in X, \\ x^{(0)} + R \frac{x - x^{(0)}}{\|x - x^{(0)}\|}, & x \notin X. \end{cases}$$

Ще раз підкреслимо, що задача проектування (6.117) суттєво залежить від множини X і в кожному з випадків повинна бути окремо досліджена.

Алгоритм $(k+1)$ – ітерації (ітерації, на якій знаходиться точка $x^{(k+1)}$) розв’язування задачі (6.103), (6.104) методом проекції градієнта передбачає такі два кроки (етапи).

Крок 1. Аналіз точки $x^{(k)}$. Якщо $\|f'(x^{(k)})\| \leq \varepsilon$ або $\|x^{(k)} - x^{(k-1)}\| \leq \varepsilon$ ($\varepsilon > 0$ – задана точність наближення), то обчислення припиняються і $x^* = x^{(k)}$, $f^* = f(x^{(k)})$. У протилежному випадку переходимо до кроку 2.

Крок 2. Визначення α_k та знаходження точки $x^{(k+1)}$ за допомогою формули (6.116).

Продемонструємо метод проекції градієнта на конкретному прикладі.

Приклад 6.9. Методом проекції градієнта знайти розв’язок задачі

$$\begin{cases} -x_1 + x_2^2 \rightarrow \min, \\ x_1^2 + x_2^2 \leq 1, \end{cases} \quad (6.118)$$

якщо $x^{(0)} = (0; 0,5)^T$, $\varepsilon = 0,01$.

◀ Очевидно, тобто $x^{(0)}$ – внутрішня точка допустимої множини. Крім того, цільова функція є квадратичною з матрицею $A = \begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix}$, яка має найбільше власне значення

$\lambda_2 = 2$. Це означає, що у (6.116) можна взяти $\alpha_k \equiv \alpha \in (0, 2/\lambda_2) = (0, 1)$ ($k = 0, 1, \dots$), наприклад $\alpha_k = 0,75$ (обґрунтування у п. 5.5.1).

Ітерація 1

Крок 1. Враховуючи, що $f'(x^{(0)}) = (-1, 1)^T$, $\|f'(x^{(0)})\| > \varepsilon$, переходимо до кроку 2.

Крок 2. Із (6.116) отримаємо ($\alpha_k = \alpha = 0,75$)

$$x^{(1)} = P_X \left(x^{(0)} - \alpha f'(x^{(0)}) \right) = P_X [(0; 0,5)^T - 0,75(-1,1)^T] = \\ = P_X \left((0,75; -0,25)^T \right).$$

Точка $(0,75; -0,25)^T$ належить допустимій множині X , тому $x^{(1)} = (0,75; -0,25)^T$.

Ітерація 2

Крок 1. Оскільки $f'(x^{(1)}) = (-1; -0,5)^T$, $x^{(1)} - x^{(0)} = (0,75; -0,25)^T - (0; 0,5)^T = (0,75; -0,75)^T$, $\|f'(x^{(1)})\| \approx 1,12 > \varepsilon$, $\|x^{(1)} - x^{(0)}\| \approx 1,06 > \varepsilon$, то бажана точність не досягнена.

Переходимо до кроку 2.

Крок 2. Знаходимо

$$x^{(2)} = P_X \left[x^{(1)} - \alpha_1 f'(x^{(1)}) \right] = P_X [(0,75; -0,25)^T - \\ - 0,75(-1; -0,5)^T] = P_X \left((1,5; 0,125)^T \right).$$

Точка $(1,5; 0,125)^T \notin X$. Оскільки допустима множина X є кругом радіуса $R=1$ з центром у точці $(0,0)^T$, то

$$x^{(2)} = P_X \left((1,5; 0,125)^T \right) = \frac{(1,5; 0,125)^T}{\sqrt{1,5^2 + 0,125^2}} = (0,996; 0,083)^T.$$

Ітерація 3

Обчислення потрібно продовжити, оскільки

$$\|f'(x^{(2)})\| = 1,013 > \varepsilon, \quad \|x^{(2)} - x^{(1)}\| = 0,414 > \varepsilon.$$

Результати цих обчислень наведено у табл. 6.1.

Таблиця 6.1

k	$x^{(k)}$	$f'(x^{(k)})$	$\ x^{(k)} - x^{(k-1)}\ $	$x^{(k)} - \alpha f'(x^{(k)})$
2	(0,9965; 0,0830)	(-1;0,1661)	0,414	(1,7465;-0,0415)
3	(0,9997; -0,0238)	(-1;-0,0476)	0,107	(1,7497;0,0119)
4	(0,99998; 0,00679)	(-1;0,0136)	0,031	(1,74998;-0,00339)
5	(0,99998; -0,00194)		0,0087	

Із табл. 6.1 випливає, що на п'ятій ітерації $\|x^{(5)} - x^{(4)}\| = 0,0087 < \varepsilon = 0,01$, тому задача (6.118) вважається розв'язаною і $x^* = x^{(5)} = (0,99998; -0,00194)^T$, $f^* = f(x^{(5)}) = -1$. Розв'язавши цю задачу графічним методом, легко перевірити, що $x^* = (1,0)^T$, $f^* = -1$. ►

6.7.2. Метод умовного градієнта

Знову повернемося до опуклої задачі умовної мінімізації (6.103), (6.104), у якій функції $f(x)$, $g_i(x)$ ($i = \overline{1, m}$) опуклі та диференційовні. Розкриємо зміст методу умовного градієнта. Нехай $x^{(k)} \in X$ – чергове наближення до розв'язку x^* задачі (6.103), (6.104). Якщо $f'(x^{(k)}) = 0$, то $x^* = x^{(k)}$, а якщо $f'(x^{(k)}) \neq 0$, то в околі точки $x^{(k)}$ згідно із (5.7)

$$f(x) = f(x^{(k)}) + \langle f'(x^{(k)}), x - x^{(k)} \rangle + o(\|x - x^{(k)}\|),$$

тому лінійна функція

$$l_k(x) = f(x^{(k)}) + \langle f'(x^{(k)}), x - x^{(k)} \rangle$$

наближає в цьому околі функцію $f(x)$ з точністю до величини $o(\|x - x^{(k)}\|)$. Вважаючи, що $\tilde{x}^{(k)}$ – розв’язок задачі

$$l_k(x) \rightarrow \min, x \in X, \quad (6.119)$$

наступне наближення $x^{(k+1)}$ до точки мінімуму x^* функції $f(x)$ на X можна знайти за формулою

$$x^{(k+1)} = x^{(k)} + \alpha_k (\tilde{x}^{(k)} - x^{(k)}), \alpha_k \in [0,1]. \quad (6.120)$$

Зрозуміло, що $x^{(k+1)} \in X$ (X – опукла множина). Довжину кроку α_k із (6.120) можна знайти, наприклад, як розв’язок одновимірної задачі мінімізації

$$f(x^{(k)} + \alpha(\tilde{x}^{(k)} - x^{(k)})) \rightarrow \min_{\alpha \in [0,1]} \quad (6.121)$$

або шляхом зменшення величини α_k (починають з $\alpha_k = 1$) до тих пір, поки не буде виконуватись умова $f(x^{(k+1)}) < f(x^{(k)})$. Отже, на кожному кроці ітераційного процесу метод умовного градієнта фактично передбачає розв’язування допоміжної задачі мінімізації (6.119) і одновимірної задачі мінімізації (6.121).

Критерієм завершення обчислень може бути виконання однієї з нерівностей: $\|x^{(k+1)} - x^{(k)}\| \leq \varepsilon$, $\|f(x^{(k+1)}) - f(x^{(k)})\| \leq \varepsilon$, $\|f'(x^{(k+1)})\| \leq \varepsilon$ ($\varepsilon > 0$ – задана точність обчислень).

Алгоритм $(k+1)$ -ої ітерації (ітерації знаходження $x^{(k+1)}$) розв’язування задачі (6.103), (6.104) методом умовного градієнта конкретизується наступними кроками.

Крок 1. Розв’язування допоміжної задачі (6.119) та знаходження $\tilde{x}^{(k)}$.

Крок 2. Розв’язування одновимірної задачі мінімізації (6.121) та знаходження $\alpha^* = \alpha_k$.

Крок 3. Визначення точки $x^{(k+1)}$ за формулою (6.120) і перевірка умови досяжності заданої точності наближення. Якщо точність забезпечена, то покладають $x^* = x^{(k+1)}$, $f^* = f(x^{(k+1)})$, у протилежному випадку повертаються до кроку 1.

Зауважимо, що метод умовного градієнта зазвичай використовується тоді, коли розв'язок допоміжної задачі (6.119) знаходиться достатньо просто. До даних випадків належать такі:

1) допустима множина X задана лінійними обмеженнями, тобто задача (6.119) є задачею лінійного програмування;

2) допустима множина $X = \{x \in \mathbb{R}^n / a_j \leq x_j \leq b_j, j = \overline{1, n}\}$ – n – вимірний паралелепіпед, тому

$$\tilde{x}_j^{(k)} = \begin{cases} a_j, & \text{якщо } \frac{\partial f(x^{(k)})}{\partial x_j} > 0, \\ b_j, & \text{якщо } \frac{\partial f(x^{(k)})}{\partial x_j} < 0, \\ \frac{a_j + b_j}{2}, & \text{якщо } \frac{\partial f(x^{(k)})}{\partial x_j} = 0; \end{cases}$$

3) допустима множина $X = \{x \in \mathbb{R}^n / \|x - z^{(0)}\| \leq R\}$ – куля радіуса R з центром у точці $z^{(0)}$, тому

$$\tilde{x}^{(k)} = z^{(0)} - R \frac{f'(x^{(k)})}{\|f'(x^{(k)})\|}.$$

Розглянемо приклад застосування методу умовного градієнта.

Приклад 6.10. Користуючись методом умовного градієнта, для задачі

$$\begin{cases} 1/2(x_1 - 3/4)^2 + 1/2(x_2 - 3/4)^2 \rightarrow \min, \\ x_1 + x_2 \leq 1, \\ x_1 \geq 0, x_2 \geq 0 \end{cases} \quad (6.122)$$

виконати одну ітерацію з початкової точки $x^{(0)} = (0, 0)^T$.

◀ Крок 1. Оскільки

$$l_0(x) = f(0, 0) + \sum_{j=1}^2 \frac{\partial f(0, 0)}{\partial x_j} x_j = \frac{9}{16} - \frac{3}{4}x_1 - \frac{3}{4}x_2,$$

то допоміжна задача (6.119) зводиться до задачі

$$\begin{cases} -3/4 x_1 - 3/4 x_2 \rightarrow \min, \\ x_1 + x_2 \leq 1, \\ x_1 \geq 0, x_2 \geq 0, \end{cases}$$

розв'язком якої буде $\tilde{x}^{(0)} = (0, 1)^T$ (у цьому легко переконатися, розв'язавши задачу графічним або симплексним методом).

Крок 2. Враховуючи, що $x^{(0)} + \alpha(\tilde{x}^{(0)} - x^{(0)}) = (0, \alpha)^T$, задача (6.121) має вигляд

$$f(0, \alpha) = 1/2(-3/4)^2 + 1/2(\alpha - 3/4)^2 \rightarrow \min_{\alpha \in [0, 1]}.$$

Очевидно, розв'язком цієї задачі є $\alpha^* = \alpha_0 = 3/4$.

Крок 3. Із (6.120) маємо $x^{(1)} = x^{(0)} + \alpha_0(\tilde{x}^{(0)} - x^{(0)}) = (0, 3/4)^T$.

Далі здійснюється перевірка умови досяжності заданої точності обчислень $\varepsilon > 0$ та реалізуються інші дії алгоритму. ►

6.8. Методи послідовної безумовної мінімізації

Частина методів умовної мінімізації використовує ідею переходу від задачі (6.1) до послідовності задач безумовної мінімізації

$$\begin{cases} f_k(x) = f(x) + \varphi_k(x) \rightarrow \min, \\ x \in \mathbb{R}^n, k = 1, 2, \dots, \end{cases} \quad (6.123)$$

де функції $\varphi_k(x)$ ($k = 1, 2, \dots$) більш повно враховують обмеження задачі (6.1) при зростанні k . Наближеним розв'язком задачі (6.1) вважається розв'язок $x^{(k)}$ задачі (6.123) при деякому (достатньо великому) k . Нижче розглянемо два методи, що належать до цієї групи методів.

6.8.1. Метод штрафних функцій

Ідея методу штрафних функцій полягає у підборі (при достатньо великих k) такої допоміжної функції $\varphi_k(x)$ із (6.123), яка б мало відрізнялася від функції $f(x)$ при $x \in X$ і швидко зростала при віддаленні точки $x \notin X$ від допустимої множини. У зв'язку з цим подамо наступне означення.

Означення 6.5. *Послідовність функцій $\{\varphi_k(x)\}$, які визначені у $\mathbb{R}^n$ і для яких*

$$\lim_{k \rightarrow \infty} \varphi_k(x) = \begin{cases} 0, & \text{якщо } x \in X, \\ +\infty, & \text{якщо } x \notin X, \end{cases} \quad (6.124)$$

називається послідовністю штрафних функцій множини $X \subset \mathbb{R}^n$.

Наприклад, послідовність функцій $\varphi_k(x) = k\varphi(x)$ ($k = 1, 2, \dots$), де $\varphi(x)$ набуває нульових значень при $x \in X$ і додатних значень при $x \notin X$, є послідовністю штрафних

функцій, оскільки для неї виконується умова (6.124). Функція $\varphi(x)$, яка фігурує у даній послідовності, може бути відстанню від точки x до множини X , тобто послідовність штрафних функцій набуде конкретного вигляду

$$\varphi_k(x) = k\rho(x, X), \quad k = 1, 2, \dots$$

Раніше вже йшлося про те, що знаходження відстані $\rho(x, X)$ – у загальному випадку непроста задача, тому на практиці штрафні функції часто вибираються такими, щоб вони враховували явне задання множини X і були зручними для використання. Якщо $X = \{x \in \mathbb{R}^n / g_i(x) \leq 0, i = \overline{1, m}\}$, де $g_i(x)$ ($i = \overline{1, m}$) – опуклі функції, то функцію $\varphi(x)$ можна взяти у вигляді

$$\varphi(x) = \sum_{i=1}^m \psi(g_i(x)), \quad (6.125)$$

де $\psi(z)$ – неперервна функція, причому $\psi(z) = 0$ при $z \leq 0$ і $\psi(z) > 0$ при $z > 0$. Очевидно, що для послідовності

$$\varphi_k(x) = a_k \sum_{i=1}^m \psi(g_i(x)), \quad k = 1, 2, \dots,$$

у якій $a_k > 0$ і $\lim_{k \rightarrow \infty} a_k = +\infty$ (наприклад, $a_k = k$), умова (6.124) виконується, тому вона є послідовністю штрафних функцій. Основним критерієм вибору функції $\psi(z)$ є максимальне спрощення задач мінімізації (6.123). У багатьох випадках функцію (6.125) вибирають у вигляді

$$\varphi(x) = \sum_{i=1}^m [g_i^+(x)]^2, \quad (6.126)$$

де

$$g_i^+(x) = \max\{0, g_i(x)\} = \frac{1}{2}(g_i(x) + |g_i(x)|).$$

Приклад 6.11. Знайти розв'язок задачі

$$\begin{cases} x_1^2 + (x_2 - 4)^2 \rightarrow \min, \\ -x_1 + x_2 - 2 \leq 0, \\ -x_1^2 + 2x_2 \leq 0, \end{cases} \quad (6.127)$$

використовуючи метод штрафних функцій.

◀ Візьмемо функцію $\varphi(x)$ у вигляді (6.126), тобто

$$\begin{aligned} \varphi(x) = & \left[g_1^+(x) \right]^2 + \left[g_2^+(x) \right]^2 = \left[\frac{1}{2}(-x_1 + x_2 - 2 + |-x_1 + x_2 - 2|) \right]^2 + \\ & + \left[\frac{1}{2}(-x_1^2 + 2x_2 + |-x_1^2 + 2x_2|) \right]^2. \end{aligned}$$

Результати розв'язування допоміжних задач [21] відображено у табл. 6.2.

Таблиця 6.2

k	$x_1^{(k)}$	$x_2^{(k)}$
0	0	4
1	0,66	3,33
5	0,91	3,09
10	0,95	3,05
100	0,995	3,004

Враховуючи, що точний розв'язок задачі (6.127) $x^* = (1,3)^T$, із табл. 6.2 випливає, що при $k = 100$ цей розв'язок практично знайдений. ►

Що ж до методу штрафних функцій, то зауважимо, що, по-перше, критерієм завершення розрахунків може бути умова $\|x^{(k)} - x^{(k/2)}\| \leq \varepsilon$, де $\varepsilon > 0$ – задана точність, k – парне число (при виконанні цієї умови вважають, що $x^* = x^{(k)}$, $f^* = f(x^{(k)})$), а, по-друге, допоміжна функція $f_k(x)$ погано обумовлена при великих k , що і є недоліком методу.

6.8.2. Метод бар'єрних функцій

Характерна особливість цього методу полягає у тому, що функції $\varphi_k(x)$ вибираються такими, щоб при великих k функції $f_k(x)$ із (6.123) мало відрізнялися від функції $f(x)$ у внутрішніх точках допустимої множини X , але при наближенні точки $x \in X$ до межі множини X ці функції необмежено зростали, чого не вимагалось у методі штрафних функцій.

Означення 6.6. *Послідовність функцій $\{\varphi_k(x)\}$, які визначені в усіх внутрішніх точках множини X , називається послідовністю бар'єрних функцій множини X , якщо виконуються умови:*

1) *для будь-якої фіксованої внутрішньої точки x цієї множини $\lim_{k \rightarrow \infty} \varphi_k(x) = 0$;*

2) *для будь-якої послідовності $\{x^{(l)}\}$ внутрішніх точок множини X , збіжної до деякої граничної точки цієї множини, $\lim_{l \rightarrow \infty} \varphi_k(x^{(l)}) = +\infty$.*

Як впливає з означення 6.6, при великих k функції $\varphi_k(x)$ створюють уздовж межі допустимої множини відповідний «бар'єр». Ці функції можна вибирати у вигляді

$$\varphi_k(x) = b_k \varphi(x), \quad (6.128)$$

де $\{b_k\}$ – деяка числова послідовність, для якої $b_k > 0$, $\lim_{k \rightarrow \infty} b_k = 0$ (на практиці часто вибирають $b_k = 1/k$ $k = 1, 2, \dots$), а функція $\varphi(x)$ визначена в усіх внутрішніх точках множини X та прямує до $+\infty$ при наближенні внутрішньої точки x до межі множини X . Очевидно, у цьому випадку послідовність (6.128) задовольняє означення 6.6. Якщо,

наприклад $X = \{x \in \mathbb{R}^n / g_i(x) \leq 0, i = \overline{1, m}\}$, то функцією $\varphi(x)$ може бути функція

$$\varphi(x) = -\sum_{i=1}^m \frac{1}{g_i(x)},$$

а функцією $f_k(x)$ – функція

$$f_k(x) = f(x) - \frac{1}{k} \sum_{i=1}^m \frac{1}{g_i(x)}.$$

Інколи бар'єрні функції (6.128) для допустимої множини $X = \{x \in \mathbb{R}^n / g_i(x) \leq 0, i = \overline{1, m}\}$ вибирають такими, що

$$\varphi(x) = \sum_{i=1}^m |g_i(x)|^{-r}, r > 0$$

або

$$\varphi(x) = -\sum_{i=1}^m \ln[-g_i(x)]. \quad (6.129)$$

Критерієм завершення обчислень у методі бар'єрних функцій також може бути умова $\|x^{(k)} - x^{(k/2)}\| \leq \varepsilon$, де $\varepsilon > 0$ – задана точність, k – парне число.

Приклад 6.12. Розв'язати задачу

$$\begin{cases} x_1^2 + 2x_2^2 \rightarrow \min, \\ -x_1 + x_2 \leq 0, \\ -x_1 - x_2 + 1 \leq 0, \end{cases} \quad (6.130)$$

використовуючи бар'єрні функції (6.129) та критерій завершення обчислень $\|x^{(k)} - x^{(k/2)}\| \leq 0,002$ (k – парне).

◀ Переходимо від задачі (6.130) до послідовності задач безумовної мінімізації

$$f_k(x) = x_1^2 + 2x_2^2 - \frac{1}{k} [\ln(x_1 - x_2) + \ln(x_1 + x_2 - 1)] \rightarrow \min, k = 1, 2, \dots$$

Вибравши початкову точку $x^{(0)} = (2, 0)^T$ і застосувавши до цих задач метод Ньютона (п. 5.5.3), отримаємо шукану послідовність точок $x^{(k)}$. Остаточні результати такі [17]: $x^{(500)} = (0,6696; 0,3319)^T$, $x^{(1000)} = (0,6682; 0,3326)^T$, $\|x^{(1000)} - x^{(500)}\| = 1,65 \cdot 10^{-3} < 0,002$. Отже, $x^* = x^{(1000)}$, $f^* = f(x^{(1000)}) = 0,6687$. Легко перевірити також, що точка $x^{(1000)}$ розміщена близько до межі допустимої множини X . ►

ЧАСТИНА III. СПЕЦІАЛЬНІ КЛАСИ МОДЕЛЕЙ ТА МЕТОДІВ УМОВНОЇ ОПТИМІЗАЦІЇ

Розділ 7. ЛІНІЙНЕ ПРОГРАМУВАННЯ

7.1. Лінійне програмування як ефективний інструментарій дослідження лінійних моделей

Оскільки проблема оптимального планування в економіці одна із центральних, а багато економічних моделей – лінійні, то історично склалося так, що саме економічні задачі привели до виникнення та розвитку лінійного програмування (ЛП), а згодом – і до основ математичної економіки взагалі. До задач ЛП зводяться більшість задач оперативного й довгострокового планування у рамках підприємства чи галузі, задачі оптимальних концентрацій, перевезень та інші. У вступній частині цієї книги наведено деякі приклади оптимізаційних економіко-математичних моделей, що зводяться до задач ЛП, тобто до задач (В.1) або (розділ 6) до задач (6.2)–(6.5), у яких цільова функція $f(x)$ та функції обмежень $g_i(x)$ ($i = \overline{1, m}$), що формують допустиму множину X , лінійні. До таких прикладів належать модель оптимального планування виробництва (В.5), модель планування найбільш економічного раціону або оптимальної дієти (В.6), модель оптимізації транспортних перевезень (Т-задача) (В.7), модель оптимізації виробничої потужності підприємства (В.8), а також інші, які або можуть бути зведені до задач ЛП (модель максимізації рентабельності виробництва (В.9)) або можуть бути формалізовані у вигляді задач ЛП

(моделі максимізації випуску продукції (В.2), максимізації прибутку однопродуктової фірми (В.3) та інші).

Звичайно, лінійні моделі, які у математичному плані описуються як задачі ЛП, зустрічаються зазвичай в усіх галузях знань і практичної діяльності. Цей факт не потребує окремого підтвердження, тому наводити нові приклади лінійних оптимізаційних моделей немає сенсу. Але підкреслимо одну концептуальну особливість, яка характерна власне самим лінійними моделями, а точніше, моделюванню як методу побудови моделей (зокрема лінійних) та дослідження на їх основі реальних об'єктів (явищ, процесів тощо). Будь-яка модель є певною ідеалізацією досліджуваного об'єкта, яка у той чи інший спосіб відображає окремі його риси, властивості, змістові характеристики. Адекватність моделі фактично вимірюється рівнем відображення лише тих її характеристик, які вивчає дослідник. Це означає, що зазвичай (особливо на перших етапах дослідження) будують відносно прості моделі, які формалізують найпростіші концептуальні схеми та принципи досліджуваної реальності, які і є конкретною дослідницькою метою. І саме у цих випадках лінійні моделі, у тому числі лінійні оптимізаційні моделі, виконують принципово важливу роль, оскільки математичний інструментарій їх розв'язування та аналізу найбільш розвинений і ефективний на сьогодні.

Значна частина економічних, технічних, технологічних і інших явищ і процесів з високим рівнем адекватності описується лінійними моделями, тому ЛП не просто потужний інструмент для дослідження лінійних оптимізаційних моделей, але й фундаментальний методологічний засіб розвитку теорії оптимізації та моделювання у цілому.

7.2. Форми запису задачі лінійного програмування та їх еквівалентність

Із наведених у вступі прикладів задач ЛП, а також математичної моделі загальної задачі математичного програмування (6.2)–(6.5) випливає, що загальну задачу ЛП можна записати, наприклад, так:

$$\left\{ \begin{array}{l} \sum_{j=1}^n c_j x_j \rightarrow \min(\max), \\ \sum_{j=1}^n a_{ij} x_j \leq b_i, \quad i = \overline{1, k}, \\ \sum_{j=1}^n a_{ij} x_j = b_i, \quad i = \overline{k+1, m}, \\ x_j \geq 0, \quad j = \overline{1, l}; \quad l \leq n, \end{array} \right. \quad (7.1)$$

де $f(x) = \sum_{j=1}^n c_j x_j$ – цільова функція або лінійна форма,

яку потрібно мінімізувати чи максимізувати, $A = (a_{ij})_{i,j=1}^{m,n}$ – матриця коефіцієнтів системи обмежень, $b = (b_1, \dots, b_m)^T$ – вектор правих частин обмежень, $c = (c_1, \dots, c_n)^T$ – вектор коефіцієнтів лінійної форми, $x = (x_1, \dots, x_n)^T$ – вектор шуканих змінних, $X_0 = \{x \in \mathbb{R}^n / x_j \geq 0, j = \overline{1, l}; l \leq n\}$. Однак найзручніші у теоретичних і прикладних дослідженнях стандартна та канонічна форми запису задачі ЛП, тобто задача ЛП з обмеженнями-нерівностями та задача ЛП з обмеженнями-рівностями. Надалі будемо вважати, що стандартною задачею на максимум буде задача (B.5), яку для зручності подальших викладень перенумеруємо на (7.2):

$$\begin{cases} \sum_{j=1}^n c_j x_j \rightarrow \max, \\ \sum_{j=1}^n a_{ij} x_j \leq b_i, \quad i = \overline{1, m}, \\ x_j \geq 0, \quad j = \overline{1, n}, \end{cases} \quad (7.2)$$

а канонічною задачею на максимум – задача

$$\begin{cases} \sum_{j=1}^n c_j x_j \rightarrow \max, \\ \sum_{j=1}^n a_{ij} x_j = b_i, \quad i = \overline{1, m}, \\ x_j \geq 0, \quad j = \overline{1, n}. \end{cases} \quad (7.3)$$

Аналогічно до (7.2) і (7.3) стандартною та канонічною задачами на мінімум вважатимемо відповідно задачу (В.6), перенумеровану у даному розділі на (7.4):

$$\begin{cases} \sum_{j=1}^n c_j x_j \rightarrow \min, \\ \sum_{j=1}^n a_{ij} x_j \geq b_i, \quad i = \overline{1, m}, \\ x_j \geq 0, \quad j = \overline{1, n} \end{cases} \quad (7.4)$$

та задачу

$$\begin{cases} \sum_{j=1}^n c_j x_j \rightarrow \min, \\ \sum_{j=1}^n a_{ij} x_j = b_i, \quad i = \overline{1, m}, \\ x_j \geq 0, \quad j = \overline{1, n}, \end{cases} \quad (7.5)$$

Очевидно, задачі (7.2) і (7.3) (чи (7.4), (7.5)) є частинними випадками задачі (7.1). Важливість стандартної задачі ЛП полягає в тому, що до неї зводяться більшість прикладних задач. Канонічна форма задачі ЛП досить поширена у теоретичних дослідженнях і практичних обчисленнях.

Довільна задача ЛП може бути зведена до однієї з вищевказаних форм за допомогою таких еквівалентних перетворень:

1) перехід від максимізації лінійної форми $\sum_{j=1}^n c_j x_j$ до мінімізації $\left(-\sum_{j=1}^n c_j x_j \right)$ і навпаки;

2) перехід від нерівності $a_{i1}x_1 + \dots + a_{in}x_n \leq b_i$ до нерівності $-a_{i1}x_1 - \dots - a_{in}x_n \geq -b_i$ і навпаки;

3) перехід від нерівності $a_{i1}x_1 + \dots + a_{in}x_n \leq b_i$ до нерівності $a_{i1}x_1 + \dots + a_{in}x_n + x_{n+1} = b_i$, де $x_{n+1} \geq 0$ – певна величина, яка називається допоміжною змінною;

4) перехід від рівності $a_{i1}x_1 + \dots + a_{in}x_n = b_i$ до системи нерівностей

$$\begin{cases} a_{i1}x_1 + \dots + a_{in}x_n \leq b_i, \\ a_{i1}x_1 + \dots + a_{in}x_n \geq b_i; \end{cases}$$

5) перехід від довільної змінної x_j до невід'ємних змінних u_j, v_j шляхом заміни: $x_j = u_j - v_j$;

6) перехід від змінної $x_j \geq d_j$ (тобто змінної, обмеженої знизу) до невід'ємної змінної y_j шляхом заміни $y_j = x_j - d_j$.

Легко переконатися, що, використовуючи наведені вище перетворення, наприклад, задачу ЛП

$$\begin{cases} 3x_1 - 2x_2 - x_3 \rightarrow \min, \\ 2x_1 - x_2 + x_3 \geq 2, \\ 3x_1 + 2x_2 + x_3 \leq 6, \\ x_1 + x_2 + x_3 = 4, \\ x_1 \geq 0, x_2 \geq 0 \end{cases} \quad (7.6)$$

можна привести до канонічної задачі

$$\begin{cases} -3x_1 - 2x_2 + u_3 - v_3 \rightarrow \max, \\ 2x_1 - x_2 + u_3 - v_3 - x_4 = 2, \\ 3x_1 + 2x_2 + u_3 - v_3 + x_5 = 6, \\ x_1 + x_2 + u_3 - v_3 = 4, \\ u_3 \geq 0, v_3 \geq 0, \\ x_j \geq 0, j = 1, 2, 4, 5. \end{cases} \quad (7.7)$$

Аналогічно канонічну задачу ЛП

$$\begin{cases} x_1 + 2x_2 - x_3 - x_4 - 2x_5 \rightarrow \max, \\ x_1 - x_2 - 2x_3 = 4, \\ 2x_2 + 4x_3 - x_4 = 8, \\ x_2 + x_3 + x_5 = 6, \\ x_j \geq 0, j = \overline{1, 5} \end{cases} \quad (7.8)$$

можна записати у симетричній формі

$$\begin{cases} 2x_2 - x_3 \rightarrow \max, \\ -x_2 - 2x_3 \leq 4, \\ -2x_2 - 4x_3 \leq -8, \\ x_2 + x_3 \leq 6, \\ x_2 \geq 0, x_3 \geq 0. \end{cases} \quad (7.9)$$

Очевидно, що задачі (7.6), (7.7), а також (7.8), (7.9) можна записати і в інших еквівалентних формах.

У теоретичних дослідженнях та багатьох прикладних ситуаціях задачі ЛП (7.2)–(7.5) зручніше використовувати у матрично-векторних формах запису (7.10)–(7.13):

$$\begin{cases} \langle c, x \rangle \rightarrow \max, \\ Ax \leq b, \\ x \geq 0; \end{cases} \quad (7.10)$$

$$\begin{cases} \langle c, x \rangle \rightarrow \max, \\ Ax = b, \\ x \geq 0; \end{cases} \quad (7.11)$$

$$\begin{cases} \langle c, x \rangle \rightarrow \min, \\ Ax \geq b, \\ x \geq 0; \end{cases} \quad (7.12)$$

$$\begin{cases} \langle c, x \rangle \rightarrow \min, \\ Ax = b, \\ x \geq 0. \end{cases} \quad (7.13)$$

Очевидно, що у задачах (7.10)–(7.13) $0 = (0, \dots, 0)^T \in \mathbb{R}^n$.

7.3. Основні властивості задач лінійного програмування

Виходячи з еквівалентності форм запису задачі ЛП, у теоретичних дослідженнях часто використовують

канонічну задачу ЛП, наприклад задачу (7.11), тобто канонічну задачу на максимум.

Множиною допустимих розв'язків або допустимою множиною задачі (7.11) є множина $\mathbb{R}^n$. Інколи її ще називають множиною допустимих планів. Це пов'язано з тим, що в прикладних задачах (наприклад, економічних) термін «програмування» й «планування» ототожнюють.

Нижче, базуючись на задачі (7.11), сформулюємо й доведемо ряд теорем, які характеризують деякі найважливіші властивості задач лінійного програмування.

Теорема 7.1. *Допустима множина задачі ЛП (7.11) опукла.*

◀ Щоб довести це твердження, візьмемо довільні два вектори $x^{(1)}, x^{(2)} \in X$ (припускається, що $X \neq \emptyset$) і число $\alpha \in [0, 1]$. Покажемо, що $x = \alpha x^{(1)} + (1 - \alpha)x^{(2)} \in X$. Справді, $x \geq 0$ і

$$\begin{aligned} Ax &= A(\alpha x^{(1)} + (1 - \alpha)x^{(2)}) = \alpha Ax^{(1)} + (1 - \alpha)Ax^{(2)} = \\ &= \alpha b + (1 - \alpha)b = b, \end{aligned}$$

що й треба було встановити. ►

Перш, ніж сформулювати наступну теорему, зазначимо, що множину $X^* = \left\{ x^* \in X / \langle c, x^* \rangle = \max_{x \in X} \langle c, x \rangle \right\}$, тобто множину розв'язків задачі ЛП (7.11) традиційно називають множиною оптимальних розв'язків або оптимальних планів задачі (7.11), що також пов'язано зі специфікою термінології у постановках економічних задач (як і у випадку з поняттями «допустимий розв'язок» чи «допустимий план»).

Не порушуючи строгості подальших викладень, також будемо користуватися поняттям «оптимальний розв'язок» чи «оптимальний план».

Теорема 7.2. *Множина X^* задачі ЛП (7.11) опукла.*

◀ Нехай $x^{(1)*}, x^{(2)*} \in X^* (X^* \neq \emptyset)$, а $\alpha \in [0, 1]$. Як і раніше, нам необхідно встановити, що $\alpha x^{(1)*} + (1 - \alpha)x^{(2)*} = x^* \in X^*$.

Оскільки $x^* \in X (X^* \subset X)$, $\langle c, x^{(1)*} \rangle = \langle c, x^{(2)*} \rangle = f_{\max}^*$ і

$$\begin{aligned} \langle c, x^* \rangle &= \langle c, \alpha x^{(1)*} + (1 - \alpha)x^{(2)*} \rangle = \\ &= \alpha \langle c, x^{(1)*} \rangle + (1 - \alpha) \langle c, x^{(2)*} \rangle = \alpha f_{\max}^* + (1 - \alpha) f_{\max}^* = f_{\max}^*, \end{aligned}$$

то справді $x^* \in X^*$ і теорема 7.2 доведена. ▶

Непорожню множину допустимих розв'язків задачі ЛП ще називають допустимим многогранником (згідно з її геометричним змістом), який може бути обмеженою або необмеженою областю у просторі змінних, зокрема вироджуватись у точку, відрізок чи промінь.

Означення 7.1. Точка $\bar{x}$ називається крайньою (або кутовою) точкою опуклої множини $\bar{X}$, якщо не існує $x^{(1)}, x^{(2)} \in \bar{X} (x^{(1)} \neq x^{(2)})$ таких, що $\bar{x} = \alpha x^{(1)} + (1 - \alpha)x^{(2)}$, де $\alpha \in (0, 1)$.

Із означення 7.1 випливає, що $\bar{x}$ буде крайньою точкою множини $\bar{X}$, якщо вона не належить внутрішності жодного з відрізків, який би повністю належав множині $\bar{X}$. Крайні або кутові точки опуклого многогранника прийнято називати вершинами, що цілком узгоджується з відомим із геометрії поняттям вершини. Теорема 7.1 фактично стверджує, що допустима множина X задачі (7.11) є опуклим многогранником, тобто множиною, яка може мати крайні точки або вершини, які, як виявляється, відіграють принципово важливу роль при знаходженні розв'язків задачі ЛП. Залишається з'ясувати, коли точка x буде вершиною множини X .

Теорема 7.3. Щоб точка x була вершиною допустимої множини X задачі (7.11), необхідно й достатньо, щоб

вектори-стовпці матриці обмежень, що відповідають її додатним координатам, були лінійно незалежні.

◀ Спочатку доведемо необхідність. Нехай x – вершина допустимого многогранника X . Не порушуючи загальності міркувань, будемо вважати, що ненульовими, тобто додатними є її перших l координат, тобто

$$x_1 A_1 + \dots + x_l A_l = b, \quad (7.14)$$

але відповідні вектори-стовпці $A_1, \dots, A_l$ матриці обмежень лінійно залежні. Це означає, що існують такі l чисел $\gamma_1, \dots, \gamma_l$, серед яких є ненульові, що

$$\gamma_1 A_1 + \dots + \gamma_l A_l = 0. \quad (7.15)$$

Помножимо рівність (7.15) на число $\delta > 0$, після чого додамо і віднімемо її від (7.14). У результаті цього отримаємо співвідношення

$$(x_1 + \delta\gamma_1)A_1 + \dots + (x_l + \delta\gamma_l)A_l = b,$$

$$(x_1 - \delta\gamma_1)A_1 + \dots + (x_l - \delta\gamma_l)A_l = b,$$

з яких випливає, що вектори

$$x^{(1)} = (x_1 + \delta\gamma_1, \dots, x_l + \delta\gamma_l, 0, \dots, 0)^T,$$

$$x^{(2)} = (x_1 - \delta\gamma_1, \dots, x_l - \delta\gamma_l, 0, \dots, 0)^T$$

є розв'язками системи $Ax = b$, причому число δ можна так підібрати, щоб числа $(x_j \pm \delta\gamma_j)$, $j = \overline{1, l}$, були невід'ємними, тобто вектори $x^{(1)}$ і $x^{(2)}$ – допустимими $(x^{(1)}, x^{(2)} \in X)$. При цьому $x = \frac{1}{2}(x^{(1)} + x^{(2)}) = \frac{1}{2}x^{(1)} + \frac{1}{2}x^{(2)}$.

Остання рівність суперечить визначенню точки x як вершини (крайньої точки) допустимого многогранника X .

Достатність випливає з таких міркувань. Нехай $x = (x_1, \dots, x_l, 0, \dots, 0)^T$ – допустимий розв'язок задачі (7.11), у якого перших l координат додатні, а відповідні вектори-

стовпці $A_1, \dots, A_l$ лінійно незалежні. Покажемо, що x – вершина X .

Якщо це не так, то знайдуться відмінні від x допустимі точки $x^{(1)}$ і $x^{(2)}$ такі, що при деякому $\alpha \in (0,1)$ маємо $x = \alpha x^{(1)} + (1-\alpha)x^{(2)}$. Із цієї рівності дійдемо висновку, що для точок $x^{(1)}$ та $x^{(2)}$ (принаймні однієї з них) ненульовими будуть ті самі координати, що й у точці x . Вважаючи, як і раніше, що це будуть перші l координат, матимемо рівності

$$x_1^{(1)} A_1 + \dots + x_l^{(1)} A_l = b,$$

$$x_1^{(2)} A_1 + \dots + x_l^{(2)} A_l = b$$

або

$$(x_1^{(1)} - x_1^{(2)}) A_1 + \dots + (x_l^{(1)} - x_l^{(2)}) A_l = 0.$$

Оскільки серед чисел $(x_1^{(1)} - x_1^{(2)}), \dots, (x_l^{(1)} - x_l^{(2)})$ є ненульові ($x^{(1)} \neq x^{(2)}$), то це означає, що вектори $A_1, \dots, A_l$ лінійно залежні, але це суперечить вихідному припущенню щодо точки x , що й завершує доведення теореми. ►

Допустиму точку x , для якої вектори-стовпці матриці обмежень, що відповідають її додатним координатам, лінійно незалежні, тобто вершину допустимої множини X задачі ЛП (у даному випадку (7.11)), ще називають базисним невід’ємним або опорним розв’язком цієї задачі, про що йтиметься пізніше. Оскільки $\text{rg} A \leq \min(m, n)$ (ранг дорівнює максимальному числу лінійно незалежних векторів-стовпців чи векторів-рядків), то число лінійно незалежних векторів-стовпців системи обмежень не перевищує m і вершина має не більше, ніж m додатних координат.

Теорема 7.4. *Якщо допустима множина задачі ЛП (7.11) непорожня ($X \neq \emptyset$), то вона має скінченне число вершин.*

◀ Спочатку покажемо, що множина X має принаймні одну вершину. При цьому скористаємося методом математичної індукції по числу змінних n . Якщо $n=1$, то система обмежень $Ax=b$ збігається з одним рівнянням $a_{11}x_1=b_1$. Не порушуючи загальності міркувань, можна вважати, що $a_{11} > 0$, тому єдиний допустимий розв'язок $x_1 \geq 0$ за означенням буде вершиною, число додатних координат якого дорівнює одиниці при $b_1 > 0$ і менше від одиниці (тобто дорівнює нулеві) при $b_1 = 0$.

Нехай теорема виконується для всіх $r < n$. Покажемо її справедливість для n , припускаючи, що $x=(x_1, \dots, x_n)^T$ – невід'ємний розв'язок системи $Ax=b$, тобто $x \in X$. Логічно вважати, що всі координати цього розв'язку додатні, інакше якби хоч одна з координат була нульовою, то був би випадок $r = n-1 < n$ (кількість невідомих буде меншою від n), а для цього випадку за припущенням індукції теорема справедлива (аналогічно при $r = n-2$ і т.д.). Отже, залишилося тільки проаналізувати допустиму точку $x \in X$, всі координати якої додатні ($x_j > 0, j = \overline{1, n}$).

Якщо вектори $A_1, \dots, A_n$ лінійно незалежні, то згідно з теоремою 7.3 точка x є вершиною, бо її додатним координатам відповідають лінійно незалежні вектористовпці матриці обмежень.

Якщо вектори $A_1, \dots, A_n$ лінійно залежні, то має місце рівність

$$\gamma_1 A_1 + \dots + \gamma_n A_n = 0, \quad (7.16)$$

причому можна вважати, що хоча б одне із чисел γ_j ($j = \overline{1, n}$) додатне. Очевидно, від (7.16) можна перейти до системи

$$\frac{1}{\delta} \left(\frac{\gamma_1}{x_1} x_1 A_1 + \dots + \frac{\gamma_n}{x_n} x_n A_n \right) = 0, \quad (7.17)$$

а від системи $Ax = b$ до системи

$$\frac{1}{\delta} (\delta x_1 A_1 + \dots + \delta x_n A_n) = b, \quad (7.18)$$

де δ – найбільше з відношень (γ_i / x_i) , наприклад $\delta = \gamma_j / x_j (1 \leq j \leq n)$.

Віднявши (7.17) від (7.18), отримаємо рівність

$$\frac{1}{\delta} \left(\delta - \frac{\gamma_1}{x_1} \right) x_1 A_1 + \dots + \frac{1}{\delta} \left(\delta - \frac{\gamma_n}{x_n} \right) x_n A_n = b. \quad (7.19)$$

Один із коефіцієнтів у (7.19) дорівнює нулеві (згідно з вибором δ), а інші – невід'ємні, причому система (7.19) рівносильна системі $Ax = b$ (обидві мають своїм розв'язком вектор x).

Оскільки кількість невідомих у системі (7.19) дорівнює $(n-1)$ (а саме $x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n$), то за припущенням індукції множина X має принаймні одну вершину, тобто точку, додатним координатам якої відповідають лінійно незалежні вектори-стовпці матриці обмежень. Нехай такою вершиною буде

$$x^{(1)} = (x_1^{(1)}, \dots, x_{j-1}^{(1)}, x_{j+1}^{(1)}, \dots, x_r^{(1)}, x_{r+1}^{(1)}, 0, \dots, 0)^T.$$

Тоді точка

$$x^{(2)} = (x_1^{(1)}, \dots, x_{j-1}^{(1)}, 0, x_{j+1}^{(1)}, \dots, x_r^{(1)}, x_{r+1}^{(1)}, 0, \dots, 0)^T$$

є базисним невід'ємним розв'язком початкової системи $Ax = b$ (це впливає із (7.19), тобто вершиною допустимого многогранника X).

Отже, існування принаймні однієї вершини у X доведено. Враховуючи, що вершина може мати не більше

ніж m додатних координат ($m \leq n$), а число всіх координат n , можна стверджувати, що число вершин не перебільшує число комбінацій C_n^m , тобто скінченне. Теорема доведена. ►

Нескладно довести також, що у випадку, коли допустима множина $X \neq \emptyset$ і обмежена, то вона є опуклою лінійною комбінацією своїх вершин.

Якщо $X \neq \emptyset$, але необмежена, то не кожну точку $x \in X$ можна подати як опуклу лінійну комбінацію вершин.

Наступна теорема розкриває природу оптимальних розв'язків задач ЛП.

Теорема 7.5. Множина X^* оптимальних розв'язків задачі ЛП (7.11) обов'язково містить принаймні одну вершину x^* допустимого многогранника X .

◀ Доведемо теорему від супротивного. Припустимо, що є така задача (7.11), серед розв'язків якої немає вершин. Нехай $\tilde{x}$ – один із таких розв'язків. Згідно з теоремою 7.3, вектори-стовпці матриці обмежень, що відповідають додатним координатам вектора $\tilde{x}$, лінійно залежні. Тоді існує такий ненульовий вектор x , що його координати, індекси яких збігаються з індексами нульових координат вектора $\tilde{x}$, будуть нульовими, причому

$$Ax = 0.$$

А це означає, що існує число $\delta > 0$, для якого вектори $(\tilde{x} + \delta x)$ та $(\tilde{x} - \delta x)$ будуть допустимими ($A\tilde{x} = b$, $Ax = 0$, тому $A(\tilde{x} + \delta x) = b$ і $A(\tilde{x} - \delta x) = b$).

Оскільки $\tilde{x}$ – розв'язок задачі, то

$$\langle c, \tilde{x} + \delta x \rangle \leq \langle c, \tilde{x} \rangle,$$

$$\langle c, \tilde{x} - \delta x \rangle \leq \langle c, \tilde{x} \rangle,$$

що є можливим лише при $\langle c, x \rangle = 0$. Отже, для будь-яких δ , для яких точки $(\tilde{x} + \delta x)$ та $(\tilde{x} - \delta x)$ допустимі, вони ж будуть і оптимальними.

Вектор x має принаймні одну ненульову координату x_q . Нехай $x_q < 0$. Тоді з того, що $(\tilde{x}_q + \delta x_q) \geq 0$ маємо нерівність $\delta \leq \tilde{x}_q / |x_q|$ (таку ж умову для δ отримаємо із нерівності $\tilde{x}_q - \delta x_q \geq 0$, припустивши, що $x_q > 0$). Якщо $\delta = \tilde{x}_q / |x_q|$, то q -та координата точки $(\tilde{x} + \delta x)$ буде нульовою, тобто $(\tilde{x}_q + \delta x_q) = 0$, а при більшому δ вона буде уже від'ємною. Отже, рухаючись уздовж вектора x , прийдемо до допустимої точки $(\tilde{x} + \delta x)$, одна з координат якої, що відповідає додатній координаті точки $\tilde{x}$, у деякий момент стане нульовою.

З урахуванням вищесказаного одержана точка $(\tilde{x} + \delta x)$ буде не тільки допустимою, але й оптимальною, тобто розв'язком задачі. При цьому вона буде мати на одну додатну координату менше, ніж у $\tilde{x}$. Застосовуючи до цієї точки попередні міркування, прийдемо до розв'язку, число додатних координат якого менше від числа додатних координат точки $\tilde{x}$ на два і т.д. Якщо припущення про те, що серед розв'язків задачі (7.11) немає вершин, справджується, то описаний нами процес буде нескінченним. Але це суперечить тому, що у точки $\tilde{x}$ є скінченне число додатних координат. Це й доводить теорему 7.5. ►

Теорема 7.6. *Нехай у задачі (7.11) множина $X \neq \emptyset$, а цільова функція $f(x) = \langle c, x \rangle$ ($c \neq 0$) обмежена зверху на X . Тоді задача (7.11) має оптимальний розв'язок.*

◀ Нехай задача (7.11) не має оптимального розв'язку. Оскільки $X \neq \emptyset$, то цільова функція $f(x) = \langle c, x \rangle$ ($c \neq 0$) повинна бути необмеженою зверху на X . Але це суперечить твердженню про обмеженість зверху на X функції $f(x)$. Отже, задача (7.11) має оптимальний розв'язок. Теорема доведена. ►

Теорема 7.7. *Якщо множина X^* містить не менше двох вершин, то довільна опукла лінійна комбінація цих вершин теж буде оптимальним розв'язком задачі (7.11).*

◀ Нехай $x^{(1)*}, \dots, x^{(l)*} \in X^*$ – вершини (їх число завжди скінченне), причому

$$f_{\max}^* = \langle c, x^{(i)*} \rangle, \quad i = \overline{1, l}.$$

Тоді для довільної точки

$$x^* = \sum_{i=1}^l \alpha_i x^{(i)*}, \quad \sum_{i=1}^l \alpha_i = 1, \quad \alpha_i \geq 0, \quad i = \overline{1, l},$$

справджуються співвідношення

$$\left\langle c, \sum_{i=1}^l \alpha_i x^{(i)*} \right\rangle = \sum_{i=1}^l \alpha_i \langle c, x^{(i)*} \rangle = \sum_{i=1}^l \alpha_i f_{\max}^* = f_{\max}^*,$$

що й треба було встановити. ►

Відзначимо, що аналогічні теореми можна сформулювати і для стандартної задачі ЛП (якщо йдеться про задачу на максимум, то відповідною до (7.11) буде задача (7.10)).

7.4. Графічний метод розв'язування задач лінійного програмування

Графічний метод знаходження розв'язків задач ЛП базується на їх геометричній інтерпретації і принципово може бути використаний тільки для розв'язування тих задач, для яких число змінних не перевищує трьох, або канонічних задач з багатьма змінними, якщо останні мають не більше трьох вільних змінних. Якщо вважати, що n – число змінних, а m – число лінійно незалежних рівнянь у канонічній задачі ЛП, то, очевидно, умови застосування графічного методу такі: $n \leq 3$ або $n - m \leq 3$ (для канонічних задач).

Нижче розглядаються тільки випадки $n \leq 2$ або $n - m \leq 2$, оскільки при $n = 3$ або $n - m = 3$ далеко не просто вдається геометрично відтворити відповідну задачу ЛП.

Отже, розглянемо задачу ЛП вигляду:

$$\left\{ \begin{array}{l} c_1 x_1 + c_2 x_2 \rightarrow \max, \\ a_{11} x_1 + a_{12} x_2 \leq b_1, \\ \dots\dots\dots \\ a_{k1} x_1 + a_{k2} x_2 \leq b_k, \\ a_{k+1,1} x_1 + a_{k+1,2} x_2 \geq b_{k+1}, \\ \dots\dots\dots \\ a_{m1} x_1 + a_{m2} x_2 \geq b_m, \\ x_j \geq 0, \quad j = \overline{1, 2}. \end{array} \right. \quad (7.20)$$

Допустиму область задачі (7.20) позначимо символом $\tilde{X}$. Процес знаходження розв'язку задачі (7.20) графічним методом включає в себе такі етапи:

1) побудова на площині координат $x_1 O x_2$ допустимої множини (допустимого многокутника) $\tilde{X}$ задачі (7.20)

(якщо $\tilde{X} = \emptyset$, то задача (7.20) не має роз'язку, тому його пошук припиняється уже на першому етапі);

2) побудова на площині координат $x_1 O x_2$ вектора $c = (c_1, c_2)^T$ (початкову точку вектора можна розмістити в початку координат) і перпендикулярно до нього однієї з ліній рівня цільової функції, тобто лінії $c_1 x_1 + c_2 x_2 = \text{const}$ (наприклад, лінії $c_1 x_1 + c_2 x_2 = 0$);

3) переміщення побудованої лінії рівня у напрямку вектора c і встановлення опорної лінії рівня (опорна лінія – це лінія, по один бік від якої знаходиться множина $\tilde{X}$, причому принаймні одна точка із $\tilde{X}$ належить цій лінії), якщо вона існує (тоді розв'язками задачі (7.20) будуть точки $x^* \in \tilde{X}$, які належать опорній лінії), або встановлення факту неіснування опорної лінії (тоді задача (7.20) не має роз'язку, оскільки цільова функція є необмеженою на $\tilde{X}$).

Зауваження 7.1. Вектор c у задачі ЛП є градієнтом лінійної форми, тобто функції $f(x) = \langle c, x \rangle$. З математичного аналізу відомо, що неперервно диференційовна функція $f(x)$ у кожній точці x з області визначення зростає в напрямку градієнта функції $f(x)$, тобто у напрямку вектора $f'(x) \equiv \text{grad} f(x)$ і спадає в напрямку антиградієнта функції $f(x)$, тобто у напрямку вектора $(-f'(x)) \equiv (-\text{grad} f(x))$. Цим і пояснюється, що при розв'язуванні задачі ЛП на максимум лінію рівня переміщують у напрямку вектора c , а при розв'язуванні задачі ЛП на мінімум – у напрямку вектора $(-c)$.

Зауваження 7.2. При розв'язуванні задачі ЛП (7.20) графічним методом можливі такі ситуації.

1) Опорній лінії належить одна точка $x^* \in \tilde{X}$. Це означає, що задача (7.20) має єдиний розв'язок.

2) Опорній лінії належить нескінченна множина точок $x^* \in \tilde{X}$, яка збігається з однією зі сторін допустимого многокутника розв'язків. У цьому випадку задача ЛП (7.20) має нескінченну множину розв'язків.

3) $\tilde{X} = \emptyset$. Задача ЛП (7.20) не має розв'язків.

4) Опорної лінії не існує. Задача (7.20) не має розв'язків, оскільки цільова функція необмежена зверху на $\tilde{X}$.

Зауваження 7.3. Якщо при розв'язуванні задач ЛП на мінімум опорної лінії не існує, то задача ЛП не має розв'язків, оскільки цільова функція необмежена знизу на $\tilde{X}$.

Для підтвердження цих випадків розглянемо ряд прикладів (7.1–7.5) розв'язування задач ЛП графічним методом.

Приклад 7.1.

$$\begin{cases} 2x_1 + x_2 \rightarrow \max, \\ x_1 + x_2 \leq 12, \\ x_1 - x_2 \leq 4, \\ x_j \geq 0, \quad j = \overline{1,2}. \end{cases} \quad (7.21)$$

◀ Геометрична ілюстрація процесу розв'язування задачі (7.21) графічним методом дана на рис. 7.1.

Чотирикутник $OBCD$ є областю допустимих розв'язків $\tilde{X}^{(1)}$ задачі (7.21). Переміщуючи лінію рівня $2x_1 + x_2 = 0$ у напрямку вектора $c = (2,1)^T$, прийдемо до опорної лінії (на рис. 7.1 вона позначена пунктиром), якій належить точка C . Очевидно, що точка C буде крайньою (вершиною) для $\tilde{X}^{(1)}$. Для знаходження координат точки C необхідно розв'язати систему рівнянь

$$\begin{cases} x_1 + x_2 = 12, \\ x_1 - x_2 = 4. \end{cases}$$

Зрозуміло, що $x_1^* = 8$, $x_2^* = 4$, $f_{\max}^* = 2 \cdot 8 + 4 \cdot 1 = 20$. Це і є шуканий розв'язок задачі (7.21). ►

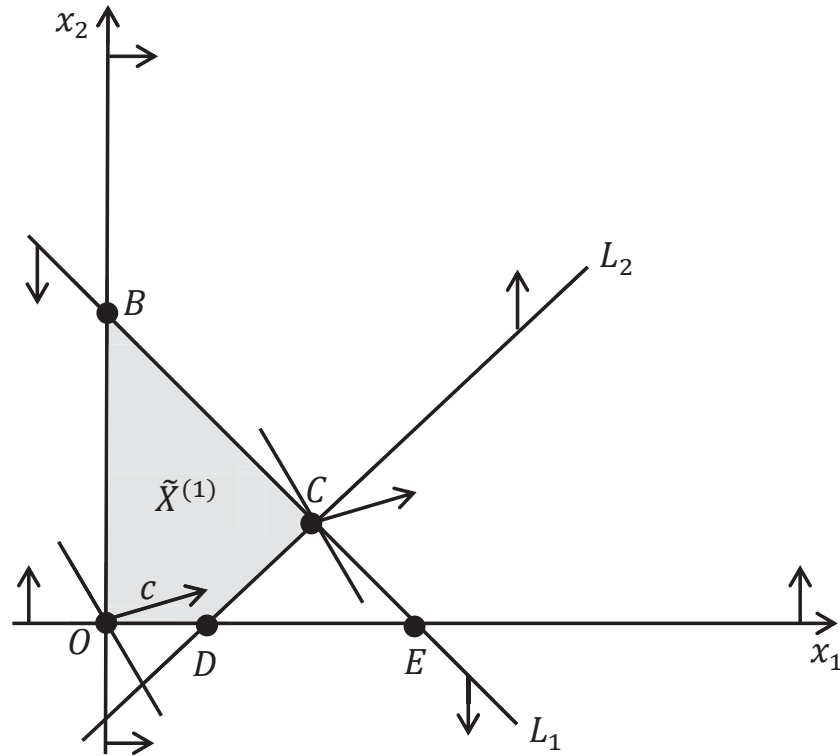


Рис. 7.1. Ілюстрація до прикладу 7.1 (точки O, B, C, D, E , лінії L_1, L_2 та вектор c такі: $O(0,0)$, $B(0,12)$, $C(8,4)$, $D(4,0)$, $E(12,0)$, $L_1: x_1 + x_2 = 12$, $L_2: x_1 - x_2 = 4$, $c = (2,1)^T$)

Приклад 7.2.

$$\begin{cases} 3x_1 + 3x_2 \rightarrow \max, \\ x_1 + x_2 \leq 2, \\ x_j \geq 0, \quad j = \overline{1,2}. \end{cases} \quad (7.22)$$

◀ Геометричне зображення задачі (7.22) дано на рис. 7.2. Очевидно, що область допустимих розв'язків

$$\tilde{X}^{(2)} = \{x = (x_1, x_2)^T \in \mathbb{R}^2 / x_1 + x_2 \leq 2, \quad x_j \geq 0, j = \overline{1,2}\}$$

збігається з трикутником OAB . Довільна лінія рівня

$$3x_1 + 3x_2 = \text{const}$$

паралельна прямій L .

Це означає, що опорна лінія рівня буде проходити через сторону AB трикутника OAB . Отже, множиною розв'язків задачі (7.22) є множина точок відрізка AB . Координати x_1^* і x_2^* довільної точки M , розміщеної на AB , визначаються як опуклі комбінації відповідних координат точок A та B :

$$x_1^* = \alpha \cdot 0 + (1 - \alpha) \cdot 2 = 2(1 - \alpha),$$

$$x_2^* = \alpha \cdot 2 + (1 - \alpha) \cdot 0 = 2\alpha, \alpha \in [0, 1].$$

Максимальне значення $f_{\max}^* = 3x_1^* + x_2^* = 3 \cdot 2 = 6$. ►

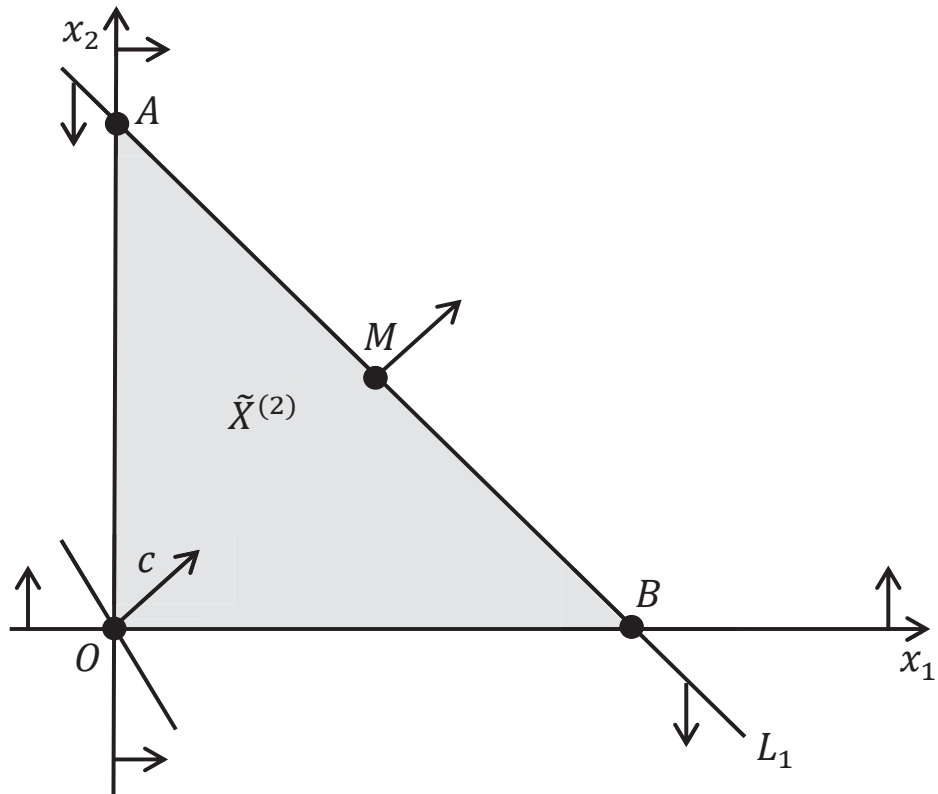


Рис. 7.2. Ілюстрація до прикладу 7.2 (точки O, A, B, M , лінія L_1 та вектор c такі: $O(0,0)$, $A(0,2)$, $B(2,0)$, $M(x_1^*, x_2^*)$, $c = (3,3)^T$)

Приклад 7.3.

$$\begin{cases} x_1 + 5x_2 \rightarrow \min, \\ x_1 + x_2 \leq 1, \\ x_1 + 2x_2 \geq 3, \\ x_j \geq 0, \quad j = \overline{1,2}. \end{cases} \quad (7.23)$$

◀ Легко побачити (рис. 7.3), що множина допустимих розв'язків задачі (7.23)

$$\tilde{X}^{(3)} = \{x = (x_1, x_2)^T \in \mathbb{R}^2 / x_1 + x_2 \leq 1,$$

$$x_1 + 2x_2 \geq 3, \quad x_j \geq 0, \quad j = \overline{1,2}\} = \emptyset.$$

Задача не має розв'язку. ►

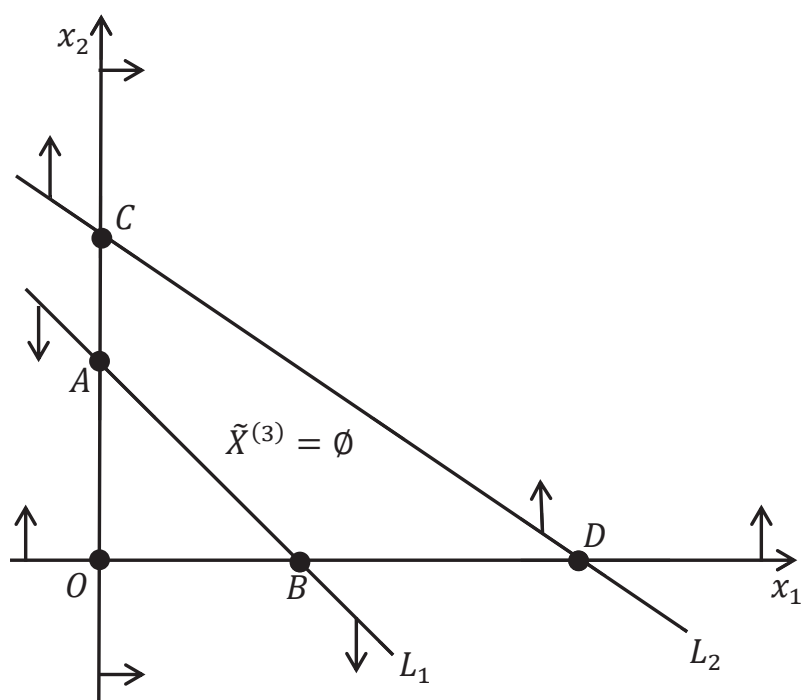


Рис. 7.3. Ілюстрація до прикладу 7.3 (точки O, A, B, C, D та лінії L_1, L_2 такі: $O(0,0)$, $A(0,1)$, $B(1,0)$, $C(0, \frac{3}{2})$, $D(3,0)$,
 $L_1 : x_1 + x_2 = 1$, $L_2 : x_1 + 2x_2 = 3$)

Приклад 7.4.

$$\begin{cases} 2x_1 + x_2 \rightarrow \max, \\ x_1 + x_2 \geq 1, \\ x_1 - x_2 \geq -1, \\ x_j \geq 0, \quad j = \overline{1,2}. \end{cases} \quad (7.24)$$

◀ Геометрична ілюстрація задачі (7.24) дана на рис. 7.4.

Нескладно переконатися, що, переміщуючи лінію рівня $2x_1 + x_2 = \text{const}$ уздовж вектора $c = (2,1)^T$, ми ніколи не

досягнемо опорного положення. Це означає, що цільова функція $f(x) = 2x_1 + x_2$ необмежена зверху на множині

$$\tilde{X}^{(4)} = \{x = (x_1, x_2)^T \in \mathbb{R}^2 / x_1 + x_2 \geq 1, x_1 - x_2 \geq -1, x_j \geq 0, j = \overline{1, 2}\},$$

тому задача (7.24) не має розв'язку. ►

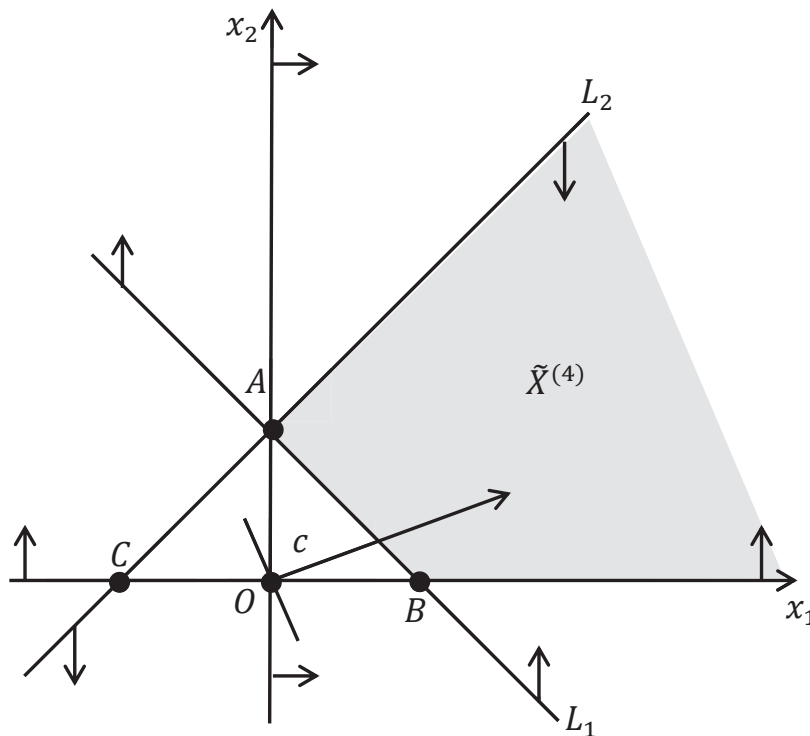


Рис. 7.4. Ілюстрація до прикладу 7.4 (точки O, A, B, C , лінії L_1, L_2 та вектор c такі: $O(0,0)$, $A(0,1)$, $B(1,0)$, $C(-1,0)$,

$$L_1 : x_1 + x_2 = 1, L_2 : x_1 - x_2 = -1, c = (2,1)^T)$$

Приклад 7.5.

$$\begin{cases} x_1 + x_2 + 2x_3 - x_4 \rightarrow \max, \\ x_1 - x_2 + x_3 - x_4 + x_5 = 4, \\ x_2 + x_3 + x_4 + x_5 = 3, \\ x_3 - x_4 - x_5 = 2, \\ x_j \geq 0, j = \overline{1, 5}. \end{cases} \quad (7.25)$$

◀ Виключимо невідому x_3 з першого та другого рівнянь (із третього рівняння визначимо x_3 через x_4 і x_5 , а потім підставимо отриманий вираз замість x_3 у перші два рівняння, після чого вони відповідно матимуть вигляд: $x_1 - x_2 + 2x_5 = 2$, $x_2 + 2x_4 + 2x_5 = 1$), а потім невідому x_2 – із отриманого у результаті першої операції першого рівняння (із другого рівняння визначаємо $x_2 = 1 - 2x_4 - 2x_5$ і підставляємо у перше рівняння $x_1 - x_2 + 2x_5 = 2$). Тоді від задачі (7.25) прийдемо до еквівалентної задачі

$$\begin{cases} x_1 + x_2 + 2x_3 - x_4 \rightarrow \max, \\ x_1 + 2x_4 + 4x_5 = 3, \\ x_2 + 2x_4 + 2x_5 = 1, \\ x_3 - x_4 - x_5 = 2, \\ x_j \geq 0, \quad j = \overline{1,5}. \end{cases} \quad (7.26)$$

Система рівнянь-обмежень у задачі (7.26) має три базисних (x_1, x_2, x_3) і дві вільних (x_4, x_5) змінних. У зв'язку з цим її можна записати так:

$$\begin{cases} x_1 = 3 - 2x_4 - 4x_5, \\ x_2 = 1 - 2x_4 - 2x_5, \\ x_3 = 2 + x_4 + x_5. \end{cases} \quad (7.27)$$

Враховуючи (7.27), цільову функцію задачі (7.26) можна записати через змінні x_4 і x_5 , а саме:

$$\begin{aligned} f(x_1, x_2, x_3, x_4, x_5) &= x_1 + x_2 + 2x_3 - x_4 = 3 - 2x_4 - 4x_5 + \\ &+ 1 - 2x_4 - 2x_5 + 2(2 + x_4 + x_5) - x_4 = 8 - 3x_4 - 4x_5. \end{aligned} \quad (7.28)$$

Співвідношення (7.27) і (7.28), а також умови невід'ємності змінних $x_j \quad j = \overline{1,5}$ дають можливість перейти від задачі (7.26) до задачі

$$\begin{cases} -3x_4 - 4x_5 + 8 \rightarrow \max, \\ 2x_4 + 4x_5 \leq 3, \\ 2x_4 + 2x_5 \leq 1, \\ x_4 + x_5 \geq -2, \\ x_4 \geq 0, x_5 \geq 0. \end{cases} \quad (7.29)$$

Розв'яжемо (7.29) графічним методом (рис. 7.5).

Областю допустимих розв'язків $\tilde{X}^{(5)}$ для задачі (7.29) є трикутник OCD . Очевидно, що $x_4^* = 0$, $x_5^* = 0$, $f_{\max}^* = -3 \cdot 0 - 4 \cdot 0 + 8 = 8$. За допомогою (7.27) знайдемо оптимальні значення $x_1^* = 3$, $x_2^* = 1$, $x_3^* = 2$.

Отже, розв'язок задачі (7.25) такий:

$$x^* = (3, 1, 2, 0, 0)^T, f_{\max}^* = 8. \blacktriangleright$$

Зауваження 7.4. Приклад 7.5 підтверджує, що в деяких указаних раніше випадках канонічна задача ЛП може бути розв'язана графічним методом. Процедура зведення таких канонічних задач до задач ЛП з двома змінними досить проста. У зв'язку з цим основний акцент потрібно зробити на графічному розв'язуванні задач типу (7.21)–(7.24).

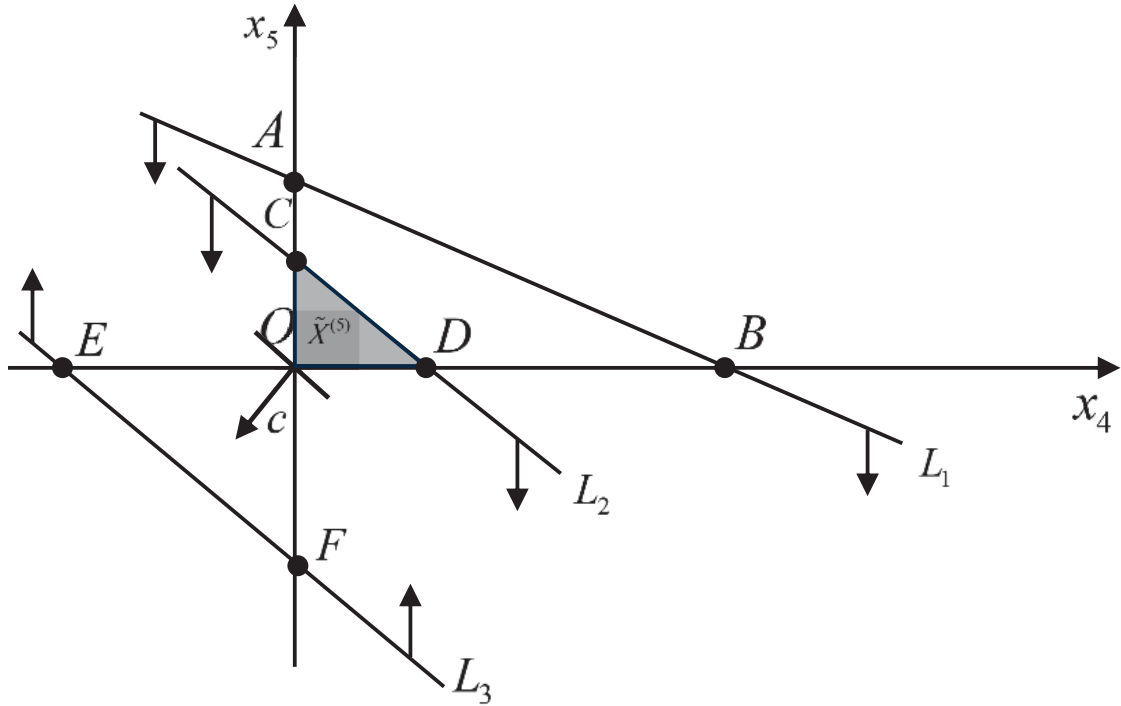


Рис. 7.5. Ілюстрація до прикладу 7.5 (точки O, A, B, C, D, E, F ,

лінії L_1, L_2, L_3 та вектор c такі: $O(0,0)$, $A\left(0, \frac{3}{4}\right)$, $B\left(\frac{3}{2}, 0\right)$,

$C\left(0, \frac{1}{2}\right)$, $D\left(\frac{1}{2}, 0\right)$, $E(-2, 0)$, $F(0, -2)$, $L_1: 2x_4 + 4x_5 = 3$,

$L_2: 2x_4 + 2x_5 = 1$, $L_3: x_4 + x_5 = -2$, $c = (-3, -4)^T$)

7.5. Симплексний метод розв'язування задач лінійного програмування

Як випливає з пункту 7.4, алгоритм геометричного розв'язування задач ЛП має загальний характер, тому теоретично його застосування можливе в довільному просторі $\mathbb{R}^n$. Однак клас задач ЛП, які вдається практично розв'язати графічним методом, дуже обмежений. Для побудови алгебраїчних методів розв'язування задач ЛП потрібно з'ясувати перш за все природу допустимих розв'язків.

У зв'язку з цим надалі розглянемо канонічну задачу ЛП (7.11). Оскільки допустима множина X задачі (7.11) складається з невід'ємних розв'язків систем лінійних алгебраїчних рівнянь (СЛАР), то першочерговим завданням є вивчення СЛАР.

7.5.1. Базисні та опорні розв'язки систем лінійних алгебраїчних рівнянь

Метод виключення Жордана-Гаусса для систем лінійних алгебраїчних рівнянь. Нехай маємо систему m лінійних рівнянь з n невідомими (m і n довільні)

[illegible]

Розв'язком системи (7.30) називають сукупність дійсних чисел (вектор) $x = (x_1, x_2, \dots, x_n)^T$, які після підстановки їх у (7.30) перетворюють рівняння системи в тотожності. Якщо система (7.30) не має жодного розв'язку, то вона називається несумісною, у протилежному випадку – сумісною. Кажуть, що сумісна система визначена, якщо вона має тільки один розв'язок, і невизначена, якщо число розв'язків більше від одиниці.

У теорії СЛАР фундаментальне місце належить теоремі Кронекера-Капеллі: система (7.30) сумісна тоді й тільки тоді, якщо ранг матриці A цієї системи дорівнює рангу її розширеної матриці $\tilde{A}$, тобто якщо $rg A = rg \tilde{A}$.

Матриці A і $\tilde{A}$ мають вигляд:

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}, \quad \tilde{A} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \dots & \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{pmatrix}.$$

Під рангом матриці розуміють найбільший порядок мінора цієї матриці, відмінного від нуля (мінор k -го порядку ($k \leq m, k \leq n$) – це визначник k -го порядку, складений з елементів матриці, розміщених на перетині виділених k рядків і k стовпчиків). Якщо ранг сумісної системи дорівнює числу невідомих n , то система визначена. Якщо ранг сумісної системи менший від числа невідомих, то система – невизначена. Надалі особливе значення для нас матимуть саме невизначені СЛАР.

Одним із найбільш ефективних методів дослідження та розв'язування СЛАР є метод виключення Жордана-Гаусса. Коротко про сутність цього методу. У матриці A системи (7.30) виберемо елемент $a_{rs} \neq 0$. Назвемо цей елемент ведучим (зауважимо, що в літературі такий елемент ще називають напрямним, розв'язувальним тощо), а r -й рядок і s -й-й стовпчик матриці A – відповідно ведучими рядком і стовпчиком. Далі перетворимо систему (7.30) до еквівалентного вигляду

[illegible]

де елементи a'_{ij} ($i = \overline{1, m}; j = \overline{1, n}$) матриці A та вільні члени b'_i ($i = \overline{1, m}$) обчислюються за формулами (так зване «правило прямокутників»):

$$b'_i = \begin{cases} b_i / a_{rs}, & i = r, \\ b_i - (b_r a_{is}) / a_{rs} & i \neq r. \end{cases} \quad (7.33)$$

Процес розв'язування системи (7.30) методом виключення Жордана-Гаусса (методом послідовного виключення невідомих) дає однозначну відповідь на запитання: чи має система (7.30) розв'язок, а якщо має, то чи він єдиний? Це дослідження системи зводиться до аналізу таких випадків.

$$\begin{cases} x_1 = d_1, \\ x_2 = d_2, \\ \dots\dots\dots \\ x_n = d_n. \end{cases} \quad (7.34)$$

2) Система (7.30) перетворена до вигляду

[illegible]

де $\text{rg}A = k \leq \min(m, n)$ (типова ситуація для випадку $m < n$). Це означає, що система (7.30) невизначена, тобто має множину розв'язків. При цьому невідомі $x_{i_1}, x_{i_2}, \dots, x_{i_k}$ називаються базисними, а невідомі $x_{i_{k+1}}, \dots, x_{i_n}$ – вільними. При конкретних значеннях вільних змінних отримуємо частинні розв'язки системи (7.30). Загальний розв'язок формалізується співвідношеннями (7.35).

3) На одній із проміжних або на кінцевій ітерації система (7.30) приведена до еквівалентної системи, в якій зустрічається рядок, де всі елементи нульові, за винятком вільного члена. Це означає, що система (7.30) несумісна, тобто не має розв'язку. Якщо зустрічаються два рядки, які відрізняються тільки вільними членами, то це теж ознака несумісності системи (7.30).

Відзначимо також, що за допомогою методу виключення Жордана-Гаусса легко знаходиться ранг вихідної матриці A : він дорівнює числу різних рівнянь кінцевої системи ((7.34), (7.35)).

Продемонструємо на конкретному прикладі застосування методу виключення Жордана-Гаусса для розв'язування СЛАР.

Приклад 7.6.

$$\begin{cases} x_1 + 2x_2 + 3x_3 + x_4 = 1, \\ 3x_1 + 13x_2 + 13x_3 + 5x_4 = 3, \\ x_1 + 5x_2 + 3x_3 + x_4 = 7, \\ 3x_1 + 7x_2 + 7x_3 + 2x_4 = 12, \\ 4x_1 + 5x_2 + 6x_3 + x_4 = 19. \end{cases} \quad (7.36)$$

◀ Результати обчислень наведені у таблиці 7.1. У ній через $\tilde{A}^{(0)}, \tilde{A}^{(1)}, \tilde{A}^{(2)}, \tilde{A}^{(3)}$ позначені розширені матриці відповідно системи (7.36) та еквівалентних до (7.36) систем, отриманих після 1-ї, 2-ї і 3-ї ітерації. Для кожної з указаних

матриць змінна x_j ($j = \overline{1,4}$) відповідає j -ому стовпчику, а змінна x_0 – стовпчику вільних членів.

Матриця $\tilde{A}^{(1)}$ формується так. У матриці $\tilde{A}^{(0)}$ виберемо ведучий елемент $a_{11}^{(0)} = 1$, після чого розділимо елементи ведучого рядка ($r=1$) на $a_{11}^{(0)}$ і отримаємо r -ий (перший) рядок матриці $\tilde{A}^{(1)}$. Елементи першого стовпчика $\tilde{A}^{(1)}$ нульові, за винятком першого $a_{11}^{(1)}$, який відповідає ведучому елементу $a_{rs}^{(0)}$ ($r=1, s=1$) матриці $\tilde{A}^{(0)}$. Інші елементи матриці $\tilde{A}^{(1)}$ обчислюються за формулами (7.32), (7.33) і набувають таких числових значень:

$$\begin{aligned} a_{22}^{(1)} &= 13 - \frac{3 \cdot 2}{1} = 7, \quad a_{23}^{(1)} = 13 - \frac{3 \cdot 3}{1} = 4, \quad a_{24}^{(1)} = 5 - \frac{3 \cdot 1}{1} = 2, \\ b_2^{(1)} &= 3 - \frac{3 \cdot 1}{1} = 0, \quad a_{32}^{(1)} = 5 - \frac{1 \cdot 2}{1} = 3, \quad a_{33}^{(1)} = 3 - \frac{1 \cdot 3}{1} = 0, \\ a_{34}^{(1)} &= 1 - \frac{1 \cdot 1}{1} = 0, \quad b_3^{(1)} = 7 - \frac{1 \cdot 1}{1} = 6, \quad a_{42}^{(1)} = 7 - \frac{3 \cdot 2}{1} = 1, \\ a_{43}^{(1)} &= 7 - \frac{3 \cdot 3}{1} = -2, \quad a_{44}^{(1)} = 2 - \frac{3 \cdot 1}{1} = -1, \quad b_4^{(1)} = 12 - \frac{3 \cdot 1}{1} = 9, \\ a_{52}^{(1)} &= 5 - \frac{4 \cdot 2}{1} = -3, \quad a_{53}^{(1)} = 6 - \frac{4 \cdot 3}{1} = -6, \quad a_{54}^{(1)} = 1 - \frac{4 \cdot 1}{1} = -3, \\ b_4^{(1)} &= 19 - \frac{4 \cdot 1}{1} = 15. \end{aligned}$$

При формуванні матриці $\tilde{A}^{(2)}$ за ведучий елемент матриці $\tilde{A}^{(1)}$ вибрано $a_{42}^{(1)} = 1$, а матриці $\tilde{A}^{(3)}$ – елемент $a_{53}^{(2)} = -12$ матриці $\tilde{A}^{(2)}$. У результаті від матриці $\tilde{A}^{(0)}$ ми прийшли до матриці $\tilde{A}^{(3)}$ тобто система (7.36) звелась до еквівалентного вигляду:

$$\left\{ \begin{array}{l} x_1 - \frac{1}{2}x_4 = \frac{15}{2}, \\ 0 = 0, \\ 0 = 0, \\ x_2 = 2, \\ x_3 + \frac{1}{2}x_4 = -\frac{7}{2}. \end{array} \right. \quad (7.37)$$

Два рівняння системи (7.37) (2-ге і 3-тє) перетворилися у тотожності ($0 \equiv 0$), тому на розв'язність системи вони не впливають. Змістовними є тільки інші три рівняння. Система є невизначеною. Одна змінна буде вільною (наприклад, змінна x_4). У зв'язку з цим загальний розв'язок системи (7.37) можна записати так:

$$x_1 = \frac{15}{2} + \frac{1}{2}x_4, \quad x_2 = 2, \quad x_3 = -\frac{7}{2} - \frac{1}{2}x_4. \quad \blacktriangleright$$

Таблиця 7.1

Матриці	x_1	x_2	x_3	x_4	x_0
$\tilde{A}^{(0)}$	1	2	3	1	1
	3	13	13	5	3
	1	5	3	1	7
	3	7	7	2	12
	4	5	6	1	19
$\tilde{A}^{(1)}$	1	2	3	1	1
	0	7	4	2	0
	0	3	0	0	6
	0	1	-2	-1	9
	0	-3	-6	-3	15
$\tilde{A}^{(2)}$	1	0	7	3	-17
	0	0	18	9	-63
	0	0	6	3	-21
	0	1	-2	-1	9
	0	0	-12	-6	42
$\tilde{A}^{(3)}$	1	0	0	$-1/2$	$15/2$
	0	0	0	0	0
	0	0	0	0	0
	0	1	0	0	2
	0	0	1	$1/2$	$-7/2$

Базисні та опорні розв'язки систем лінійних алгебраїчних рівнянь. Розглянемо знову систему (7.30). Надалі нас буде цікавити випадок $m < n$. Не порушуючи загальності міркувань,

можна вважати, що система (7.30) методом виключення Жордана-Гаусса приведена до вигляду (7.35), причому $k = m$. Якщо всі вільні змінні набувають нульових значень (тобто $x_{i_{m+1}} = \dots = x_{i_n} = 0$), то базисні змінні визначаються так:

$$x_{i_1} = b_{i_1}'', \dots, x_{i_m} = b_{i_m}''.$$

Отриманий частинний розв'язок системи (7.30) називається базисним. Число базисних розв'язків не перевищує

$$C_n^m = \frac{n!}{m!(n-m)!}.$$

Перебравши всі комбінації з m базисних і $(n-m)$ вільних змінних, ми знайдемо всі можливі базисні розв'язки системи (7.30).

Приклад 7.7.

$$\begin{cases} 3x_1 + x_2 - 9x_3 = 4, \\ x_1 - 3x_3 + x_4 = 2. \end{cases} \quad (7.38)$$

◀ Оскільки $n = 4$, $m = 2$, $C_4^2 = 6$, то всі можливі базисні розв'язки можуть бути такими: $(x_1, x_2, 0, 0)^T$, $(x_1, 0, x_3, 0)^T$, $(x_1, 0, 0, x_4)^T$, $(0, x_2, x_3, 0)^T$, $(0, x_2, 0, x_4)^T$, $(0, 0, x_3, x_4)^T$. Обчислювальний процес знаходження всіх базисних розв'язків зафіксований у таблиці 7.2, де, як і раніше, змінна x_j ($j = \overline{1, 4}$) відповідає j -ому стовпчику, а змінна x_0 – стовпчику вільних членів. Для системи (7.38) базисними змінними є x_2 і x_4 , вільними – x_1 і x_3 . При $x_1 = x_3 = 0$ базисні змінні $x_2 = 4$, $x_4 = 2$, тобто $x^{(6,0)} = (0, 4, 0, 2)^T$. Далі вибирається ведучий елемент $a_{21}^{(0)} = 1$ і здійснюються перетворення Жордана-Гаусса, в результаті чого змінна x_1 вводиться в базис замість змінної x_4 . Знайдений наступний базисний розв'язок $x^{(6,1)} = (2, -2, 0, 0)^T$. Пошук інших базисних

розв'язків відбувається аналогічно. Усі вони виписані в таблиці 7.2. Зауважимо тільки, що змінні x_1 і x_3 не можуть бути базисними, оскільки відповідні їм перший і третій стовпчики матриці $A = A^{(0)}$ системи (7.38) лінійно залежні (табл. 7.2). ►

Отже, метод виключення Жордана-Гаусса дозволяє перетворити вихідну СЛАР до більш зручного еквівалентного вигляду. При цьому отримана кінцева система рівнянь дає явний опис множини всіх розв'язків, якщо вони існують. Із цієї множини розв'язків легко виділити базисні розв'язки, присвоївши вільним змінним нульові значення.

Алгоритм знаходження розв'язків (у тому числі базисних) СЛАР за допомогою методу виключення Жордана-Гаусса передбачає такі два етапи:

- 1) вибір відмінного від нуля ведучого елемента у поточній розширеній матриці системи;*
- 2) зведення поточної розширеної матриці системи до еквівалентної матриці за допомогою правила прямокутників.*

Виконання етапів 1), 2) здійснюється до тих пір, поки не будуть знайдені шукані розв'язки системи або встановлена її несумісність.

Тепер уточнимо поняття опорного розв'язку системи (7.30). *Опорними, або допустимими базисними розв'язками системи лінійних алгебраїчних рівнянь називають її невід'ємні базисні розв'язки.*

Попередній приклад переконує, що далеко не всі базисні розв'язки є опорними. Очевидно, що опорні розв'язки можна виділити, коли знайдені всі базисні розв'язки, але в загальному випадку такий шлях складний.

Таблиця 7.2

Номер ітерації	Базисні змінні	x_1	x_2	x_3	x_4	x_0	Базисні роз'язки
0	x_2	3	1	-9	0	4	$x^{(\delta,0)} = (0, 4, 0, 2)^T$
	x_4	1	0	-3	1	2	
1	x_2	0	1	0	-3	-2	$x^{(\delta,1)} = (2, -2, 0, 0)^T$
	x_1	1	0	-3	1	2	
2	x_4	0	-1/3	0	1	2/3	$x^{(\delta,2)} = (4/3, 0, 0, 2/3)^T$
	x_1	1	1/3	-3	0	4/3	
3	x_4	0	-1/3	0	1	2/3	$x^{(\delta,3)} = (0, 0, -4/9, 2/3)^T$
	x_3	-1/3	-1/9	1	0	-4/9	
4	x_2	0	1	0	-3	-2	$x^{(\delta,4)} = (0, -2, -2/3, 0)^T$
	x_3	-1/3	0	1	-1/3	-2/3	

Однак, якщо вибирати ведучий елемент так, щоб він задовольняв деякі додаткові обмеження (назвемо їх симплексними перетвореннями), то ті ж перетворення Жордана-Гаусса забезпечать нам знаходження тільки опорних розв'язків.

Ідеться про такі додаткові обмеження:

а) *ведучий s -й стовпчик вибирається такий, у якому є принаймні один додатний елемент;*

б) *ведучий r -й рядок вибирається такий, для якого виконується умова*

$$\frac{\tilde{b}_r}{\tilde{a}_{rs}} = \min_{\{i: \tilde{a}_{is} > 0\}} \left(\frac{\tilde{b}_i}{\tilde{a}_{is}} \right), \quad (7.39)$$

де $\tilde{a}_{is}, \tilde{a}_{rs}$ – додатні елементи s -го стовпчика, а $\tilde{b}_i, \tilde{b}_r$ – відповідні елементи стовпчика вільних членів системи (7.30) або еквівалентної до неї системи, отриманої на тому чи іншому кроці в результаті перетворень Жордана-Гаусса.

Отже, алгоритм знаходження опорних розв'язків СЛАР за допомогою методу виключення Жордана-Гаусса відрізняється від попереднього алгоритму знаходження довільних розв'язків СЛАР лише вибором ведучого елемента, тобто його основні етапи такі:

1) вибір додатного ведучого елемента у поточній розширеній матриці системи з урахуванням додаткових обмежень а) та б);

2) зведення поточної розширеної матриці системи до еквівалентної матриці за допомогою правила прямокутників.

Теорема 7.8. *Система (7.30) з невід'ємними вільними членами за допомогою симплексних перетворень завжди зводиться до еквівалентної системи з невід'ємними вільними членами.*

◀ Нехай у системі (7.30) всі $b_i \geq 0, i = \overline{1, m}$. Виберемо ведучий елемент $a_{rs} > 0$ і, користуючись формулою (7.33), обчислимо вільні члени $\tilde{b}_i$ ($i = \overline{1, m}$) еквівалентної системи, яку отримаємо за допомогою симплексних перетворень. Очевидно, $\tilde{b}_r = b_r / a_{rs} \geq 0$. Покажемо, що $\tilde{b}_i \geq 0, i = \overline{1, m} (i \neq r)$.

Справді, для $a_{is} < 0$ за формулою (7.33)

$$\tilde{b}_i = b_i - b_r a_{is} / a_{rs} = b_i + b_r (-a_{is}) / a_{rs} \geq 0.$$

Якщо $a_{is} > 0$, то

$$\tilde{b}_i = b_i - b_r a_{is} / a_{rs} = a_{is} (b_i / a_{is} - b_r / a_{rs}) \geq 0,$$

оскільки $b_i / a_{is} \geq b_r / a_{rs}$ згідно з правилом вибору ведучого рядка. Нарешті, якщо $a_{is} = 0$, то $\tilde{b}_i = b_i \geq 0$. Теорема доведена. ►

Приклад 7.8.

$$\begin{cases} -2x_1 + 3x_2 + x_3 = 9, \\ x_1 + x_2 + x_4 = 8, \\ 3x_1 - 2x_2 + x_5 = 9. \end{cases} \quad (7.40)$$

◀ Знайдемо всі опорні розв'язки системи (7.40). Результати обчислень наведені у таблиці 7.3. Як бачимо, система (7.40) має початковий базисний розв'язок

$$x^{(6,0)} = (0, 0, 9, 8, 9)^T,$$

який і є опорним. Далі користуємося симплексними перетвореннями. Вибираємо перший стовпчик ведучим, оскільки в ньому є два додатних елементи. Враховуючи (7.39), тобто те, що

$$\min(\tilde{b}_2 / \tilde{a}_{21} = 8, \tilde{b}_3 / \tilde{a}_{31} = 3) = 3,$$

ведучим рядком вибираємо третій рядок. На перетині ведучих стовпчика й рядка знаходиться ведучий елемент. Використовуючи вже відомі перетворення Жордана-Гаусса,

отримуємо еквівалентну форму запису системи (7.40) та відповідний їй опорний розв'язок $x^{(6,1)} = (3, 0, 15, 5, 0)^T$. Ітераційний процес знаходження інших опорних розв'язків продовжується аналогічно (табл. 7.3). ►

Зауваження 7.5. При знаходженні опорних розв'язків у базис треба вводити ту змінну, яка приводить до нового опорного розв'язку, тобто до того, який ще не знайдений.

Процес знаходження опорних розв'язків скінченний. Легко переконатися (табл. 7.3), що інших опорних розв'язків система (7.40) не має.

Таблиця 7.3

Номер ітерації	бз	x_1	x_2	x_3	x_4	x_5	x_0	Розв'язки
0	x_3	-2	3	1	0	0	9	$x^{(6,0)} = (0,0,9,8,9)^T$
	x_4	1	1	0	1	0	8	
	x_5	3	-2	0	0	1	9	
1	x_3	0	$\frac{5}{3}$	1	0	$\frac{2}{3}$	15	$x^{(6,1)} = (3,0,15,5,0)^T$
	x_4	0	$\frac{5}{3}$	0	1	$-\frac{1}{3}$	5	
	x_1	1	$-\frac{2}{3}$	0	0	$\frac{1}{3}$	3	
2	x_3	0	0	1	-1	1	10	$x^{(6,2)} = (5,3,10,0,0)^T$
	x_2	0	1	0	$\frac{3}{5}$	$-\frac{1}{5}$	3	
	x_1	1	0	0	$\frac{2}{5}$	$\frac{1}{5}$	5	
3	x_5	0	0	1	-1	1	10	$x^{(6,3)} = (3,5,0,0,10)^T$
	x_2	0	1	$\frac{1}{5}$	$\frac{2}{5}$	0	5	
	x_1	1	0	$-\frac{1}{5}$	$\frac{3}{5}$	0	3	
4	x_5	$\frac{5}{3}$	0	$\frac{2}{3}$	0	1	15	$x^{(6,4)} = (0,3,0,5,15)^T$
	x_2	$-\frac{2}{3}$	1	$\frac{1}{3}$	0	0	3	
	x_4	$\frac{5}{3}$	0	$-\frac{1}{3}$	1	0	5	

7.5.2.Основи симплекс-методу

Симплексний метод (симплекс-метод) основний при розв'язуванні задач лінійного програмування. Він точний, гарантує знаходження розв'язку за скінченне число ітерацій і належить до класу методів послідовного поліпшення допустимого базисного розв'язку (опорного плану).

Для розкриття сутності симплекс-методу на початковому етапі будемо припускати, що для задачі ЛП виконуються такі умови:

- 1) задача записана у канонічній формі;
- 2) матриця A системи обмежень $Ax = b$ має m рядків і n стовпців ($m \leq n$), причому всі m рядків лінійно незалежні ($\text{rg } A = m$);
- 3) вектор вільних членів у системі обмежень $Ax = b$ додатний, тобто $b = (b_1, \dots, b_m)^T > 0 = (0, \dots, 0)^T$;
- 4) відомий початковий опорний розв'язок системи $Ax = b$;
- 5) усі опорні розв'язки (опорні плани) системи обмежень не вироджені, тобто серед їх базисних змінних немає нульових.

Умови 1)-5) інколи називають стандартними початковими умовами.

Нехай задача ЛП

$$C_1 X_1 + \dots + C_n X_n \rightarrow \max, \quad (7.41)$$

[illegible]

$$x_j \geq 0, j = \overline{1, n} \quad (7.43)$$

власне і є задачею, що задовольняє умови 1)-5). Початковим опорним (невід'ємним базисним) розв'язком системи (7.42) задачі ЛП (7.41)–(7.43) є $x^{(6,0)} = (b_1, \dots, b_m, 0, \dots, 0)^T$, що повністю узгоджується з умовою 4) та не порушує загальності подальших викладень.

Очевидно задача (7.41)–(7.43) – це у певному сенсі частинний випадок задачі (7.11), тому, згідно з теоремою 7.5, хоча б один її оптимальний розв'язок (якщо він існує) є крайньою точкою або вершиною області допустимих розв'язків X , яка описується співвідношеннями (7.42), (7.43).

Оскільки поняття крайньої точки, вершини й опорного розв'язку є еквівалентними, причому знаходити допустимі опорні розв'язки канонічної задачі ЛП (тобто опорні розв'язки СЛАР) ми вже вміємо (підпункт 7.5.1), то можна запропонувати такий алгоритм знаходження оптимального розв'язку задачі (7.41)–(7.43):

а) відшукати всі опорні розв'язки;

б) виділити з опорних розв'язків ті, яким відповідає максимальне значення цільової функції $f(x) = c_1x_1 + \dots + c_nx_n$.

Такий алгоритм вимагає скінченного числа операцій, однак реалізувати його в більшості випадків неможливо, бо величина C_n^m астрономічна навіть при невеликих значеннях n і m . Очевидно, потрібний упорядкований перебір опорних розв'язків, що й становить суть симплекс-методу. Щоб організувати такий перебір, необхідно з'ясувати критерій оптимальності опорного розв'язку, який сформулюємо для загальної канонічної задачі ЛП (7.11).

Теорема 7.9 (критерій оптимальності опорного розв'язку задачі ЛП). Нехай $x^{(6,l)} = (x_1^{(6,l)}, \dots, x_m^{(6,l)}, 0, \dots, 0)^T$ – деякий допустимий опорний розв'язок задані (7.11),

одержаний за допомогою симплексних перетворень системи обмежень після l ітерацій, а

$$f_j = \sum_{i=1}^m c_i a_{ij}^{(l)}, \quad (7.44)$$

де $a_{ij}^{(l)}$ – це відповідні коефіцієнти розкладу j -го вектора-стовпця A_j по базисних векторах-стовпцях $A_1, \dots, A_m$ після l ітерацій. Тоді $x^{(\bar{\sigma}, l)}$ буде оптимальним розв'язком задачі (7.11), якщо

$$\sigma_j = f_j - c_j \geq 0, \quad j = \overline{1, n}. \quad (7.45)$$

◀ Для допустимого опорного розв'язку $x^{(\bar{\sigma}, l)}$ маємо рівності:

$$x_1^{(\bar{\sigma}, l)} A_1 + \dots + x_m^{(\bar{\sigma}, l)} A_m = A_0, \quad (7.46)$$

$$f(x^{(\bar{\sigma}, l)}) = \sum_{i=1}^m c_i x_i^{(\bar{\sigma}, l)}. \quad (7.47)$$

Для довільного вектора A_j з системи векторів $\{A_1, \dots, A_n\}$ мають місце аналогічні рівності

$$a_{1j}^{(l)} A_1 + \dots + a_{mj}^{(l)} A_m = A_j \quad (7.48)$$

і (7.44) (є зміст ці рівності розглядати тільки для $j \neq 1, \dots, m$).

Помножимо (7.44) і (7.48) на деяке число $\rho > 0$ і віднімемо отримані співвідношення відповідно від (7.47) і (7.46). Матимемо:

$$\sum_{i=1}^m c_i (x_i^{(\bar{\sigma}, l)} - \rho a_{ij}^{(l)}) = f(x^{(\bar{\sigma}, l)}) - \rho f_j, \quad (7.49)$$

$$(x_1^{(\bar{\sigma}, l)} - \rho a_{1j}^{(l)}) A_1 + \dots + (x_m^{(\bar{\sigma}, l)} - \rho a_{mj}^{(l)}) A_m + \rho A_j = A_0. \quad (7.50)$$

Якщо до (7.49) зліва і справа додати величину ρc_j , то одержимо

$$\sum_{i=1}^m c_i (x_i^{(\delta,l)} - \rho a_{ij}^{(l)}) + \rho c_j = f(x^{(\delta,l)}) - \rho(f_j - c_j). \quad (7.51)$$

Очевидно, число $\rho > 0$ можна підібрати так, щоб усі коефіцієнти при векторах $A_1, \dots, A_m, A_j$ у (7.50) були невід'ємними, тобто отримаємо новий опорний розв'язок

$$x^{(\delta,l+1)} = (x_1^{(\delta,l)} - \rho a_{1j}^{(l)}, \dots, x_m^{(\delta,l)} - \rho a_{mj}^{(l)}, \rho, 0, \dots, 0)^T.$$

У лівій частині (7.51) маємо значення цільової функції у точці $x^{(\delta,l+1)}$, тобто

$$f(x^{(\delta,l+1)}) = f(x^{(\delta,l)}) - \rho(f_j - c_j) = f(x^{(\delta,l)}) - \rho\sigma_j. \quad (7.52)$$

Проаналізуємо співвідношення (7.52).

Якщо $\sigma_j < 0$, то із (7.52) випливає, що $f(x^{(\delta,l+1)}) > f(x^{(\delta,l)})$, тобто $x^{(\delta,l)}$ не буде оптимальним розв'язком задачі (7.11).

Якщо $\sigma_j \geq 0$, то $f(x^{(\delta,l+1)}) \leq f(x^{(\delta,l)})$. Це означає, що $x^{(\delta,l)}$ – оптимальний розв'язок задачі (7.11), оскільки збільшити значення цільової функції вже не вдається. Теорема доведена. ►

Величини σ_j , які обчислюються за формулами (7.45), називаються симплекс-різницями, або оцінками заміщення.

Якщо при перевірці на оптимальність опорного розв'язку $x^{(\delta,l)}$ виявилось, що він не оптимальний, то знаходять новий опорний розв'язок задачі ЛП. Для цього симплексні перетворення треба доповнити таким обмеженням: *ведучим вибирається стовпчик з номером s , у якого не тільки існує хоча б один додатний елемент, але й для якого оцінка $\sigma_s < 0$ найменша.*

Зауваження 7.6. При розв'язуванні задачі ЛП на мінімум критерієм оптимальності служать умови: $\sigma_j \leq 0$, $j = \overline{1, n}$. У процесі симплексних перетворень ведучий s -й стовпчик

вибирається таким, щоб у ньому був хоч один додатний елемент, а величина $\sigma_s > 0$ – найбільшою.

Отже, алгоритм симплекс-методу для задачі ЛП зі стандартними початковими умовами (умовами 1)-5)) такий:

1) вибір додатного ведучого елемента у поточній розширеній матриці (поточній симплекс-таблиці) з урахуванням критерію оптимальності та додаткових обмежень а) та б), які забезпечують знаходження опорного плану задачі;

2) зведення поточної розширеної матриці (поточної симплекс-таблиці) до еквівалентної матриці (симплекс-таблиці) за допомогою правила прямокутників.

Симплексний метод є одночасно й методом дослідження задачі ЛП на розв'язність у цілому. Він дає можливість установити, чи має задача розв'язок, чи цей розв'язок єдиний і як знайти інші розв'язки у випадку неєдиності.

Сформулюємо основні результати такого дослідження:

1. Розв'язок єдиний, якщо серед оцінок σ_j ($j = \overline{1, n}$) оптимального розв'язку нульові тільки базисні.

2. Розв'язок неєдиний, якщо серед оцінок σ_j ($j = \overline{1, n}$) оптимального розв'язку нульові не тільки базисні. Для знаходження множини розв'язків задачі ЛП вводимо в базис змінну, якій відповідає нульова оцінка заміщення. Довільна опукла комбінація знайдених двох розв'язків також буде розв'язком даної задачі ЛП.

3. Канонічна задача ЛП не має розв'язку, бо допустима множина порожня. Цей факт або зразу впливає з постановки задачі (СЛАР є апріорі несумісною), або встановлюється у процесі пошуку початкового базисного розв'язку.

4. В одній із симплекс-таблиць для деякого значення $\sigma_k < 0$ (при розв'язуванні задачі на мінімум – для $\sigma_k > 0$) усі елементи k -го стовпчика недодатні. Це означає, що задача ЛП не має розв'язку, бо цільова функція необмежена на допустимій множині.

Розглянемо ряд прикладів (7.9–7.14), у яких задачу ЛП треба розв'язати симплекс-методом.

Приклад 7.9.

$$\begin{cases} 2x_1 + x_2 \rightarrow \max, \\ x_1 + x_2 + x_3 = 12, \\ x_1 - x_2 + x_4 = 4, \\ x_j \geq 0, j = \overline{1,4}. \end{cases} \quad (7.53)$$

◀ Як бачимо, задача (7.53) є задачею типу (7.41)–(7.43). Для неї виконуються всі п'ять перелічених вище припущень. Процес знаходження оптимального розв'язку починаємо з $x^{(6,l)} = (0, 0, 12, 4)^T$.

Результати обчислень наведені у таблиці 7.4. Вона складається з трьох симплекс-таблиць, у кожній з яких символами «бз», « c_i^6 », « x_i^6 » позначені відповідно стовпчики базисних змінних, коефіцієнтів цільової функції при базисних змінних, числових значень базисних змінних. Симплекс-різниці σ_j , $j = \overline{1, n}$, записані в $(m+1)$ -му рядку навпроти відповідної змінної x_j , $j = \overline{1, n}$, значення цільової функції – на перетині $(m+1)$ -го рядка і стовпчика « x_i^6 ».

Таблиця 7.4

i	Бз	c_i^{σ}	x_i^{σ}	2	1	0	0
				x_1	x_2	x_3	x_4
1	x_3	0	12	1	1	1	0
2	x_4	0	4	1	-1	0	1
$m+1$	σ_j		0	-2	-1	0	0
1	x_3	0	8	0	2	1	-1
2	x_1	2	4	1	-1	0	1
$m+1$	σ_j		8	0	-3	0	2
1	x_2	1	4	0	1	1/2	-1/2
2	x_1	2	8	1	0	1/2	1/2
$m+1$	σ_j		20	0	0	3/2	1/2

У першій симплекс-таблиці є дві від'ємних симплекс-різниці $\sigma_1 = -2$, $\sigma_2 = -1$. Оскільки $\sigma_1 < \sigma_2$ то ведучим стовпчиком вибирається перший. Згідно зі співвідношенням (7.39), $\min(12/1, 4/1) = 4$, тому другий рядок буде ведучим. Отже, в базис уводимо змінну x_1 замість змінної x_4 . З другої симплекс-таблиці випливає, що $x^{(\sigma, 2)} = (4, 0, 8, 0)^T$ також не є оптимальним, бо $\sigma_2 = -3 < 0$. У другому стовпчику один додатний елемент, тому в базис вводиться змінна x_2 замість змінної x_3 . Отриманий у третій симплекс-таблиці опорний розв'язок

$x^{(\bar{6},3)} = (8,4,0,0)^T$ оптимальний, бо всі $\sigma_j \geq 0$, $j = \overline{1,4}$.

Оптимальне значення цільової функції $f_{\max}^* = 2 \cdot 8 + 1 \cdot 4 + 0 \cdot 0 + 0 \cdot 0 = 20$. ►

Приклад 7.10.

$$\begin{cases} x_1 + x_2 \rightarrow \max, \\ x_1 + x_2 + x_3 = 2, \\ 4x_1 + 3x_2 + x_4 = 12, \\ x_j \geq 0, j = \overline{1,4}. \end{cases} \quad (7.54)$$

◀ Результати розв'язування задачі (7.54) симплекс-методом наведені у таблиці 7.5. Уже на другій ітерації ми отримали оптимальний розв'язок задачі (7.54):

$$x^* \equiv x^{(\bar{6},1)} = (2,0,0,4)^T, f_{\max}^* = 2 + 0 = 2.$$

Однак серед оцінок σ_j , $j = \overline{1,4}$ нульові не тільки базисні σ_1 і σ_4 , але й небазисна σ_2 . У зв'язку з цим обчислимо ще одну симплекс-таблицю, увівши в базис змінну x_2 замість змінної x_1 .

У результаті такої операції отримаємо ще один оптимальний розв'язок задачі (7.54):

$$x^{(\bar{6},2)} = (0,2,0,6)^T, f_{\max}^* = 0 + 2 = 2.$$

Довільна опукла лінійна комбінація $x^{(\bar{6},1)}$ і $x^{(\bar{6},2)}$ теж буде розв'язком задачі (7.54). Інакше кажучи, множина оптимальних розв'язків складається з допустимих точок, які визначаються так:

$$x^* = \alpha x^{(\bar{6},1)} + (1 - \alpha) x^{(\bar{6},2)}, \alpha \in [0,1]. \quad \blacktriangleright$$

Таблиця 7.5

i	бз	$c_i^{\bar{o}}$	$x_i^{\bar{o}}$	1	1	0	0
				x_1	x_2	x_3	x_4
1	x_3	0	2	1	1	1	0
2	x_4	0	12	4	3	0	1
$m+1$	σ_j		0	-1	-1	0	0
1	x_1	1	2	1	1	1	0
2	x_4	0	4	0	-1	-4	1
$m+1$	σ_j		2	0	0	1	0
1	x_2	1	2	1	1	1	0
2	x_4	0	6	1	0	-3	1
$m+1$	σ_j		2	0	0	1	0

Приклад 7.11.

$$\begin{cases} 2x_1 + x_3 \rightarrow \min, \\ x_1 + x_2 + x_3 = 1, \\ 2x_1 + 2x_2 + 2x_3 = 4, \\ x_j \geq 0, j = \overline{1,3}. \end{cases} \quad (7.55)$$

◀ У задачі (7.55) СЛАР несумісна, бо рівняння суперечать одне одному. Отже, допустима множина порожня, тому задача (7.55) апіорі не має розв'язку. ▶

Приклад 7.12.

$$\begin{cases} 2x_1 + 5x_2 \rightarrow \min, \\ x_1 + x_2 + x_3 = 1, \\ x_1 - x_2 - x_4 = 2, \\ x_j \geq 0, j = \overline{1,4}. \end{cases} \quad (7.56)$$

◀ Очевидно, що в задачі (7.56) невідомий невід’ємний початковий базисний (опорний) розв’язок, тому до цієї задачі ми повернемося пізніше у наступному пункті. ▶

Приклад 7.13.

$$\begin{cases} x_1 + 2x_2 \rightarrow \max, \\ x_1 - x_2 + x_3 = 2, \\ -x_1 + x_2 + x_4 = 2, \\ x_j \geq 0, j = \overline{1, 4}. \end{cases} \quad (7.57)$$

Таблиця 7.6

i	бз	$c_i^{\bar{b}}$	$x_i^{\bar{b}}$	1	2	0	0
				x_1	x_2	x_3	x_4
1	x_3	0	2	1	-1	1	0
2	x_4	0	2	-1	1	0	1
$m+1$	σ_j		0	-1	-2	0	0
1	x_3	0	4	0	0	1	1
2	x_2	2	2	-1	1	0	1
$m+1$	σ_j		4	-3	0	0	2

◀ У другій симплекс-таблиці (таблиця 7.6) перший стовпчик (стовпчик, що відповідає оцінці $\sigma_1 = -3 < 0$) складається з недодатних елементів. Це означає, що цільова функція $f(x) = x_1 + 2x_2$ необмежена зверху на допустимій множині розв’язків, тому оптимального розв’язку задача (7.57) не має. ▶

Приклад 7.14.

$$\begin{cases} 2x_1 - 2x_2 \rightarrow \max, \\ x_1 - x_2 \leq 1, \\ -x_1 + x_2 \leq 0, \\ x_j \geq 0, j = \overline{1, 2}. \end{cases} \quad (7.58)$$

◀ Від задачі (7.58) переходимо до канонічної задачі ЛП:

$$\begin{cases} 2x_1 - 2x_2 \rightarrow \max, \\ x_1 - x_2 + x_3 = 1, \\ -x_1 + x_2 + x_4 = 0, \\ x_j \geq 0, j = \overline{1, 4}. \end{cases} \quad (7.59)$$

Процес розв'язування задачі (7.59) зафіксований у таблиці 7.7. Оптимальний розв'язок отриманий у другій симплекс-таблиці, причому

$$x^* \equiv x^{(6,1)} = (1, 0, 0, 1)^T, f_{\max}^* = 2.$$

Цей розв'язок не єдиний, оскільки $\sigma_2 = 0$ (а другий стовпчик не базисний). Отримати множину оптимальних розв'язків не можемо запропонованим раніше шляхом (приклад 7.10), бо змінну x_2 неможливо ввести в базис (другий стовпчик складається з недодатних елементів).

Така ситуація означає, що множина оптимальних розв'язків (як і допустимих) задачі (7.58) необмежена. Крім того, оскільки лінія рівня $2x_1 - 2x_2 = \text{const}$ паралельна до лінії $x_1 - x_2 = 1$, то множина оптимальних розв'язків задачі (7.58) збігається з ребром

$$\left\{ x^* = (x_1^*, x_2^*)^T \in \mathbb{R}^2 / x_1^* - x_2^* = 1, x_1^* \geq 1, x_2^* \geq 0 \right\}.$$

У цьому легко переконатися, розв'язавши задачу (7.58) графічним методом. ►

Таблиця 7.7

i	бз	$c_i^{\bar{0}}$	$x_i^{\bar{0}}$	2	-2	0	0
				x_1	x_2	x_3	x_4
1	x_3	0	1	1	-1	1	0
2	x_4	0	0	-1	1	0	1
$m+1$	σ_j		0	-2	2	0	0
1	x_1	2	1	1	-1	1	0
2	x_4	0	1	0	0	1	1
$m+1$	σ_j		2	0	0	2	0

7.5.3. Знаходження початкового опорного розв'язку канонічної задачі лінійного програмування

У попередньому підпункті сформульовані п'ять припущень, які спрощували вихідну задачу ЛП та її розв'язування симплекс-методом. Очевидно, що перші дві з них не принципові. Довільну задачу ЛП можна звести до еквівалентної канонічної задачі ЛП, для якої умова 2) буде виконуватися. Зовсім інший характер мають четверта і п'ята умови (умова 3) у певному сенсі пов'язана з умовою 5)). У цьому пункті зупинимося детальніше на знаходженні початкового опорного розв'язку канонічної задачі ЛП.

Метод еквівалентних перетворень розширеної матриці системи обмежень задачі

Цей метод насправді уже викладений у підпункті 7.5.1. Зміст його полягає в еквівалентному переході (за допомогою відомих нам симплекс-перетворень) від початкової розширеної матриці, яка не має відповідної кількості базисних стовпців, до такої розширеної матриці,

яка має цю кількість базисних стовпців. Інакше кажучи, початкову систему рівнянь-обмежень зводимо до такої еквівалентної системи рівнянь-обмежень, яка має початковий опорний розв'язок або є несумісною.

Метод розв'язування допоміжної задачі лінійного програмування

Нехай для задачі ЛП (7.3) (чи (7.11)), тобто задачі

[illegible]

невідомий початковий опорний розв'язок (тобто немає рівно m базисних векторів-стовпців у матриці A). Для визначеності будемо вважати, що матриця A не має жодного базисного вектора-стовпця. Поряд із цією задачею розглянемо задачу ЛП

[illegible]

Легко перекоонатися, що задача (7.60) має оптимальний розв'язок, оскільки, згідно з теоремою 7.6, про досяжність розв'язку, допустима множина цієї задачі непорожня (їй належить, наприклад, вектор $y = (0, \dots, 0, b_1, \dots, b_m)^T$), а цільова функція обмежена зверху числом 0. Нехай F^* –

оптимальне (максимальне) значення цільової функції задачі (7.60). Можливі два випадки: або $F^* = 0$, або $F^* < 0$.

Справедлива наступна теорема.

Теорема 7.10. Якщо $F^* = 0$, тобто оптимальним розв'язком задачі (7.60) є вектор $y^* = (x_1^*, \dots, x_n^*, 0, \dots, 0)^T$, то задача (7.3) має допустимий опорний розв'язок $x^* = (x_1^*, \dots, x_n^*)^T$. Якщо $F^* < 0$, то задача (7.3) не має розв'язку, оскільки її допустима множина порожня.

◀ Нехай $F^* = 0$. Покажемо, що задача (7.3) має допустимий розв'язок. Справді, якщо $F^* = 0$ і $x_j \geq 0$ ($j = \overline{n+1, n+m}$), то в оптимальному розв'язку

$$y^* = (x_1^*, x_2^*, \dots, x_n^*, x_{n+1}^*, \dots, x_{n+m}^*)^T$$

компоненти $x_{n+j}^* = 0$ ($j = \overline{1, m}$). Оскільки вектор

$$y^* = (x_1^*, x_2^*, \dots, x_n^*, 0, \dots, 0)^T$$

задовольняє обмеження допоміжної задачі (7.60), то вектор $x^* = (x_1^*, x_2^*, \dots, x_n^*)^T$ задовольняє обмеження задачі (7.3). Отже, вектор x^* – допустимий опорний розв'язок задачі (7.3).

Нехай $F^* < 0$. Оскільки всі змінні задачі (7.60) невід'ємні, то це означає, що принаймні одна зі змінних $x_{n+j}^* > 0$ ($j = \overline{1, m}$). Отже, як мінімум, одне обмеження (j -те) задачі (7.3) не буде задовольнятися вектором x^* , тобто допустима множина порожня. Теорема доведена. ►

Із теореми 7.10 випливає, що для знаходження початкового опорного (невід'ємного базисного) розв'язку задачі (7.3) досить розв'язати допоміжну задачу ЛП (7.60).

Продемонструємо це на прикладі.

Приклад 7.15.

$$\begin{cases} x_1 + 3x_2 \rightarrow \max, \\ x_1 + x_2 \leq 1, \\ x_1 - x_2 \geq 1, \\ x_j \geq 0, j = \overline{1, 2}. \end{cases} \quad (7.61)$$

◀ Від задачі (7.61) переходимо до еквівалентної канонічної задачі

$$\begin{cases} x_1 + 3x_2 \rightarrow \max, \\ x_1 + x_2 + x_3 = 1, \\ x_1 - x_2 - x_4 = 1, \\ x_j \geq 0, j = \overline{1, 4}. \end{cases} \quad (7.62)$$

Оскільки у задачі (7.62) немає початкового допустимого невід'ємного базисного (тобто опорного) розв'язку (є тільки один базисний вектор-стовпець), то потрібно розв'язати допоміжну задачу

$$\begin{cases} -x_5 \rightarrow \max, \\ x_1 + x_2 + x_3 = 1, \\ x_1 - x_2 - x_4 + x_5 = 1, \\ x_j \geq 0, j = \overline{1, 5}. \end{cases} \quad (7.63)$$

Процес розв'язування задач (7.63) і (7.62) симплекс-методом зафіксований у таблиці 7.8.

Таблиця 7.8 складається з чотирьох симплекс-таблиць. З другої симплекс-таблиці випливає, що задача (7.63) розв'язана, причому $y^* = (1, 0, 0, 0, 0)^T$, $F^* = 0$. А це означає, що $x^* = (1, 0, 0, 0)^T$ – допустимий опорний розв'язок задачі (7.62).

У другій симплекс-таблиці прямокутником виділена розширена матриця системи лінійних рівнянь-обмежень (вона еквівалентна початковій системі задачі (7.62), з якої й починається у третій симплекс-таблиці процес знаходження оптимального розв'язку задачі (7.62)). У кінцевому результаті з останньої симплекс-таблиці знаходимо оптимальний розв'язок задачі (7.62): $x^* = (1, 0, 0, 0)^T$, $f_{\max}^* = 1 \cdot 1 + 3 \cdot 0 = 1$. ►

Таблиця 7.8

i	бз	$c_i^{\bar{c}}$	$x_i^{\bar{c}}$	0	0	0	0	-1
				x_1	x_2	x_3	x_4	x_5
1	x_3	0	1	1	1	1	0	0
2	x_5	-1	1	1	-1	0	-1	1
$m+1$	σ_j		-1	-1	1	0	1	0
1	x_3	0	0	0	2	1	1	-1
2	x_1	0	1	1	-1	0	-1	1
$m+1$	σ_j		0	0	0	0	0	1
i	бз	$c_i^{\bar{c}}$	$x_i^{\bar{c}}$	1	3	0	0	
				x_1	x_2	x_3	x_4	
1	x_3	0	0	0	2	1	1	
2	x_1	1	1	1	-1	0	-1	
$m+1$	σ_j		1	0	-4	0	-1	
1	x_2	3	0	0	1	$\frac{1}{2}$	$\frac{1}{2}$	
2	x_1	1	1	1	0	$\frac{1}{2}$	$-\frac{1}{2}$	
$m+1$	σ_j		1	0	0	2	1	

Зауваження 7.7. Якщо вихідна канонічна задача ЛП має вигляд (7.5), тобто є задачею

[illegible]

то для знаходження початкового опорного розв'язку (якщо він невідомий) потрібно розв'язати допоміжну задачу ЛП (для визначеності припускається, що матриця A не має жодного базисного вектора-стовпця)

[illegible]

Позначивши, як і раніше, через F^* оптимальне (мінімальне) значення цільової функції задачі (7.64), відзначимо, що можливі два випадки: $F^* = 0$ або $F^* > 0$.

Аналогічно, як і в задачі на максимум, при $F^* > 0$ задача (7.5) не має розв'язку, а при $F^* = 0$ вона має допустимий опорний розв'язок. Повернемося до прикладу 7.12.

◀ Поставимо у відповідність до задачі (7.56) задачу ЛП

$$\begin{cases} x_5 \rightarrow \min, \\ x_1 + x_2 + x_3 = 1, \\ x_1 - x_2 - x_4 + x_5 = 2, \\ x_j \geq 0, j = \overline{1, 5}. \end{cases} \quad (7.65)$$

Процес розв'язування задачі (7.65) зафіксований у таблиці 7.9.

Таблиця 7.9

i	Бз	$c_i^{\bar{0}}$	$x_i^{\bar{0}}$	0	0	0	0	1
				x_1	x_2	x_3	x_4	x_5
1	x_3	0	1	1	1	1	0	0
2	x_5	1	2	1	-1	0	-1	1
$m+1$	σ_j		2	1	-1	0	-1	0
1	x_1	0	1	1	1	1	0	0
2	x_5	1	1	0	-2	-1	-1	1
$m+1$	σ_j		1	0	-2	-1	-1	0

У другій симплекс-таблиці отриманий оптимальний розв'язок задачі (7.65): $y^* = (1, 0, 0, 0, 1)^T$, $F^* = 1$. Оскільки $F^* > 0$, то задача (7.56) не має розв'язку, бо її допустима множина порожня. ►

М-метод (метод розв'язування розширеної задачі лінійного програмування)

Часто для знаходження початкового опорного розв'язку використовується так званий М-метод, або метод штучного базису. За своєю суттю це процес одночасного розв'язування задач (7.3) і (7.60) (або (7.5) і (7.64)), який у випадку, коли початкова матриця A не має жодного базисного стовпця, зводиться до розв'язування так званої розширеної задачі ЛП (або М-задачі):

(7.66)

додатне число), а x_{n+j} , $j = \overline{1, m}$ – так звані штучні змінні.

Має місце така теорема.

Теорема 7.11. *Якщо в оптимальному розв'язку $y^* = (x_1^*, x_2^*, \dots, x_n^*, x_{n+1}^*, \dots, x_{n+m}^*)^T$ розширеної задачі ЛП (7.66) штучні змінні $x_{n+j}^* = 0$ ($j = \overline{1, m}$), то розв'язок $x^* = (x_1^*, x_2^*, \dots, x_n^*)^T$ оптимальний для задачі (7.3).*

◀ Якщо вектор y^* задовольняє обмеження розширеної задачі, то вектор x^* задовольняє обмеження вихідної задачі (7.3), тобто він допустимий. Крім того, значення цільових функцій цих задач ((7.3) і (7.66)) $f(x^*)$ і $F(y^*)$ однакові, тобто $f(x^*) = F(y^*)$. Доведемо, що x^* – оптимальний розв'язок задачі (7.3). Доведення проведемо від супротивного. Нехай x^* – не оптимальний розв'язок задачі (7.3). Тоді існує розв'язок $\tilde{x} = (\tilde{x}_1, \tilde{x}_2, \dots, \tilde{x}_n)^T$, для якого виконується нерівність $f(\tilde{x}) > f(x^*)$, а розв'язок розширеної задачі $\tilde{y} = (\tilde{x}_1, \tilde{x}_2, \dots, \tilde{x}_n, 0, \dots, 0)^T$ такий, що $F(\tilde{y}) = f(\tilde{x}) > f(x^*) = F(y^*)$, а це суперечить тому, що y^* є оптимальним розв'язком розширеної задачі (7.66). Теорема доведена. ▶

Продемонструємо отриманий результат на конкретному прикладі.

Приклад 7.16.

$$\begin{cases} 5x_1 + 3x_2 + 4x_3 - x_4 \rightarrow \max, \\ x_1 + 3x_2 + 2x_3 + 2x_4 = 3, \\ 2x_1 + 2x_2 + x_3 + x_4 = 3, \\ x_j \geq 0, j = \overline{1,4}. \end{cases} \quad (7.67)$$

◀ Система обмежень даної задачі не має початкового базису, тому, ввівши в обмеження штучні змінні x_5 і x_6 , перейдемо від задачі (7.67) до розширеної задачі (М-задачі):

$$\begin{cases} 5x_1 + 3x_2 + 4x_3 - x_4 - Mx_5 - Mx_6 \rightarrow \max, \\ x_1 + 3x_2 + 2x_3 + 2x_4 + x_5 = 3, \\ 2x_1 + 2x_2 + x_3 + x_4 + x_6 = 3, \\ x_j \geq 0, j = \overline{1,6}. \end{cases} \quad (7.68)$$

Задачу (7.68) розв'яжемо симплекс-методом (таблиця 7.10).

Таблиця 7.10 складається з чотирьох симплекс таблиць. У перших двох симплекс-таблицях (у яких є штучні змінні) рядок для симплекс-різниць розбивається на два рядки. У $(m+1)$ -му рядку записуються вільні члени, а в $(m+2)$ -му – коефіцієнти при M у цільовій функції та оцінки $\sigma_j, j = \overline{1,6}$. Та доки не виключені з базису штучні змінні, критерієм оптимальності служить $(m+2)$ -ий рядок. У процесі симплексних перетворень вектори-стовпці, які відповідають штучним змінним, при виводі їх із базису можна не перераховувати. Що ж стосується прикладу 7.16, то після перших двох симплекс-таблиць штучні змінні x_5 і x_6 виключені з базису, тому подальші розрахунки ведуться тільки з орієнтацією на $(m+1)$ -ий рядок (критерій оптимальності).

Таблиця 7.10

i	бз	$c_i^{\bar{\theta}}$	$x_i^{\bar{\theta}}$	5	3	4	-1	$-M$	$-M$
				x_1	x_2	x_3	x_4	x_5	x_6
1	x_5	$-M$	3	1	3	2	2	1	0
2	x_6	$-M$	3	2	2	1	1	0	1
$m+1$	σ_j		0	-5	-3	-4	1	0	0
$m+2$			-6	-3	-5	-3	-3	0	0
1	x_2	3	1	$\frac{1}{3}$	1	$\frac{2}{3}$	$\frac{2}{3}$	$\frac{1}{3}$	0
2	x_6	$-M$	1	$\frac{4}{3}$	0	$-\frac{1}{3}$	$-\frac{1}{3}$	$-\frac{2}{3}$	1
$m+1$	σ_j		3	-4	0	-2	3	1	0
$m+2$			-1	$-\frac{4}{3}$	0	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{5}{3}$	0
1	x_2	3	$\frac{3}{4}$	0	1	$\frac{3}{4}$	$\frac{3}{4}$		
2	x_1	5	$\frac{3}{4}$	1	0	$-\frac{1}{4}$	$-\frac{1}{4}$		
$m+1$	σ_j		6	0	0	-3	2		
1	x_3	4	1	0	$\frac{4}{3}$	1	1		
2	x_1	5	1	1	$\frac{1}{3}$	0	0		
$m+1$	σ_j		9	0	4	0	5		

У четвертій симплекс-таблиці знайдений оптимальний розв'язок задачі (7.67): $x^* = (1, 0, 1, 0)^T$, $f_{\max}^* = 5 \cdot 1 + 3 \cdot 0 + 4 \cdot 1 - 1 \cdot 0 = 9$. ►

Справедливі такі теореми.

Теорема 7.12. Якщо в оптимальному розв'язку розширеної задачі (7.66) хоча б одна зі штучних змінних

додатна, то множина допустимих розв'язків задачі (7.3) порожня (система обмежень несумісна).

◀ Нехай $y^* = (x_1^*, \dots, x_n^*, x_{n+1}^*, 0, \dots, 0)^T$ – оптимальний розв'язок розширеної задачі, причому будемо вважати, що компонента $x_{n+1}^* > 0$. Доведення проведемо від супротивного. Нехай задача (7.3) має оптимальний розв'язок $x = (x_1, \dots, x_n)^T$. Тоді вектор $y = (x_1, \dots, x_n, 0, \dots, 0)^T$ буде допустимим розв'язком розширеної задачі (7.66), бо задовольняються її обмеження. За рахунок вибору досить великого додатного числа M можна домогтися виконання нерівності

$$F(y) = f(x) = \sum_{j=1}^n c_j x_j > \sum_{j=1}^n c_j x_j^* - M x_{n+1}^* = F(y^*).$$

Отримана нерівність ($F(y) > F(y^*)$) означає, що вектор y^* не є оптимальним розв'язком розширеної задачі. Отже, з одного боку, розширена задача має оптимальний розв'язок y^* , а з другого – цей розв'язок не є оптимальним. Дана суперечність показує, що вихідне припущення помилкове. Значить, задача (7.3) не має оптимальних розв'язків. Тепер доведемо, що множина допустимих розв'язків задачі (7.3) порожня, тобто її система обмежень несумісна. Задача ЛП на максимум не має оптимальних розв'язків у двох випадках: 1) система обмежень несумісна, тобто множина допустимих розв'язків порожня; 2) множина допустимих розв'язків необмежена зверху.

Другий випадок виключається. Справді, множина допустимих розв'язків задачі (7.3) міститься у множині допустимих розв'язків розширеної задачі (7.66), тому, якщо вона є необмеженою зверху множиною, то множина допустимих розв'язків розширеної задачі (7.66) також буде необмеженою зверху, тобто необмеженою зверху буде

цільова функція на допустимій множині. Отже, розширена задача (7.66) не має оптимального розв'язку, що суперечить твердженню теореми про те, що розширена задача має оптимальний розв'язок.

Отже, залишається справедливим перший випадок – система обмежень несумісна, тобто множина допустимих розв'язків задачі (7.3) порожня. Теорема доведена. ►

Теорема 7.13. *Якщо M -задача (7.66) не має оптимального розв'язку, то і вихідна задача (7.3) також не має оптимального розв'язку.*

◄ Розширена задача (7.66) не має оптимальних розв'язків у двох випадках:

1) система обмежень несумісна – множина допустимих розв'язків порожня;

2) множина допустимих розв'язків не обмежена зверху.

Розглянемо перший випадок. Доведення будемо проводити від супротивного. Нехай вихідна задача (7.3) має оптимальний розв'язок $x^* = (x_1^*, \dots, x_n^*)^T$. Тоді розв'язок $y = (x_1^*, \dots, x_n^*, 0, \dots, 0)^T$ буде належати множині допустимих розв'язків розширеної задачі, тобто ця множина буде непорожньою. Це суперечить початковому припущенню. Отже, у першому випадку вихідна задача не має оптимальних розв'язків.

Другий випадок – це випадок, коли множина допустимих розв'язків розширеної задачі (7.66) не обмежена зверху. Припустимо, як і раніше, що вихідна задача має оптимальний розв'язок $x^* = (x_1^*, \dots, x_n^*)^T$. Тоді $y = (x_1^*, \dots, x_n^*, 0, \dots, 0)^T$ належить допустимій множині задачі (7.66). Значення цільових функцій вихідної та розширеної задач будуть рівними (позначимо їх через f^*):

$$f(x^*) = \sum_{j=1}^n c_j x_j^* = F(y) = f^* .$$

Візьмемо довільний допустимий вектор

$$z = (y_1, \dots, y_n, y_{n+1}, \dots, y_{n+m})^T$$

розширеної задачі і знайдемо значення цільової функції

$$F(z) = \sum_{j=1}^n c_j y_j - M \sum_{i=1}^m y_{m+i} .$$

Оскільки M – досить велике додатне число, $y_j \geq 0$ ($j = 1, \overline{n+m}$), а вихідна задача (7.3) має оптимальний розв'язок, то

$$F(z) \leq \sum_{j=1}^n c_j y_j = f^* ,$$

звідки випливає, що множина допустимих розв'язків розширеної задачі обмежена, але одночасно бути необмеженою і обмеженою вона не може. Отже, припущення про існування у вихідної задачі (7.3) оптимального розв'язку неправильне, тобто задача (7.3) не має оптимальних розв'язків. Теорема доведена. ►

На завершення цього пункту відзначимо, що при розв'язуванні задачі ЛП на мінімум (у якої невідомий початковий базис) у цільовій функції відповідної розширеної задачі ЛП штучні змінні, помножені на коефіцієнт M ($M \gg 0$), додаються. Інакше кажучи, M -задачею для задачі (7.5), у якій матриця A не має жодного базисного стовпця, буде задача (7.69):

(7.69)

б) Якщо матриця обмежень A містить $k < m$ різних базисних векторів-стовпців, то при формуванні M -задачі вводять $(m - k)$ штучних змінних у відповідні рівняння системи обмежень. Випадок $k = m$ розглянений вище.

Антициклін

295

всі базисні розв'язки (а значить, усі опорні розв'язки) задачі ЛП вироджені. У такому випадку доцільно виключити з розгляду лінійно залежні рядки матриці обмежень.

Якщо $rgA = m$, то серед базисних розв'язків також можуть бути вироджені. Для нас актуальний тільки останній випадок. При практичному розв'язуванні задачі ЛП виродженість означає, що процес послідовного перебору вершин на якійсь із них може зациклитись.

Довільне правило вибору ведучого елемента, за допомогою якого можна позбутися зациклювання, називається антицикліном.

Один із найпростіших антициклінів базується на понятті так званої лексикографічної впорядкованості векторів. Нехай маємо два вектори $x = (x_0, x_1, \dots, x_n)^T$ і $y = (y_0, y_1, \dots, y_n)^T$. Вектор x лексикографічно менший від вектора y ($x \prec y$), якщо при порівнянні їх координат зліва направо перша неспівпадаюча координата вектора x менша від відповідної координати вектора y .

При перерозрахунку симплекс-таблиць треба дотримуватися такого антицикліну: ведучий r -ий рядок вибирається так, щоб

$$\frac{1}{\tilde{a}_{rs}}(x_r^\delta, \tilde{a}_{r1}, \dots, \tilde{a}_{rn}) \prec \frac{1}{\tilde{a}_{is}}(x_i^\delta, \tilde{a}_{i1}, \dots, \tilde{a}_{in}), \quad (7.70)$$

де $\tilde{a}_{rs} > 0$, $\tilde{a}_{is} > 0$ ($\tilde{a}_{rs}, \tilde{a}_{is}$ – елементи поточної розширеної матриці обмежень).

Антициклін (7.70) теоретично обґрунтований, тому зупинятися на його детальному аналізі не будемо.

Зауважимо тільки ту обставину, що алгоритм розв'язування вироджених задач ЛП повністю збігається з відповідним алгоритмом для неvirоджених задач ЛП, але з урахуванням правила (7.70). Для ілюстрації зациклення та

виходу з нього за допомогою антицикліну (7.70) розглянемо наступний приклад знаходження розв'язку задачі ЛП за допомогою симплекс-методу.

Приклад 7.17.

$$\begin{cases} \frac{3}{4}x_1 - 150x_2 + \frac{1}{50}x_3 - 6x_4 \rightarrow \max, \\ \frac{1}{4}x_1 - 60x_2 - \frac{1}{25}x_3 + 9x_4 + x_5 = 0, \\ \frac{1}{2}x_1 - 90x_2 - \frac{1}{50}x_3 + 3x_4 + x_6 = 0, \\ x_3 + x_7 = 0, \\ x_j \geq 0, \quad j = \overline{1, 7}. \end{cases} \quad (7.71)$$

◀ Якщо процес знаходження оптимального розв'язку задачі (7.71) (без урахування антицикліну (7.70)) почати з початкового опорного розв'язку із базисними змінними x_5 , x_6 , x_7 і ведучим вважати рядок з найменшим номером, то симплекс-алгоритм зациклиться у вихідній точці: на 1-й і 7-й ітераціях базиси співпадуть (таблиця 7.11).

Тепер розв'яжемо задачу (7.71), користуючись описаним вище антицикліном.

Результати обчислень наведені у таблиці 7.12, яка складається з чотирьох симплекс-таблиць. Оскільки початковий опорний розв'язок у задачі (7.71) вироджений, то вибір ведучого рядка здійснюємо за допомогою правила (7.70). У першому ведучому стовпчику є два додатних елементи (перший і другий рядок). Згідно з правилом (7.70), перший рядок лексикографічно менший від другого рядка, тобто

$$(0, 1/4, -60, -1/25, 9, 1, 0, 0) / (1/4) \prec \\ \prec (0, 1/2, -90, -1/50, 3, 0, 1, 0) / (1/2).$$

У зв'язку з цим ведучим вибирається перший рядок.

Таблиця 7.11

Номер ітерації	Оцінки $\sigma_j \ (j = \overline{1, 7})$							Базисні змінні	Значення функції
1	-3/4	150	-1/50	6	0	0	0	x_5, x_6, x_7	0
2	0	-30	-7/50	33	3	0	0	x_1, x_6, x_7	0
3	0	0	-2/25	18	1	1	0	x_1, x_2, x_7	0
4	1/4	0	0	-3	-2	3	0	x_3, x_2, x_7	0
5	-1/2	120	0	0	-1	1	0	x_3, x_4, x_7	0
6	-7/4	330	1/50	0	0	-2	0	x_5, x_4, x_7	0
7	-3/4	150	-1/50	6	0	0	0	x_5, x_6, x_7	0

У другій симплекс-таблиці вибір ведучого елемента однозначний. Легко перекоонатися, що ведучим рядком у третій симплекс-таблиці буде третій, оскільки він лексикографічно найменший. Четверта симплекс-таблиця – остання. У ній знайдений оптимальний розв'язок. ►

Зауваження 7.9. При практичному розв'язуванні задач ЛП (і вироджених також) зациклювання відбувається досить рідко. Однак у випадку виродженого опорного розв'язку все-таки треба користуватись антицикліном (7.70).

Таблица 7.12

i	$\bar{b}_i$	$\bar{c}_i$	$\bar{x}_i$	$\frac{3}{4}$	-150	$\frac{1}{50}$	-6	0	0	0
				x_1	x_2	x_3	x_4	x_5	x_6	x_7
1	x_5	0	0	$\frac{1}{4}$	-60	$-\frac{1}{25}$	9	1	0	0
2	x_6	0	0	$\frac{1}{2}$	-90	$-\frac{1}{50}$	3	0	1	0
3	x_7	0	0	0	0	1	0	0	0	1
$m+1$	σ_j		0	$-\frac{3}{4}$	150	$-\frac{1}{50}$	6	0	0	0
1	x_1	$\frac{3}{4}$	0	1	-240	$-\frac{4}{25}$	36	4	0	0
2	x_6	0	0	0	30	$\frac{3}{50}$	-15	-2	1	0
3	x_7	0	0	0	0	1	0	0	0	1
$m+1$	σ_j		0	0	-30	$-\frac{7}{50}$	33	3	0	0
1	x_1	$\frac{3}{4}$	0	1	0	$\frac{8}{25}$	-84	-12	8	0
2	x_2	-150	0	0	1	$\frac{1}{500}$	$-\frac{1}{2}$	$-\frac{1}{15}$	$\frac{1}{30}$	0
3	x_7	0	0	0	0	1	0	0	0	1
$m+1$	σ_j		0	0	0	$-\frac{2}{25}$	18	1	1	0

Продовження таблиці 7.12

1	x_1	$\frac{3}{4}$	0	1	0	0	-84	-12	8	$-\frac{8}{25}$
2	x_2	-150	0	0	1	0	$-\frac{1}{2}$	$-\frac{1}{15}$	$\frac{1}{30}$	$-\frac{1}{500}$
3	x_3	$\frac{1}{50}$	0	0	0	1	0	0	0	1
$m+1$	σ_j		0	0	0	0	18	1	1	$\frac{3}{50}$

7.6. Двоїсті задачі лінійного програмування

7.6.1. Правила побудови двоїстих задач

Довільній задачі ЛП ставиться у відповідність деяка інша задача ЛП, яка називається двоїстою (або спряженою). Наприклад, загальній задачі ЛП (7.1) на максимум, тобто задачі

$$\left\{ \begin{array}{l} c_1 x_1 + \dots + c_n x_n \rightarrow \max, \\ a_{1l} x_1 + \dots + a_{ln} x_n \leq b_l, \\ \dots\dots\dots \\ a_{kl} x_1 + \dots + a_{kn} x_n \leq b_k, \\ a_{k+1,l} x_1 + \dots + a_{k+1,n} x_n = b_{k+1}, \\ \dots\dots\dots \\ a_{ml} x_1 + \dots + a_{mn} x_n = b_m, \\ x_j \geq 0, j = \overline{1, l}; \quad l \leq n \end{array} \right. \quad (7.72)$$

ставиться у відповідність двоїста задача ЛП

(7.73)

На основі задач (7.72) і (7.73) сформулюємо правила побудови двоїстих задач, які впливають із загальних принципів двоїстості в математиці:

2. Якщо A – матриця обмежень прямої задачі, то A^T – матриця обмежень двоїстої задачі.

4. Якщо $f(x) = c_1x_1 + \dots + c_nx_n$ – цільова функція прямої задачі, то $\tilde{f}(\lambda) = b_1\lambda_1 + \dots + b_m\lambda_m$ – цільова функція двоїстої задачі.

6. Якщо у прямій задачі змінна $x_j \geq 0$, то j -те обмеження у двоїстій задачі є нерівністю типу « $\geq$ » (або « $\leq$ »). Якщо змінна x_j може набувати довільних значень (як

невід'ємних, так і від'ємних), то j – те обмеження у двоїстій задачі є рівністю («=»). Аналогічний зв'язок існує між змінними двоїстої задачі та обмеженнями прямої задачі.

Задачі (7.72) і (7.73) є взаємно двоїстими задачами зі змішаними обмеженнями. Часто розглядають симетричні й несиметричні двоїсті задачі ЛП.

Коли система обмежень кожної з пари двоїстих задач ЛП є системою нерівностей відповідного типу, то це пара симетричних двоїстих задач. Коли одна з пари двоїстих задач канонічна, то це пара несиметричних двоїстих задач.

Наведемо деякі приклади побудови двоїстих задач ЛП.

Приклад 7.18.

$$\begin{cases} x_1 + 5x_2 \rightarrow \max, \\ -x_1 + 4x_2 \leq 2, \\ 2x_1 + 3x_2 \leq 4, \\ x_1 - x_2 = 1, \\ x_1 \geq 0. \end{cases} \quad (7.74)$$

◀ Користуючись сформульованими вище правилами побудови двоїстих задач, легко переконатися, що двоїстою до (7.74) буде задача ЛП:

$$\begin{cases} 2\lambda_1 + 4\lambda_2 + \lambda_3 \rightarrow \min, \\ -\lambda_1 + 2\lambda_2 + \lambda_3 \geq 1, \\ 4\lambda_1 + 3\lambda_2 - \lambda_3 = 5, \\ \lambda_1 \geq 0, \lambda_2 \geq 0. \end{cases} \quad (7.75)$$

Двоїстою до задачі (7.75) буде задача (7.74). ►

Розглянемо також приклад побудови двоїстих задач ЛП у тих випадках, коли у прямій задачі ЛП на максимум є нерівності типу « $\geq$ », або в прямій задачі на мінімум – нерівності типу « $\leq$ ». У цьому випадку скористаємось наступним зауваженням.

Зауваження 7.10. Якщо у прямій задачі ЛП на максимум i -те обмеження є нерівністю типу « $\geq$ » або у прямій задачі на мінімум i -те обмеження є нерівністю типу « $\leq$ », то відповідна двоїста змінна $\lambda_i \leq 0$. Аналогічне твердження справедливе і в зворотньому напрямку.

Це доповнення до правил побудови двоїстих задач має суттєвий, але похідний характер, оскільки без нього можна обійтися, що й продемонструємо у наступному прикладі.

Приклад 7.19.

$$\begin{cases} 7x_1 - 2x_2 \rightarrow \max, \\ -5x_1 + x_2 \leq 7, \\ x_1 - 2x_2 = 3, \\ 3x_1 + 4x_2 \geq 10, \\ x_1 \geq 0. \end{cases} \quad (7.76)$$

◀ Пряме застосування до задачі (7.76) правил побудови двоїстих задач ЛП із урахуванням зауваження 7.10 приводить нас до такої двоїстої задачі:

$$\begin{cases} 7\lambda_1 + 3\lambda_2 + 10\lambda_3 \rightarrow \min, \\ -5\lambda_1 + \lambda_2 + 3\lambda_3 \geq 7, \\ \lambda_1 - 2\lambda_2 + 4\lambda_3 = -2, \\ \lambda_1 \geq 0, \lambda_3 \leq 0. \end{cases} \quad (7.77)$$

Часто розглядають інший варіант побудови двоїстих задач. Запишемо задачу (7.76) в еквівалентній формі (7.78):

$$\begin{cases} 7x_1 - 2x_2 \rightarrow \max, \\ -5x_1 + x_2 \leq 7, \\ x_1 - 2x_2 = 3, \\ -3x_1 - 4x_2 \leq -10, \\ x_1 \geq 0. \end{cases} \quad (7.78)$$

Зі сформульованих вище шести правил побудови двоїстих задач ЛП випливає, що двоїстою до задачі (7.78) є задача

$$\begin{cases} 7y_1 + 3y_2 - 10y_3 \rightarrow \min, \\ -5y_1 + y_2 - 3y_3 \geq 7, \\ y_1 - 2y_2 - 4y_3 = -2, \\ y_1 \geq 0, y_3 \geq 0. \end{cases} \quad (7.79)$$

Порівняння задач (7.77) та (7.79) показує, що задачу (7.77) ми отримуємо із задачі (7.79), якщо зробимо заміну змінних: $\lambda_1 = y_1$, $\lambda_2 = y_2$, $\lambda_3 = -y_3$. Задачі (7.77) та (7.79) еквівалентні двоїсті задачі до задачі (7.76). ►

7.6.2. Основні властивості пари взаємно двоїстих задач

Не порушуючи цілісності та загальності міркувань, надалі будемо розглядати пару несиметричних двоїстих задач ЛП, а саме задачу (7.11), тобто задачу

$$\begin{cases} \langle c, x \rangle \rightarrow \max, \\ Ax = b, \\ x \geq 0 \end{cases}$$

та двоїсту до неї задачу

$$\begin{cases} \langle b, \lambda \rangle \rightarrow \min, \\ A^T \lambda \geq c. \end{cases} \quad (7.80)$$

Оскільки довільна задача ЛП може бути приведена до еквівалентної канонічної задачі ЛП, то задачі (7.11) і (7.80) можна вважати основними для вивчення питання про властивості пари двоїстих задач ЛП та їх сумісне розв'язування.

Зв'язок між задачами (7.11) і (7.80), а також їх розв'язками стверджується у наступних теоремах ($f(x) = \langle c, x \rangle$, $\tilde{f}(x) = \langle b, \lambda \rangle$).

Теорема 7.14. Для довільних допустимих розв'язків задач (7.11), (7.80), тобто для довільних

$$x \in X = \{x \in \mathbb{R}^n / Ax = b, x \geq 0\},$$

$$\lambda \in \Lambda = \{\lambda \in \mathbb{R}^m / A^T \lambda \geq c\}$$

виконується нерівність $f(x) \leq \tilde{f}(\lambda)$.

◀ Помножимо скалярно матрично-векторне рівняння $Ax = b$ на довільне $\lambda \in \Lambda$, а нерівність $A^T \lambda \geq c$ на довільне $x \in X$. У результаті цього отримаємо співвідношення:

$$\langle \lambda, Ax \rangle = \langle \lambda, b \rangle = \langle b, \lambda \rangle, \quad \langle A^T \lambda, x \rangle \geq \langle c, x \rangle.$$

З даних співвідношень випливає, що

$$\langle c, x \rangle \leq \langle A^T \lambda, x \rangle = \langle \lambda, Ax \rangle = \langle b, \lambda \rangle,$$

тобто $f(x) \leq \tilde{f}(\lambda)$, що й потрібно встановити. Теорема доведена. ►

Теорема 7.15. Якщо для деяких допустимих $x^* \in X$ і $\lambda^* \in \Lambda$ виконується рівність $f(x^*) = \tilde{f}(\lambda^*)$, то x^* і λ^* – оптимальні розв'язки пари двоїстих задач ЛП (7.11) і (7.80).

◀ Справді, для довільного допустимого $x \in X$ (за теоремою 7.14) справджується нерівність

$$\langle c, x \rangle \leq \langle b, \lambda^* \rangle,$$

але ж $\langle b, \lambda^* \rangle = \langle c, x^* \rangle$. Отже, $x^* \in X$ – оптимальний розв'язок задачі (7.11). Аналогічно для довільного $\lambda \in \Lambda$ маємо співвідношення

$$\langle b, \lambda \rangle \geq \langle c, x^* \rangle = \langle b, \lambda^* \rangle,$$

тобто $\lambda^* \in \Lambda$ – оптимальний розв'язок задачі (7.80). Теорема доведена. ►

Теорема 7.16 (перша основна теорема двоїстості).
Пряма і двоїста задачі ЛП ((7.11) і (7.80)) або обидві мають оптимальні розв'язки $x^ \in X$, $\lambda^* \in \Lambda$, причому $f(x^*) = \tilde{f}(\lambda^*)$, або обидві не мають оптимальних розв'язків.*

◀ Нехай $x^* \in X$ – оптимальний розв'язок задачі (7.11). Не порушуючи загальності міркувань, можна вважати, що оптимальний базис відповідної матриці обмежень складається з m перших векторів-стовпців $A_1, \dots, A_m$, які й утворюють базисну матрицю B ($rg A = m$). Очевидно, визначник квадратної матриці B відмінний від 0, тому існує обернена матриця B^{-1} . Помножимо рівність $Ax^* = b$ зліва на B^{-1} і перейдемо до еквівалентної рівності $B^{-1}Ax^* = B^{-1}b$ або $Dx^* = \beta$, де $\beta = B^{-1}b$, $D = B^{-1}A$. Звідси однозначно випливає, що $x^* = (\beta_1, \dots, \beta_m, 0, \dots, 0)^T$ (бо базисні компоненти довільного опорного розв'язку x завжди співпадають з відповідними компонентами вектора $A_x^{-1}b$, де A_x^{-1} – матриця, обернена до матриці A_x , побудованої з базисних векторів-стовпців).

Оскільки x^* – оптимальний розв'язок (7.11), то

$$\sigma_k = \sum_{i=1}^m c_i d_{ik} - c_k = \langle c^*, D_k \rangle - c_k \geq 0,$$

$$\text{де } c^* = (c_1, \dots, c_m)^T, \quad D_k = (d_{1k}, \dots, d_{mk})^T.$$

У матрично-векторній формі попередня нерівність запишеться так (транспонування « T » указує на вектор-рядок):

$$\sigma^T = (c^*)^T D - c^T = (c^*)^T B^{-1} A - c^T \geq 0^T.$$

Позначимо $(\lambda^*)^T = (c^*)^T B^{-1}$. Тоді із попередніх співвідношень матимемо: $(\lambda^*)^T A \geq c^T$ або $A^T \lambda^* \geq c$. Очевидно, що вектор $\lambda^* \in \Lambda$ (тобто допустимий). Покажемо,

що він оптимальний для двоїстої задачі (7.80). Справді, оскільки $\lambda^* = (B^{-1})^T c^*$ і $\langle (B^{-1})^T c^*, b \rangle = \langle c^*, B^{-1}b \rangle$, то

$$\begin{aligned} \langle b, \lambda^* \rangle &= \langle \lambda^*, b \rangle = \langle (B^{-1})^T c^*, b \rangle = \langle c^*, B^{-1}b \rangle = \\ &= \langle c^*, \beta \rangle = \sum_{i=1}^m c_i \beta_i = \langle c, x^* \rangle. \end{aligned}$$

Отже, для допустимих розв'язків $x^* \in X$ і $\lambda^* \in \Lambda$ значення цільових функцій збігаються. Згідно з теоремою 7.15, можна стверджувати, що λ^* – оптимальний розв'язок задачі (7.80).

Якщо допустити, що має оптимальний розв'язок задача (7.80), то аналогічно доводиться, що має оптимальний розв'язок і задача (7.11). Залишилося тільки довести, що якщо не має оптимального розв'язку задача (7.11), то не має оптимального розв'язку і задача (7.80) (та навпаки).

Нехай (7.11) має необмежену зверху цільову функцію $f(x)$, тобто

$$\sup_X f(x) = \infty.$$

Оскільки для довільних $x \in X$, $\lambda \in \Lambda$

$$\langle b, \lambda \rangle \geq \langle c, x \rangle,$$

то, очевидно,

$$\inf_{\Lambda} \tilde{f}(\lambda) \geq \sup_X f(x) = \infty,$$

звідки й робимо висновок: задача (7.80) не має оптимального розв'язку. Правильно і навпаки. Якщо $X = \emptyset$, то, припустивши, що (7.80) має розв'язок, одержимо, що має розв'язок і (7.11), а це суперечить вихідному припущенню. Аналогічний висновок зробимо і тоді, коли припустимо, що $\Lambda = \emptyset$. Теорема доведена. ►

Теорема 7.17 (друга основна теорема двоїстості). Для того, щоб допустимі розв'язки $x^* \in X$ і $\lambda^* \in \Lambda$ задач (7.11) і

(7.80) були оптимальними, необхідно й достатньо, щоб виконувалися співвідношення:

$$\left(b_i - \sum_{j=1}^n a_{ij} x_j^* \right) \lambda_i^* = 0, \quad i = \overline{1, m}, \quad (7.81)$$

$$\left(c_j - \sum_{i=1}^m a_{ij} \lambda_i^* \right) x_j^* = 0, \quad j = \overline{1, n}. \quad (7.82)$$

◀ *Необхідність.* Нехай $x^* \in X$ і $\lambda^* \in \Lambda$ – оптимальні розв'язки пари двоїстих задач (7.11) і (7.80). Тоді (за теоремою 7.16)

$$f(x^*) = \sum_{j=1}^n c_j x_j^* = \sum_{i=1}^m b_i \lambda_i^* = \tilde{f}(\lambda^*).$$

Доведемо справедливість співвідношень (7.81) і (7.82). Нехай, наприклад, співвідношення (7.81) при деякому $i = i_0$ не виконується, тобто

$$\left(b_{i_0} - \sum_{j=1}^n a_{i_0 j} x_j^* \right) \lambda_{i_0}^* \neq 0.$$

Тоді просумуємо (7.81) по i від 1 до m , а (7.82) по j від 1 до n , після чого одержимо:

$$0 \neq \sum_{i=1}^m b_i \lambda_i^* - \sum_{i=1}^m \sum_{j=1}^n a_{ij} x_j^* \lambda_i^*,$$

$$0 = \sum_{j=1}^n c_j x_j^* - \sum_{i=1}^m \sum_{j=1}^n a_{ij} x_j^* \lambda_i^*.$$

Очевидно, що

$$\tilde{f}(\lambda^*) = \sum_{i=1}^m b_i \lambda_i^* \neq \sum_{j=1}^n c_j x_j^* = f(x^*).$$

Згідно з теоремою 7.16, один із розв'язків x^* або λ^* не є оптимальним. Виникло суперечливе твердження: доведено, що один із розв'язків x^* або λ^* не є оптимальним, а за

умовою теореми x^* і λ^* – оптимальні. Отже, співвідношення (7.81) і (7.82) теореми справедливі.

Достатність. Нехай для деяких $x^* \in X$ і $\lambda^* \in \Lambda$ виконуються співвідношення (7.81), (7.82). Доведемо, що ці розв'язки оптимальні. Просумуємо рівності (7.81) по i від 1 до m , а рівності (7.82) по j від 1 до n . Одержимо:

$$\sum_{i=1}^m b_i \lambda_i^* = \sum_{i=1}^m \lambda_i^* \sum_{j=1}^n a_{ij} x_j^* = \sum_{i=1}^m \sum_{j=1}^n a_{ij} x_j^* \lambda_i^*,$$

$$\sum_{j=1}^n c_j x_j^* = \sum_{j=1}^n x_j^* \sum_{i=1}^m a_{ij} \lambda_i^* = \sum_{i=1}^m \sum_{j=1}^n a_{ij} x_j^* \lambda_i^*.$$

Оскільки праві частини одержаних співвідношень рівні між собою, то і ліві частини рівні між собою, тобто

$$\tilde{f}(\lambda^*) = \sum_{i=1}^m b_i \lambda_i^* = \sum_{j=1}^n c_j x_j^* = f(x^*).$$

За теоремою 7.15 розв'язки x^* і λ^* оптимальні для пари двоїстих задач (7.11) і (7.80). Теорема доведена. ►

Теорема 7.18. Для існування розв'язку однієї з пари двоїстих задач (а значить, і обидвох) необхідно й достатньо, щоб множини X і Λ і були непорожніми ($X \neq \emptyset$, $\Lambda \neq \emptyset$).

◄ *Необхідність* очевидна.

Достатність. Для довільних допустимих $x \in X$ і $\lambda \in \Lambda$ (теорема 7.14)

$$\langle c, x \rangle \leq \langle b, \lambda \rangle.$$

Отже, $X \neq \emptyset$ і $f(x) = \langle c, x \rangle$ обмежена зверху на X . За теоремою 7.6 задача (7.11) має оптимальний розв'язок. Аналогічно встановлюється існування оптимального розв'язку і для задачі (7.80). Теорема доведена. ►

7.6.3. Сумісне розв'язування взаємно двоїстих задач

Кожна із задач (7.11) і (7.80) може бути розв'язана незалежно одна від одної. Однак сформульовані вище теореми (тобто результати теорії двоїстості) дають можливість знаходити розв'язки обох задач, розв'язавши тільки одну з них. Для цього є два підходи. Перший з них продемонструємо на прикладі 7.20.

Приклад 7.20. Розглянемо задачу ЛП (7.53). Двоїстою до неї буде задача

$$\begin{cases} 12\lambda_1 + 4\lambda_2 \rightarrow \min, \\ \lambda_1 + \lambda_2 \geq 2, \\ \lambda_1 - \lambda_2 \geq 1, \\ \lambda_1 \geq 0, \lambda_2 \geq 0. \end{cases} \quad (7.83)$$

◀ Розв'язавши задачу (7.83) графічним методом, одержимо: $\lambda_1^* = \frac{3}{2}$, $\lambda_2^* = \frac{1}{2}$, $\tilde{f}_{\min}^* = 12 \cdot \frac{3}{2} + 4 \cdot \frac{1}{2} = 20$. Далі скористаємось другою теоремою двоїстості, тобто теоремою 7.17. Випишемо для даного випадку рівності (7.82). Маємо:

$$\begin{cases} (2 - \lambda_1^* - \lambda_2^*)x_1^* = 0, \\ (1 - \lambda_1^* + \lambda_2^*)x_2^* = 0, \\ (0 - \lambda_1^*)x_3^* = 0, \\ (0 - \lambda_2^*)x_4^* = 0, \end{cases} \quad \text{або} \quad \begin{cases} 0 \cdot x_1^* = 0, \\ 0 \cdot x_2^* = 0, \\ -\frac{3}{2} \cdot x_3^* = 0, \\ -\frac{1}{2} \cdot x_4^* = 0. \end{cases}$$

З останніх співвідношень однозначно випливає, що $x_3^* = 0$, $x_4^* = 0$. Оптимальні значення x_1^* і x_2^* знаходимо, розв'язавши систему рівнянь

$$\begin{cases} x_1^* + x_2^* = 12, \\ x_1^* - x_2^* = 4. \end{cases}$$

Очевидно, що ця система має такий розв'язок: $x_1^* = 8$, $x_2^* = 4$. Оптимальне (максимальне) значення цільової функції у задачі (7.53) $f_{\max}^* = 2 \cdot 8 + 1 \cdot 4 = 20$. Отже, пара двоїстих задач ЛП (7.53) і (7.83) розв'язана. ►

Однак далеко не завжди одну із пари двоїстих задач ЛП можна розв'язати графічним методом. Тоді використовується інший (більш загальний) підхід. Нехай задача (7.11) розв'язана симплекс-методом, причому x^* – її оптимальний розв'язок, базисними змінними якого є $x_{i_1}, x_{i_2}, \dots, x_{i_m}$ (вони розміщені зверху вниз у стовпчику « x_i^0 » останньої симплекс-таблиці, тобто симплекс-таблиці, у якій і знайдений цей розв'язок).

Позначимо через $c^* = (c_{i_1}, \dots, c_{i_m})^T$ – вектор із коефіцієнтів цільової функції при відповідних базисних змінних, а через B^{-1} – матрицю розмірності $(m \times m)$, обернену до матриці B , складеної зі стовпчиків початкової матриці A задачі (7.11), які відповідають базисним змінним $x_{i_1}, \dots, x_{i_m}$ (при цьому порядок стовпців строго відповідає порядку базисних змінних у рядках симплекс-таблиці). Справедлива наступна теорема.

Теорема 7.19. *Якщо $x^* \in X$ – оптимальний розв'язок задачі (7.11), то*

$$\lambda^* = (B^{-1})^T c^* \quad (7.84)$$

– оптимальний розв'язок задачі (7.80).

◄ Це твердження доведено раніше у теоремі 7.16. ►

Звичайно, теорема 7.19 справедлива і для випадку, коли розглядається пара двоїстих задач ЛП з прямою канонічною задачею на мінімум.

Повернемося до задачі (7.53) і переконаємось у справедливості вище сформульованої теореми. Справді, з останньої симплекс-таблиці (таблиця 7.4) знаходимо

$$c^* = (1, 2)^T, \quad B = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}, \quad B^{-1} = \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}, \quad (B^{-1})^T = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix}.$$

$$\text{Отже, згідно із (7.84) } \lambda^* = (B^{-1})^T c^* = \left(\frac{3}{2}, \frac{1}{2}\right)^T, \quad \tilde{f}_{\min}^* = 20,$$

що і треба було підтвердити. ►

Приклад 7.21.

$$\begin{cases} 3x_1 - x_2 \rightarrow \min, \\ 2x_1 - x_2 + x_3 = 5, \\ x_1 + 2x_2 + x_4 = 8, \\ x_j \geq 0, \quad j = \overline{1, 4}. \end{cases} \quad (7.85)$$

◄ Двоїстою до задачі (7.85) буде задача ЛП

$$\begin{cases} 5\lambda_1 + 8\lambda_2 \rightarrow \max, \\ 2\lambda_1 + \lambda_2 \leq 3, \\ -\lambda_1 + 2\lambda_2 \leq -1, \\ \lambda_1 \leq 0, \quad \lambda_2 \leq 0. \end{cases} \quad (7.86)$$

Розв'яжемо задачу (7.85) симплекс-методом (таблиця 7.13).

Таблиця 7.13

i	Бз	$c_i^{\bar{\theta}}$	$x_i^{\bar{\theta}}$	3	-1	0	0
				x_1	x_2	x_3	x_4
1	x_3	0	5	2	-1	1	0
2	x_4	0	8	1	2	0	1
$m+1$	σ_j		0	-3	1	0	0
1	x_3	0	9	$\frac{5}{2}$	0	1	$\frac{1}{2}$
2	x_2	-1	4	$\frac{1}{2}$	1	0	$\frac{1}{2}$
$m+1$	σ_j		-4	$-\frac{7}{2}$	0	0	$-\frac{1}{2}$

У другій симплекс-таблиці 7.13 знайдений оптимальний розв'язок задачі (7.85):

$$x^* = (0, 4, 9, 0)^T, \quad f_{\min}^* = -4.$$

Уточнимо розв'язок двоїстої задачі (7.86). Оскільки

$$c^* = (0, -1)^T, \quad B = \begin{pmatrix} 1 & -1 \\ 0 & 2 \end{pmatrix}, \quad B^{-1} = \begin{pmatrix} 1 & 1/2 \\ 0 & 1/2 \end{pmatrix}, \quad (B^{-1})^T = \begin{pmatrix} 1 & 0 \\ 1/2 & 1/2 \end{pmatrix},$$

то із (7.84) випливає, що

$$\lambda^* = (B^{-1})^T c^* = \left(0, -\frac{1}{2}\right)^T, \quad \tilde{f}_{\max}^* = 8 \cdot \left(-\frac{1}{2}\right) = -4. \quad \blacktriangleright$$

Приклад 7.22.

$$\begin{cases} 3x_1 - x_2 + x_4 \rightarrow \min, \\ 2x_1 - x_2 + x_3 = 5, \\ x_1 + 2x_2 + x_4 = 8, \\ x_j \geq 0, \quad j = \overline{1, 4}. \end{cases} \quad (7.87)$$

◀ Двоїстою до задачі (7.87) буде задача ЛП

$$\begin{cases} 5\lambda_1 + 8\lambda_2 \rightarrow \max, \\ 2\lambda_1 + \lambda_2 \leq 3, \\ -\lambda_1 + 2\lambda_2 \leq -1, \\ \lambda_1 \leq 0, \lambda_2 \leq 1. \end{cases} \quad (7.88)$$

Процес розв'язування задачі (7.87) проілюстровано у таблиці 7.14.

Таблиця 7.14

i	бз	$c_i^{\bar{c}}$	$x_i^{\bar{c}}$	3	-1	0	1
				x_1	x_2	x_3	x_4
1	x_3	0	5	2	-1	1	0
2	x_4	1	8	1	2	0	1
$m+1$	σ_j		8	-2	3	0	0
1	x_3	0	9	$\frac{5}{2}$	0	1	$\frac{1}{2}$
2	x_2	-1	4	$\frac{1}{2}$	1	0	$\frac{1}{2}$
$m+1$	σ_j		-4	$-\frac{7}{2}$	0	0	$-\frac{3}{2}$

Оптимальний розв'язок задачі (7.87) такий:

$$x^* = (0, 4, 9, 0)^T, \quad f_{\min}^* = -4.$$

Аналогічно, як і в попередньому прикладі, знаходимо розв'язок двоїстої задачі:

$$c^* = (0, -1)^T, \quad B = \begin{pmatrix} 1 & -1 \\ 0 & 2 \end{pmatrix}, \quad B^{-1} = \begin{pmatrix} 1 & 1/2 \\ 0 & 1/2 \end{pmatrix}, \quad (B^{-1})^T = \begin{pmatrix} 1 & 0 \\ 1/2 & 1/2 \end{pmatrix},$$

$$\lambda^* = (B^{-1})^T c^* = \left(0, -\frac{1}{2}\right)^T, \quad \tilde{f}_{\max}^* = 8 \cdot \left(-\frac{1}{2}\right) = -4. \quad \blacktriangleright$$

Зауваження 7.11. Якщо у задачі (7.11) відомий початковий опорний розв'язок, то матрицю B^{-1}

утворюють стовпчики останньої симплекс-таблиці, які відповідають базисним змінним першої симплекс-таблиці, причому компоненти оптимального розв'язку задачі (7.80) збігаються з відповідними елементами $(m+1)$ -го рядка, якщо відповідні $c_j = 0$ та дорівнюють сумі відповідного елемента та c_j , якщо $c_j \neq 0$ (таблиці 7.4, 7.13, 7.14).

Зауваження 7.12. Якщо вихідна (пряма) задача ЛП задана не в канонічній формі, але відповідна їй канонічна задача має початковий опорний розв'язок, то оптимальний розв'язок двоїстої задачі знаходиться аналогічно.

Наприклад, пара симетричних двоїстих задач ЛП (7.21) та (7.83) розв'язується у такий спосіб:

Задача (7.21) зводиться до канонічної задачі (7.53), у якій відомий початковий опорний розв'язок.

Дальніші міркування та висновки нам уже відомі (приклад 7.20). У випадку двоїстих задач

$$\begin{cases} 3x_1 - x_2 \rightarrow \min, \\ -2x_1 + x_2 \geq -5, \\ -x_1 - 2x_2 \geq -8, \\ x_j \geq 0, j = \overline{1,2} \end{cases} \quad (7.89)$$

та

$$\begin{cases} -5\lambda_1 - 8\lambda_2 \rightarrow \max, \\ -2\lambda_1 - \lambda_2 \leq 3, \\ \lambda_1 - 2\lambda_2 \leq -1, \\ \lambda_i \geq 0, i = \overline{1,2} \end{cases} \quad (7.90)$$

також розв'язують пряму задачу (7.89). Вона зводиться до канонічної задачі (7.85), розв'язок якої знайдений у

таблиці 7.13. Однак розв'язком задачі (7.90) буде вектор $\lambda^* = (0, 1/2)^T$, а не $\lambda^* = (0, -1/2)^T$ (розв'язок задачі (7.86)).

Тепер розглянемо випадок, коли канонічна задача ЛП (7.11) не має початкового опорного розв'язку (не має повного базису). Питання про те, як знаходиться у даній ситуації розв'язок двоїстої задачі (7.80), вивчимо на конкретному прикладі.

Приклад 7.23.

◀ Розглянемо задачу (7.67) (приклад 7.16). Двоїстою до неї буде задача:

$$\begin{cases} 3\lambda_1 + 3\lambda_2 \rightarrow \min, \\ \lambda_1 + 2\lambda_2 \geq 5, \\ 3\lambda_1 + 2\lambda_2 \geq 3, \\ 2\lambda_1 + \lambda_2 \geq 4, \\ 2\lambda_1 + \lambda_2 \geq -1. \end{cases} \quad (7.91)$$

Від задачі (7.67) переходимо до розширеної М-задачі (7.68), розв'язок якої знайдений у таблиці 7.10, а саме:

$$x^* = (1, 0, 1, 0)^T, \quad f_{\max}^* = 9.$$

У останній симплекс-таблиці задачі (7.67) знайдений також розв'язок двоїстої задачі (7.91) (($m+1$)-й рядок навпроти штучних змінних x_5 і x_6):

$$\lambda^* = (1, 2)^T, \quad \tilde{f}_{\min}^* = 3 \cdot 1 + 3 \cdot 2 = 9.$$

Очевидно, що $\tilde{f}_{\min}^* = f_{\max}^*$. Це легко також перевірити згідно зі сформульованою вище теоремою 7.19:

$$c^* = (4, 5)^T, \quad B = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}, \quad B^{-1} = \begin{pmatrix} 2/3 & -1/3 \\ -1/3 & 2/3 \end{pmatrix},$$

$$(B^{-1})^T = \begin{pmatrix} 2/3 & -1/3 \\ -1/3 & 2/3 \end{pmatrix}, \quad \lambda^* = (B^{-1})^T c^* = (1, 2)^T. \blacktriangleright$$

7.6.4. Економічна інтерпретація двоїстих задач

Повернемося до задачі оптимального планування виробництва (В.5) або (7.2). Двоїстою до неї буде задача

$$\begin{cases} \sum_{i=1}^m b_i \lambda_i \rightarrow \min, \\ \sum_{i=1}^m a_{ij} \lambda_i \geq c_j, j = \overline{1, n}, \\ \lambda_i \geq 0, i = \overline{1, m}. \end{cases} \quad (7.92)$$

Обидві задачі мають змістову економічну інтерпретацію. Справді, нехай вони є задачами відповідно двох виробничих фірм (підприємств) Π_1 і Π_2 . Задача (7.92) є математичною моделлю ситуації, коли Π_2 прагне закупити у Π_1 усю сировину відповідно за цінами $\lambda_i \geq 0$ (λ_i – ціна одиниці i -тої сировини S_i) грошових одиниць (г.о.). Від продажу одиниці продукції P_j підприємство Π_1 має c_j г.о. доходу.

Оскільки на виготовлення одиниці продукції P_j ($j = \overline{1, n}$) витрачається відповідно a_{ij} ($i = \overline{1, m}$) одиниць сировини S_i ($i = \overline{1, m}$), а від продажу даної кількості сировини Π_1 отримає $\sum_{i=1}^m a_{ij} \lambda_i$ г.о. доходу, то Π_1 може продати цю сировину тільки за умови, що $\sum_{i=1}^m a_{ij} \lambda_i \geq c_j$ ($j = \overline{1, n}$). Якщо Π_2 закупить усю сировину у Π_1 , то сумарна вартість сировини становить $\sum_{i=1}^m b_i \lambda_i$ г.о. Зрозуміло, що Π_2

(організація-покупець) прагне мінімізувати сумарну вартість, у результаті чого маємо задачу (7.92).

Нехай задачі (B.5) (чи (7.2)) і (7.92) розв'язані, причому

$$x^* = (x_1^*, x_2^*, \dots, x_n^*)^T, \quad \lambda^* = (\lambda_1^*, \lambda_2^*, \dots, \lambda_m^*)^T$$

– оптимальні розв'язки цих задач. Виявляється, що λ^* є вектором об'єктивно зумовлених оцінок сировини (ресурсів) у задачі (B.5). Справді, якщо $\lambda_i^* > 0$, то для довільного оптимального розв'язку x^* задачі (B.5) виконується рівність

$$\sum_{j=1}^n a_{ij} x_j^* = b_i,$$

яка впливає із (7.81) (теорема 7.17).

Це означає, що сировина S_i (її запаси становлять b_i одиниць) при довільному x^* використовується повністю, тобто S_i дефіцитна. Якщо $\lambda_i^* = 0$, то знайдеться такий оптимальний розв'язок x^* , для якого

$$\sum_{j=1}^n a_{ij} x_j^* < b_i,$$

тобто сировина S_i не буде повністю використовуватися (не буде дефіцитною).

Отже, λ^* є вектором показників дефіцитності кожного типу сировини (поняття дефіцитності визначається внутрішніми властивостями технології виробничого процесу).

7.7. Транспортні задачі (Т-задачі)

7.7.1. Теоретичні основи Т-задач

Транспортні задачі (Т-задачі) належать до спеціального класу задач лінійного програмування. Зрозуміло, що Т-задачі можна вивчати й розв'язувати як традиційні задачі

ЛП. Однак специфіка Т-задач дозволила розробити надзвичайно ефективні методи їх розв'язування. Розглянемо однопродуктову (однотоварну) транспортну задачу (В.7) – одну з найпростіших із класу Т-задач.

Т-задача (В.7) називається закритою, або збалансованою, якщо

$$\sum_{i=1}^m a_i = \sum_{j=1}^n b_j. \quad (7.93)$$

Надалі будемо розглядати тільки закриті задачі. Зазначимо, що постановку й розв'язування транспортної задачі зручно здійснювати за допомогою таблиці 7.15.

У таблиці 7.15 матриці $C = (c_{ij})_{i,j=1}^{m,n}$ та $X = (x_{ij})_{i,j=1}^{m,n}$ – це матриці вартостей та обсягів транспортних перевезень продукції від постачальників до споживачів. Ще раз підкреслимо, що у цьому пункті через X позначено відповідну матрицю, а не допустиму множину задачі.

Таблиця 7.15

Споживачі Постачальники	S_1	S_2	...	S_n	Обсяги Постачання
P_1	c_{11} x_{11}	c_{12} x_{12}	...	c_{1n} x_{1n}	a_1
P_2	c_{21} x_{21}	c_{22} x_{22}	...	c_{2n} x_{2n}	a_2
...	...	...	...	...	...
P_m	c_{m1} x_{m1}	c_{m2} x_{m2}	...	c_{mn} x_{mn}	a_m
Обсяги споживання	b_1	b_2	...	b_n	

Основні теоретичні результати для задачі (B.7) сформулюємо у наступних твердженнях.

Теорема 7.20. *Щоб Т-задача мала розв'язок, необхідно й достатньо, щоб вона була закритою, тобто, щоб виконувалося співвідношення (7.93).*

◀ **Необхідність.** Нехай Т-задача має розв'язок, тобто має допустимий план X . Тоді, очевидно, справедливі рівності

$$\sum_{i=1}^m \sum_{j=1}^n x_{ij} = \sum_{i=1}^m a_i, \quad \sum_{j=1}^n \sum_{i=1}^m x_{ij} = \sum_{j=1}^n b_j,$$

з яких і випливає співвідношення (7.93). Отже, Т-задача замкнута (збалансована). Необхідність доведена.

Достатність. Нехай виконується рівність (7.93). Покажемо, що Т-задача має розв'язок (допустимий план). Сформуємо матрицю X , елементи якої визначаються так: $x_{ij} = (a_i b_j) / d$,

$i = \overline{1, m}; \quad j = \overline{1, n} \left(d = \sum_{i=1}^m a_i = \sum_{j=1}^n b_j \right)$. Легко переконатись, що X є

допустимим розв'язком Т-задачі, оскільки

$$\sum_{i=1}^m x_{ij} = \sum_{i=1}^m \frac{a_i b_j}{d} = \frac{b_j}{d} \sum_{i=1}^m a_i = b_j, \quad j = \overline{1, n},$$

$$\sum_{j=1}^n x_{ij} = \sum_{j=1}^n \frac{a_i b_j}{d} = \frac{a_i}{d} \sum_{j=1}^n b_j = a_i, \quad i = \overline{1, m}.$$

Крім того, значення x_{ij} – невід'ємні (за побудовою). Отже, допустима множина непорожня. Оскільки цільова функція у Т-задачі обмежена знизу (наприклад, нулем), то це означає, що Т-задача має розв'язок. Достатність (а значить і теорема у цілому) доведена. ►

Теорема 7.21. *Ранг системи обмежень задачі (B.7) дорівнює числу $r = m + n - 1$, тобто будь-який опорний план Т-задачі має не більше, ніж $(m + n - 1)$ змінних x_{ij} , які набувають відмінних від нуля значень.*

◀ Спочатку доведемо, що ранг системи обмежень не перевищує величини $r = m + n - 1$. Для цього сформуємо матрицю A розмірності $[(m + n) \times mn]$ системи обмежень задачі ЛП (В.7). Її стовпці (зліва направо) відповідають змінним (елементам матриці X) $x_{11}, \dots, x_{1n}, x_{21}, \dots, x_{2n}, \dots, x_{m1}, \dots, x_{mn}$. Отже,

$$A = \begin{pmatrix} 11\dots1 & 00\dots0 & \dots & 00\dots0 \\ 00\dots0 & 11\dots1 & \dots & 00\dots0 \\ \dots & \dots & \dots & \dots \\ 00\dots0 & 00\dots0 & \dots & 11\dots1 \\ 10\dots0 & 10\dots0 & \dots & 10\dots0 \\ 01\dots0 & 01\dots0 & \dots & 01\dots0 \\ \dots & \dots & \dots & \dots \\ 00\dots1 & 00\dots1 & \dots & 00\dots1 \end{pmatrix}.$$

Нехай $A_1, A_2, \dots, A_m, A_{m+1}, \dots, A_{m+n}$ – вектор-рядки матриці A . Очевидно, що для цих векторів справджується рівність

$$A_1 + \dots + A_m = A_{m+1} + \dots + A_{m+n},$$

тобто множина векторів-рядків лінійно залежна і $rgA \leq r = m + n - 1$. Доведемо, що $rgA = r = m + n - 1$. Для цього досить показати, що система з $(m + n - 1)$ векторів-рядків або векторів-стовпців матриці A лінійно незалежна, тобто що їх лінійна комбінація дорівнює нульовому вектору тільки тоді, коли всі сталі коефіцієнти дорівнюють нулю.

Нехай $\alpha_2, \alpha_3, \dots, \alpha_m, \beta_1, \beta_2, \dots, \beta_n$ – деякі сталі. Розглянемо рівність

$$\alpha_2 A_2 + \dots + \alpha_m A_m + \beta_1 A_{m+1} + \dots + \beta_n A_{m+n} = 0 \quad (7.94)$$

і з'ясуємо, при яких значеннях сталих коефіцієнтів $\alpha_2, \dots, \alpha_m, \beta_1, \dots, \beta_n$ вона можлива.

Рівність (7.94) еквівалентна такій системі співвідношень:

(7.95)

векторів-рядків $A_2, \dots, A_m, A_{m+1}, \dots, A_{m+n}$

є лінійно незалежною. Це й означає, що $rgA = r = m + n - 1$.

Теорема доведена. ►

які задовольняють умови

(7.96)

◀ **Необхідність.** До задачі (B.7) двоїста набуває вигляду:

$$(7.97)$$

задовольняють умови

$$\begin{cases} \alpha_i + \beta_j = c_{ij} & \text{для } x_{ij}^* > 0, \\ \alpha_i + \beta_j \leq c_{ij} & \text{для } x_{ij}^* = 0. \end{cases}$$

або

$$d_{ij} = \alpha_i + \beta_j - c_{ij} \begin{cases} = 0 & \text{для } x_{ij} > 0, \\ \leq 0 & \text{для } x_{ij} = 0. \end{cases}$$

Достатність. Нехай виконуються умови (7.96), тобто

$$d_{ij} = \alpha_i + \beta_j - c_{ij} \begin{cases} = 0 & \text{для } x_{ij} > 0, \\ \leq 0 & \text{для } x_{ij} = 0. \end{cases}$$

Покажемо, що план X^* – оптимальний для Т-задачі. Величини d_{ij} є симплекс-різницями або оцінками заміщення Т-задачі як задачі ЛП на мінімум. З теорії ЛП відомо, що коли всі оцінки d_{ij} набувають недодатних значень, то задача ЛП, тобто Т-задача розв'язана і X^* – її оптимальний план (оптимальний розв'язок). Теорема доведена. ►

7.7.2. Методи розв'язування Т-задач

Процес розв'язування Т-задачі розбивається на два основних етапи:

1) знаходження початкового опорного плану (початкового невід'ємного базисного розв'язку);

2) знаходження оптимального плану задачі.

При цьому на кожному з цих етапів існує свій пакет методів, якими можна користуватися.

Для конкретизації процесу розв'язування Т-задачі надалі використаємо методи північно-західного кута (перший етап) і потенціалів (другий етап). Суть цих методів (як і інші теоретичні обґрунтування) розкриємо у такому алгоритмі розв'язування Т-задач.

I. Перевірка типу (закрита чи відкрита) Т-задачі (В.7). Якщо вона замкнута, то переходять до пункту II. Якщо вона відкрита, то поступають так. У випадку, коли

$$\sum_{i=1}^m a_i < \sum_{j=1}^n b_j,$$

вводять у розгляд фіктивного постачальника P_{m+1} із обсягом постачання

$$a_{m+1} = \sum_{j=1}^n b_j - \sum_{i=1}^m a_i$$

і вартостями $c_{m+1,j} = 0, \quad j = \overline{1, n}$.

У випадку, коли

$$\sum_{i=1}^m a_i > \sum_{j=1}^n b_j,$$

вводять фіктивного споживача S_{n+1} із обсягом споживання

$$b_{n+1} = \sum_{i=1}^m a_i - \sum_{j=1}^n b_j$$

та відповідними вартостями $c_{i,n+1} = 0, \quad i = \overline{1, m}$.

Отже, у кожному з цих випадків будують замкнуту Т-задачу, яка, згідно з теоремою 7.20, завжди має розв'язок.

II. Знаходження за допомогою методу північно-західного кута початкового опорного плану $X = X^{(1)}$ для заданої (або отриманої) замкненої Т-задачі. Суть методу північно-західного кута полягає в тому, що на кожному кроці розглядається перший із пунктів P_i і перший із пунктів S_j , що залишились. Заповнення клітинок Т-таблиці 7.15 починається з лівої верхньої для x_{11} і завершується для x_{mn} (реалізація методу проілюстрована на прикладі 7.24).

III. Аналіз знайденого плану X на виродженість. Нехай k – число відмінних від нуля величин x_{ij} у плані X .

Якщо $k = r = m + n - 1$ (теорема 7.21), то Т-задача не вироджена, а якщо $k < r$, то Т-задача вироджена. Перехід від виродженої Т-задачі до не виродженої здійснюється так: фіксується $(r - k)$ клітинок Т-таблиці (як правило, вздовж північно-західної діагоналі), після чого вважають ці клітинки заповненими «фіксованими» нулями.

IV. Перевірка методом потенціалів поточного опорного плану Т-задачі на оптимальність. Теоретичною основою методу потенціалів служить теорема 7.22. Для знаходження оптимального плану методом потенціалів необхідно виконати послідовність операцій:

1) розв'язати для заповнених клітинок Т-таблиці систему $(m + n - 1)$ лінійних алгебраїчних рівнянь $\alpha_i + \beta_j = c_{ij}$ відносно $(m + n)$ невідомих потенціалів $\alpha_1, \dots, \alpha_m, \beta_1, \dots, \beta_n$ (оскільки така система має безліч розв'язків, то для вибору конкретного частинного розв'язку досить вибрати одну вільну змінну, наприклад α_1 , нульовою);

2) для вільних клітинок Т-таблиці (клітинок, для яких $x_{ij} = 0$) визначити величини $d_{ij} = \alpha_i + \beta_j - c_{ij}$;

3) перевірити умови (7.96).

Якщо ці умови виконуються, то знайдений опорний план (позначимо його X^*) буде оптимальним і Т-задача буде розв'язана. Якщо умови (7.96) не виконуються, то переходять до пункту V.

V. Побудова для вільної клітинки з найбільшою величиною $d_{ij} > 0$ циклу перерозрахунку, знаходження нового опорного плану $X = X^{(2)}$ і перехід до пункту III.

Зробимо деякі уточнення стосовно циклу перерозрахунку.

Циклом перерозрахунку називається замкнутий багатокутник (не обов'язково опуклий), сторонами якого є горизонтальні та вертикальні відрізки й одна вершина якого збігається з вільною клітинкою (для якої утворюється цикл), а всі інші – із зайнятими клітинками.

Цикл завжди єдиний. Надалі при розв'язуванні Т-задач домовимося позначати вільну клітинку циклу символом «□». Переміщення продукції по клітинках циклу здійснюється за правилом:

а) вільній клітинці приписують знак «+», а всім іншим по черзі «-» і «+»;

б) у вільну клітинку заносять найменше з чисел, які розміщені у клітинках зі знаком «-», крім того, це число додають до чисел у «+» – клітинках і віднімають від чисел у «-» – клітинках, у результаті чого знаходять новий опорний план.

Продемонструємо роботу алгоритму на конкретному прикладі.

Приклад 7.24. Для Т-задачі (В.7) з двома постачальниками та трьома споживачами задані обсяги постачання і споживання, а також матриця вартостей транспортних перевезень (таблиця 7.16). Визначити оптимальний план Т-задачі.

Таблиця 7.16

$a_i \backslash b_j$	5	10	5
10	1	3	2
10	3	1	2

◀ Оскільки $\sum_{i=1}^2 a_i = \sum_{j=1}^3 b_j = 20$, то Т-задача замкнута.

Методом північно-західного кута знаходимо початковий опорний план $X^{(1)}$. Матриця $X^{(1)}$ формується так. Потреби споживача S_1 можна повністю задовольнити за рахунок постачальника P_1 . Із $a_1 = 10$ одиниць продукції P_1 віддає необхідні йому $b_1 = 5$ одиниць продукції ($x_{11} = 5$), у результаті чого в постачальника P_1 залишиться ще 5 одиниць продукції. Їх він віддає (відправляє) наступному споживачеві S_2 ($x_{12} = 5$). Резерви P_1 повністю вичерпані. Тепер розподілимо продукцію постачальника P_2 . Оскільки споживач S_2 потребує ще 5 одиниць продукції, то їх він і одержить від P_2 , тобто $x_{22} = 5$. Потреби споживача S_2 вичерпані. У постачальника P_2 залишилось ще 5 одиниць продукції, які він відправляє споживачу S_3 ($x_{23} = 5$). На цьому процес знаходження початкового опорного плану методом північно-західного кута завершується. Отже, початковий опорний план

$$X^{(1)} = \begin{pmatrix} 5 & 5 & 0 \\ 0 & 5 & 5 \end{pmatrix},$$

а значення цільової функції

$$F(X) = \sum_{i=1}^2 \sum_{j=1}^3 a_{ij} x_{ij}$$

для цього плану

$$F(X^{(1)}) = 1 \cdot 5 + 3 \cdot 5 + 2 \cdot 0 + 3 \cdot 0 + 1 \cdot 5 + 2 \cdot 5 = 35.$$

Проаналізуємо план $X^{(1)}$ на виродженість. Число ненульових базисних змінних $k = 4$ дорівнює рангу $r = m + n - 1 = 2 + 3 - 1 = 4$. Отже, дана Т-задача не вироджена. Усі подальші розрахунки здійснюємо, користуючись таблицею 7.17.

Таблиця 7.17

$a_i \backslash b_j$	5	10	5	
10	1 5	3 5 -	2 + □	$\alpha_1 = 0$
10	3	1 5 +	2 5 -	$\alpha_2 = -2$
	$\beta_1 = 1$	$\beta_2 = 3$	$\beta_3 = 4$	

Для знаходження потенціалів $\alpha_1, \alpha_2, \beta_1, \beta_2, \beta_3$ розв'язуємо для зайнятих клітинок таблиці 7.17 систему рівнянь

$$\begin{cases} \alpha_1 = 0, \\ \alpha_1 + \beta_1 = 1, \\ \alpha_1 + \beta_2 = 3, \\ \alpha_2 + \beta_2 = 1, \\ \alpha_2 + \beta_3 = 2. \end{cases}$$

Очевидно, що, користуючись таблицею 7.17, таку систему рівнянь можна розв'язати усно. Результати (розв'язки системи) записуємо відповідно у правому стовпці та нижньому рядку.

Далі для вільних клітинок визначаємо

$$\begin{cases} d_{13} = \alpha_1 + \beta_3 - c_{13} = 0 + 4 - 2 = 2, \\ d_{21} = \alpha_2 + \beta_1 - c_{21} = -2 + 1 - 3 = -4. \end{cases}$$

Цикл перерозрахунку будуємо для клітинки, яка розміщена у першому рядку та третьому стовпці. Циклом буде прямокутник (таблиця 7.17). Оскільки в обидвох «-»-клітинках знаходиться число 5, то у вільну клітинку заноситься число 5. Що стосується змін в інших клітинках-вершинах прямокутника, то число 5 додається до відповідного числа, розміщеного у «+»-клітинці і віднімається від числа, розміщеного у «-»-клітинках.

Зауважимо, що число 0 записується у клітинках Т-таблиці (як сказано вище) лише тоді, коли воно є фіксованим числом.

Таблиця 7.18

$a_i \backslash b_j$	5	10	5	
10	1 5	3 0	2 5	$\alpha_1 = 0$
10	3	1 10	2	$\alpha_2 = -2$
	$\beta_1 = 1$	$\beta_2 = 3$	$\beta_3 = 2$	

Після перерозподілу продукції у вершинах циклу одержимо новий опорний план Т-задачі (таблиця 7.18)

$$X^{(2)} = \begin{pmatrix} 5 & 0 & 5 \\ 0 & 10 & 0 \end{pmatrix}.$$

Значення цільової функції $F(X^{(2)}) = 25$. Оскільки план вироджений ($k = 3 < r = 4$), то для знаходження потенціалів зафіксуємо елемент $x_{12} = 0$. Для вільних клітинок $d_{21} = -4$, $d_{23} = -2$. Отже, план $X^* = X^{(2)}$ буде оптимальним, $F^* = 25$. Задача розв'язана. ►

Зупинимось на деяких особливостях процесу розв'язування Т-задач.

1. У Т-таблиці серед максимальних оцінок $d_{ij} > 0$ зустрічається дві або й більше однакових оцінок. Тоді вільною треба вибирати ту клітинку, якій відповідає мінімальна вартість c_{ij} , але серед реальних P_i та S_j .

2. У Т-таблиці з оптимальним планом X^* нульовими є тільки ті величини d_{ij} , які відповідають зайнятим клітинкам (базисним змінним). Це означає, що Т-задача має єдиний розв'язок.

3. У Т-таблиці з оптимальним планом X^* серед оцінок d_{ij} нульовими є також ті оцінки, які відповідають вільним клітинкам (не базисним змінним). Тоді Т-задача має безліч розв'язків. Якщо для деякої вільної клітинки з нульовою оцінкою d_{ij} побудувати цикл перерозрахунку, то можна одержати ще один оптимальний план (позначимо його через X^{**}). Довільна лінійна опукла комбінація $\tilde{X} = \gamma X^* + (1 - \gamma) X^{**}$ також буде оптимальним планом Т-задачі ($0 \leq \gamma \leq 1$).

На завершення пункту 7.7 зазначимо, що Т-задачі або задачі транспортного типу мають широке застосування на практиці та можуть бути формалізовані не тільки у вигляді (В.7). У більш загальному випадку сумарний обсяг продукції, яка ввозиться від постачальника, не обов'язково дорівнює запасу продукції, але не перевищує його, тобто

$$\sum_{j=1}^n x_{ij} \leq a_i, \quad i = \overline{1, m}.$$

Аналогічно обсяги постачання продукції споживачу не обов'язково дорівнюють обсягу його потреб, але не менші від них, що приводить до обмежень

$$\sum_{i=1}^m x_{ij} \geq b_j, \quad j = \overline{1, n}.$$

Для розв'язності Т-задачі з такими обмеженнями необхідно й достатньо, щоб виконувалася нерівність

$$\sum_{i=1}^m a_i \geq \sum_{j=1}^n b_j.$$

Зауважимо також, що для побудови опорних (допустимих) планів Т-задач, крім методу північно-західного кута, який є одним з найпростіших за процедурою, часто використовуються методи мінімального елемента (метод

мінімальної вартості), метод подвійної переваги та інші методи.

Згідно із методом мінімального елемента вибирається клітинка Т-таблиці з найменшою вартістю, і у цю клітинку записують найменше з чисел a_i ($i = \overline{1, m}$), b_j ($j = \overline{1, n}$). Після цього з розгляду виключають або рядок, або стовпчик, або рядок і стовпчик, для яких вичерпані запаси постачальника або (і) задоволені потреби споживача. Далі ця процедура повторюється із частиною Т-таблиці, яка залишилася після виключення з розгляду відповідних рядка, стовпчика чи рядка і стовпчика, до тих пір, поки всі запаси будуть вичерпані, а потреби задоволені.

Метод подвійної переваги доцільно використовувати, коли Т-таблиця велика. Суть методу полягає у тому, що у кожному стовпчику і рядку Т-таблиці позначають клітинку з найменшою вартістю, наприклад значком « $\vee$ ». У результаті цього у деяких клітинках появиться помітка « $\vee \vee$ », оскільки у цих клітинках вартість мінімальна як по стовпчику, так і по рядку. Потім у зазначені клітинки записують максимально можливі обсяги та виключають з розгляду відповідні стовпчики або рядки. Далі розподіл здійснюється по клітинках, які мають помітку « $\vee$ ». У залишеній частині Т-таблиці розподіл відбувається за найменшою вартістю.

Крім класичної постановки Т-задачі, якій і присвячений пункт 7.7, зустрічаються також задачі транспортного типу з деякими специфічними обмеженнями, наприклад:

- 1) за деякими маршрутами обсяги перевезень фіксовані;
- 2) на обсяги перевезень задані обмеження знизу;
- 3) пропускна здатність окремих маршрутів обмежена;
- 4) перевезення за деякими маршрутами заборонені.

Виявляється, що задачі з такими та іншими обмеженнями за допомогою спеціальних перетворень можна звести до класичної Т-задачі. Для детального знайомства із задачами такого типу рекомендується відповідна література [12, 22].

7.8. Задачі цілочислового лінійного програмування

Задачі лінійного програмування, у яких шукані змінні (усі або деякі) можуть набувати лише цілих чисел, називаються задачами цілочислового лінійного програмування (ЦЛП). Актуальність задач ЦЛП визначається тим, що вони є математичними моделями багатьох реальних економічних (та ін.) задач. Така, зокрема, задача про оптимізацію виробничої потужності підприємства (або задача про придбання нового обладнання), тобто задача (В.8).

Перейдемо до аналізу задачі ЦЛП й опису методу її розв'язання. Розглянемо канонічну задачу ЦЛП

$$\langle c, x \rangle \rightarrow \max, \quad (7.98)$$

$$Ax = b, \quad (7.99)$$

$$x \geq 0, \quad (7.100)$$

$$x_j \in \mathbb{Z}, \quad j = \overline{1, n_1}, \quad (7.101)$$

де $n_1 \leq n$, $\mathbb{Z}$ – множина цілих чисел.

Якщо у задачі (7.98)–(7.101) $n_1 = n$, то вона буде задачею повністю цілочислового лінійного програмування (ПЦЛП), у протилежному випадку ($n_1 < n$) – задачею частково цілочислового лінійного програмування (ЧЦЛП).

Одним із ефективних методів розв'язування задач ЦЛП є метод Гоморі, який полягає у послідовному відокремленні від допустимої множини нецілочислової задачі ЛП (7.98)–(7.100) підмножин, які не містять точок з цілими

координатами. Таке відокремлення здійснюється введенням у задачу деяких додаткових обмежень.

Опишемо алгоритм методу Гоморі.

1. Розв'язування за допомогою симплекс-методу задачі (7.98)–(7.100) і знаходження оптимального розв'язку x^* (якщо він існує). Якщо для x^* виконується обмеження (7.101), то задача (7.98)–(7.101) розв'язана. У протилежному випадку здійснюється перехід до наступного кроку.

2. Вибір серед нецілих компонент оптимального плану задачі (7.98)–(7.100) компоненти, яка має найбільше дробове значення, наприклад компоненти з номером i (тобто x_i), та побудова на основі рядка, що відповідає змінній x_i , останньої симплекс-таблиці (у якій знайдений оптимальний розв'язок задачі (7.98)–(7.100)), додаткового обмеження. Вигляд цього обмеження різний для задач ПЦЛП і ЧЦЛП.

У випадку задачі ПЦЛП додаткове обмеження має вигляд

$$\sum_{j=1}^n \{a_{ij}^*\} x_j \geq \{b_i^*\}, \quad (7.102)$$

де $b_i^*, a_{ij}^*, j = \overline{1, n}$ – елементи рядка, що відповідає змінній x_i , останньої симплекс-таблиці для задачі (7.98)–(7.100).

У випадку задачі ЧЦЛП додаткове обмеження має вигляд

$$\sum_{j=1}^n \gamma_{ij} x_j \geq \{b_i^*\}, \quad (7.103)$$

де величини γ_{ij} визначаються так:

а) для змінних x_j , які можуть набувати нецілочислових значень,

$$\gamma_{ij} = \begin{cases} a_{ij}^*, & a_{ij}^* \geq 0, \\ \frac{\{b_i^*\}}{1 - \{b_i^*\}} |a_{ij}^*|, & a_{ij}^* < 0; \end{cases} \quad (7.104)$$

б) для змінних x_j , які можуть набувати тільки цілочислових значень,

$$\gamma_{ij} = \begin{cases} \{a_{ij}^*\}, & \{a_{ij}^*\} \leq \{b_i^*\}, \\ \frac{\{b_i^*\}}{1 - \{b_i^*\}} (1 - \{a_{ij}^*\}), & \{a_{ij}^*\} > \{b_i^*\}. \end{cases} \quad (7.105)$$

3. Розв'язування за допомогою симплекс-методу задачі ЛП, одержаної у результаті приєднання до задачі (7.98)–(7.100), записаній на основі останньої симплекс-таблиці, побудованого додаткового обмеження. Якщо отриманий оптимальний розв'язок задовольняє обмеження (7.101), то задача (7.98)–(7.101) розв'язана. У протилежному випадку здійснюється перехід до пункту 2.

Зауваження 7.13. Якщо задача ЦЛП записана не в канонічній формі (наприклад, як задача (B.8)), то перехід до канонічної форми не гарантує цілочисловості допоміжних змінних. У загальному випадку канонічна задача буде задачею ЧЦЛП. Але, якщо всі $a_{ij}, b_i \in \mathbb{Z}$, то допоміжні змінні будуть цілочисловими.

Зауваження 7.14. Якщо не має розв'язку задача (7.98)–(7.100), то не буде мати розв'язку і задача (7.98)–(7.101).

Наведемо деякі приклади задач ЦЛП.

Приклад 7.25.

$$\begin{cases} x_1 + 2x_2 \rightarrow \max, \\ 2x_1 + 2x_2 + x_3 = 13, \\ x_1 - x_2 + x_4 = 6, \\ x_j \geq 0, j = \overline{1,4}, \\ x_j \in \mathbb{Z}, j = \overline{1,4}. \end{cases} \quad (7.106)$$

◀ Задача (7.106) є задачею ПЦЛП. Розв'язуючи цю задачу без умови цілочисловості, одержимо оптимальний розв'язок: $\tilde{x} = (0, 13/2, 0, 25/2)^T$, $f(\tilde{x}) = 0$ (табл. 7.19). Звичайно, цей розв'язок не оптимальний для задачі (7.106), оскільки $x_2 = 13/2$, $x_4 = 25/2$. Дробові частини для x_2 і x_4 однакові, тому додаткове обмеження можна формувати для кожної з цих компонент. Побудуємо таке обмеження для x_2 . Із таблиці 7.19 випливає, що $x_1 + x_2 + 1/2x_3 = 13/2$. Обмеження (7.102) для попередньої рівності має вигляд: $1/2x_3 \geq 1/2$ або $x_3 \geq 1$. Далі розв'язуємо задачу ЛП

$$\begin{cases} x_1 + 2x_2 \rightarrow \max, \\ x_1 + x_2 + 1/2x_3 = 13/2, \\ 2x_1 + 1/2x_3 + x_4 = 25/2, \\ x_3 - y = 1, \\ x_j \geq 0, j = \overline{1,4}, y \geq 0. \end{cases} \quad (7.107)$$

Увівши у задачу (7.107) штучну змінну z , перейдемо від задачі (7.107) до М-задачі

$$\begin{cases} x_1 + 2x_2 - Mz \rightarrow \max, \\ x_1 + x_2 + 1/2x_3 = 13/2, \\ 2x_1 + 1/2x_3 + x_4 = 25/2, \\ x_3 - y + z = 1, \\ x_j \geq 0, j = \overline{1,4}, \\ y \geq 0, z \geq 0. \end{cases} \quad (7.108)$$

Процес розв'язування М-задачі (7.108) зафіксований у табл. 7.20, з якої знаходимо розв'язок задачі (7.107): $x^* = (0, 6, 1, 12)^T$, $f_{\max}^* = 12$. Оскільки цей розв'язок цілочисловий, то він і є оптимальним розв'язком вихідної задачі (7.106). ►

Таблиця 7.19

i	бз	$c_i^{\bar{0}}$	$x_i^{\bar{0}}$	1	2	0	0
				x_1	x_2	x_3	x_4
1	x_3	0	13	2	2	1	0
2	x_4	0	6	1	-1	0	1
$m+1$	σ_j		0	-1	-2	0	0
1	x_2	2	$\frac{13}{2}$	1	1	$\frac{1}{2}$	0
2	x_4	0	$\frac{25}{2}$	2	0	$\frac{1}{2}$	1
$m+1$	σ_j		13	1	0	1	0

Таблиця 7.20

i	бз	$c_i^{\bar{o}}$	$x_i^{\bar{o}}$	1	2	0	0	0	$-M$
				x_1	x_2	x_3	x_4	y	z
1	x_2	2	$\frac{13}{2}$	1	1	$\frac{1}{2}$	0	0	0
2	x_4	0	$\frac{25}{2}$	2	0	$\frac{1}{2}$	1	0	0
3	z	$-M$	1	0	0	1	0	-1	1
$m+1$	σ_j		13	1	0	1	0	0	0
$m+2$			-1	0	0	-1	0	1	0
1	x_2	2	6	1	1	0	0	$\frac{1}{2}$	
2	x_4	0	12	2	0	0	1	$\frac{1}{2}$	
3	x_3	0	1	0	0	1	0	-1	
$m+1$	σ_j		12	1	0	0	0	1	

Приклад .7.26.

$$\begin{cases} x_1 + 2x_2 \rightarrow \max, \\ 2x_1 + 2x_2 + x_3 = 13, \\ x_1 - x_2 + x_4 = 19/3, \\ x_j \geq 0, \quad j = \overline{1,4}, \\ x_j \in \mathbb{Z}, \quad j = \overline{1,2}. \end{cases} \quad (7.109)$$

◀ Задача (7.109) є задачею ЧЦЛП, оскільки змінні x_3 і x_4 можуть набувати нецілочислових значень. Спочатку розв'яжемо задачу ЛП.

$$\begin{cases} x_1 + 2x_2 \rightarrow \max, \\ 2x_1 + 2x_2 + x_3 = 13, \\ x_1 - x_2 + x_4 = 19/3, \\ x_j \geq 0, j = \overline{1,4}. \end{cases} \quad (7.110)$$

Із таблиці 7.21 випливає, що оптимальним вектором для задачі (7.110) є вектор $\tilde{x} = (0, 13/2, 0, 77/6)^T$. Оскільки змінна x_2 не є цілочисловою, то треба сформулювати додаткове обмеження (7.103). Зі співвідношень (7.105) і (7.104) випливає, що додаткове обмеження має вигляд

$$0 \cdot x_1 + 0 \cdot x_2 + 1/2 x_3 + 0 \cdot x_4 \geq 1/2 \text{ або } x_3 \geq 1 \text{ або } x_3 - y = 1,$$

де $y \geq 0$ – допоміжна змінна.

Далі розв'яжемо М-методом задачу ЛП

$$\begin{cases} x_1 + 2x_2 \rightarrow \max, \\ x_1 + x_2 + 1/2 x_3 = 13/2, \\ 2x_1 + 1/2 x_3 + x_4 = 77/6, \\ x_3 - y = 1, \\ x_j \geq 0, j = \overline{1,4}, y \geq 0. \end{cases} \quad (7.111)$$

Процес розв'язування задачі (7.111) відображений у таблиці 7.22, з якої випливає, що $x^* = (0, 6, 1, 37/3)^T$ є оптимальним розв'язком задачі (7.109). При цьому оптимальне значення функції $f_{\max}^* = 12$. ►

Таблица 7.21

i	бз	c_i^{δ}	x_i^{δ}	1	2	0	0
				x_1	x_2	x_3	x_4
1	x_3	0	13	2	2	1	0
2	x_4	0	$\frac{19}{3}$	1	-1	0	1
$m+1$	σ_j		0	-1	-2	0	0
1	x_2	2	$\frac{13}{2}$	1	1	$\frac{1}{2}$	0
2	x_4	0	$\frac{77}{6}$	2	0	$\frac{1}{2}$	1
$m+1$	σ_j		13	1	0	1	0

Таблица 7.22

i	Бз	c_i^{δ}	x_i^{δ}	1	2	0	0	0	$-M$
				x_1	x_2	x_3	x_4	y	z
1	x_2	2	$\frac{13}{2}$	1	1	$\frac{1}{2}$	0	0	0
2	x_4	0	$\frac{77}{6}$	2	0	$\frac{1}{2}$	1	0	0
3	z	$-M$	1	0	0	1	0	-1	1
$m+1$	σ_j		13	1	0	1	0	0	0
$m+2$			-1	0	0	-1	0	1	0
1	x_2	2	6	1	1	0	0	$\frac{1}{2}$	
2	x_4	0	$\frac{37}{3}$	2	0	0	1	$\frac{1}{2}$	
3	x_3	0	1	0	0	1	0	-1	
$m+1$	σ_j		12	1	0	0	0	1	

Розділ 8. НЕЛІНІЙНЕ ПРОГРАМУВАННЯ

8.1. Нелінійні оптимізаційні моделі та їх формалізація задачами нелінійного програмування

Задачі лінійного програмування як лінійні моделі планування економічних (та інших) процесів складають всього-на-всього вузький клас задач математичного програмування і прийняття рішень у цілому. Багато реальних явищ і процесів не можуть бути адекватно описані лінійними моделями, оскільки ті чи інші характеристики мають явно нелінійний характер. До того ж, у певному сенсі нелінійність – це завжди узагальнення лінійності, а лінійність – можливий наслідок нелінійності. Практично кожену модель, яка описується задачею лінійного програмування, можна розширити чи модифікувати задачею нелінійного програмування. Наприклад, оптимізаційні економіко-математичні моделі (B.2)–(B.4), (B.9)–(B.11) нелінійні, оскільки в усіх цих моделях цільова функція нелінійна, хоча вони можуть бути також узагальненнями деяких лінійних моделей. Щось подібне можна стверджувати і про лінійні моделі (B.5)–(B.8), від яких легко перейти до нелінійних, змінивши, наприклад, цільові функції. У будь-якому разі, реальний світ, тобто реальні об'єкти дослідження, зокрема економічні явища, процеси тощо, досить складні та нелінійні, тому їх формалізація лінійними моделями є надто ідеалізованою та може бути використана дослідниками лише для початкового (досить загального) уявлення про досліджувані об'єкти. Більш досконалі знання про ці об'єкти можна отримати тільки за допомогою нелінійних моделей, у тому числі нелінійних оптимізаційних моделей, що формалізуються задачами нелінійного програмування.

Отже, для дослідження складних явищ і процесів потрібний інструментарій методів і моделей нелінійної математики, зокрема нелінійного програмування. Зазначимо також, що сучасний рівень розвитку комп'ютерно-інформаційних технологій дозволяє ефективно використовувати нелінійні методи та моделі практично в усіх галузях знань і практичної діяльності.

Що стосується математичної формалізації задач нелінійного програмування, то вони (як і задачі лінійного програмування) можуть бути представлені у різних еквівалентних формах запису. Перехід від однієї форми запису до іншої можна здійснити за допомогою еквівалентних перетворень, про які вже йшлося у попередньому розділі. Інакше кажучи, загальну задачу нелінійного програмування можна записати, наприклад, у вигляді

$$\begin{cases} f(x) \rightarrow \max, \\ g_i(x) \leq b_i, \quad i = \overline{1, m}, \\ x_j \geq 0, \quad j = \overline{1, n}, \end{cases} \quad (8.1)$$

де хоча б одна із функцій $f(x)$, $g_i(x)$ ($i = \overline{1, m}$) нелінійна, b_i ($i = \overline{1, m}$) – задані сталі величини, $x = (x_1, \dots, x_n)^T \in \mathbb{R}^n$.

Очевидно, що від задачі нелінійного (математичного) програмування (6.2)–(6.5) чи (6.22) можна перейти до задачі (8.1) та навпаки, тому надалі у цьому розділі задачу (8.1) як загальну задачу нелінійного програмування (НЛП) будемо вважати базовою для вивчення НЛП.

Що ж до методів розв'язування задач НЛП, то зазначимо, що аналогічного до симплекс-методу у ЛП універсального методу розв'язування задач НЛП немає. Крім того, більшість методів розв'язування задач НЛП, як правило, наближені та

визначають тільки локальні розв'язки, про що і свідчать методи, викладені у розділі 6 цієї книги. Пошук глобального екстремуму – набагато складніша задача. Очевидно, що глобальний розв'язок задачі (8.1) є і локальним. Обернене твердження неправильне.

Нижче розглянемо тільки деякі спеціальні класи задач НЛП і найпростіші методи їх розв'язування.

8.2. Графічний метод розв'язування задач нелінійного програмування

Найбільш ефективно графічний метод використовується для розв'язування задач НЛП з обмеженнями-нерівностями, якщо останні містять не більше двох вільних змінних. Інколи вдається розв'язати задачу НЛП графічним методом і у випадку трьох змінних, але для спрощення геометричної інтерпретації задачі НЛП будемо розглядати тільки задачі НЛП з двома змінними.

Отже, розглянемо задачу НЛП

$$\begin{cases} f(x_1, x_2) \rightarrow \max, \\ g_i(x_1, x_2) \leq b_i, \quad i = \overline{1, m}, \end{cases} \quad (8.2)$$

яка є частинним випадком задачі (8.1).

Знаходження оптимальних розв'язків задачі (8.2) зводиться до визначення таких точок області допустимих розв'язків $\tilde{X} = \{x \in \mathbb{R}^n / g_i(x_1, x_2) \leq b_i, i = \overline{1, m}\}$, через які проходить лінія найвищого рівня (у задачі НЛП на мінімум – лінія найнижчого рівня) $f(x_1, x_2) = h$, де h – деяка постійна величина. Указані точки можуть знаходитись як на границі, так і всередині області допустимих розв'язків.

Процес знаходження оптимального розв'язку задачі (8.2) графічним методом розбивається на ряд етапів

(нагадаємо, що як і в ЛП, поняття «оптимальний розв'язок», а не просто «розв'язок», використовується у зв'язку з прикладними поняттями «оптимальне планування» чи «оптимальне програмування»). Конкретизуємо їх.

1. Побудова на площині координат x_1Ox_2 допустимої множини $\tilde{X}$ задачі (8.2). Якщо $\tilde{X} = \emptyset$, то задача НЛП (8.2) не має розв'язку.

2. Побудова на площині координат x_1Ox_2 ліній рівня $f(x_1, x_2) = h$, де h – довільна стала.

3. Установлення лінії найвищого рівня (якщо задача НЛП є задачею на мінімум, то встановлення лінії найнижчого рівня), тобто лінії, яка містить оптимальні розв'язки, або встановлення факту неіснування такої лінії, а отже, і нерозв'язності задачі у зв'язку з тим, що цільова функція необмежена зверху (у задачі на мінімум знизу) на допустимій множині.

4. Уточнення оптимального розв'язку (чи розв'язків у випадку, коли їх більше, ніж один) задачі, тобто точки (чи точок), яка належить (які належать) лінії найвищого (або найнижчого для задачі на мінімум) рівня, та обчислення оптимального значення цільової функції.

Розглянемо ряд прикладів.

Приклад 8.1. Знайти максимальне та мінімальне значення функції

$$f(x_1, x_2) = (x_1 - 2)^2 + (x_2 - 3)^2 \quad (8.3)$$

на множині $\tilde{X}^{(1)}$, яка задається системою нерівностей

$$\begin{cases} x_1 + 2x_2 \leq 12, \\ x_1 + x_2 \leq 9, \\ x_j \geq 0, \quad j = \overline{1, 2}. \end{cases} \quad (8.4)$$

◀ Область допустимих розв'язків $\tilde{X}^{(1)}$ задачі (8.3), (8.4) є чотирикутником $OABC$ (рис. 8.1), а рівняння лінії рівня $(x_1 - 2)^2 + (x_2 - 3)^2 = h$ – рівнянням кола з центром у точці $E(2, 3)$ і радіусом $\sqrt{h}$ ($h \geq 0$). Оскільки центр кола розміщений всередині чотирикутника $OABC$, то колом мінімального радіуса (лінією найнижчого рівня), яке має спільну точку з чотирикутником $OABC$, є коло з центром у точці $E(2, 3)$ і радіусом $\sqrt{h} = 0$, тобто коло, вироджене у точку E . Зрозуміло, що мінімальне значення функції (8.3) на множині (8.4) $f_{\min}^* = f(2, 3) = (2 - 2)^2 + (3 - 3)^2 = 0$.

Колом максимального радіуса з центром у точці $E(2, 3)$, яке має спільну точку з чотирикутником $OABC$ (лінією найвищого рівня), буде коло з радіусом $\sqrt{h} = EC$. Отже, максимальне значення функції (8.3) досягається у точці $C(9, 0)$, тобто $f_{\max}^* = f(9, 0) = (9 - 2)^2 + (0 - 3)^2 = 58$. ►

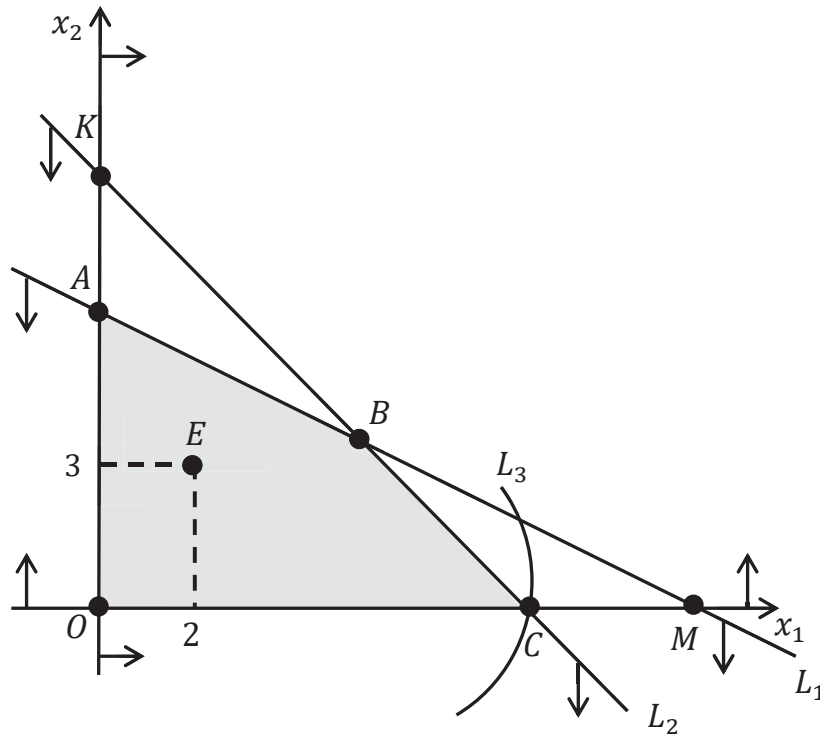


Рис. 8.1. Ілюстрація до прикладу 8.1 (точки O, A, B, C, K, M, E та лінії L_1, L_2, L_3 такі: $O(0,0)$, $A(0,6)$, $B(6,3)$, $C(9,0)$, $K(0,9)$, $M(12,0)$, $E(2,3)$, $L_1: x_1 + 2x_2 = 12$, $L_2: x_1 + x_2 = 9$, L_3 : фрагмент кола $f(x_1, x_2) = h$, що проходить через точку C)

Приклад 8.2. Знайти максимальне та мінімальне значення функції

$$f(x_1, x_2) = 2(x_1 - 5)^2 + (x_2 - 7)^2 \quad (8.5)$$

на допустимій множині (8.4).

◀ Нехай, як і раніше, h – довільна стала величина ($h \geq 0$). Тоді лініями рівня для даної задачі будуть еліпси, канонічні рівняння яких мають вигляд

$$\frac{(x_1 - 5)^2}{h/2} + \frac{(x_2 - 7)^2}{h} = 1.$$

Центром кожного з цих еліпсів буде точка $E(5,7)$, а велика і мала півосі мають відповідно довжини $\sqrt{h/2}$ і $\sqrt{h}$.

Точка $E(5,7)$ знаходиться за межами чотирикутника $OABC$, тому лінією найнижчого рівня, яка має спільну точку з чотирикутником $OABC$, буде еліпс із точкою дотику P (рис. 8.2). Для визначення координат точки P використаємо умову рівності $k_1 = k_2$, де k_1 – кутовий коефіцієнт прямої, яка проходить через точки A, B, M (прямої L_1), а k_2 – кутовий коефіцієнт дотичної до еліпса у точці P . Очевидно, що $k_2 = -1/2$. Позначимо через x_1^* і x_2^* – невідомі координати точки P , а через $\varphi(x_1, x_2) = (x_1 - 5)^2 / (h/2) + (x_2 - 7)^2 / h - 1 = 0$ – неявну функцію залежності змінної x_2 від x_1 . Оскільки

$$\frac{dx_2}{dx_1} = -\frac{\partial \varphi / \partial x_1}{\partial \varphi / \partial x_2} = -\frac{2(x_1 - 5)}{x_2 - 7},$$

$$\text{то } k_2 = -2(x_1^* - 5) / (x_2^* - 7).$$

Враховуючи останню рівність, а також той факт, що точка лежить на прямій, яка проходить через точки A, B, M , випишемо для знаходження невідомих x_1^* і x_2^* таку систему рівнянь:

$$\begin{cases} -2(x_1^* - 5) / (x_2^* - 7) = -1/2, \\ x_1^* + 2x_2^* = 12. \end{cases}$$

Розв'язавши цю систему, отримаємо $x_1^* = 38/9$, $x_2^* = 35/9$, $f_{\min}^* = 2(38/9 - 5)^2 + (35/9 - 7)^2 = 98/9$.

Лінією найвищого рівня, яка має спільну точку з чотирикутником $OABC$ (рис. 8.2), буде еліпс, який проходить через точку $O(0,0)$ або точку $C(9,0)$. Оскільки $f(0,0) = 99$, а

$f(9,0) = 81$, то лінія найвищого рівня проходить через точку $O(0,0)$ і $f_{\max}^* = 99$. ►

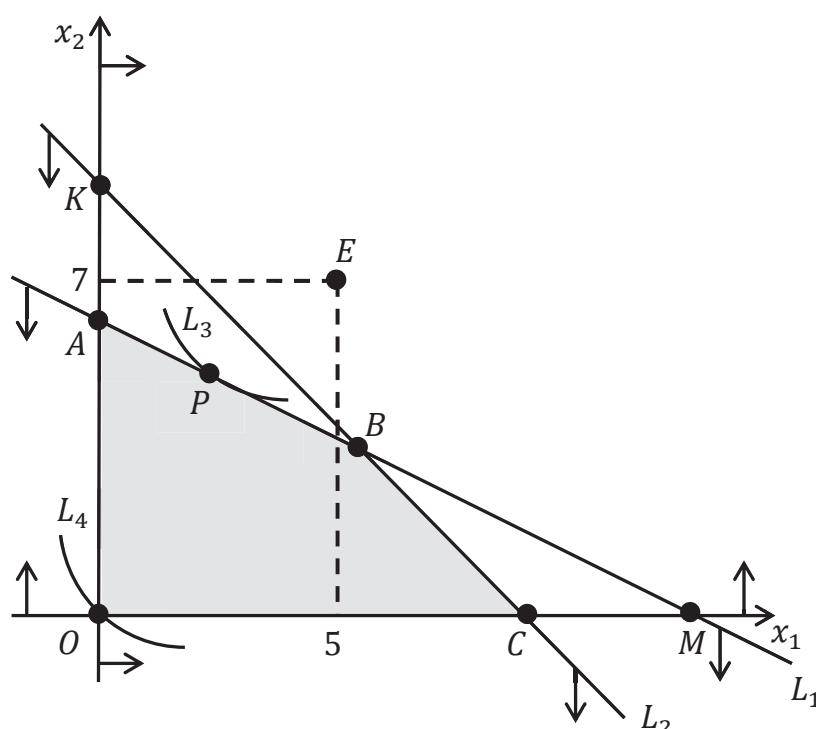


Рис. 8.2. Ілюстрація до прикладу 8.2 (точки O, A, B, C, K, M, E та лінії L_1, L_2, L_3, L_4 такі: $O(0,0)$, $A(0,6)$, $B(6,3)$, $C(9,0)$, $K(0,9)$, $M(12,0)$, $E(5,7)$, $L_1: x_1 + 2x_2 = 12$, $L_2: x_1 + x_2 = 9$, L_3 : фрагмент еліпса $f(x_1, x_2) = h$, що проходить через точку P , L_4 : фрагмент еліпса $f(x_1, x_2) = h$, що проходить через точку O)

Приклад 8.3. Знайти максимальне та мінімальне значення функції

$$f(x_1, x_2) = (x_1 - 7)(x_2 - 1) \quad (8.6)$$

на допустимій множині (8.4).

◀ Лініями рівня для функції (8.6) будуть гіперболи $(x_1 - 7)(x_2 - 1) = h$, асимптотами яких є прямі $x_1 = 7$ і $x_2 = 1$ (рис. 8.3). Зі зростанням h гіперболи віддаляються від точки

перетину асимптот $E(7,1)$. Лінією найнижчого рівня, очевидно, буде гіпербола, яка вироджується у точку $E(7,1)$, що є внутрішньою точкою чотирикутника $OABC$. Отже, $f_{\min}^* = f(7,1) = (7-7)(1-1) = 0$. Лінією найвищого рівня буде гіпербола (її нижня вітка), яка проходить через точку $O(0,0)$, тому $f_{\max}^* = f(0,0) = (0-7)(0-1) = 7$. ►

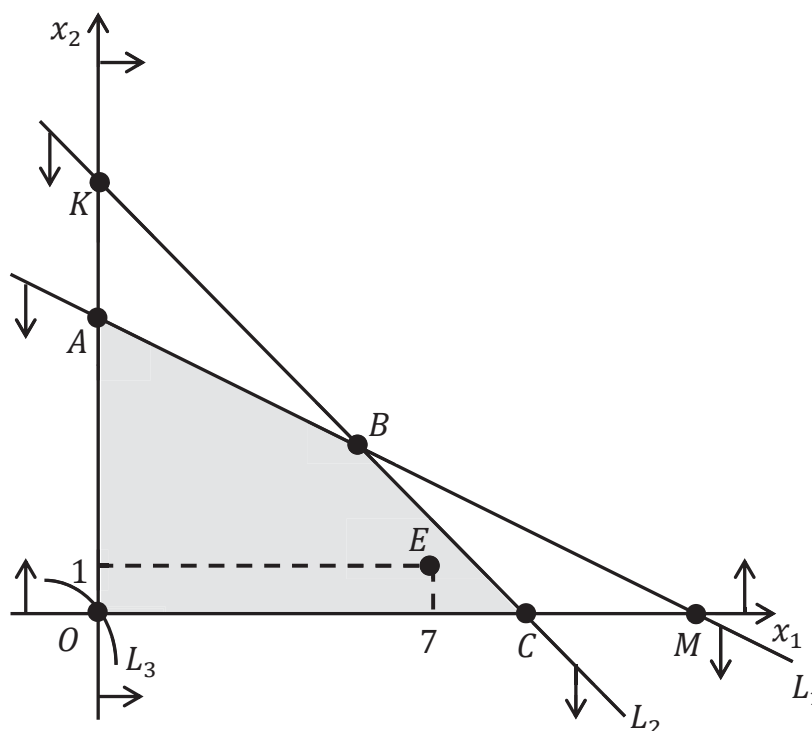


Рис. 8.3. Ілюстрація до прикладу 8.3 (точки O, A, B, C, K, M, E та лінії L_1, L_2, L_3 такі: $O(0,0)$, $A(0,6)$, $B(6,3)$, $C(9,0)$, $K(0,9)$, $M(12,0)$, $E(7,1)$, $L_1: x_1 + 2x_2 = 12$, $L_2: x_1 + x_2 = 9$, L_3 : фрагмент гіперболи $f(x_1, x_2) = h$, яка проходить через точку O)

Приклад 8.4. Знайти максимальне значення функції

$$f(x_1, x_2) = 3x_1 + 4x_2 \quad (8.7)$$

на множині $\tilde{X}^{(2)}$, яка задається системою нерівностей

$$\begin{cases} x_1^2 + x_2^2 \leq 25, \\ x_1 x_2 \geq 4, \\ x_j \geq 0, j = \overline{1, 2}. \end{cases} \quad (8.8)$$

◀ Множина $\tilde{X}^{(2)}$, яка описана співвідношеннями (8.8), зображена на рис. 8.4. Лініями рівня для функції (8.7) є прямі $3x_1 + 4x_2 = h$, де h – довільна стала величина. Лінією найвищого рівня є пряма, яка дотикається до кола $x_1^2 + x_2^2 = 25$ у точці $E(x_1^*, x_2^*)$. Як і раніше, для визначення координат точки E скористаємося рівністю $k_1 = k_2$, де k_1 – кутовий коефіцієнт прямої $3x_1 + 4x_2 = h$, яка проходить через точку E , а k_2 – кутовий коефіцієнт дотичної до кола $x_1^2 + x_2^2 = 25$ у точці E . Оскільки для неявної функції $\varphi(x_1, x_2) = x_1^2 + x_2^2 - 25 = 0$ похідна

$$\frac{dx_2}{dx_1} = -\frac{\partial \varphi / \partial x_1}{\partial \varphi / \partial x_2} = -\frac{2x_1}{2x_2} = -\frac{x_1}{x_2},$$

а $k_1 = -3/4$, то для знаходження x_1^* і x_2^* маємо систему рівнянь

$$\begin{cases} -\frac{x_1^*}{x_2^*} = -\frac{3}{4}, \\ (x_1^*)^2 + (x_2^*)^2 = 25. \end{cases}$$

Розв'язавши дану систему, одержимо $x_1^* = 3$, $x_2^* = 4$,
 $f_{\max}^* = f(3, 4) = 3 \cdot 3 + 4 \cdot 4 = 25$. ►

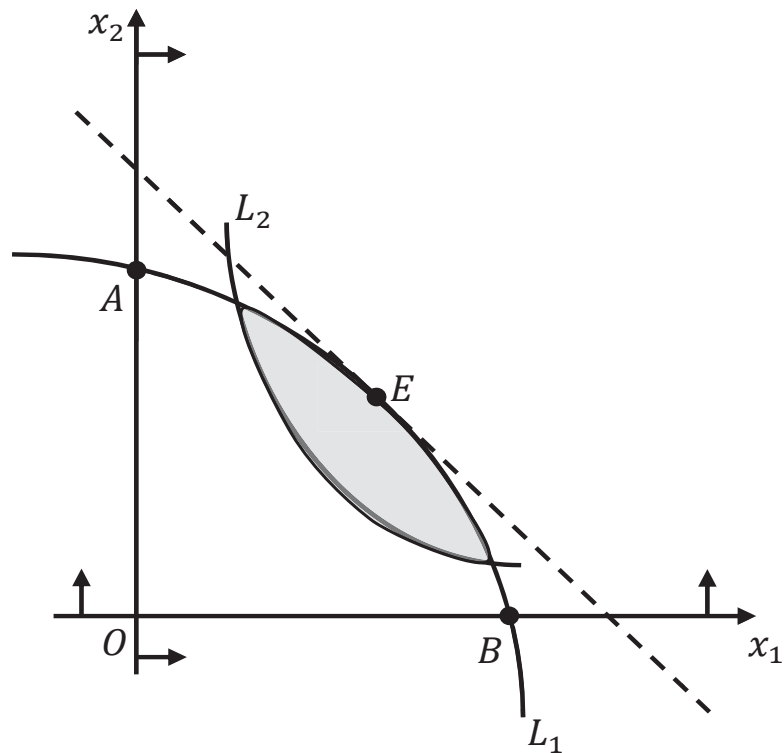


Рис. 8.4. Ілюстрація до прикладу 8.4 (точки O, A, B та лінії L_1, L_2 такі: $O(0,0)$, $A(0,5)$, $B(5,0)$, $L_1 : x_1^2 + x_2^2 = 25$, $L_2 : x_1x_2 = 4$)

8.3. Задачі дробово-лінійного програмування

Деякі задачі НЛП зводяться до задач ЛП. До таких задач належить, наприклад, задача дробово-лінійного програмування (В.9), якою формалізується модель максимізації рентабельності виробництва. Розглянемо задачу дробово-лінійного програмування у більш загальній формі:

$$\left\{ \begin{array}{l} \left(\sum_{j=1}^n c_j x_j + c_0 \right) / \left(\sum_{j=1}^n d_j x_j + d_0 \right) \rightarrow \max, \\ \sum_{j=1}^n a_{ij} x_j \leq b_i, \quad i = \overline{1, m_1}, \\ \sum_{j=1}^n a_{ij} x_j = b_i, \quad i = \overline{m_1 + 1, m}, \\ x_j \geq 0, \quad j = \overline{1, n}. \end{array} \right. \quad (8.9)$$

Будемо вважати, що на допустимій множині задачі (8.9) знаменник цільової функції

$$\sum_{j=1}^n d_j x_j + d_0 > 0.$$

Введемо нові змінні:

$$y_0 = 1 / \left(\sum_{j=1}^n d_j x_j + d_0 \right), \quad y_j = y_0 x_j, \quad j = \overline{1, n}.$$

Очевидно, що $y_0 > 0$, $y_j \geq 0$, $j = \overline{1, n}$.

У нових змінних задача (8.9) набуває вигляду

$$\left\{ \begin{array}{l} \sum_{j=0}^n c_j y_j \rightarrow \max, \\ \sum_{j=1}^n a_{ij} y_j - b_i y_0 \leq 0, \quad i = \overline{1, m_1}, \\ \sum_{j=1}^n a_{ij} y_j - b_i y_0 = 0, \quad i = \overline{m_1 + 1, m}, \\ \sum_{j=0}^n d_j y_j = 1, \\ y_j \geq 0, \quad j = \overline{0, n}. \end{array} \right. \quad (8.10)$$

Задача (8.10) є задачею ЛП, тому її можна розв'язати симплекс-методом. Якщо $y^* = (y_0^*, \dots, y_n^*)^T$ – оптимальний розв'язок задачі (8.10), а $\bar{f}^*$ – оптимальне значення її цільової функції, то оптимальний розв'язок задачі (8.9) знаходиться так: $x_j^* = y_j^* / y_0^*$, $j = \overline{1, n}$; $f^* = \bar{f}^*$.

Та коли $y_0^* = 0$, то це означає, що задача (8.9) не має оптимального розв'язку, оскільки її цільова функція необмежена зверху на допустимій множині.

У випадку, коли задача (8.10) має безліч оптимальних розв'язків, для знаходження оптимальних розв'язків задачі (8.9) потрібно розглядати тільки ті розв'язки задачі (8.10), у яких $y_0^* \neq 0$.

Продемонструємо процес розв'язування задачі дробово-лінійного програмування на наступному прикладі.

Приклад 8.5.

$$\begin{cases} \frac{3x_1 + x_2}{x_1 + x_2} \rightarrow \max, \\ x_1 + 2x_2 \leq 4, \\ 5x_1 + 3x_2 \leq 15, \\ x_j \geq 0, \quad j = \overline{1, 2}. \end{cases} \quad (8.11)$$

◀ Введемо нові змінні $y_0 = 1/(x_1 + x_2)$, $y_j = y_0 x_j$; $j = \overline{1, 2}$.

Відповідна до (8.11) задача ЛП матиме вигляд

$$\begin{cases} 3y_1 + y_2 \rightarrow \max, \\ y_1 + 2y_2 - 4y_0 \leq 0, \\ 5y_1 + 3y_2 - 15y_0 \leq 0, \\ y_1 + y_2 = 1, \\ y_j \geq 0, \quad j = \overline{0, 2}. \end{cases} \quad (8.12)$$

Запишемо задачу (8.12) у канонічній формі

$$\begin{cases} 3y_1 + y_2 \rightarrow \max, \\ y_1 + 2y_2 - 4y_0 + y_3 = 0, \\ 5y_1 + 3y_2 - 15y_0 + y_4 = 0, \\ y_1 + y_2 = 1, \\ y_j \geq 0, \quad j = \overline{0, 4}. \end{cases} \quad (8.13)$$

Оскільки канонічна задача (8.13) не має повного початкового базису, то поставимо їй у відповідність і розв'яжемо наступну розширену М-задачу ЛП:

$$\begin{cases} 3y_1 + y_2 - My_5 \rightarrow \max, \\ -4y_0 + y_1 + 2y_2 + y_3 = 0, \\ -15y_0 + 5y_1 + 3y_2 + y_4 = 0, \\ y_1 + y_2 + y_5 = 1, \\ y_j \geq 0, \quad j = \overline{0, 5}. \end{cases} \quad (8.14)$$

Процес розв'язування задачі (8.14) зафіксований у таблиці 8.1. Таблиця 8.1 складається з п'яти симплекс-таблиць. В останній із них знайдений оптимальний розв'язок задачі (8.13): $y^* = (y_0^*, y_1^*, y_2^*, y_3^*, y_4^*)^T = (1/3, 1, 0, 1/3, 0)^T$. Оптимальним розв'язком вихідної задачі дробово-лінійного програмування є $x^* = (x_1^*, x_2^*)^T = (y_1^*/y_0^*, y_2^*/y_0^*)^T = (3, 0)^T$, $f_{\max}^* = (3x_1^* + x_2^*)/(x_1^* + x_2^*) = 3$. ►

Таблица 8.1

i	ζ_3	$c_i^{\bar{\sigma}}$	$y_i^{\bar{\sigma}}$	0	3	1	0	0	$-M$
				y_0	y_1	y_2	y_3	y_4	y_5
1	y_3	0	0	-4	1	2	1	0	0
2	y_4	0	0	-15	5	3	0	1	0
3	y_5	$-M$	1	0	1	1	0	0	1
$m+1$	σ_j		0	0	-3	-1	0	0	0
$m+2$			-1	0	-1	-1	0	0	0
1	y_1	3	0	-4	1	2	1	0	0
2	y_4	0	0	5	0	-7	-5	1	0
3	y_5	$-M$	1	4	0	-1	-1	0	1
$m+1$	σ_j		0	-12	0	5	3	0	0
$m+2$			-1	-4	0	1	1	0	0
1	y_1	3	0	0	1	$-\frac{18}{5}$	-3	$\frac{4}{5}$	0
2	y_0	0	0	1	0	$-\frac{7}{5}$	-1	$\frac{1}{5}$	0
3	y_5	$-M$	1	0	0	$\frac{23}{5}$	3	$-\frac{4}{5}$	1
$m+1$	σ_j		0	0	0	$-\frac{59}{5}$	-9	$\frac{12}{5}$	0
$m+2$			-1	0	0	$-\frac{23}{5}$	-3	$\frac{4}{5}$	0

Продовження таблиці 8.1

1	y_1	3	$\frac{18}{23}$	0	1	0	$-\frac{15}{23}$	$\frac{4}{23}$
2	y_0	0	$\frac{7}{23}$	1	0	0	$-\frac{2}{23}$	$-\frac{1}{23}$
3	y_2	1	$\frac{5}{23}$	0	0	1	$\frac{15}{23}$	$-\frac{4}{23}$
$m+1$	σ_j		$\frac{59}{23}$	0	0	0	$-\frac{30}{23}$	$\frac{8}{23}$
1	y_1	3	1	0	1	1	0	0
2	y_0	0	$\frac{1}{3}$	1	0	$\frac{2}{15}$	0	$-\frac{1}{15}$
3	y_3	0	$\frac{1}{3}$	0	0	$\frac{23}{15}$	1	$-\frac{4}{15}$
$m+1$	σ_j		3	0	0	2	0	0

8.4. Опукле та квадратичне програмування

Важливими підкласами задач НЛП є опукле та квадратичне програмування. Якщо у задачі (8.1) функції $(-f(x))$ та $g_i(x)$ ($i = \overline{1, m}$) опуклі (опуклі вниз), то задача (8.1) буде задачею опуклого або угнутого програмування (у зв'язку з тим, що функція $f(x)$ угнута, тобто опукла вгору). Перш ніж сформулювати центральний результат теорії опуклого програмування, уточнимо деякі поняття.

Означення 8.1. Допустима множина задачі (8.1) задовольняє умову регулярності Слейтера, якщо існує хоча б одна точка $\tilde{x}$ із цієї множини, у якій виконуються строгі нерівності: $g_i(\tilde{x}) < b_i$, $i = \overline{1, m}$.

Виконання умови регулярності Слейтера (як і будь-якої іншої умови регулярності) для задачі (8.1) означає, що вона регулярна, тобто регулярною є її функція Лагранжа

$$L(x, \lambda) = f(x) + \sum_{i=1}^m \lambda_i (b_i - g_i(x)), \quad (8.15)$$

де $\lambda_1, \dots, \lambda_m$ – множники Лагранжа.

Зазначимо, що функція Лагранжа (8.15) задачі (8.1) угнута по x і опукла по λ (по λ функція (8.15) лінійна, тобто одночасно є опуклою та угнутою).

Означення 8.2. Точка $(x^*, \lambda^*)^T = (x_1^*, \dots, x_n^*, \lambda_1^*, \dots, \lambda_m^*)^T$ називається сідловою точкою функції Лагранжа (8.15) на множині $\mathbb{R}_{++} = \mathbb{R}_+^n \times \mathbb{R}_+^m$, якщо для всіх $x \in \mathbb{R}_+^n$ і $\lambda \in \mathbb{R}_+^m$ виконуються нерівності

$$L(x, \lambda^*) \leq L(x^*, \lambda^*) \leq L(x^*, \lambda). \quad (8.16)$$

Теорема 8.1 (теорема Куна-Таккера). Якщо у задачі (8.1) множина допустимих розв'язків задовольняє умову регулярності Слейтера, то точка $x^* \in \mathbb{R}_+^n$ буде оптимальним розв'язком цієї задачі тоді і тільки тоді, коли існує така точка $\lambda^* \in \mathbb{R}_+^m$, що $(x^*, \lambda^*)^T$ – сідлова точка функції Лагранжа (8.15) на множині $\mathbb{R}_{++}$.

Не зупиняючись на доведенні теореми 8.1 (його можна знайти у рекомендованій літературі), зазначимо лише, що вона повністю узгоджується з результатами, описаними у розділі 6.

У випадку, коли функції $f(x)$, $g_i(x)$ ($i = \overline{1, m}$) – неперервно диференційовні, можна виписати аналітичні умови існування сідлової точки, що має принципове значення для практичного розв'язування задач опуклого програмування. Сформулюємо ці результати у вигляді такої теореми.

Теорема 8.2. Нехай функції $f(x)$, $g_i(x)$ ($i = \overline{1, m}$) – неперервно диференційовні в області свого визначення. Для того, щоб точка $(x^*, \lambda^*)^T \in \mathbb{R}_{++}$ була сідловою точкою функції Лагранжа (8.15) задачі (8.1), необхідне й достатнє виконання таких умов:

$$\frac{\partial L^*}{\partial x_j} \leq 0, \quad j = \overline{1, n}, \quad (8.17)$$

$$x_j^* \frac{\partial L^*}{\partial x_j} = 0, \quad j = \overline{1, n}, \quad (8.18)$$

$$x_j^* \geq 0, \quad j = \overline{1, n}, \quad (8.19)$$

$$\frac{\partial L^*}{\partial \lambda_i} \geq 0, \quad i = \overline{1, m}, \quad (8.20)$$

$$\lambda_i^* \frac{\partial L^*}{\partial \lambda_i} = 0, \quad i = \overline{1, m}, \quad (8.21)$$

$$\lambda_i^* \geq 0, \quad i = \overline{1, m}, \quad (8.22)$$

де

$$\frac{\partial L^*}{\partial x_j} = \frac{\partial L(x^*, \lambda^*)}{\partial x_j}, \quad \frac{\partial L^*}{\partial \lambda_i} = \frac{\partial L(x^*, \lambda^*)}{\partial \lambda_i} \quad (j = \overline{1, n}; i = \overline{1, m}).$$

◀ **Необхідність.** Нехай $(x^*, \lambda^*)^T$ – сідлова точка функції Лагранжа (8.15). Тоді для всіх $(x^*, \lambda^*)^T \in \mathbb{R}_{++}$ виконуються нерівності (8.16), з яких випливає, що по x функція Лагранжа має максимум, а по λ – мінімум.

Інакше кажучи, $x^* \in \mathbb{R}_+^n$ є розв'язком задачі $L(x, \lambda) \rightarrow \max$, $x \in \mathbb{R}_+^n$, а $\lambda^* \in \mathbb{R}_+^m$ – розв'язком задачі $L(x, \lambda) \rightarrow \min$, $\lambda \in \mathbb{R}_+^m$ (ці задачі є аналогами задач (6.76), (6.77)), тому, якщо точка $(x^*, \lambda^*)^T \in \mathbb{R}_{++}$, то необхідною умовою екстремуму є дорівнюваність нулю частинних похідних від функції Лагранжа, тобто

$$\begin{cases} \frac{\partial L^*}{\partial x_j} = 0, & j = \overline{1, n}, \\ \frac{\partial L^*}{\partial \lambda_i} = 0, & i = \overline{1, m}, \end{cases} \quad (8.23)$$

а якщо x^* або λ^* знаходиться на одній з осей координат, то із теореми 6.5 випливає, що

$$\frac{\partial L^*}{\partial x_j} \leq 0, \text{ при } x_j^* = 0, \quad j = \overline{1, n}, \quad (8.24)$$

$$\frac{\partial L^*}{\partial \lambda_i} \geq 0, \text{ при } \lambda_i^* = 0, \quad i = \overline{1, m}. \quad (8.25)$$

Зі співвідношень (8.23)–(8.25) випливають умови (8.17)–(8.22), які є аналогами умов (6.79), (6.78). Необхідність доведена.

Достатність. Нехай виконуються умови (8.17)–(8.22). Треба довести, що точка $(x^*, \lambda^*)^T$ – сідлова точка функції Лагранжа (8.15).

Умови (8.17)–(8.22) для функції Лагранжа гарантують її угнутість по x і опуклість по λ для всіх $(x, \lambda)^T \in \mathbb{R}_{++}$ (а це означає, що функція (8.15) по x має максимум, а по λ – мінімум на множині $\mathbb{R}_{++}$). Справді, оскільки функція $L(x, \lambda)$ угнута по x , то із теореми 5.15, а точніше, із умови (5.14) випливає, що для $x, x^* \in \mathbb{R}_+^n$ та $\lambda, \lambda^* \in \mathbb{R}_+^m$

$$L(x, \lambda) \leq L(x^*, \lambda) + \left\langle \frac{\partial L(x^*, \lambda)}{\partial x}, x - x^* \right\rangle,$$

або із урахуванням (8.17)–(8.19) при $\lambda = \lambda^*$

$$L(x, \lambda^*) \leq L(x^*, \lambda^*).$$

Аналогічно встановлюється, що

$$L(x^*, \lambda^*) \leq L(x^*, \lambda).$$

Отже, існує точка $(x^*, \lambda^*)^T \in \mathbb{R}_{++}$, у якій для всіх $(x, \lambda) \in \mathbb{R}_{++}$ виконуються нерівності

$$L(x^*, \lambda) \geq L(x^*, \lambda^*), \quad L(x^*, \lambda^*) \geq L(x, \lambda^*).$$

Це означає, що виконуються нерівності (8.16), тобто $(x^*, \lambda^*)^T$ – сідлова точка функції Лагранжа (8.15). Достатність, а значить, і теорема у цілому доведена. ►

Серед задач опуклого програмування особливе місце займають задачі квадратичного програмування.

У розділі 5 уже йшлося про функції, які називаються квадратичними формами (функції (5.12), (5.21) чи (5.22)), а також їх знаковизначеність, квазізнаковизначеність (напіввизначеність) і невизначеність (знакозмінність). Був сформований також критерій знаковизначеності Сильвестра (теорема 5.7). Для квадратичної форми

$$Q(x_1, \dots, x_n) = \sum_{k=1}^n \sum_{j=1}^n c_{kj} x_k x_j, \quad (8.26)$$

де $C = (c_{kj})_{k,j=1}^n$ – симетрична матриця,

$$C_1 = c_{11}, \quad C_2 = \begin{vmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{vmatrix}, \dots, C_n = \begin{vmatrix} c_{11} & \dots & c_{1n} \\ \dots & \dots & \dots \\ c_{n1} & \dots & c_{nn} \end{vmatrix},$$

на основі критерію Сильвестра та інших досліджень установлені такі ознаки.

Ознака 1. Якщо всі визначники $C_1, \dots, C_n$ – додатні, то квадратична форма (8.26) – додатно визначена.

Ознака 2. Якщо у послідовності чисел $1, C_1, C_2, \dots, C_n$ знаки чергуються, то (8.26) – від'ємно визначена.

Ознака 3. Нехай $r = \text{rg} C < n$, то форма (8.26) додатно напіввизначена, якщо $C_1 > 0, \dots, C_r > 0, C_{r+1} = 0, \dots, C_n = 0$, і від'ємно

напіввизначена, якщо у послідовності $1, C_1, \dots, C_r$ знаки чергуються, а $C_{r+1} = 0, \dots, C_n = 0$.

Ознака 4. Якщо у послідовності чисел $1, C_1, C_2, \dots, C_n$ є від'ємні, але немає чергування знаків, то форма (8.26) є невизначеною.

Теорема 8.3. Квадратична форма (8.26) є угнутою (опуклою вгору) функцією, якщо вона від'ємно напіввизначена й опуклою (опуклою вниз), якщо вона додатно напіввизначена.

◀ Спочатку покажемо справедливості першої частини теореми. Для цього необхідно показати, що для довільних $x, y \in \mathbb{R}^n$ і $\alpha \in [0, 1]$ виконується нерівність

$$Q(\alpha x + (1 - \alpha)y) \geq \alpha Q(x) + (1 - \alpha)Q(y).$$

Справді, за означенням від'ємно-напіввизначеної квадратичної форми (8.26) відносно змінних $(x_1 - y_1), \dots, (x_n - y_n)$ має місце нерівність

$$\sum_{k=1}^n \sum_{j=1}^n c_{kj} (x_k - y_k)(x_j - y_j) \leq 0$$

або

$$\sum_{k=1}^n \sum_{j=1}^n c_{kj} (x_k y_j + x_j y_k) \geq \sum_{k=1}^n \sum_{j=1}^n c_{kj} (x_k x_j + y_k y_j).$$

Враховуючи попередню нерівність, отримаємо

$$\begin{aligned} Q(\alpha x + (1 - \alpha)y) &= \sum_{k=1}^n \sum_{j=1}^n c_{kj} [\alpha x_k + (1 - \alpha)y_k][\alpha x_j + (1 - \alpha)y_j] = \\ &= \alpha^2 \sum_{k=1}^n \sum_{j=1}^n c_{kj} x_k x_j + \alpha(1 - \alpha) \sum_{k=1}^n \sum_{j=1}^n c_{kj} [x_k y_j + x_j y_k] + \\ &+ (1 - \alpha)^2 \sum_{k=1}^n \sum_{j=1}^n c_{kj} y_k y_j \geq \alpha^2 Q(x) + \alpha(1 - \alpha) \sum_{k=1}^n \sum_{j=1}^n c_{kj} (x_k x_j + y_k y_j) + \\ &+ (1 - \alpha)^2 Q(y) = \alpha^2 Q(x) + \alpha(1 - \alpha)[Q(x) + Q(y)] + (1 - \alpha)^2 Q(y) = \end{aligned}$$

$= \alpha Q(x)[\alpha + 1 - \alpha] + (1 - \alpha)Q(y)[\alpha + 1 - \alpha] = \alpha Q(x) + (1 - \alpha)Q(y)$,
що й потрібно було встановити.

Аналогічним шляхом доводиться опуклість (опуклість вниз) квадратичної форми $Q(x)$ у випадку, коли вона додатно-напіввизначена. Теорема доведена. ►

Задачею квадратичного програмування на максимум називається задача вигляду

$$\begin{cases} \sum_{j=1}^n d_j x_j + \sum_{k=1}^n \sum_{j=1}^n c_{kj} x_k x_j \rightarrow \max, \\ \sum_{j=1}^n a_{ij} x_j \leq b_i, \quad i = \overline{1, m}, \\ x_j \geq 0, \quad j = \overline{1, n}, \end{cases} \quad (8.27)$$

де квадратична форма є угнутою функцією, а d_j, b_i, a_{ij} – задані величини ($i, j = \overline{1, n}$).

Виявляється, що задачу (8.27) можна розв'язати методами лінійного програмування. Запишемо спочатку функцію Лагранжа для задачі (8.27) (для допустимої множини цієї задачі виконується умова регулярності Слейтера):

$$\begin{aligned} L(x_1, \dots, x_n, \lambda_1, \dots, \lambda_m) = & \sum_{j=1}^n d_j x_j + \sum_{k=1}^n \sum_{j=1}^n c_{kj} x_k x_j + \\ & + \sum_{i=1}^m \lambda_i \left(b_i - \sum_{j=1}^n a_{ij} x_j \right). \end{aligned} \quad (8.28)$$

Умови існування сідлової точки $(x^*, \lambda^*)^T$ (тобто умови (8.17)–(8.22) для функції (8.28) можна записати у такому вигляді:

$$\frac{\partial L^*}{\partial x_j} + \nu_j = 0, \quad j = \overline{1, n}, \quad (8.29)$$

$$\frac{\partial L^*}{\partial \lambda_i} - \omega_i = 0, \quad i = \overline{1, m}, \quad (8.30)$$

$$x_j^* \nu_j = 0, \quad j = \overline{1, n}, \quad (8.31)$$

$$\lambda_i^* \omega_i = 0, \quad i = \overline{1, m}, \quad (8.32)$$

$$x_j^* \geq 0, \lambda_i^* \geq 0, \nu_j \geq 0, \omega_i \geq 0 \quad (j = \overline{1, n}; i = \overline{1, m}), \quad (8.33)$$

де $\nu_j, \omega_i, j = \overline{1, n}; i = \overline{1, m}$ – допоміжні змінні.

Отже, щоб знайти розв'язок задачі (8.27), потрібно знайти невід'ємний розв'язок системи (8.29), (8.30), який задовольняє умови (8.31), (8.32). Розв'язати систему (8.29), (8.30) при виконанні умов (8.33) – це розв'язати деяку допоміжну задачу ЛП. Продемонструємо це на конкретному прикладі.

Приклад 8.6.

$$\begin{cases} -x_1^2 - x_2^2 + 10x_1 + 15x_2 \rightarrow \max, \\ 2x_1 + 3x_2 \leq 13, \\ 2x_1 + x_2 \leq 10, \\ x_j \geq 0, \quad j = \overline{1, 2}. \end{cases} \quad (8.34)$$

◀ Цільова функція задачі (8.34) угнута, оскільки вона є сумою лінійної та угнутої функцій. Функція Лагранжа має вигляд

$$L(x_1, x_2, \lambda_1, \lambda_2) = -x_1^2 - x_2^2 + 10x_1 + 15x_2 + \lambda_1(13 - 2x_1 - 3x_2) + \lambda_2(10 - 2x_1 - x_2).$$

Випишемо співвідношення (8.29)–(8.33) для задачі (8.34):

$$\begin{cases} 2x_1 + 2\lambda_1 + 2\lambda_2 - \nu_1 = 10, \\ 2x_2 + 3\lambda_1 + \lambda_2 - \nu_2 = 15, \\ 2x_1 + 3x_2 + \omega_1 = 13, \\ 2x_1 + x_2 + \omega_2 = 10, \end{cases} \quad (8.35)$$

$$\begin{cases} x_1 \nu_1 = 0, x_2 \nu_2 = 0, \\ \lambda_1 \omega_1 = 0, \lambda_2 \omega_2 = 0, \end{cases} \quad (8.36)$$

$$x_j \geq 0, \lambda_i \geq 0, \nu_j \geq 0, \omega_i \geq 0, i = \overline{1,2}; j = \overline{1,2}. \quad (8.37)$$

Розв'яжемо систему (8.35)–(8.37), поставивши їй у відповідність таку задачу ЛП:

$$\begin{cases} -z_1 - z_2 \rightarrow \max, \\ 2x_1 + 2\lambda_1 + 2\lambda_2 - \nu_1 + z_1 = 10, \\ 2x_2 + 3\lambda_1 + \lambda_2 - \nu_2 + z_2 = 15, \\ 2x_1 + 3x_2 + \omega_1 = 13, \\ 2x_1 + x_2 + \omega_2 = 10, \\ x_j \geq 0, \lambda_i \geq 0, \nu_j \geq 0, \omega_i \geq 0, z_i \geq 0, i = \overline{1,2}; j = \overline{1,2}. \end{cases} \quad (8.38)$$

Процес розв'язування задачі (8.38) проілюстрований у таблиці 8.2.

Оптимальним розв'язком задачі (8.38) є вектор $(x_1^*, x_2^*, \lambda_1^*, \lambda_2^*, \nu_1^*, \nu_2^*, \omega_1^*, \omega_2^*, z_1^*, z_2^*)^T = (0, 0, 5, 0, 0, 0, 13, 10, 0, 0)^T$. Із останньої симплекс-таблиці випливає, що цей розв'язок не єдиний. Отже, ми знайшли один із невід'ємних розв'язків системи (8.35):

$$\begin{cases} x_1^* = 0, x_2^* = 0, \lambda_1^* = 5, \lambda_2^* = 0, \\ \nu_1^* = 0, \nu_2^* = 0, \omega_1^* = 13, \omega_2^* = 10. \end{cases} \quad (8.39)$$

Однак розв'язок (8.39) не відповідає умовам (8.36). Знайдемо інший розв'язок задачі (8.38) (а отже, і системи (8.35)), для чого введемо у базис змінну x_1 замість ω_1 . Результати перерозрахунку зафіксовані у таблиці 8.3.

У таблиці 8.3 знайдений ще один оптимальний розв'язок задачі (8.38): $(x_1^*, x_2^*, \lambda_1^*, \lambda_2^*, \nu_1^*, \nu_2^*, \omega_1^*, \omega_2^*, z_1^*, z_2^*)^T = (2, 3, 3, 0, 0, 0, 0, 3, 0, 0)^T$. Відповідний йому розв'язок системи (8.35) такий:

$$x_1^* = 2, x_2^* = 3, \lambda_1^* = 3, \lambda_2^* = 0, \nu_1^* = 0, \nu_2^* = 0, \omega_1^* = 0, \omega_2^* = 3.$$

Очевидно, що цей розв'язок задовольняє умови (8.36), тому $(x_1^*, x_2^*, \lambda_1^*, \lambda_2^*)^T = (2, 3, 3, 0)^T$ – сідлова точка функції Лагранжа для задачі (8.34). Отже, $x^* = (2, 3)^T$ – оптимальний розв'язок задачі (8.34) і $f_{\max}^* = 52$. ►

Таблица 8.2

i	б3	c_i^δ	x_i^δ	0	x_2	λ_1	0	0	λ_2	ν_1	ν_2	ω_1	0	ω_2	z_1	-1
1	z_1	-1	10	2	0	2	2	-1	2	-1	0	0	0	0	1	0
2	z_2	-1	15	0	2	3	1	0	1	0	-1	0	0	0	0	1
3	ω_1	0	13	2	3	0	0	0	0	0	0	1	0	0	0	0
4	ω_2	0	10	2	1	0	0	0	0	0	0	0	1	0	0	0
$m+1$	σ_j		-25	-2	-2	-5	-3	1	1	1	1	0	0	0	0	0
1	λ_1	0	5	1	0	1	1	$-\frac{1}{2}$	1	$-\frac{1}{2}$	0	0	0	0	$\frac{1}{2}$	0
2	z_2	-1	0	-3	2	0	-2	$\frac{3}{2}$	-2	$\frac{3}{2}$	-1	0	0	0	$-\frac{3}{2}$	1
3	ω_1	0	13	2	3	0	0	0	0	0	0	1	0	0	0	0
4	ω_2	0	10	2	1	0	0	0	0	0	0	0	1	0	0	0
$m+1$	σ_j		0	3	-2	0	2	$-\frac{3}{2}$	2	$-\frac{3}{2}$	1	0	0	0	$\frac{3}{2}$	0

Продовження таблиці 8.2

1	λ_1	0	5	1	0	1	1	$-\frac{1}{2}$	0	0	0	$\frac{1}{2}$	0
2	x_2	0	0	$-\frac{3}{2}$	1	0	-1	$\frac{3}{4}$	$-\frac{1}{2}$	0	0	$-\frac{3}{4}$	$\frac{1}{2}$
3	ω_1	0	13	$\frac{13}{2}$	0	0	3	$-\frac{9}{4}$	$\frac{3}{2}$	1	0	$\frac{9}{4}$	$\frac{3}{2}$
4	ω_2	0	10	$\frac{7}{2}$	0	0	1	$-\frac{3}{4}$	$\frac{1}{2}$	0	1	$\frac{3}{4}$	$-\frac{1}{2}$
$m+1$	σ_j	0	0	0	0	0	0	0	0	0	0	1	1

Таблица 8.3

i	β_3	c_i^{β}	x_i^{β}	0	0	0	0	0	0	0	0	-1	-1
				x_1	x_2	λ_1	λ_2	v_1	v_2	ω_1	ω_2	z_1	z_2
1	λ_1	0	3	0	0	1	$\frac{7}{13}$	$-\frac{2}{3}$	$-\frac{3}{13}$	$-\frac{2}{13}$	0	$\frac{2}{13}$	$\frac{3}{13}$
2	x_2	0	3	0	1	0	$-\frac{4}{13}$	$\frac{3}{13}$	$-\frac{2}{13}$	$\frac{3}{13}$	0	$-\frac{3}{13}$	$\frac{2}{13}$
3	x_1	0	2	1	0	0	$\frac{6}{13}$	$-\frac{9}{26}$	$\frac{3}{13}$	$\frac{2}{13}$	0	$\frac{9}{26}$	$-\frac{3}{13}$
4	ω_2	0	3	0	0	0	$-\frac{8}{13}$	$\frac{6}{13}$	$-\frac{4}{13}$	$-\frac{7}{13}$	1	$-\frac{6}{13}$	$\frac{4}{13}$
$m+1$	σ_j	0	0	0	0	0	0	0	0	0	0	1	1

Зауваження 8.1. Якщо задача квадратичного програмування має обмеження-рівності, то щоб звести таку задачу до задачі вигляду (8.27), треба обмеження-рівності замінити еквівалентними нерівностями або виразити з обмежень-рівностей деякі базисні змінні через інші змінні та записати умову невід'ємності для базисних змінних.

Продемонструємо це на прикладі.

Приклад 8.7.

$$\begin{cases} -x_1^2 - x_2^2 + x_1 + 2x_3 \rightarrow \max, \\ x_1 - x_2 + 2x_3 = 6, \\ x_j \geq 0, \quad j = \overline{1,3}. \end{cases} \quad (8.40)$$

◀ Представимо задачу (8.40) у такій формі запису:

$$\begin{cases} -x_1^2 - x_2^2 + x_1 + 2x_3 \rightarrow \max, \\ x_1 - x_2 + 2x_3 \leq 6, \\ -x_1 + x_2 - 2x_3 \leq -6, \\ x_j \geq 0, \quad j = \overline{1,3}. \end{cases} \quad (8.41)$$

Функція Лагранжа для задачі (8.41) має вигляд

$$L(x_1, x_2, x_3, \lambda_1, \lambda_2) = -x_1^2 - x_2^2 + x_1 + 2x_3 + \lambda_1(6 - x_1 + x_2 - 2x_3) + \lambda_2(-6 + x_1 - x_2 + 2x_3).$$

Необхідні й достатні умови існування сідлової точки для даної функції, тобто співвідношення (8.29)–(8.33), запишуться так:

$$\begin{cases} 2x_1 + \lambda_1 - \lambda_2 - \nu_1 = 1, \\ 2x_2 - \lambda_1 + \lambda_2 - \nu_2 = 0, \\ 2\lambda_1 - 2\lambda_2 - \nu_3 = 2, \\ x_1 - x_2 + 2x_3 + \omega_1 = 6, \\ x_1 - x_2 + 2x_3 - \omega_2 = 6, \end{cases} \quad (8.42)$$

$$\begin{cases} x_1 v_1 = 0, x_2 v_2 = 0, x_3 v_3 = 0, \\ \lambda_1 \omega_1 = 0, \lambda_2 \omega_2 = 0, \end{cases} \quad (8.43)$$

$$x_j \geq 0, \lambda_i \geq 0, v_j \geq 0, \omega_i \geq 0, i = \overline{1,2}; j = \overline{1,3}. \quad (8.44)$$

Системі (8.42) поставимо у відповідність задачу ЛП

$$\begin{cases} -z_1 - z_2 - z_3 - z_4 \rightarrow \max, \\ 2x_1 - \lambda_1 + \lambda_2 - v_1 + z_1 = 1, \\ 2x_2 - \lambda_1 + \lambda_2 - v_2 + z_2 = 0, \\ 2\lambda_1 - 2\lambda_2 - v_3 + z_3 = 2, \\ x_1 - x_2 + 2x_3 + \omega_1 = 6, \\ x_1 - x_2 + 2x_3 - \omega_2 + z_4 = 6, \\ x_j \geq 0, \lambda_i \geq 0, v_j \geq 0, \omega_i \geq 0, \\ z_k \geq 0, j = \overline{1,3}; i = \overline{1,2}; k = \overline{1,4}. \end{cases} \quad (8.45)$$

Розв'язавши задачу (8.45) симплекс-методом, отримаємо її оптимальний розв'язок:

$$(x_1^*, x_2^*, x_3^*, \lambda_1^*, \lambda_2^*, v_1^*, v_2^*, v_3^*, \omega_1^*, \omega_2^*, z_1^*, z_2^*, z_3^*, z_4^*)^T = \\ = (0, 1/2, 13/4, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0)^T,$$

тобто одержимо невід'ємний розв'язок системи (8.42):

$$x_1^* = 0, x_2^* = 1/2, x_3^* = 13/4, \lambda_1^* = 1, \lambda_2^* = 0, v_1^* = 0, v_2^* = 0, \\ v_3^* = 0, \omega_1^* = 0, \omega_2^* = 0.$$

Цей розв'язок задовольняє умови (8.43), тому $(x_1^*, x_2^*, x_3^*, \lambda_1^*, \lambda_2^*)^T = (0, 1/2, 13/4, 1, 0)^T$ – сідлова точка функції Лагранжа для задачі (8.41), звідки випливає, що $x^* = (0, 1/2, 13/4)^T$ – оптимальний розв'язок задачі (8.41) (а отже, і (8.40)) і $f_{\max}^* = 25/4$. Можливий також інший шлях розв'язування задачі (8.40). Розв'яжемо рівність $x_1 - x_2 + 2x_3 = 6$ відносно x_1 . Отримаємо $x_1 = x_2 - 2x_3 + 6$. Виключимо змінну x_1 із цільової функції й запишемо умову

невід'ємності x_1 , після чого задача (8.40) зведеться до еквівалентної задачі

$$\begin{cases} -(x_2 - 2x_3 + 6)^2 - x_2^2 + (x_2 - 2x_3 + 6) + 2x_3 \rightarrow \max, \\ -x_2 + 2x_3 \leq 6, \\ x_j \geq 0, \quad j = \overline{2,3}. \end{cases} \quad (8.46)$$

Розв'язуючи задачу (8.46) аналогічно, як і задачу (8.41), отримаємо вже відомий нам оптимальний розв'язок задачі (8.40). ►

Зауваження 8.2. Якщо система (8.29)–(8.33) не має розв'язку, то це означає, що функція Лагранжа (8.28) не має сідлової точки і задача (8.27) не має розв'язку.

8.5. Поняття двоїстої задачі опуклого програмування

Для нелінійних задач математичного програмування, а саме для задач опуклого чи угнутого програмування можна побудувати так звані двоїсті задачі.

Розглянемо, як і раніше, задачу угнутого програмування (8.1), у якій функція $f(x)$ є угнутою, а функції $g_i(x)$ ($i = \overline{1,m}$) – опуклими в області визначення. Уведемо у $\mathbb{R}_+^n$ функцію

$$\psi(x) = \min_{\lambda \in \mathbb{R}_+^m} L(x, \lambda) = \begin{cases} f(x), & x \in X_n, \\ -\infty, & x \in \mathbb{R}_+^n / X_n, \end{cases}$$

де X_n – допустима множина задачі (8.1), $L(x, \lambda)$ – її функція Лагранжа, тобто функція (8.15).

Очевидно, що задача (8.1) може бути записана у вигляді еквівалентної задачі

$$\begin{cases} \psi(x) \rightarrow \max, \\ x \in \mathbb{R}_+^n. \end{cases} \quad (8.47)$$

Поряд із функцією $\psi(x)$ розглянемо також функцію

$$\tilde{\psi}(\lambda) = \max_{x \in \mathbb{R}_+^n} L(x, \lambda).$$

Задачу

$$\begin{cases} \tilde{\psi}(\lambda) \rightarrow \min, \\ \lambda \in \mathbb{R}_+^m. \end{cases} \quad (8.48)$$

називають двоїстою до задачі (8.47) (або (8.1)).

Такий принцип побудови двоїстих задач тісно пов'язаний з уведенням вище поняттям сідлової точки функції $L(x, \lambda)$. Розв'язування пари взаємно двоїстих задач (8.47), (8.48) (як і у задачах ЛП) зводиться до знаходження точки $(x^*, \lambda^*)^T \in \mathbb{R}_{++}$, для якої має місце рівність:

$$L(x^*, \lambda^*) = \max_{x \in \mathbb{R}_+^n} \min_{\lambda \in \mathbb{R}_+^m} L(x, \lambda) = \min_{\lambda \in \mathbb{R}_+^m} \max_{x \in \mathbb{R}_+^n} L(x, \lambda).$$

Принцип побудови двоїстих задач дозволяє для кожної конкретної задачі побудувати двоїсту, яка, як відомо, має важливе прикладне значення. Нижче розглянемо приклад побудови двоїстої задачі до задачі ЛП (відзначимо, що лінійне програмування належить до найпростішого класу задач опуклого чи угнутого програмування). Щодо правил побудови двоїстих задач ЛП, то легко переконатись, що вони є реалізацією принципу побудови двоїстих задач для конкретних задач ЛП.

Приклад 8.8. Використовуючи розглянутий вище підхід, побудувати двоїсту задачу ЛП до задачі (7.76) (приклад 7.19).

◀ Зауважимо, що коли у задачі ЛП є обмеження-рівності, то у побудованій для неї функції Лагранжа відповідні змінні λ_i , можуть набувати довільних значень.

Отже, функція Лагранжа для задачі (7.76)

$$\begin{aligned} L(x_1, x_2, \lambda_1, \lambda_2, \lambda_3) &= (7x_1 - 2x_2) + \lambda_1[7 - (-5x_1 + x_2)] + \\ &+ \lambda_2[3 - (x_1 - 2x_2)] + \lambda_3[-10 - (-3x_1 - 4x_2)] = (7\lambda_1 + 3\lambda_2 - 10\lambda_3) + \\ &+ x_1(7 + 5\lambda_1 - \lambda_2 + 3\lambda_3) + x_2(-2 - \lambda_1 + 2\lambda_2 + 4\lambda_3). \end{aligned}$$

Обмеження для аргументів цієї функції такі:

$$x \in X_L = \{x \in \mathbb{R}^2 / x_1 \geq 0\}, \lambda \in \Lambda_L = \{\lambda \in \mathbb{R}^3 / \lambda_1 \geq 0, \lambda_3 \geq 0\}.$$

Функція

$$\tilde{\psi}_1(\lambda) = \max_{x \in X_L} L(x, \lambda) = \begin{cases} 7\lambda_1 + 3\lambda_2 - 10\lambda_3, & \lambda \in \Lambda_n, \\ +\infty, & \lambda \in \Lambda_L / \Lambda_n, \end{cases}$$

$$\text{де } \Lambda_n = \{\lambda \in \Lambda_L / -5\lambda_1 + \lambda_2 - 3\lambda_3 \geq 7, \lambda_1 - 2\lambda_2 - 4\lambda_3 = -2\}.$$

Очевидно, що двоїста задача

$$\begin{cases} \tilde{\psi}_1(\lambda) \rightarrow \min, \\ \lambda \in \Lambda_L \end{cases}$$

і буде задачею ЛП (7.79), якщо формально перепозначити λ_i на y_i ($i = \overline{1,3}$). Установлено, що знаходження розв'язків для пари двоїстих задач ЛП еквівалентне знаходженню сідлової точки відповідної функції Лагранжа. Цей висновок має місце і для деяких класів задач НЛП. ►

Розділ 9. Динамічне програмування

9.1. Вступ до динамічного програмування

В арсеналі методів оптимізації особливе місце належить динамічному програмуванню (ДП). Підкреслимо, що, на відміну від термінів «лінійне програмування», «нелінійне програмування» тощо, які належать до певного типу задач, термін «динамічне програмування» належить до обчислювального методу, а не певного типу задач. Отже, якщо навіть у літературі з математичного чи конкретно динамічного програмування зустрічається словосполучення «задача динамічного програмування», то його змістове навантаження стосується методу розв'язання цієї задачі, а не її типу. Виникнення методу ДП пов'язано з іменем американського математика Р. Беллмана, який на початку 50-х років XX століття (1950-1953 рр.) при розв'язуванні деяких оптимізаційних задач застосовував прийом, названий пізніше принципом Беллмана, суть якого полягає у заміні задачі з багатьма змінними сукупністю послідовно розв'язуваних задач із меншим числом змінних. Інакше кажучи, запропонований Р. Беллманом метод часто дозволяє звести задачу оптимізації функції n змінних ($n > 1$) до n -крокового розв'язування задач оптимізації функції однієї змінної.

Задачі, для яких можливе застосування методу ДП, утворюють певні класи задач, для яких характерними є:

- можливість розбиття початкової задачі на окремі підзадачі, кожна з яких містить меншу кількість змінних (наприклад, одну змінну);
- однотипність виділених підзадач;

- вимірювання ефекту від прийнятого рішення (на основі розв'язку кожної із підзадач) в однакових одиницях виміру;

- обчислення загального ефекту як суми ефектів окремих підзадач.

Ці та інші властивості задач, які ефективно розв'язуються методом ДП, будуть уточнені пізніше. Наразі відзначимо лише, що один із класів таких задач – це задачі, які є математичними моделями динамічних процесів, зазвичай залежних від часу багатокрокових чи багатетапних процесів. Очевидно, що, на відміну від задач математичного програмування (лінійного, нелінійного), постановки яких мають статичний характер і розв'язок яких знаходиться за один крок чи етап, ці задачі за постановкою багатокрокові чи багатетапні. Будь-який процес характеризує зміну станів досліджуваної системи (об'єкта) у часі, а тому залежно від того, вивчається зміна станів у дискретні моменти часу чи у довільні моменти часу із деякого часового проміжку, процеси бувають дискретними (зі скінченним або нескінченним числом кроків (етапів)) і неперервними. Дискретні та неперервні процеси тісно взаємопов'язані між собою: дискретні процеси отримують із неперервних шляхом квантування змінної часу, а неперервні із дискретних – шляхом граничного переходу. Зрозуміло, що стани системи утворюють траєкторію її руху, причому таких траєкторій може бути багато, що приводить дослідника до задачі вибору «найкращої» у певному сенсі траєкторії, адже за словами Р. Беллмана «Математик може вважати свою проблему розв'язаною, лише коли збагне суть оптимального підходу до розв'язання». Ці слова стосуються кожного дослідника. Задачу вибору оптимальної траєкторії або оптимального процесу називають не просто

оптимізаційною задачею, а саме задачею керування (або оптимального керування) процесами.

Велика кількість прикладних, зокрема економічних задач (приклади деяких будуть наведені нижче) вписуються у рамки дискретних процесів зі скінченним числом кроків (етапів). У зв'язку з цим для розкриття змісту методу ДП розглянемо тільки задачу керування дискретним процесом зі скінченним числом кроків (етапів).

9.2. Постановка задачі керування дискретним процесом зі скінченним числом кроків

Розглянемо деяку систему S (економічну, соціальну, фізичну тощо), стан якої змінюється у дискретні моменти часу $t_0, t_1, \dots, t_m$, утворюючи деякий процес, розбитий на m етапів: $[t_0, t_1], \dots, [t_{m-1}, t_m]$.

Стан S_i системи S у момент часу t_i ($i = \overline{0, m}$) опишемо деяким вектором $x^{(i)} = (x_1^{(i)}, \dots, x_n^{(i)})^T \equiv (x_1^{(i)}(t_i), \dots, x_n^{(i)}(t_i))^T \in \mathbb{R}^n$, координати якого будемо називати фазовими координатами. Зміну стану системи S будемо розуміти як її переміщення у фазовому просторі $\mathbb{R}^n$. Сукупність відповідних станів характеризує процес: $x^{(0)} \rightarrow \dots \rightarrow x^{(i-1)} \rightarrow x^{(i)} \rightarrow \dots \rightarrow x^{(m)}$.

Передбачається, що на кожному i -му етапі (тобто на відрізку часу $[t_{i-1}, t_i]$) процес визначається як власною динамікою, так і зовнішніми впливами $u^{(i)} = (u_1^{(i)}, \dots, u_r^{(i)})^T \in \mathbb{R}^r$, $i = \overline{1, m}$, які називаються **керуваннями**.

Отже, у результаті i -го етапу система прийде у стан $x^{(i)}$, який визначається тільки станом $x^{(i-1)}$ та вибраним на i -му

етапі керуванням $u^{(i)}$ і не залежить від «передісторії» процесу до i -го етапу. Зв'язок між станами можна описати рівнянням стану

$$x^{(i)} = f^{(i)}(x^{(i-1)}, u^{(i)}), \quad (9.1)$$

де $f^{(i)}(x, u)$ ($i = \overline{1, m}$) – деякі задані функції. Будемо вважати, що ефективність i -го етапу оцінюється цільовою функцією $k^{(i)}(x^{(i-1)}, u^{(i)})$, а ефективність усього процесу – аддитивною цільовою функцією

$$k(x^{(0)}, \dots, x^{(m-1)}, u^{(1)}, \dots, u^{(m)}) = \sum_{i=1}^m k^{(i)}(x^{(i-1)}, u^{(i)}). \quad (9.2)$$

На фазову траєкторію та вибір керування можуть бути накладені обмеження

$$x^{(i)} \in X^{(i)}, i = \overline{1, m-1}, \quad (9.3)$$

$$u^{(i)} \in U^{(i)}(x^{(i-1)}), i = \overline{1, m}, \quad (9.4)$$

де $X^{(i)} \subset \mathbb{R}^n$, $U^{(i)}(x^{(i-1)}) \subset \mathbb{R}^r$ – задані множини (оскільки $U^{(i)}$ може залежати від початкового стану $x^{(i-1)}$, то ми її позначаємо $U^{(i)}(x^{(i-1)})$).

Окремо виділимо обмеження на початковий та кінцевий стани процесу, тобто обмеження

$$x^{(0)} \in X^{(0)}, \quad x^{(m)} \in X^{(m)}, \quad (9.5)$$

де $X^{(0)}, X^{(m)} \in \mathbb{R}^n$. У багатьох випадках множини станів $X^{(0)}$ і $X^{(m)}$ містять тільки по одній точці, тобто початкову та кінцеву точку фазової траєкторії.

Задача оптимізації (або оптимального керування) описаним вище процесом ставиться так: серед усіх можливих керувань вибрати таке керування $\{u^{(1)*}, \dots, u^{(m)*}\}$, яке при наявності умов (9.3)–(9.5) переводить систему S із

стану $x^{(0)}$ у стан $x^{(m)}$ і для якого цільова функція (9.2) досягає свого екстремального значення (максимального або мінімального залежно від змісту задачі). Керування $\{u^{(1)*}, \dots, u^{(m)*}\}$ і відповідна йому траєкторія $\{x^{(0)*}, x^{(1)*}, \dots, x^{(m)*}\}$ називаються **оптимальними**.

Математична модель цієї задачі на максимум має вигляд

$$\begin{cases} k(x^{(0)}, \dots, x^{(m-1)}, u^{(1)}, \dots, u^{(m)}) \rightarrow \max, \\ x^{(i)} = f^{(i)}(x^{(i-1)}, u^{(i)}), i = \overline{1, m}, \\ x^{(i)} \in X^{(i)}, i = \overline{1, m-1}, \\ u^{(i)} \in U^{(i)}(x^{(i-1)}), i = \overline{1, m}, \\ x^{(0)} \in X^{(0)}, x^{(m)} \in X^{(m)}. \end{cases} \quad (9.6)$$

Однак у задачі (9.6) обмеження на стани фазової траєкторії $x^{(1)}, x^{(2)}, \dots, x^{(m)}$, тобто обмеження (9.3) і (9.5), можна включити в обмеження (9.4). При цьому (9.3) і (9.4) будуть замінені еквівалентним співвідношенням

$$\begin{aligned} u^{(i)} \in \tilde{U}^{(i)}(x^{(i-1)}) &= \{u^{(i)} \in U^{(i)}(x^{(i-1)}) / \\ x^{(i)} = f^{(i)}(x^{(i-1)}, u^{(i)}) &\in X^{(i)}, i = \overline{1, m}\}. \end{aligned} \quad (9.7)$$

З урахуванням (9.7) задача (9.6) набуде вигляду

$$\begin{cases} k(x^{(0)}, \dots, x^{(m-1)}, u^{(1)}, \dots, u^{(m)}) \rightarrow \max, \\ x^{(i)} = f^{(i)}(x^{(i-1)}, u^{(i)}), i = \overline{1, m}, \\ u^{(i)} \in \tilde{U}^{(i)}(x^{(i-1)}), i = \overline{1, m}, \\ x^{(0)} \in X^{(0)}. \end{cases} \quad (9.8)$$

Задача (9.8) (або (9.6)) може описувати як неперервну, так і дискретну модель багатокрокового процесу

оптимізації, залежно від того, вектори $x^{(i)}, u^{(i)}$ змінюються у часі неперервно чи дискретно.

Багато економічних задач формалізуються дискретними моделями. Функціональні залежності у таких задачах задаються не аналітичним, а табличним способом. Однак, незалежно від того, якою є модель багатокрокового процесу оптимізації, схема їх розв'язування методом динамічного програмування залишається без змін.

9.3. Принцип оптимальності Беллмана й алгоритм методу динамічного програмування

В основі методу динамічного програмування лежить принцип оптимальності Беллмана. Його можна сформулювати так:

оптимальне керування у довільний момент часу (чи період часу для дискретного процесу) не залежить від передісторії системи і визначається тільки станом системи у цей момент часу (чи станом системи на початок періоду) та метою керування.

По-іншому принцип оптимальності Беллмана формулюється ще так:

яким би не був стан системи перед наступним кроком, керування на цьому кроці потрібно вибрати так, щоб сума ефективності процесу на цьому кроці й оптимальної ефективності процесу на всіх наступних кроках була максимальною.

З даного принципу випливає, що довільна частина $\{x^{(i-1)*}, \dots, x^{(m)*}\}$ ($i = \overline{2, m}$) оптимальної траєкторії $\{x^{(0)*}, \dots, x^{(m)*}\}$ завжди буде оптимальною для відповідних завершальних $(m - i + 1)$ кроків (етапів) процесу. Опишемо метод ДП більш

детально. Нехай через $(i-1)$ кроків система S перейшла зі стану $x^{(0)}$ у стан $x^{(i-1)}$. Розглянемо задачу оптимізації (9.8) для наступних $(m-i+1)$ кроків. Припустимо, що знайдені $\{u^{(i)*}, \dots, u^{(m)*}\}$ і $\{x^{(i-1)}, x^{(i)*}, \dots, x^{(m)*}\}$ – відповідно оптимальне керування для вказаних $(m-i+1)$ кроків і оптимальна траєкторія з початковим станом $x^{(i-1)}$. Оптимальне значення цільової функції буде також залежати від $x^{(i-1)}$. Позначимо його так:

$$B_i(x^{(i-1)}) = k^{(i)}(x^{(i-1)}, u^{(i)*}) + \sum_{j=i+1}^m k^{(j)}(x^{(j-1)*}, u^{(j)*}). \quad (9.9)$$

Функція $B_i(x^{(i-1)})$ називається функцією Беллмана останніх $(m-i+1)$ кроків.

Для останнього кроку

$$B_m(x^{(m-1)}) = \max_{u^{(m)} \in \bar{U}^{(m)}(x^{(m-1)})} k^{(m)}(x^{(m-1)}, u^{(m)}). \quad (9.10)$$

Із принципу оптимальності випливає також рекурентне співвідношення між функціями Беллмана:

$$B_i(x^{(i-1)}) = \max_{u^{(i)} \in \bar{U}^{(i)}(x^{(i-1)})} \{k^{(i)}(x^{(i-1)}, u^{(i)}) + B_{i+1}(f^{(i)}(x^{(i-1)}, u^{(i)}))\}, \quad i = \overline{1, m-1}. \quad (9.11)$$

Співвідношення (9.10) і (9.11), які дозволяють послідовно знайти функції Беллмана $B_m(x^{(m-1)})$, $B_{m-1}(x^{(m-2)})$, ..., $B_1(x^{(0)})$, називаються рівняннями Беллмана.

Знаходячи функції $B_i(x^{(i-1)})$, $i = \overline{m, 1}$, ми одночасно знаходимо і відповідні керування $u^{(m)*}(x^{(m-1)})$, ..., $u^{(1)*}(x^{(0)})$, які називаються умовними оптимальними керуваннями, а процес їх знаходження – умовною оптимізацією. Після цього відбувається зворотній процес – безумовна оптимізація,

допомогою співвідношень (9.13) оптимального керування $u^{(1)*}$ та стану $x^{(1)*}$.

Крок 2. Визначення за допомогою співвідношень (9.13) оптимального керування $u^{(2)*}$ та стану $x^{(2)*}$.

.....
Крок m. Визначення за допомогою співвідношень (9.13) оптимального керування $u^{(m)*}$ та стану $x^{(m)*}$.

Зробимо деякі зауваження щодо методу ДП.

Зауваження 9.1. У багатьох задачах, які можна розв'язати методом ДП, цільова функція не є аддитивною. У такому випадку початкову постановку задачі треба так видозмінити, щоб цільова функція була аддитивною.

Зауваження 9.2. Принциповою умовою застосування методу ДП є умова відсутності післядії, яка описується співвідношенням (9.1). Якщо (9.1) не виконується, то треба збільшити число координат фазового простору, ввівши в нього ті параметри з минулого (тобто з «передісторії» процесу), від яких залежить майбутнє.

Зауваження 9.3. При розв'язуванні задач методом ДП часто виникає складність, пов'язана з розумним розбиттям процесу керування на скінченну кількість кроків (етапів). Треба так розбити (розчленувати) процес переходу системи S зі стану $x^{(0)}$ у стан $x^{(m)}$ на кроки, щоб вони допускали зручну нумерацію й чітку послідовність дій. Така задача, як правило, непростя.

Зауваження 9.4. Основні переваги методу ДП полягають у тому, що цільові функції багатоетапних задач оптимізації можуть бути задані як аналітично, так і таблично, а змінні можуть набувати цілочислових значень. До недоліків методу ДП можна віднести значний обсяг обчислень, особливо для задач великої розмірності. Саме у зв'язку з цим свого часу

Р. Беллман говорив про так зване «прокляття багатовимірності».

9.4. Приклади розв'язування оптимізаційних задач методом динамічного програмування

Приклад 9.1 (задача про розподіл капітальних вкладів (КВ) заданого обсягу між m підприємствами).

Нехай для збільшення потужності виробництва продукції, яка користується підвищеним попитом m , підприємствам виділяється КВ загальним обсягом s одиниць (од.). Використання i -м підприємством $u^{(i)}$ од. КВ дає прибуток цьому підприємству $k^{(i)}(u^{(i)})$ од. Знайти такий розподіл КВ між підприємствами, який забезпечить максимальний сумарний прибуток.

◀ Математична модель цієї задачі має вигляд:

$$\left\{ \begin{array}{l} \sum_{i=1}^m k^{(i)}(u^{(i)}) \rightarrow \max, \\ \sum_{i=1}^m u^{(i)} = s, \\ u^{(i)} \geq 0, i = \overline{1, m}. \end{array} \right. \quad (9.14)$$

Очевидно, що в загальному випадку задача (9.14) є задачею НЛП, але далеко не завжди її вдається розв'язати методами НЛП (усе залежить від функцій $k^{(i)}(u^{(i)})$, $i = \overline{1, m}$). Проте задачу (9.14) можна розв'язати за допомогою методу ДП. Для цього задачу (9.14) будемо розглядати як багатостадійну або багатокрокову, тобто будемо розглядати не допустимі варіанти розподілу КВ між m підприємствами та оцінювати їх ефективність, а досліджувати ефективність

вкладених КВ на одному підприємстві, на двох підприємствах і т.д., на m підприємствах. У результаті цього одержимо m етапів, на кожному з яких стан системи описується обсягом КВ, виділених для освоєння i підприємствам з номерами від 1 до i . Позначимо цей обсяг КВ через $x^{(i)}$. Величина $u^{(i)}$ і є керуванням на відповідному етапі. Отже, задача (9.14) зведеться до багатокрокової задачі оптимізації типу (9.8), тобто до задачі

$$\begin{cases} \sum_{i=1}^m k^{(i)}(u^{(i)}) \rightarrow \max, \\ x^{(i)} = x^{(i-1)} + u^{(i)}, i = \overline{1, m}, \\ u^{(i)} \in [0, s - x^{(i-1)}], i = \overline{1, m}, \\ x^{(0)} = 0, x^{(m)} = s. \end{cases} \quad (9.15)$$

Задача (9.15) розв'язується методом ДП. Продемонструємо процес розв'язування задачі (9.15) для конкретних числових даних.

Нехай $s = 5$ од., $m = 3$, а значення $u^{(i)}$ та $k^{(i)}(u^{(i)}) = k^{(i)}$ є цілочисловими (тобто дискретними) та наведені у таблиці 9.1.

Таблиця 9.1

$k^{(i)} \backslash u^{(i)}$	0	1	2	3	4	5
$k^{(1)}$	0	4	8	10	11	13
$k^{(2)}$	0	3	6	11	13	14
$k^{(3)}$	0	6	7	9	10	11

Умовна оптимізація

Крок 1. Розв'язування задачі (9.10), тобто знаходження

$$B_3(x^{(2)}) = \max_{u^{(3)} \in [0, 5 - x^{(2)}]} k^{(3)}(u^{(3)})$$

та $u^{(3)*}(x^{(2)})$ при різних значеннях $x^{(2)}$ (табл. 9.2).

Таблиця 9.2

$x^{(2)}$	0	1	2	3	4	5
$B_3(x^{(2)})$	11	10	9	7	6	0
$u^{(3)*}(x^{(2)})$	5	4	3	2	1	0

Крок 2. Розв'язування задачі (9.11), тобто знаходження $B_2(x^{(1)}) = \max_{u^{(2)} \in [0, 5-x^{(1)}]} \{k^{(2)}(u^{(2)}) + B_3(x^{(1)} + u^{(2)})\}$ та $u^{(2)*}(x^{(1)})$.

Для цього, використовуючи дані таблиць 9.1 і 9.2, сформуємо таблицю 9.3 для знаходження значень суми

$$\{k^{(2)}(u^{(2)}) + B_3(x^{(1)} + u^{(2)})\}.$$

Таблиця 9.3

$u^{(2)} \backslash x^{(1)}$	0	1	2	3	4	5
0	11	13	15	18	19	14
1	10	12	13	17	13	—
2	9	10	12	11	—	—
3	7	9	6	—	—	—
4	6	3	—	—	—	—
5	0	—	—	—	—	—

У таблиці 9.3 рамками обведені максимальні значення суми $\{k^{(2)}(u^{(2)}) + B_3(x^{(1)} + u^{(2)})\}$ по $u^{(2)}$ при різних значеннях $x^{(1)}$. На основі таблиці 9.3 побудуємо табл. 9.4.

Таблиця 9.4

$x^{(1)}$	0	1	2	3	4	5
$B_2(x^{(1)})$	19	17	12	9	6	0
$u^{(2)*}(x^{(1)})$	4	3	2	1	0	0

Крок 3. Розв'язування задачі (9.11), тобто знаходження

$$B_1(x^{(0)}) = \max_{u^{(1)} \in [0,5]} \{k^{(1)}(u^{(1)}) + B_2(u^{(1)})\} \quad \text{та} \quad u^{(1)*}(x^{(0)}) \quad \text{з}$$

урахуванням того, що $x^{(0)} = 0$ ($X^{(0)} = \{0\}$). Результати виконання цього кроку відображені у табл. 9.5.

Таблиця 9.5

$u^{(1)}$	0	1	2	3	4	5
$k^{(1)}(u^{(1)}) + B_2(u^{(1)})$	19	21	20	19	17	13

З таблиці 9.5 випливає, що $B_1(x^{(0)}) = B_1(0) = 21$.

Безумовна оптимізація

Використовуючи послідовно формули (9.13), отримаємо кінцевий розв'язок задачі.

Крок 1. Оскільки $x^{(0)*} = 0$ і $u^{(1)*}(0) = u^{(1)*} = 1$, то $x^{(1)*} = x^{(0)*} + u^{(1)*} = 0 + 1 = 1$.

Крок 2. $x^{(1)*} = 1$, тому із таблиці 9.4 випливає, що $u^{(2)*} = u^{(2)*}(1) = 3$, тобто $x^{(2)*} = x^{(1)*} + u^{(2)*} = 1 + 3 = 4$.

Крок 3. Для $x^{(2)*} = 4$ із табл. 9.2 знаходимо $u^{(3)*} = u^{(3)*}(4) = 1$, тобто $x^{(3)*} = x^{(2)*} + u^{(3)*} = 4 + 1 = 5$.

Отже, остаточно знайдено оптимальне керування $\{u^{(1)*}, u^{(2)*}, u^{(3)*}\} = \{1, 3, 1\}$ і відповідну цьому керуванню оптимальну траєкторію $\{x^{(0)*}, x^{(1)*}, x^{(2)*}, x^{(3)*}\} = \{0, 1, 4, 5\}$.

При цьому $k^* \equiv B_1(0) = 21$.

Тож, оптимальний варіант розподілу КВ між трьома підприємствами такий: першому підприємству треба виділити 1 од., другому – 3 од., третьому – 1 од. ►

Приклад 9.2 (задача про t -етапний внесок у дане підприємство КВ заданого обсягу).

КВ обсягом s од. виділяються деякому підприємству протягом t років. Укладені у розвиток підприємства протягом i -го року $u^{(i)}$ од. дають прибуток $k^{(i)}(u^{(i)})$ од.

Вимагається s од. КВ розподілити по роках у такий спосіб, щоб сумарний прибуток підприємства за t років був максимальним.

◄ Нехай $x^{(i)}$ – обсяг КВ, укладений у розвиток підприємства протягом перших i років ($i = \overline{1, t}$). Якщо обмеження $x^i \in [0, s]$, $i = \overline{1, t-1}$; $u^{(i)} \in [0, s]$, $i = \overline{1, t}$, замінити еквівалентними співвідношеннями

$$u^{(i)} \in U^{(i)}(x^{(i-1)}) = [0, s - x^{(i-1)}], \quad i = \overline{1, t},$$

то задача про t -етапний вклад у дане підприємство КВ обсягом s од. знову зведеться до задачі (9.15). Практичне розв'язування задачі (9.15) продемонстроване в попередньому прикладі. ►

Приклад 9.3 (задача про t -етапний розподіл КВ заданого обсягу між двома підприємствами Π_1 і Π_2).

Планується розподіл КВ обсягом s од. між двома підприємствами Π_1 і Π_2 на t років планового періоду.

Будемо вважати, що u од. КВ, вкладених у підприємство Π_j приносять йому у кінці року прибуток $k^{(j)}(u)$ од. i зменшуються до $f^{(j)}(u)$ од. ($j=\overline{1,2}$), які знову використовуються для інвестування підприємств на наступний рік.

Потрібно знайти такий варіант розподілу КВ між підприємствами Π_1 і Π_2 , при якому сумарний прибуток двох підприємств за m років буде максимальним.

◀ Нехай $u^{(i)}$ – обсяг КВ, виділений підприємству Π_1 у i -ому році планового періоду, а $x^{(i-1)}$ – обсяг КВ, який залишився у кінці $(i-1)$ -го року (або на початку i -го року) для інвестування підприємств у i -му році. Тоді обсяг КВ, виділений підприємству Π_2 у i -му році, становить $(x^{(i-1)} - u^{(i)})$ од., а сумарний прибуток – $[k^{(1)}(u^{(i)}) + k^{(2)}(x^{(i-1)} - u^{(i)})]$ од. Оскільки $0 \leq u^{(i)} \leq x^{(i-1)}$, а рівняння станів описуються співвідношеннями $x^{(i)} = f^{(1)}(u^{(i)}) + f^{(2)}(x^{(i-1)} - u^{(i)})$ ($i=\overline{1,m}$), то математична модель задачі про m -етапний розподіл КВ обсягом s од. між двома підприємствами Π_1 і Π_2 матиме вигляд

$$\begin{cases} \sum_{i=1}^m [k^{(1)}(u^{(i)}) + k^{(2)}(x^{(i-1)} - u^{(i)})] \rightarrow \max, \\ x^{(i)} = f^{(1)}(u^{(i)}) + f^{(2)}(x^{(i-1)} - u^{(i)}), i=\overline{1,m}, \\ u^{(i)} \in [0, x^{(i-1)}], i=\overline{1,m}, \\ x^{(0)} = s, x^{(m)} \geq 0. \end{cases} \quad (9.16)$$

Зауважимо, що умова $x^{(m)} \geq 0$ означає, що в кінці планового періоду (m -го року) може ще залишитись деякий обсяг КВ для подальшого інвестування підприємств Π_1 і Π_2 .

Розв'яжемо задачу (9.16) методом ДП для таких конкретних вихідних даних:

$$m = 4, \quad s = 200, \quad k^{(1)}(u) = 0,3u, \quad k^{(2)}(u) = 0,4u, \\ f^{(1)}(u) = 0,8u, \quad f^{(2)}(u) = 0,5u.$$

Очевидно, в цьому випадку ми маємо справу з неперервною моделлю задачі (9.16).

Умовна оптимізація

Крок 1. Із (9.10) випливає, що для $m = 4$

$$B_4(x^{(3)}) = \max_{u^{(4)} \in [0, x^{(3)}]} [k^{(1)}(u^{(4)}) + k^{(2)}(x^{(3)} - u^{(4)})] = \\ = \max_{u^{(4)} \in [0, x^{(3)}]} [0,3u^{(4)} + 0,4(x^{(3)} - u^{(4)})] = \\ = \max_{u^{(4)} \in [0, x^{(3)}]} [0,4x^{(3)} - 0,1u^{(4)}] = 0,4x^{(3)}.$$

Умовне оптимальне керування $u^{(4)*}(x^{(3)}) = 0$.

Крок 2. Для $i = m - 1 = 3$ згідно із (9.11)

$$B_3(x^{(2)}) = \max_{u^{(3)} \in [0, x^{(2)}]} [k^{(1)}(u^{(3)}) + k^{(2)}(x^{(2)} - u^{(3)}) + B_4(x^{(3)})] = \\ = \max_{u^{(3)} \in [0, x^{(2)}]} [k^{(1)}(u^{(3)}) + \\ + k^{(2)}(x^{(2)} - u^{(3)}) + B_4(f^{(1)}(u^{(3)}) + f^{(2)}(x^{(2)} - u^{(3)}))] = \\ = \max_{u^{(3)} \in [0, x^{(2)}]} [0,3u^{(3)} + \\ + 0,4(x^{(2)} - u^{(3)}) + 0,4(0,8u^{(3)} + 0,5(x^{(2)} - u^{(3)}))] = \\ = \max_{u^{(3)} \in [0, x^{(2)}]} [0,02u^{(3)} + 0,6x^{(2)}] = 0,62x^{(2)}.$$

Умовне оптимальне керування $u^{(3)*}(x^{(2)}) = x^{(2)}$.

Крок 3. Для $i = m - 2 = 2$ згідно із (9.11)

$$B_2(x^{(1)}) = \max_{u^{(2)} \in [0, x^{(1)}]} [k^{(1)}(u^{(2)}) + k^{(2)}(x^{(1)} - u^{(2)}) + B_3(x^{(2)})] =$$

$$\begin{aligned}
&= \max_{u^{(2)} \in [0, x^{(1)}]} \left[k^{(1)}(u^{(2)}) + \right. \\
&\quad \left. + k^{(2)}(x^{(1)} - u^{(2)}) + B_3(f^{(1)}(u^{(2)}) + f^{(2)}(x^{(1)} - u^{(2)})) \right] = \\
&= \max_{u^{(2)} \in [0, x^{(1)}]} \left[0, 3u^{(2)} + \right. \\
&\quad \left. + 0,4(x^{(1)} - u^{(2)}) + 0,62(0,8u^{(2)} + 0,5(x^{(1)} - u^{(2)})) \right] = \\
&= \max_{u^{(2)} \in [0, x^{(1)}]} \left[0,71x^{(1)} + 0,086u^{(2)} \right] = 0,796x^{(1)}.
\end{aligned}$$

Умовне оптимальне керування $u^{(2)*}(x^{(1)}) = x^{(1)}$.

Крок 4. Для $i = m - 3 = 1$ згідно із (9.11)

$$\begin{aligned}
B_1(x^{(0)}) &= \max_{u^{(1)} \in [0, x^{(0)}]} \left[k^{(1)}(u^{(1)}) + k^{(2)}(x^{(0)} - u^{(1)}) + B_2(x^{(1)}) \right] = \\
&= \max_{u^{(1)} \in [0, x^{(0)}]} \left[k^{(1)}(u^{(1)}) + \right. \\
&\quad \left. + k^{(2)}(x^{(0)} - u^{(1)}) + B_2(f^{(1)}(u^{(1)}) + f^{(2)}(x^{(0)} - u^{(1)})) \right] = \\
&= \max_{u^{(1)} \in [0, x^{(0)}]} \left[0, 3u^{(1)} + \right. \\
&\quad \left. + 0,4(x^{(0)} - u^{(1)}) + 0,796(0,8u^{(1)} + 0,5(x^{(0)} - u^{(1)})) \right] = \\
&= \max_{u^{(1)} \in [0, x^{(0)}]} \left[0,798x^{(0)} + 0,1388u^{(1)} \right] = 0,9368x^{(0)}.
\end{aligned}$$

Умовне оптимальне керування $u^{(1)*}(x^{(0)}) = x^{(0)}$.

Безумовна оптимізація

Крок 1. Оскільки $x^{(0)} = s = 200$, то розв'язком задачі (9.12) є $x^{(0)*} = 200$ і

$$\begin{aligned}
B_1(200) &= 0,9368 \cdot 200 = 187,36, \\
u^{(1)*} &= 200, \quad x^{(1)*} = f^{(1)}(u^{(1)*}) + f^{(2)}(x^{(0)*} - u^{(1)*}) = \\
&= 0,8 \cdot 200 + 0,5 \cdot (200 - 200) = 160.
\end{aligned}$$

Крок 2. Із (9.13) випливає, що $u^{(2)*} = u^{(2)*}(x^{(1)*}) = x^{(1)*} = 160$,

$$x^{(2)*} = f^{(1)}(u^{(2)*}) + f^{(2)}(x^{(1)*} - u^{(2)*}) = \\ = 0,8 \cdot 160 + 0,5 \cdot (160 - 160) = 128.$$

Крок 3. Підставимо у (9.13) знайдене $x^{(2)*}$, отримаємо $u^{(3)*} = u^{(3)*}(x^{(2)*}) = x^{(2)*} = 128$,

$$x^{(3)*} = f^{(1)}(u^{(3)*}) + f^{(2)}(x^{(2)*} - u^{(3)*}) = \\ = 0,8 \cdot 128 + 0,5 \cdot (128 - 128) = 102,4.$$

Крок 4. На основі попередніх результатів із (9.13) отримуємо $u^{(4)*} = u^{(4)*}(x^{(3)*}) = 0$,

$$x^{(4)*} = f^{(1)}(u^{(4)*}) + f^{(2)}(x^{(3)*} - u^{(4)*}) = \\ = 0,8 \cdot 0 + 0,5 \cdot (102,4 - 0) = 51,2.$$

Отже, $\{u^{(1)*}, u^{(2)*}, u^{(3)*}, u^{(4)*}\} = \{200, 160, 128, 0\}$ і

$\{v^{(1)*}, v^{(2)*}, v^{(3)*}, v^{(4)*}\} = \{0; 0; 0; 102,4\}$ – оптимальні розподіли КВ для першого та другого підприємств відповідно. Оптимальна траєкторія

$$\{x^{(0)*}, x^{(1)*}, x^{(2)*}, x^{(3)*}, x^{(4)*}\} = \{200; 160; 128; 102,4; 51,2\},$$

а оптимальний прибуток $k^* = 187,36$ од. ►

Приклад 9.4 (задача про t -етапний розподіл КВ заданого обсягу між l підприємствами $\Pi_1, \Pi_2, \dots, \Pi_l$).

Планується розподіл КВ обсягом s од. між підприємствами $\Pi_1, \Pi_2, \dots, \Pi_l$ ($l > 2$) на t років планового періоду. Відомо, що u од. КВ, укладених у підприємство Π_j в i -ому році, дають цьому підприємству прибуток $k^{(j,i)}(u)$ од. і до кінця року зменшуються до $f^{(j,i)}(u)$ од., які знову можуть бути використані для інвестування підприємств у майбутньому. Потрібно знайти такий варіант розподілу КВ

між підприємствами, при якому сумарний прибуток l підприємств за t років буде максимальним.

◀ Очевидно, приклад 9.4 є узагальненням прикладу 9.3. Нехай $u^{(j,i)}$ – обсяг КВ, виділений підприємству Π_j у i -ому році планового періоду, а $x^{(i-1)}$ (як і раніше) – обсяг КВ, який залишився у кінці $(i-1)$ -го року для інвестування підприємств у i -му році. Тоді рівняння стану процесу розподілу КВ матиме вигляд

$$x^{(i)} = x^{(i-1)} - \sum_{j=1}^l u^{(j,i)} + \sum_{j=1}^l f^{(j,i)}(u^{(j,i)}).$$

Показником ефективності процесу буде сумарний прибуток $\sum_{i=1}^m \sum_{j=1}^l k^{(j,i)}(u^{(j,i)})$, а множиною допустимих керувань (тобто вкладів) у i -му році – множина, яка описується співвідношенням

$$\sum_{j=1}^l u^{(j,i)} = x^{(i-1)}.$$

Отже, математична модель задачі про t -етапний розподіл КВ заданого обсягу s од. між підприємствами $\Pi_1, \Pi_2, \dots, \Pi_l$ набуде вигляду

$$\left\{ \begin{array}{l} \sum_{i=1}^m \sum_{j=1}^l k^{(j,i)}(u^{(j,i)}) \rightarrow \max, \\ x^{(i)} = x^{(i-1)} - \sum_{j=1}^l u^{(j,i)} + \sum_{j=1}^l f^{(j,i)}(u^{(j,i)}), \quad i = \overline{1, m}, \\ \sum_{j=1}^l u^{(j,i)} = x^{(i-1)}, \quad i = \overline{1, m}, \\ x^{(0)} = s, \quad x^{(m)} \geq 0. \end{array} \right. \quad (9.17)$$

Зрозуміло, що задача (9.17) складніша, ніж задача (9.16), але розв'язуються обидві задачі аналогічно. ►

Приклад 9.5 (сепарабельна задача цілочислового математичного програмування).

Розв'язати методом ДП задачу математичного програмування

$$\begin{cases} \sum_{i=1}^m k^{(i)}(u^{(i)}) \rightarrow \max, \\ \sum_{i=1}^m a^{(j,i)} u^{(i)} \leq b_j, \quad j = \overline{1, l}, \\ u^{(i)} \geq 0, \quad u^{(i)} \in \mathbb{Z}, \quad i = \overline{1, m}, \end{cases} \quad (9.18)$$

де $a^{(j,i)} \geq 0$, $b_j \geq 0$, $j = \overline{1, l}$; $i = \overline{1, m}$; $\mathbb{Z}$ – множина цілих чисел.

◀ Задача (9.18) називається сепарабельною, оскільки цільова функція й обмеження задачі сепарабельні, тобто функції, які входять до умов задачі, є функціями вигляду

$$\varphi(x_1, \dots, x_n) = \sum_{i=1}^n \varphi_i(x_i).$$

Задача (9.18) зводиться до багатокрокової задачі оптимізації, яку можна розв'язати методом ДП. Справді, нехай

$$x^{(j,i)} = b_j - \sum_{r=1}^i a^{(j,r)} u^{(r)}, \quad i = \overline{1, m}, \quad j = \overline{1, l};$$

$$x^{(i)} = (x^{(1,i)}, \dots, x^{(l,i)})^T, \quad a^{(i)} = (a^{(1,i)}, \dots, a^{(l,i)})^T,$$

$$b = (b_1, \dots, b_l)^T, \quad 0 = \underbrace{(0, \dots, 0)}_l^T.$$

Тоді задачу (9.18) можна записати у вигляді

$$\left\{ \begin{array}{l} \sum_{i=1}^m k^{(i)}(u^{(i)}) \rightarrow \max, \\ x^{(i)} = x^{(i-1)} - u^{(i)} a^{(i)}, \quad i = \overline{1, m}, \\ x^{(i)} \geq 0, \quad i = \overline{1, m-1}, \\ u^{(i)} \geq 0, \quad u^{(i)} \in \mathbb{Z}, \quad i = \overline{1, m}, \\ x^{(0)} = b, \quad x^{(m)} \geq 0. \end{array} \right. \quad (9.19)$$

Відзначимо також, що задача (9.18) є математичною моделлю багатьох реальних задач, зокрема задач економіки. Та обставина, що задачу (9.18) можна звести до задачі (9.19), надзвичайно важлива, оскільки в багатьох випадках це спрощує процес розв'язування задачі (9.18).

Розглянемо наступну конкретну задачу типу (9.18):

$$\left\{ \begin{array}{l} (4u^{(1)} + 3u^{(2)}) \rightarrow \max, \\ 4u^{(1)} + u^{(2)} \leq 10, \\ 2u^{(1)} + 3u^{(2)} \leq 8, \\ u^{(1)}, u^{(2)} \geq 0; \quad u^{(1)}, u^{(2)} \in \mathbb{Z}. \end{array} \right. \quad (9.20)$$

Задача (9.20) є задачею цілочислового лінійного програмування. Вона зводиться до задачі (9.19) з такими конкретними вихідними даними:

$$\begin{aligned} m &= 2, \quad l = 2, \quad 0 = (0, 0)^T, \quad b = (10, 8)^T, \\ x^{(0)} &= b, \quad a^{(1)} = (4, 2)^T, \quad a^{(2)} = (1, 3)^T, \\ x^{(1)} &= x^{(0)} - u^{(1)} a^{(1)} = (10, 8)^T - u^{(1)} (4, 2)^T = \\ &= (10 - 4u^{(1)}, 8 - 2u^{(1)})^T = (x^{(1,1)}, x^{(2,1)})^T, \\ x^{(2)} &= x^{(1)} - u^{(2)} a^{(2)} = (10 - 4u^{(1)}, 8 - 2u^{(1)})^T - \end{aligned}$$

$$-u^{(2)}(1,3)^T = (10 - 4u^{(1)} - u^{(2)}, 8 - 2u^{(1)} - 3u^{(2)})^T = (x^{(1,2)}, x^{(2,2)})^T,$$

$$k^{(1)}(u^{(1)}) = 4u^{(1)}, k^{(2)}(u^{(2)}) = 3u^{(2)}.$$

Розв'яжемо сформульовану задачу методом ДП.

Умовна оптимізація

Крок 1. З рівняння стану при $i=2$ випливає, що $x^{(1)} - u^{(2)}a^{(2)} \geq 0$ або $u^{(2)}a^{(2)} \leq x^{(1)}$. Це означає, що

$$\tilde{U}^{(2)}(x^{(1)}) = \{u^{(2)} \geq 0, u^{(2)} \in \mathbb{Z} / u^{(2)} \leq x^{(1,1)}, u^{(2)} \leq 1/3x^{(2,1)}\}.$$

Отже, згідно із (9.10)

$$B_2(x^{(1)}) = \max_{u^{(2)} \in \tilde{U}^{(2)}(x^{(1)})} k^{(2)}(u^{(2)}) =$$

$$= \max_{u^{(2)} \in \tilde{U}^{(2)}(x^{(1)})} \{3u^{(2)}\} = 3 \min \{ \lceil x^{(1,1)} \rceil, \lceil 1/3x^{(2,1)} \rceil \},$$

де $\lceil x^{(1,1)} \rceil$ – ціла частина $x^{(1,1)}$, $\lceil 1/3x^{(2,1)} \rceil$ – ціла частина $1/3x^{(2,1)}$.

Відповідне умовне оптимальне керування

$$u^{(2)*}(x^{(1)}) = \min \{ \lceil x^{(1,1)} \rceil, \lceil 1/3x^{(2,1)} \rceil \}.$$

Крок 2. Допустима множина керувань

$$\tilde{U}^{(1)}(x^{(0)}) = \{u^{(1)} \geq 0, u^{(1)} \in \mathbb{Z} / u^{(1)}a^{(1)} \leq x^{(0)}\} =$$

$$= \{u^{(1)} \geq 0, u^{(1)} \in \mathbb{Z} / u^{(1)} \leq 2,5\}.$$

Із (9.11) випливає, що

$$B_1(x^{(0)}) = \max_{u^{(1)} \in \tilde{U}^{(1)}(x^{(0)})} \{k^{(1)}(u^{(1)}) + B_2(x^{(1)})\} =$$

$$= \max_{0 \leq u^{(1)} \leq 2,5} \left\{ 4u^{(1)} + 3 \min \left\{ \lceil 10 - 4u^{(1)} \rceil, \lceil 1/3(8 - 2u^{(1)}) \rceil \right\} \right\} =$$

$$= 4 \cdot 2 + 3 \min \{ \lceil 2 \rceil, \lceil 4/3 \rceil \} = 8 + 3 \cdot 1 = 11.$$

Відповідне оптимальне керування $u^{(1)*}(x^{(0)}) = [2, 5] = 2$.

Безумовна оптимізація

Крок 1. Оскільки $x^{(0)*} = (10, 8)^T$, а $u^{(1)*} = 2$, то

$$x^{(1)*} = x^{(0)} - 2a^{(1)} = (10, 8)^T - 2(4, 2)^T = (2, 4)^T.$$

Крок 2. Для відомого $x^{(1)*}$ із (9.13) маємо

$$u^{(2)*} = u^{(2)*}(x^{(1)*}) = \min\{[2], [4/3]\} = 1,$$

$$x^{(2)*} = x^{(1)*} - u^{(2)*}a^{(2)} = (2, 4)^T - 1 \cdot (1, 3)^T = (1, 1)^T.$$

Отже, оптимальний розв'язок задачі (9.20) такий:

$$\{u^{(1)*}, u^{(2)*}\} = \{2, 1\}, \quad f^* = 4 \cdot 2 + 3 \cdot 1 = 11. \quad \blacktriangleright$$

Аналогічно до задачі (9.18) можна сформулювати канонічну сепарабельну задачу, яка також зводиться до багатокрокової задачі оптимізації типу (9.19). Останню успішно розв'язують методом ДП.

Приклад 9.6. Розв'язати методом ДП задачу ПЦЛП (7.106).

◀ Зробимо перепозначення змінних: $x_i = u^{(i)}, i = \overline{1, 4}$.

Вихідні дані еквівалентної динамічної задачі такі:

$$m = 4, l = 2, b = (13, 6)^T, x^{(0)} = b, a^{(1)} = (2, 1)^T,$$

$$a^{(2)} = (2, -1)^T, a^{(3)} = (1, 0)^T, a^{(4)} = (0, 1)^T,$$

$$x^{(1)} = x^{(0)} - u^{(1)}a^{(1)} = (13 - 2u^{(1)}, 6 - u^{(1)})^T = (x^{(1,1)}, x^{(2,1)})^T,$$

$$x^{(2)} = x^{(1)} - u^{(2)}a^{(2)} = (13 - 2u^{(1)} - 2u^{(2)}, 6 - u^{(1)} + u^{(2)})^T =$$

$$= (x^{(1,2)}, x^{(2,2)})^T,$$

$$x^{(3)} = x^{(2)} - u^{(3)}a^{(3)} = (13 - 2u^{(1)} - 2u^{(2)} - u^{(3)}, 6 - u^{(1)} + u^{(2)})^T =$$

$$\begin{aligned}
&= \left(x^{(1,3)}, x^{(2,3)} \right)^T, \\
x^{(4)} &= x^{(3)} - u^{(4)} a^{(4)} = \\
&= \left(13 - 2u^{(1)} - 2u^{(2)} - u^{(3)}, 6 - u^{(1)} + u^{(2)} - u^{(4)} \right)^T = \left(x^{(1,4)}, x^{(2,4)} \right)^T, \\
k^{(1)}(u^{(1)}) &= u^{(1)}, k^{(2)}(u^{(2)}) = 2u^{(2)}, k^{(3)}(u^{(3)}) = k^{(4)}(u^{(4)}) = 0.
\end{aligned}$$

Умовна оптимізація

Крок 1. З рівняння стану для $i = 4$ випливає, що $x^{(4)} = 0$, тобто $x^{(3)} = u^{(4)} a^{(4)}$ або $u^{(4)} = x^{(2,3)}$. Це означає, що

$$\tilde{U}^{(4)}(x^{(3)}) = \{u^{(4)} \geq 0, u^{(4)} \in \mathbb{Z} / x^{(3)} - u^{(4)} a^{(4)} = 0\} = \left\{ \left\lceil x^{(2,3)} \right\rceil \right\},$$

тому, згідно із (9.10),

$$B_4(x^{(3)}) = \max_{u^{(4)} \in \tilde{U}^{(4)}(x^{(3)})} k^{(4)}(u^{(4)}) = 0.$$

Відповідне умовне оптимальне керування

$$u^{(4)*}(x^{(3)}) = \left\lceil x^{(2,3)} \right\rceil.$$

Крок 2. Оскільки $x^{(1,3)} = 0$, $x^{(2,3)} \geq 0$, то з рівняння стану для $i = 3$ випливає, що $u^{(3)} = x^{(1,2)}$. Отже,

$$\tilde{U}^{(3)}(x^{(2)}) = \{u^{(3)} \geq 0, u^{(3)} \in \mathbb{Z} / x^{(1,3)} = 0, x^{(2,3)} \geq 0\} = \left\{ \left\lceil x^{(1,2)} \right\rceil \right\},$$

тобто відповідно до (9.11)

$$B_3(x^{(2)}) = \max_{u^{(3)} \in \tilde{U}^{(3)}(x^{(2)})} \{k^{(3)}(u^{(3)}) + B_4(x^{(3)})\} = 0.$$

Умовне оптимальне керування $u^{(3)*}(x^{(2)}) = \left\lceil x^{(1,2)} \right\rceil$.

Крок 3. З рівняння стану для $i = 2$ випливає, що $x^{(1)} - u^{(2)} a^{(2)} \geq 0$, або $u^{(2)} a^{(2)} \leq x^{(1)}$. Отже,

$$\tilde{U}^{(2)}(x^{(1)}) = \{u^{(2)} \geq 0, u^{(2)} \in \mathbb{Z} / -x^{(2,1)} \leq u^{(2)} \leq 1 / 2x^{(1,1)}\},$$

тому із (9.11) випливає

$$B_2(x^{(1)}) = \max_{u^{(2)} \in U^{(2)}(x^{(1)})} \{k^{(2)}(u^{(2)}) + B_3(x^{(2)})\} = 2 \lceil 1/2x^{(1,1)} \rceil.$$

Відповідне умовне оптимальне керування

$$u^{(2)*}(x^{(1)}) = \lceil 1/2x^{(1,1)} \rceil.$$

Крок 4. З рівняння стану для $i=1$ випливає, що $x^{(0)} \geq u^{(1)}a^{(1)}$. Отже,

$$\begin{aligned} \tilde{U}^{(1)}(x^{(0)}) &= \{u^{(1)} \geq 0, u^{(1)} \in \mathbb{Z} / u^{(1)} \leq 1/2x^{(1,0)}, u^{(1)} \leq x^{(2,0)}\} = \\ &= \{u^{(1)} \in \mathbb{Z} / 0 \leq u^{(1)} \leq \min(1/2x^{(1,0)}, x^{(2,0)})\} = \\ &= \{u^{(1)} \in \mathbb{Z} / 0 \leq u^{(1)} \leq 6\}. \end{aligned}$$

Із (9.11) випливає, що

$$\begin{aligned} B_1(x^{(0)}) &= \max_{u^{(1)} \in U^{(1)}(x^{(0)})} \{k^{(1)}(u^{(1)}) + B_2(x^{(1)})\} = \\ &= \max_{u^{(1)} \in U^{(1)}(x^{(0)})} \{u^{(1)} + 2 \lceil 1/2x^{(1,1)} \rceil\} = \\ &= \max_{u^{(1)} \in U^{(1)}(x^{(0)})} \{u^{(1)} + 2 \lceil 1/2(13 - 2u^{(1)}) \rceil\} = 12. \end{aligned}$$

Відповідне умовне оптимальне керування $u^{(1)*}(x^{(0)}) = 0$.

Безумовна оптимізація

Крок 1. Оскільки $u^{(1)*} = 0$, то

$$x^{(1)*} = (13 - 2 \cdot 0, 6 - 1 \cdot 0)^T = (13, 6)^T.$$

Крок 2.

$$u^{(2)*} = \lceil 1/2(13 - 2u^{(1)*}) \rceil = \lceil 13/2 \rceil = 6,$$

$$x^{(2)*} = (13 - 2 \cdot 0 - 2 \cdot 6, 6 - 0 + 6)^T = (1, 12)^T.$$

Крок 3.

$$u^{(3)*} = \lceil x^{(1,2)} \rceil = \lceil 13 - 2u^{(1)*} - 2u^{(2)*} \rceil = \lceil 1 \rceil = 1,$$

$$x^{(3)*} = (13 - 2 \cdot 0 - 2 \cdot 6 - 1, 6 - 0 + 6)^T = (0, 12)^T.$$

Крок 4.

$$u^{(4)*} = [x^{(2,3)}] = [6 - u^{(1)*} + u^{(2)*}] = [12] = 12,$$

$$x^{(4)*} = (13 - 2 \cdot 0 - 2 \cdot 6 - 1, 6 - 0 + 6 - 12)^T = (0, 0)^T.$$

Отже, маємо оптимальне керування

$$\{u^{(1)*}, u^{(2)*}, u^{(3)*}, u^{(4)*}\} = \{0, 6, 1, 12\} \text{ (при цьому } k^* = 12),$$

тобто розв'язок задачі (7.106). ►

Розділ 10. Комп'ютерні технології розв'язування оптимізаційних задач

Сучасні можливості застосування комп'ютерних систем для розв'язування оптимізаційних задач та їх аналізу достатньо широкі. Нижче для демонстрації знаходження розв'язків оптимізаційних задач використаємо загальновідомі і популярні електронні таблиці MS Excel, мову програмування R та середовище Matlab.

10.1. Стандартна задача лінійного програмування та двоїста до неї

Приклад 10.1 (П.10.1). Підприємство виготовляє і продає два типи продукції P_1 та P_2 . Для їх виробництва використовується сировина A та B , витрати якої на одиницю продукції A та B , а також їх запас на складі наведені у таблиці 10.1.

Таблиця 10.1

Сировина	Витрата продуктів на одиницю продукції		Запас сировини на складі
	P_1	P_2	
A	2	3	10
B	3	2	15

Ціна за одиницю продукції P_1 та P_2 складає 30 та 10 гр. од. відповідно. Крім того відомо, що попит на продукцію P_2 ніколи не перевищував 2 од. за добу, а попит на продукцію P_1 – не перевищував попит на продукцію P_2 більше, ніж на одну одиницю.

Потрібно визначити, яку кількість продукції кожного типу слід виробляти підприємству, щоб одержати максимальний дохід.

Для розв'язування цієї задачі спочатку сформуємо таблицю 10.2 із описом окремих елементів моделі задачі (П.10.1) та побудуємо її математичну модель.

Таблиця 10.2

№ п/п	Елементи моделі	Економічний Зміст	Математична формалізація
1.	Шукані змінні	Кількість продукції типу Π_1 та Π_2	x_1, x_2
2.	Цільова функція	Максимальний дохід від реалізації продукції	$30x_1 + 10x_2 \rightarrow \max$
3.	Основні обмеження моделі	Максимальні витрати сировини $A - 10$ од.	$2x_1 + 3x_2 \leq 10$
		Максимальні витрати сировини $B - 15$ од.	$3x_1 + 2x_2 \leq 15$
		Попит на прод. Π_2 не перевищує 2 од.	$x_2 \leq 2$
		Попит на прод. Π_1 не перевищує попит на продукцію Π_2 більше, ніж на 1 од.	$x_1 - x_2 \leq 1$
4.	Прямі обмеження	Невід'ємна кількість продукції Π_1 та Π_2	$x_1 \geq 0, x_2 \geq 0.$

Із урахуванням таблиці 10.2 отримуємо відповідну оптимізаційну модель. Очевидно, що її формалізацією є стандартна задача лінійного програмування:

$$\left\{ \begin{array}{l} f(x_1, x_2) = 30x_1 + 10x_2 \rightarrow \max, \\ 2x_1 + 3x_2 \leq 10, \\ 3x_1 + 2x_2 \leq 15, \\ x_1 - x_2 \leq 1, \\ x_2 \leq 2, \\ x_1 \geq 0, x_2 \geq 0. \end{array} \right. \quad (10.1)$$

Далі продемонструємо побудову комп'ютерної моделі задачі (10.1) та її розв'язування в MS Excel, R та Matlab.

Розв'язування задачі у MS Excel

Спочатку для задачі (10.1) сформуємо таблицю із коефіцієнтів цільової функції та обмежень (рис. 10.1):

- для змінних задачі x_1 та x_2 відведемо комірки B3 і C3 (вони називаються робочими або змінними комірками, причому змінні комірки спочатку залишаються порожніми, а після завершення розв'язування задачі тут будуть знаходитися оптимальні значення змінних);
- у комірки B4 і C4 занесемо коефіцієнти цільової функції;
- у комірки B5:C8 введемо коефіцієнти основних обмежень.

	А	В	С
1	Елементи моделі	Продукція	
2		П1	П2
3	План випуску продукції		
4	Дохід, гр. од.	30	10
5	Витрати сировини А, од. сировини	2	3
6	Витрати сировини В, од. сировини	3	2
7	Співвідношення між обсягами виробництва	1	-1
8	Обсяг продукту П2	0	1

Рис. 10.1. Коефіцієнти цільової функції та обмежень

Далі у комірки D4:D8 вводимо розрахункові формули для цільової функції та обмежень (рис. 10.2).

	A	B	C	D
1	Елементи моделі	Продукція		
2		П1	П2	
3	План випуску продукції			По плану
4	Дохід, гр. од.	30	10	=СУММПРОИЗВ(\$B\$3:\$C\$3;B4:C4)
5	Витрати сировини А, од. сировини	2	3	=СУММПРОИЗВ(\$B\$3:\$C\$3;B5:C5)
6	Витрати сировини В, од. сировини	3	2	=СУММПРОИЗВ(\$B\$3:\$C\$3;B6:C6)
7	Співвідношення між обсягами виробництва	1	-1	=СУММПРОИЗВ(\$B\$3:\$C\$3;B7:C7)
8	Обсяг продукту П2	0	1	=СУММПРОИЗВ(\$B\$3:\$C\$3;B8:C8)

Рис. 10.2. Розрахункові формули для цільової функції та обмежень

Після цього, уточнимо критерій оптимізації, знаки обмежень і граничні значення задачі, у результаті чого фактично отримаємо її відповідну комп'ютерну модель (рис. 10.3).

	A	B	C	D	E	F
1	Елементи моделі	Продукція				
2		П1	П2			
3	План випуску продукції			По плану	Умова	Граничні значення
4	Дохід, гр. од.	30	10	0	max	
5	Витрати сировини А, од. сировини	2	3	0	<=	10
6	Витрати сировини В, од. сировини	3	2	0	<=	15
7	Співвідношення між обсягами виробництва	1	-1	0	<=	1
8	Обсяг продукту П2	0	1	0	<=	2

Рис. 10.3. Комп'ютерна модель задачі (10.1)

Зауважимо, що символи, що знаходяться у комірках E5:E8 у розрахунках не використовуються, проте дозволяють краще орієнтуватися у моделі.

Для знаходження розв'язку оптимізаційної задачі (10.1) у меню **Дані** вибираємо процедуру **Пошук рішення**. Під час першого звернення до неї необхідно активізувати режим **Параметри Excel** ⇒ **Надбудова**, в діалоговому вікні вибрати неактивну надбудову **Пошук рішення** та натиснути **ОК**.

У вікні, що з'явилося, додаємо обмеження (рис. 10.4), вводимо необхідні посилання (рис. 10.5), фіксуємо галочку для забезпечення виконання невід'ємності обсягів продукції P_1 і P_2 та вибираємо метод розв'язування задачі у випадяючому списку.

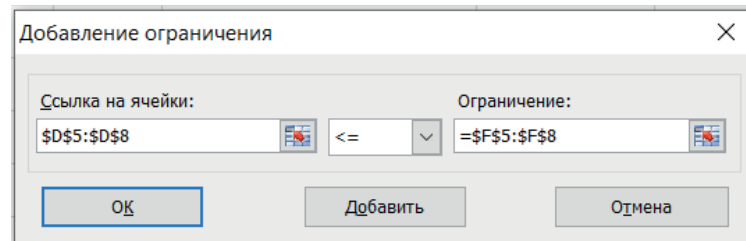


Рис. 10.4. Додавання обмежень

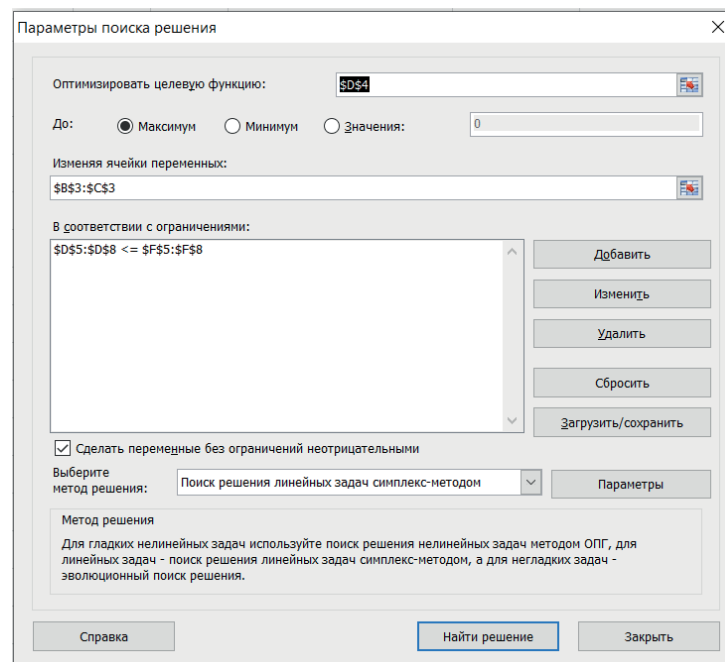


Рис. 10.5. Параметры нахождения розв'язку задачі (10.1)

Для знаходження розв'язку задачі необхідно натиснути кнопку «Знайти розв'язок». При цьому можливі такі ситуації:

1. Отримано повідомлення «Розв'язок знайдено» (рис. 10.6). Тоді, після натискання кнопки ОК, буде отримано оптимальний розв'язок. У даному випадку розв'язок

$x^* = (2,6; 1,6)^T$ (T – транспонування) знаходиться у комірках B3 та C3, а максимальний дохід $f_{\max}^* = 94$ – у комірці D4. На основі даних з робочого листа можна зробити висновок, що ресурс A дефіцитний, оскільки його запас та витрати збігаються, а ресурс B недефіцитний, оскільки його запас перевищує витрати.

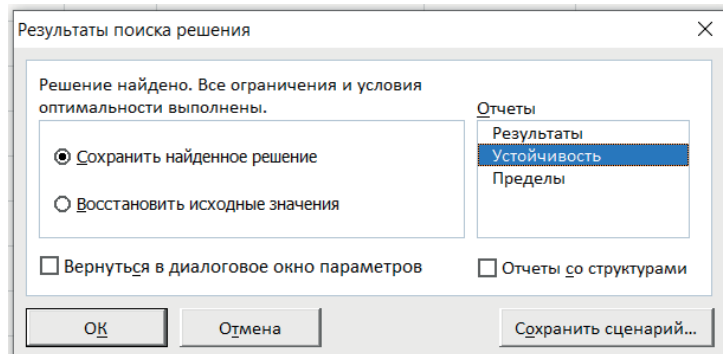


Рис. 10.6. Результати роботи надбудови «Пошук рішення»

2. Отримано повідомлення «Значення цільової комірки не збігаються». Тоді це означає, що задача лінійного програмування не має розв'язку, оскільки цільова функція не є обмеженою зверху (знизу) на допустимій множині задачі.

3. Отримано повідомлення «Пошук не може знайти допустимого розв'язку». У цьому випадку область допустимих розв'язків є порожньою множиною, а, отже, задача розв'язків не має.

Зупинимось коротко також на двоїстій задачі до задачі (10.1).

Двоїстою до (10.1) буде така задача ЛП:

$$\begin{cases} \tilde{f} = 10\lambda_1 + 15\lambda_2 + \lambda_3 + 2\lambda_4 \rightarrow \min, \\ 2\lambda_1 + 3\lambda_2 + \lambda_3 \geq 30, \\ 3\lambda_1 + 2\lambda_2 - \lambda_3 + \lambda_4 \geq 10, \\ \lambda_j \geq 0, j = \overline{1,4}. \end{cases} \quad (10.2)$$

За допомогою надбудови «Пошук рішення» можна отримати оптимальний розв'язок $\lambda^* = (8; 0; 14; 0)^T$, $\tilde{f}_{\min}^* = 94$. У плані економічного змісту числа $\lambda_1^* = 8$ та $\lambda_2^* = 0$ є оптимальними цінами на ресурси А та В, що, зокрема, підтверджує той факт, що ресурс А є дефіцитним (цінним для виробництва), а ресурс В не є дефіцитним. Числа 14 та 0 із розв'язку λ^* є оцінками чутливості оптимального значення цільової функції до зміни вимог щодо співвідношення між обсягами виробництва першого та другого типів продукції, а також обсягу другого типу продукції відповідно.

Разом з тим, розв'язок задачі (10.2), двоїстої до задачі (10.1), можна знайти автоматично, побудувавши звіт про стійкість для задачі (10.1) (рис. 10.7). Цей розв'язок знаходиться у стовпці «Тіньова ціна» (рис. 10.7).

	A	B	C	D	E	F	G	H
1	Microsoft Excel 14.0 Отчет об устойчивости							
2	Лист: [Книга1.xlsx]Лист1							
3	Отчет создан: 05.01.2022 12:12:35							
4								
5								
6	Ячейки переменных							
7								
8	Ячейка	Имя	Окончательное Значение	Приведенн. Стоимость	Целевая функция Коэффициент	Допустимое Увеличение	Допустимое Уменьшение	
9	\$B\$3	План выпуска продукции П1	2,6	0	30	1E+30	23,33333333	
10	\$C\$3	План выпуска продукции П2	1,6	0	10	35	40	
11								
12	Ограничения							
13								
14	Ячейка	Имя	Окончательное Значение	Тень Цена	Ограничение Правая сторона	Допустимое Увеличение	Допустимое Уменьшение	
15	\$D\$5	Витрати сировини А, од. сировини По плану	10	8	10	2	8	
16	\$D\$6	Витрати сировини В, од. сировини По плану	11	0	15	1E+30	4	
17	\$D\$7	Співвідношення між обсягами виробництва По плану	1	14	1	4	1	
18	\$D\$8	Обсяг продукту П2 По плану	1,6	0	2	1E+30	0,4	
19								

Рис. 10.7. Звіт про стійкість задачі (10.1)

Звіт про стійкість задачі використовується для аналізу чутливості знайденого розв'язку до змін параметрів задачі. Він не створюється для моделей, змінні яких повинні набувати цілочислових значень.

У першому блоці цього звіту у стовпці «Кінцеве значення» знаходиться розв'язок задачі (10.1). Стовпець «Зведена вартість» показує на скільки зміниться цільова функція, якщо обсяг продукції, який у оптимальному розв'язку дорівнює нулеві, збільшити на одиницю. У нашому випадку оптимальний розв'язок (оптимальний план) передбачає виробництво обох типів продукції, тому їхня зведена вартість дорівнює нулю. Якби оптимальне значення будь-якої з невідомих дорівнювало б нулю, а зведена вартість дорівнювала б, наприклад, (-3) , то примусовий випуск двох одиниць цієї продукції призвів би до зміни цільової функції на $2 \cdot (-3) = (-6)$ одиниць. Зазначимо, що із того, що оптимальне значення деякої змінної дорівнює нулю, не випливає автоматично, що її зведена вартість відмінна від нуля.

У стовпцях «Допустиме збільшення» і «Допустиме зменшення» першого блоку звіту показані допустимі межі зміни коефіцієнтів цільової функції, при яких зберігається даний оптимальний план. Наприклад, із рис. 10.7 випливає, що даний оптимальний план не зміниться, якщо ціну на продукцію P_2 , що дорівнює 10 грошових одиниць (гр. од.), збільшити на 35 гр. од., а ціну на продукцію P_1 , що дорівнює 30 гр. од. – поставити як завгодно великою.

Другий блок звіту про стійкість містить інформацію про чутливість до змін обмежень даної задачі. У стовпці «Кінцеве значення» знаходяться значення лівих частин обмежень задачі (10.1), розрахованих для оптимального

розв'язку. Столпчик «Тіньова ціна» містить значення двоїстих змінних, що відповідають заданим обмеженням і характеризують внутрішню вартість одиниці ресурсів. Значення у стовпцях «Допустиме збільшення» і «Допустиме зменшення» показують допустимі межі зміни приросту ресурсів, при яких зберігається даний оптимальний план. Наприклад, якщо запас першого типу сировини (який дорівнює 10) збільшити більше, ніж на 2 одиниці, то оптимальний план (розв'язок) зміниться, а якщо запас другого типу сировини (який дорівнює 15) збільшити на як завгодно велике значення, то оптимальний план (розв'язок) залишиться без змін.

Розв'язування задачі за допомогою R language

R, як мова програмування і програмне середовище для статистичних обчислень, аналізу та зображення даних у графічному вигляді, є добрим інструментом для розв'язування оптимізаційних задач завдяки своїй гнучкості та великій кількості спеціалізованих пакетів. Пакети в R дають додаткові функції та інструменти, що розширюють базові можливості R і дозволяють ефективно розв'язувати специфічні задачі.

Одним з найпопулярніших пакетів для лінійного програмування є *lpSolve*. Він розроблений для розв'язування широкого спектру задач лінійного програмування, включаючи ті, що мають лінійні обмеження та цілочислові змінні.

Для того, щоб встановити пакет *lpSolve*, потрібно виконати команду `install.packages("lpSolve")`, а потім, щоб завантажити її функції в оперативну пам'ять – команду `library(lpSolve)`. Варто пам'ятати, що `install.packages ( )` треба

виконати один раз, а *library()* – кожного разу перед початком роботи з цією бібліотекою.

Для того, щоб розв'язати задачу (10.1), потрібно реалізувати такий алгоритм:

1) завантажити пакет lpSolve:

```
library(lpSolve)
```

2) ввести вхідні дані:

```
f = c(30,10), # цільова функція
A = matrix(c(2, 3, 3, 2,
             1,-1, 0, 1),
           nrow = 4,
           byrow = TRUE) # матриця обмежень
dir = c("<=", "<=", "<=", "<=") # знаки обмежень
b = c(10,15,1,2) # праві частини нерівностей
```

3) розв'язати задачу за допомогою пакету lpSolve:

```
primalsol = lp("max", f, A, dir, b).
```

4) вивести результати:

```
> print(primalsol)
Success: the objective function is 94
> print(primalsol$solution)
[1] 2.6 1.6
```

Отже, розв'язок задачі (10.1) такий: $x^* = (2,6; 1,6)$ і оптимальне значення цільової функції $f^* = 94$.

Знайти розв'язки прямої задачі (10.1) та двоїстої до неї задачі (10.2) можна також за допомогою пакету «Rglpk». Для цього необхідно виконати такі операції:

1) встановити та завантажити пакет:

```
install.packages("Rglpk")  
library(Rglpk)
```

2) ввести вхідні дані (повторити крок 2 наведеного вище алгоритму);

3) розв'язати пару взаємно двоїстих задач ЛП (10.1) та (10.2):

```
primalsol = Rglpk_solve_LP(f, A, dir, b, max = TRUE)
```

4) вивести результати:

```
> primalsol$optimum  
[1] 94  
> primalsol$solution  
[1] 2.6 1.6  
> primalsol$auxiliary$dual  
[1] 8 0 14 0
```

У результаті виконання цього алгоритму отримаємо уже відомий нам розв'язок прямої задачі (10.1) та розв'язок $\lambda^* = (8; 0; 14; 0)^T$ двоїстої задачі (10.2). При цьому значення цільових функцій обидвох задач збігаються. Зрозуміло, що розв'язок задачі (10.2) можна знайти і за допомогою пакету lpSolve, якщо реалізувати алгоритм:

1) завантажити пакет lpSolve:

```
library(lpSolve)
```

2) ввести вхідні дані:

```
f_dual = b  
A_dual = t(A)  
dir_dual = rep(">=", ncol(A))  
b_dual = f
```

3) розв'язати двоїсту задачу:

```
dualsol = lp("min", f_dual, A_dual, dir_dual, b_dual)
```

4) вивести результати:

```
> print(dualsol)
Success: the objective function is 94
> print(dualsol$solution)
[1] 8 0 14 0
```

Мова програмування R також має інструменти для розв'язування задач ЛП графічним методом. Для отримання графічного розв'язку задачі (10.1) встановлюємо пакет `install.packages("gMOIP")` та реалізовуємо процедуру:

```
library(gMOIP)
plotPolytope( A, b, f,
  type = rep("c", ncol(A)),
  crit = "max",
  faces = rep("c", ncol(A)),
  plotFaces = TRUE,
  plotFeasible = TRUE,
  plotOptimum = TRUE,
  labels = "coord"
)
```

Графічна ілюстрація знаходження розв'язку продемонстрована на рис. 10.8.

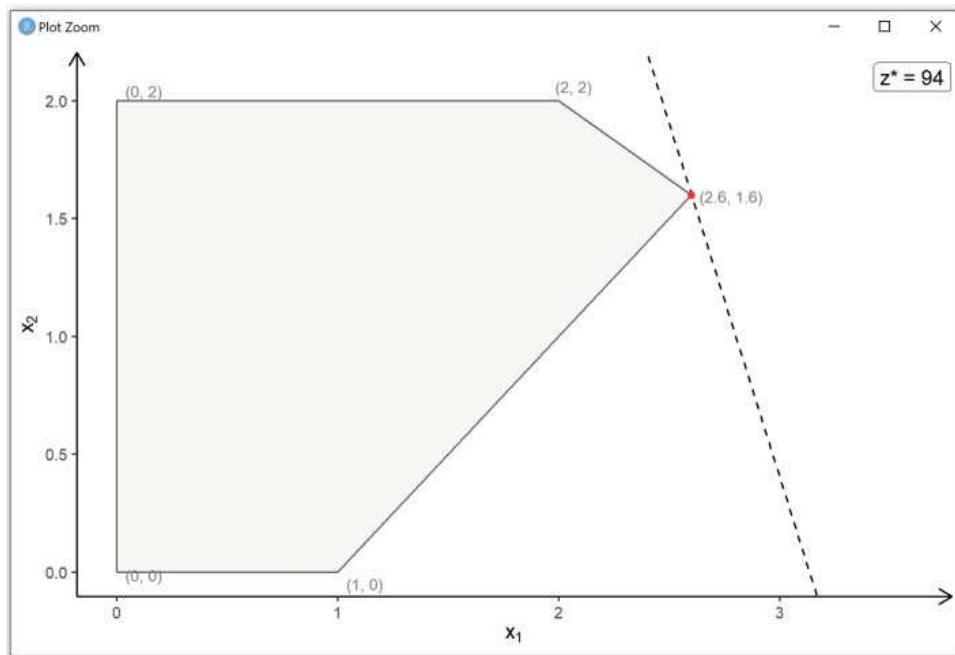


Рис. 10.8. Графічна ілюстрація знаходження розв'язку задачі (10.1)

Розв'язування задачі за допомогою Matlab

Пакет Matlab широко використовується також для побудови та розв'язування оптимізаційних економіко-математичних моделей.

Для розв'язування задач ЛП в MATLAB використовується функція *linprog*.

Алгоритм розв'язування задачі (10.1) у цьому випадку такий:

1) введення вхідних даних (коефіцієнтів цільової функції f зі знаком мінус, оскільки задача (10.1) є задачею на максимум; матриці обмежень A ; вектора обмежень b ; нижніх обмежень (умов невід'ємності на змінні):

$$f = [-30, -10];$$

$$A = \begin{bmatrix} 2 & 3 \\ 3 & 2 \\ 0 & 1 \end{bmatrix} \begin{array}{l} \% \text{ для витрат сировини A} \\ \% \text{ для витрат сировини B} \\ \% \text{ для попиту на продукцію П2} \end{array}$$


```

1, -1 % для співвідношення попиту на П1 та П2];
b = [10; 15; 2; 1];
lb = [0; 0];

```

2) застосування функції мінімізації:

```
[x, fval, exitflag, output] = linprog(f, A, b, [], [], lb);
```

3) виведення кінцевих результатів:

```

disp(['Оптимальна кількість Продукції 1: ',
num2str(x(1))]);
disp(['Оптимальна кількість Продукції 2: ',
num2str(x(2))]);
disp(['Максимальний дохід: ', num2str(-fval)]);

```

Аналогічні процедури слід здійснити для знаходження розв'язку двоїстої задачі:

1) введення вхідних даних (коефіцієнтів цільової функції $\tilde{f}$; матриці обмежень A^T ; вектора обмежень c ; нижніх обмежень (умов невід'ємності на двоїсті змінні):

```

f = [10; 15; 1; 2];
Aeq = [2, 3, 1, 0; 3, 2, -1, 1];
beq = [30; 10];
lb = [0; 0; 0; 0];
ub = [];

```

2) застосування функції мінімізації:

```
[l, fval, exitflag] = linprog(f, [], [], Aeq, beq, lb, ub);
```

3) виведення результатів:

```

disp('Оптимальні значення змінних:');
disp(l);
disp('Мінімальне значення цільової функції:');
disp(fval)

```

Отже, результати розв'язування прямої задачі (10.1) та двоїстої до неї задачі (10.2) такі:

оптимальна кількість Продукції 1: 2.6

оптимальна кількість Продукції 2: 1.6

максимальний дохід: 94

оптимальні значення змінних:

8.0000

0.0000

14.0000

0.0000

i

мінімальне значення цільової функції:

94.0000.

10.2. Транспортні задачі (Т-задачі)

10.2.1. Стандартна Т-задача

Приклад 10.2.1. (П.10.2.1). Постачальники P_1, P_2, P_3 перевозять однорідну продукцію чотирьом споживачам S_1, S_2, S_3, S_4 . Обсяги постачання та споживання, а також витрати на перевезення одиниці вантажу від постачальників до споживачів наведені у таблиці 10.3.

Таблиця 10.3

Споживачі Постачальники	S_1	S_2	S_3	S_4	Обсяги Постачання
P_1	3	1	4	2	30
P_2	2	4	3	6	20
P_3	7	5	8	4	15
Обсяги споживання	10	20	30	15	

Потрібно побудувати такий план перевезення товару, при якому сумарні транспортні витрати будуть мінімальні.

Насамперед відзначимо, що сумарний обсяг споживання (75 од.) перевищує сумарний обсяг постачання (65 од.). Це означає, що дана транспортна задача є відкритою і необхідно звести її до закритого (збалансованого) типу (табл. 10.4).

Таблиця 10.4

Споживачі Постачальники	S_1	S_2	S_3	S_4	Обсяги Постачання
P_1	3	1	4	2	30
P_2	2	4	3	6	20
P_3	7	5	8	4	15
P_4	0	0	0	0	10
Обсяги споживання	10	20	30	15	

Модель закритої транспортної задачі спочатку конкретизуємо у вигляді таблиці 10.5.

Таблиця 10.5

№ п/п	Елементи моделі	Економічний Зміст	Математична формалізація
1.	Шукані змінні	Кількість продукції, перевезеної від постачальників до споживачів	$x_{11}, x_{12}, x_{13}, x_{14},$ $x_{21}, x_{22}, x_{23}, x_{24},$ $x_{31}, x_{32}, x_{33}, x_{34},$ $x_{41}, x_{42}, x_{43}, x_{44}.$
2.	Цільова функція	Мінімальні витрати на перевезення продукції	$3x_{11} + x_{12} + 4x_{13} + 2x_{14} + 2x_{21} +$ $+4x_{22} + 3x_{23} + 6x_{24} + 7x_{31} +$ $+5x_{32} + 8x_{33} + 4x_{34} \rightarrow \min$
3.	Основні обмеження моделі	за запасами постачальників	$\begin{cases} x_{11} + x_{12} + x_{13} + x_{14} = 30, \\ x_{21} + x_{22} + x_{23} + x_{24} = 20, \\ x_{31} + x_{32} + x_{33} + x_{34} = 15, \\ x_{41} + x_{42} + x_{43} + x_{44} = 10. \end{cases}$
		за потребами споживачів	$\begin{cases} x_{11} + x_{21} + x_{31} + x_{41} = 10, \\ x_{12} + x_{22} + x_{32} + x_{42} = 20, \\ x_{13} + x_{23} + x_{33} + x_{43} = 30, \\ x_{14} + x_{24} + x_{34} + x_{44} = 15. \end{cases}$
4.	Прямі обмеження	невід'ємність обсягів перевезеної продукції	$x_{ij} \geq 0, i = \overline{1,4}, j = \overline{1,4}.$

За допомогою таблиці 10.5 уточнюємо математичну формалізацію моделі задачі (П.10.2.1):

$$\left\{ \begin{array}{l} F(X) = 3x_{11} + x_{12} + 4x_{13} + 2x_{14} + 2x_{21} + 4x_{22} + \\ + 3x_{23} + 6x_{24} + 7x_{31} + 5x_{32} + 8x_{33} + 4x_{34} \rightarrow \min, \\ x_{11} + x_{12} + x_{13} + x_{14} = 30, \\ x_{21} + x_{22} + x_{23} + x_{24} = 20, \\ x_{31} + x_{32} + x_{33} + x_{34} = 15, \\ x_{41} + x_{42} + x_{43} + x_{44} = 10, \\ x_{11} + x_{21} + x_{31} + x_{41} = 10, \\ x_{12} + x_{22} + x_{32} + x_{42} = 20, \\ x_{13} + x_{23} + x_{33} + x_{43} = 30, \\ x_{14} + x_{24} + x_{34} + x_{44} = 15, \\ x_{ij} \geq 0, i = \overline{1,4}, j = \overline{1,4}. \end{array} \right. \quad (10.3)$$

Далі продемонструємо процедури розв'язування задачі (10.3) у MS Excel, R та Matlab.

Розв'язування задачі у MS Excel

Введемо відповідні дані для задачі (10.3) у комірки MS Excel (рис. 10.9). При цьому:

- комірки C4:F7 порожні, зарезервовані для шуканих змінних задачі;
- комірки L4:O7 містять коефіцієнти цільової функції Т-задачі (тарифи перевезень одиниці продукції);
- у комірку L9 введено розрахункову формулу для цільової функції: =СУММПРОИЗВ(L4:O7;C4:F7);
- у комірках C9:F9 введені обсяги споживання, а у комірках H4:H7 – запаси споживачів;

- комірки G4:G7 містять формули для розрахунку сумарного вивезеного товару від постачальників P_1, P_2, P_3, P_4 . Наприклад, у комірці G4 введена формула =СУММ(C4:F4);

- комірки C8:F8 містять формули для розрахунку сумарного доставленого товару для споживачів S_1, S_2, S_3, S_4 . Наприклад, у комірці C8 введена формула =СУММ(C4:C7).

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	Перевезення									Тарифи					
2			Споживачі				вивезено	запаси				Споживачі			
3			S1	S2	S3	S4						S1	S2	S3	S4
4	постачальники	P1					0	30		постачальники	P1	3	1	4	2
5		P2					0	20			P2	2	4	3	6
6		P3					0	15			P3	7	5	8	4
7		P4					0	10			P4	0	0	0	0
8	Доставка		0	0	0	0									
9	потреби		10	20	30	15					затрати	0	min		

Рис. 10.9. Значення коефіцієнтів цільової функції та обсягів постачання і споживання

Під час заповнення вікна процедури «Пошук рішення» використано інформацію:

- цільова функція: L9;
- значення цільової функції: min;
- змінні комірки: C4:F7;
- обмеження: C8:F8 = C9:F9, G4:G7 = H4:H7.

У цьому ж вікні фіксуємо галочку для забезпечення виконання невід'ємності змінних задачі, вибираємо метод розв'язування задачі, а саме «Пошук розв'язку задач симплекс-методом» та натискаємо кнопку «Знайти розв'язок». У результаті з'явиться діалогове вікно «Результати пошуку розв'язку», у якому потрібно виділити звіти по результатах та стійкості для проведення післяоптимізаційного аналізу (рис. 10.10).

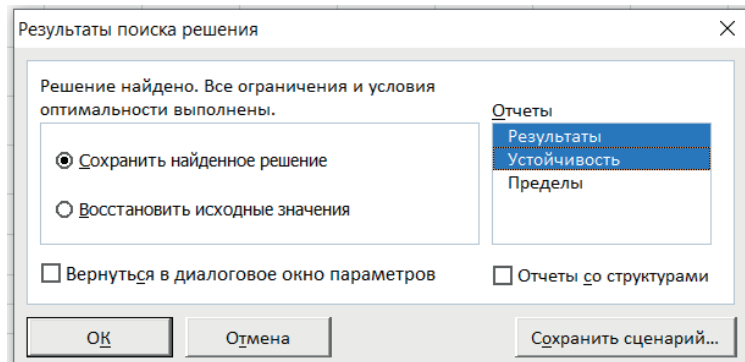


Рис. 10.10. Діалогове вікно «Результати пошуку розв’язку»

У підсумку (рис. 10.11) у комірках C4:F6 будуть отримані значення оптимальних обсягів перевезень, а у комірках C7:F7 записані фіктивні перевезення, якими під час аналізу нехтують. Мінімальна вартість перевезення відображена у комірці L9 і дорівнює 170 гр. од.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	Перевезення									Тарифи					
2			Споживачі				вивезено	запаси				Споживачі			
3			S1	S2	S3	S4						S1	S2	S3	S4
4	постачальники	P1	0	20	10	0	30	30		постачальники	P1	3	1	4	2
5		P2	10	0	10	0	20	20			P2	2	4	3	6
6		P3	0	0	0	15	15	15			P3	7	5	8	4
7		P4	0	0	10	0	10	10			P4	0	0	0	0
8	Доставка		10	20	30	15									
9	потреби		10	20	30	15				затрати	170	min			

Рис. 10.11. Результати розв’язування задачі (10.3)

Для проведення післяоптимізаційного аналізу спочатку отримаємо звіт за результатами розв’язування задачі (10.3) (рис. 10.12).

	A	B	C	D	E	F	G
14	Ячейка целевой функции (Минимум)						
15	Ячейка	Имя	Исходное значение	Окончательное значение			
16	\$L\$9	затраты S1	0	170			
17							
18							
19	Ячейки переменных						
20	Ячейка	Имя	Исходное значение	Окончательное значение	Целочисленное		
21	\$C\$4	P1 S1	0	0	Продолжить		
22	\$D\$4	P1 S2	0	20	Продолжить		
23	\$E\$4	P1 S3	0	10	Продолжить		
24	\$F\$4	P1 S4	0	0	Продолжить		
25	\$C\$5	P2 S1	0	10	Продолжить		
26	\$D\$5	P2 S2	0	0	Продолжить		
27	\$E\$5	P2 S3	0	10	Продолжить		
28	\$F\$5	P2 S4	0	0	Продолжить		
29	\$C\$6	P3 S1	0	0	Продолжить		
30	\$D\$6	P3 S2	0	0	Продолжить		
31	\$E\$6	P3 S3	0	0	Продолжить		
32	\$F\$6	P3 S4	0	15	Продолжить		
33	\$C\$7	P4 S1	0	0	Продолжить		
34	\$D\$7	P4 S2	0	0	Продолжить		
35	\$E\$7	P4 S3	0	10	Продолжить		
36	\$F\$7	P4 S4	0	0	Продолжить		
37							
38							
39	Ограничения						
40	Ячейка	Имя	Значение ячейки	Формула	Состояние	Допуск	
41	\$C\$8	Доставка S1	10	\$C\$8=\$C\$9	Привязка	0	
42	\$D\$8	Доставка S2	20	\$D\$8=\$D\$9	Привязка	0	
43	\$E\$8	Доставка S3	30	\$E\$8=\$E\$9	Привязка	0	
44	\$F\$8	Доставка S4	15	\$F\$8=\$F\$9	Привязка	0	
45	\$G\$4	P1 вивезено	30	\$G\$4=\$H\$4	Привязка	0	
46	\$G\$5	P2 вивезено	20	\$G\$5=\$H\$5	Привязка	0	
47	\$G\$6	P3 вивезено	15	\$G\$6=\$H\$6	Привязка	0	
48	\$G\$7	P4 вивезено	10	\$G\$7=\$H\$7	Привязка	0	
49							

Рис. 10.12. Звіт за результатами розв'язування задачі (10.3)

Цей звіт складається з трьох таблиць. У першій з них міститься інформація про цільову функцію, у другій – про значення змінних, отриманих після знаходження розв'язку задачі, а у третій – про значення обмежень, отриманих для оптимального розв'язку задачі.

Звіт про стійкість задачі (10.3) складається з двох таблиць (рис. 10.13). У першій таблиці у стовпці «Кінцеве значення» знаходиться розв'язок задачі (10.3). Стовпець «Зведена вартість» показує на скільки зміниться цільова функція, якщо обсяг перевезень, який у оптимальному розв'язку дорівнює нулю, збільшити на одиницю. Крім того, зведена (нормована) вартість, взята за абсолютною величиною, показує на яку мінімальну величину слід

змінити коефіцієнт у цільовій функції, щоб у оптимальному плані відповідний обсяг перевезень став відмінним від нуля. Наприклад, для комунікації (2;4) (тобто від другого постачальника до четвертого споживача) це значення дорівнює 5 (рис. 10.13) і означає, що витрати на перевезення одиниці вантажу з 2-го пункту відправлення до 4-го пункту призначення повинні бути скорочені на 5 гр. од., щоб цей напрямок став привабливим.

	A	B	C	D	E	F	G	H
7	Окончательное Приведенн. Целевая функция Допустимое Допустимое							
8	Ячейка	Имя	Значение	Стоимость	Коэффициент	Увеличение	Уменьшение	
9	\$C\$4	P1 S1	0	0	3	1E+30	0	
10	\$D\$4	P1 S2	20	0	1	2	1E+30	
11	\$E\$4	P1 S3	10	0	4	0	2	
12	\$F\$4	P1 S4	0	0	2	2	2	
13	\$C\$5	P2 S1	10	0	2	0	1E+30	
14	\$D\$5	P2 S2	0	4	4	1E+30	4	
15	\$E\$5	P2 S3	10	0	3	4	0	
16	\$F\$5	P2 S4	0	5	6	1E+30	5	
17	\$C\$6	P3 S1	0	2	7	1E+30	2	
18	\$D\$6	P3 S2	0	2	5	1E+30	2	
19	\$E\$6	P3 S3	0	2	8	1E+30	2	
20	\$F\$6	P3 S4	15	0	4	2	1E+30	
21	\$C\$7	P4 S1	0	1	0	1E+30	1	
22	\$D\$7	P4 S2	0	3	0	1E+30	3	
23	\$E\$7	P4 S3	10	0	0	1	1E+30	
24	\$F\$7	P4 S4	0	2	0	1E+30	2	
25								
26	Ограничения							
27								
28	Ячейка	Имя	Окончательное Значение	Тень Цена	Ограничение Правая сторона	Допустимое Увеличение	Допустимое Уменьшение	
29	\$C\$8	Доставка S1	10	5	10	0	10	
30	\$D\$8	Доставка S2	20	3	20	0	15	
31	\$E\$8	Доставка S3	30	6	30	0	10	
32	\$F\$8	Доставка S4	15	4	15	0	15	
33	\$G\$4	P1 вивезено	30	-2	30	15	0	
34	\$G\$5	P2 вивезено	20	-3	20	10	0	
35	\$G\$6	P3 вивезено	15	0	15	0	1E+30	
36	\$G\$7	P4 вивезено	10	-6	10	10	0	

Рис. 10.13. Звіт про стійкість задачі (10.3)

У стовпцях «Допустиме збільшення» і «Допустиме зменшення» показані допустимі межі зміни коефіцієнтів цільової функції, при яких зберігається даний оптимальний план. Наприклад, із рис. 10.13 випливає, що даний оптимальний план не зміниться, якщо вартість перевезення від 1-го постачальника до 2-го споживача збільшити на 2 гр. од. або зменшити на максимальне значення (тобто вважати нульовим).

Другий блок звіту про стійкість містить інформацію про чутливість до змін обмежень даної задачі. У стовпці «Кінцеве значення» знаходяться сумарні обсяги перевезень від постачальників до споживачів, розрахованих для оптимального розв'язку. Стовпчик «Тіньова ціна» містить значення потенціалів оптимального розв'язку. Значення у стовпцях «Допустиме збільшення» і «Допустиме зменшення» показують допустимі межі зміни запасів постачальників та потреб споживачів, при яких зберігається даний оптимальний план.

Варто зауважити, що:

- якщо перевезення вантажу характеризуються не витратами, а виручкою або прибутком, то Т-задача розв'язується аналогічно, але цільова функція досліджується на максимум.
- якщо деякі маршрути перевезень недопустимі, то витрати на перевезення слід вважати такими, що дорівнюють досить великому числу, яке значно перевищує витрати на інші перевезення.
- якщо обсяги запасів вантажів у пунктах відправлення та потреб у пунктах призначення виражаються цілими числами, то й отриманий розв'язок буде цілочисловим.

Розв'язування задачі за допомогою R language

Для того, щоб розв'язати дану задачу за допомогою R language, потрібно реалізувати наведений нижче алгоритм:

- 1) завантажити пакет lpSolve:

```
library(lpSolve)
```

- 2) ввести вхідні дані:

```
C = matrix(c(3, 1, 4, 2,
            2, 4, 3, 6,
            7, 5, 8, 4,
            0, 0, 0, 0),
          ncol = 4, byrow = TRUE)

a = c(30, 20, 15, 10) # обсяги постачання
b = c(10, 20, 30, 15) # обсяги споживання
```

3) знайти розв'язок Т-задачі за допомогою функції `lp.transport` з пакету `lpSolve`:

```
solution = lp.transport(cost.mat = C,
                        direction = "min",
                        row.signs = rep("=", 4),
                        col.signs = rep("=", 4),
                        row.rhs = a,
                        col.rhs = b)
```

4) записати відповідь у змінну `solution_matrix` та задати імена рядків і стовпців матриці:

```
solution_matrix = matrix(solution$solution, ncol = ncol(C))
rownames(solution_matrix) = paste0("P", 1:nrow(C))
colnames(solution_matrix) = paste0("S", 1:ncol(C))
```

5) вивести результати:

```
> print(solution_matrix)
  S1 S2 S3 S4
P1 10 20 0 0
P2 0 0 20 0
P3 0 0 0 15
P4 0 0 10 0
> print(paste("Мінімальні витрати:", solution$objval))
[1] "Мінімальні витрати: 170"
```

Отже, розв'язком задачі (10.3) є матриця

$$X^* = \begin{pmatrix} 10 & 20 & 0 & 0 \\ 0 & 0 & 20 & 0 \\ 0 & 0 & 0 & 15 \\ 0 & 0 & 10 & 0 \end{pmatrix}$$

та оптимальне значення цільової функції $f^* = 170$.

Розв'язування задачі за допомогою Matlab

Оскільки Т-задача є задачею лінійного програмування, то від початкової задачі можна перейти до класичної задачі лінійного програмування з одноіндексними змінними, що власне і спростить розрахунки. Тоді алгоритм розв'язування Т-задачі буде таким:

1) ввести вхідні дані (вектор коефіцієнтів (вартостей) цільової функції c ; матрицю обмежень рівностей A ; вектор обсягів постачання та споживання; нижніх обмежень (умов невід'ємності на змінні):

```
d = [3 1 4 2 2 4 3 6 7 5 8 4 0 0 0 0];
Aeq = [1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0; 0 0 0 0 1 1 1 1 0 0 0 0 0
0 0 0; 0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 0; 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1;
1 0 0 0 1 0 0 0 1 0 0 0 1 0 0 0; 0 1 0 0 0 1 0 0 0 1 0 0 0 1 0 0; 0 0 1
0 0 0 1 0 0 0 1 0 0 0 1 0; 0 0 0 1 0 0 0 1 0 0 0 1 0 0 0 1];
beq = [30; 20; 15; 10; 10; 20; 30; 15];
lb = zeros (16,1);
```

2) застосувати функцію:

```
[y, fval] = linprog (d, [], [], Aeq, beq, lb, [])
```

3) вивести результати:

```
y =
3.3064 20.0000 6.6936 0.0000
6.6936 0.0000 13.3064 0.0000
```

0.0000 0.0000 0.0000 15.0000
0.0000 0.0000 10.0000 0.0000

fval =170.0000.

На завершення цього підпункту зауважимо, що розв'язки (плани перевезення товару) Т-задачі (10.3) у MS Excel, R та Matlab різні, але мінімальні значення цільової функції однакові. Це свідчить про ситуацію, коли Т-задача має неєдиний розв'язок.

10.2.2. Цілочислова задача транспортного типу

Приклад 10.2.2. (П.10.2.2, задача про призначення).

Підприємство має дві вакантні посади, на які претендують три нових спеціалісти (С1, С2 і С3) та два працівники (С4 і С5), які працюють у інших відділах даного підприємства. При цьому керівництво підприємства не вважає за потрібне переводити на місця один одного або звільняти з роботи існуючих працівників.

Оцінка компетентності усіх претендентів за десятибальною шкалою для роботи на двох вакантних посадах і двох посадах працівників, які бажають отримати нову роботу, наведена у таблиці 10.6.

Таблиця 10.6

Спеціалісти		С1	С2	С3	С4	С5
Посади						
Вакантні посади	П1	5	5	8	8	9
	П2	6	8	6	7	7
Посади працівників	П3	6	6	7	8	–
	П4	7	4	6	–	6

Необхідно розподілити працівників за посадами у такий спосіб, щоб сумарна ефективність їх роботи була максимальною.

Перш за все уточнимо, що у таблицях 10.6, 10.7 через П1, П2, П3, П4 позначено відповідні посади.

Як і раніше, спочатку побудуємо математичну модель задачі (П.10.2.2). При цьому відзначимо, що кількість кандидатів перевищує кількість вакантних посад. Це означає, що дана задача як задача транспортного типу (підприємство – постачальник посад, спеціалісти – споживачі посад) є відкритою і її необхідно звести до закритого (збалансованого) типу. Для цього введемо фіктивну посаду, для якої у нових спеціалістів покладемо однакову, досить низьку компетентність, а в існуючих працівників покладемо нульову компетентність, яка дозволить не призначити їх на цю посаду, а, отже, звільнити з роботи (табл. 10.7).

Таблиця 10.7

Спеціалісти Посади		С1	С2	С3	С4	С5
Вакантні посади	П1	5	5	8	8	9
	П2	6	8	6	7	7
Посади співробітників	П3	6	6	7	8	–
	П4	7	4	6	–	6
Фіктивна посада	П5	2	2	2	0	0

Модель задачі спочатку представимо поелементно таблицею 10.8.

Таблиця 10.8

Елементи моделі	Економічний Зміст	Математична формалізація
Шукані змінні	Призначення j -го спеціаліста на i -ту посаду	$x_{ij} = \begin{cases} 1, & j \text{ призначено на } i\text{-у посаду,} \\ 0, & j \text{ не признач. на } i\text{-у посаду,} \end{cases}$ $i = \overline{1,5}, j = \overline{1,5}.$
Цільова функція	Максимальна ефективність роботи спеціалістів	$5x_{11} + 5x_{12} + 8x_{13} + 8x_{14} + 9x_{15} +$ $+6x_{21} + 8x_{22} + 6x_{23} + 7x_{24} + 7x_{25} +$ $+6x_{31} + 6x_{32} + 7x_{33} + 8x_{34} +$ $+7x_{41} + 4x_{42} + 6x_{43} + 6x_{45}$ $+2x_{51} + 2x_{52} + 2x_{53} \rightarrow \max$
Основні обмеження моделі	за кількістю посад	$\begin{cases} x_{11} + x_{12} + x_{13} + x_{14} + x_{15} = 1, \\ x_{21} + x_{22} + x_{23} + x_{24} + x_{25} = 1, \\ x_{31} + x_{32} + x_{33} + x_{34} + x_{35} = 1, \\ x_{41} + x_{42} + x_{43} + x_{44} + x_{45} = 1, \\ x_{51} + x_{52} + x_{53} + x_{54} + x_{55} = 1. \end{cases}$
Основні обмеження моделі	за кількістю спеціалістів	$\begin{cases} x_{11} + x_{21} + x_{31} + x_{41} + x_{51} = 1, \\ x_{12} + x_{22} + x_{32} + x_{42} + x_{52} = 1, \\ x_{13} + x_{23} + x_{33} + x_{43} + x_{53} = 1, \\ x_{14} + x_{24} + x_{34} + x_{44} + x_{54} = 1, \\ x_{15} + x_{25} + x_{35} + x_{45} + x_{55} = 1. \end{cases}$
Прямі обмеження	змінні набувають лише значень 0 або 1	$x_{ij} \geq 0, x_{ij} \leq 1, x_{ij} \in \mathbb{Z},$ $i = \overline{1,5}, j = \overline{1,5}.$

Об'єднавши відображені у таблиці 10.8 співвідношення, отримаємо таку математичну модель:

$$\left\{ \begin{array}{l} 5x_{11} + 5x_{12} + 8x_{13} + 8x_{14} + 9x_{15} + 6x_{21} + 8x_{22} + 6x_{23} \\ + 7x_{24} + 7x_{25} + 6x_{31} + 6x_{32} + 7x_{33} + 8x_{34} + \\ + 7x_{41} + 4x_{42} + 6x_{43} + 6x_{45} + 2x_{51} + 2x_{52} + 2x_{53} \rightarrow \max \\ \sum_{j=1}^5 x_{ij} = 1, \quad i = \overline{1,5}, \\ \sum_{i=1}^5 x_{ij} = 1, \quad j = \overline{1,5}, \\ x_{ij} \geq 0, x_{ij} \leq 1, x_{ij} \in \mathbb{Z}, \quad i = \overline{1,5}, j = \overline{1,5}. \end{array} \right. \quad (10.4)$$

Розв'язування задачі у MS Excel

Початкові дані для задачі (10.4) (задачі про призначення) введемо у комірки Excel (рис. 10.14), які мають таке призначення:

- комірки C4:G8 порожні, зарезервовані для шуканих змінних задачі, що можуть набувати лише значень 0 або 1;
- комірки M4:Q8 містять коефіцієнти цільової функції задачі (10.4);
- у комірку M10 введено розрахункову формулу для цільової функції: =СУММПРОИЗВ(M4:Q8;C4:G8);
- оскільки кожний спеціаліст може бути призначений лише на одну роботу, то у кожному рядку може бути лише одне значення, що дорівнює 1 (інші значення при цьому дорівнюють 0) і у комірках H4:H8 будуть знаходитися суми елементів рядків з діапазону C4:G8, а у комірках I4:I8 – одиниці;
- оскільки на кожную роботу потрібно призначити лише одного спеціаліста, то у кожному стовпці може бути лише одне значення, що дорівнює 1 (інші значення при цьому дорівнюють 0) і у комірках C9:G9 будуть знаходитися суми

елементів стовпців з діапазону C4:G8, а у комірках C10:G10 – одиниці.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	
1	Призначення на посади										Компетентність претендентів							
2			Спеціалісти					Кількість посад	Потреба		Спеціалісти							
C1			C2	C3	C4	C5	C1			C2	C3	C4	C5					
3											Посади		P1	5	5	8	8	9
4	Посади	P1						0	1	P2			6	8	6	7	7	
5		P2						0	1	P3			6	6	7	8	0	
6		P3						0	1	P4			7	4	6	0	6	
7		P4						0	1	P5			2	2	2	0	0	
8		P5						0	1									
9	Кількість спеціалістів		0	0	0	0	0											
10	Потреба		1	1	1	1	1				Ефективність	0	max					
11																		

Рис. 10.14. Початкові дані для задачі (10.4)

Для знаходження розв'язку оптимізаційної задачі (10.4) у меню **Дані** вибираємо процедуру **Пошук рішення**. У вікні, що з'явилося, додаємо обмеження (рис. 10.15), вводимо необхідні посилання, фіксуємо галочку для забезпечення виконання невід'ємності кількості спеціалістів та вибираємо метод розв'язування задачі у випадяючому списку.

Рис. 10.15. Параметри знаходження розв'язку задачі (10.4)

У результаті виконання операцій у комірках C4:F8 будуть отримані рекомендовані призначення на вакантні посади, а у комірках G4:G8 –призначення на фіктивну роботу, що під час аналізу інтерпретується як відмова від посади (рис. 10.16).

		Призначення на посади									Компетентність претендентів									
		Спеціалісти					Кількість посад	Потреба				Спеціалісти								
		C1	C2	C3	C4	C5						C1	C2	C3	C4	C5				
1	Посади	P1	0	0	0	0	1	1	1			Посади	P1	5	5	8	8	9		
2		P2	0	1	0	0	0	1	1				P2	6	8	6	7	7		
3		P3	0	0	0	1	0	1	1				P3	6	6	7	8	0		
4		P4	1	0	0	0	0	1	1				P4	7	4	6	0	6		
5		P5	0	0	1	0	0	1	1				P5	2	2	2	0	0		
6	Кількість спеціалістів		1	1	1	1	1													
7	Потреба		1	1	1	1	1					Ефективність		34	max					
8																				
9																				
10																				

Рис. 10.16. Результати розв'язування задачі (10.4)

Згідно з отриманим розв'язком, на перше вакантне місце слід взяти існуючого працівника C5, на друге вакантне місце – нового спеціаліста C2, на місце колишнього працівника C5 – нового спеціаліста C1, існуючого працівника C4 залишити на своєму місці, а спеціалісту C3 – відмовити у роботі. При цьому сумарний показник компетентності буде найвищим (він дорівнює 34), що відображено у комірці M10.

Розв'язування задачі за допомогою R language

Для того, щоб розв'язати задачу про призначення (10.4) за допомогою R, потрібно виконати такі дії:

- 1) підключити пакет lpSolve:

```
library(lpSolve)
```

- 2) ввести початкові дані задачі:

```
C <- matrix(c(5, 5, 8, 8, 9,
              6, 8, 6, 7, 7,
              6, 6, 7, 8, 0,
```

```
7, 4, 6, 0, 6,  
2, 2, 2, 0, 0),  
nrow = 5,  
byrow = TRUE)
```

3) розв'язати задачу за допомогою функції *lp.assign*:

```
solution = lp.assign(cost.mat = C, direction = "max")
```

4) виконати додаткові перетворення для отримання результату у вигляді матриці:

```
solution_matrix = matrix(solution$solution,  
                           ncol = nrow(C))  
rownames(solution_matrix) <- paste0("P", 1:nrow(C))  
colnames(solution_matrix) <- paste0("C", 1:ncol(C))
```

5) вивести результати:

```
> print(solution_matrix)
```

	C1	C2	C3	C4	C5
P1	0	0	0	0	1
P2	0	1	0	0	0
P3	0	0	0	1	0
P4	1	0	0	0	0
P5	0	0	1	0	0

```
> print(paste("Сумарний показник компетентності:",  
solution$objval))  
[1] "Сумарний показник компетентності: 34"
```

Після виконання цього алгоритму отримаємо розв'язок, який повністю збігається з тим, що був отриманим за допомогою Excel.

Розв'язування задачі за допомогою Matlab

Алгоритм розв'язування задачі у Matlab:

1) ввести початкові дані:

```
f = [5;6;6;7;2; 5;8;6;4;2; 8;6;7;6;2; 8;7;8;0;0; 9;7;0;6;0];  
intcon = 1:25; % вектор цілих змінних від 1 до 25  
lb = zeros(25,1); % нижня межа для змінних - нулі  
ub = ones(25,1); % верхня межа змінних - одиниці  
Aeq = [1 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0;  
       0 0 0 0 0 1 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0;  
       0 0 0 0 0 0 0 0 0 0 1 1 1 1 1 0 0 0 0 0 0 0 0;  
       0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 1 0 0 0 0;  
       0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 1;  
       1 0 0 0 0 1 0 0 0 0 1 0 0 0 0 1 0 0 0 0 1 0 0 0 0;  
       0 1 0 0 0 0 1 0 0 0 0 1 0 0 0 0 1 0 0 0 0 1 0 0 0;  
       0 0 1 0 0 0 0 1 0 0 0 0 1 0 0 0 0 1 0 0 0 0 1 0 0;  
       0 0 0 1 0 0 0 0 1 0 0 0 0 1 0 0 0 0 1 0 0 0 0 1 0;  
       0 0 0 0 1 0 0 0 0 1 0 0 0 0 1 0 0 0 0 1 0 0 0 0 1];  
beq = ones(10,1);
```

2) застосувати функцію:

```
x = intlinprog(-f, intcon, [], [], Aeq, beq, lb, ub);  
x = reshape(x, [5,5]);  
disp(x);
```

3) вивести результати:

```
0  0  0  0  1  
0  1  0  0  0  
0  0  0  1  0  
1  0  0  0  0  
0  0  1  0  0
```

Значення функції дорівнює 34.

10.3. Стандартна задача цілочислового лінійного програмування

Приклад 10.3. (П.10.3). Парк відпочинку обслуговує сім груп працівників, які мають різні вихідні дні. При цьому кожний працівник входить тільки в одну групу і має два вихідних дні, які у таблиці 10.9 позначено нулем. Також на основі статистичних даних про денну кількість відвідувачів парку відпочинку, встановлено мінімальну потребу у працівниках у кожний із днів тижня (таблиця 10.9).

Тижнева оплата праці одного працівника складає 2500 грн і не залежить від графіку роботи. Необхідно підібрати таку кількість працівників у кожній групі, щоб мінімізувати тижневі витрати на оплату праці.

Таблиця 10.9

День тижня Група	Нд	Пн	Вт	Ср	Чт	Пт	Сб
Група 1	0	0	1	1	1	1	1
Група 2	1	0	0	1	1	1	1
Група 3	1	1	0	0	1	1	1
Група 4	1	1	1	0	0	1	1
Група 5	1	1	1	1	0	0	1
Група 6	1	1	1	1	1	0	0
Група 7	0	1	1	1	1	1	0
Мін. потреба у працівниках	22	17	13	14	15	18	24

Спочатку побудуємо математичну модель задачі (П.10.3)
 Її поелементна формалізація відображена у таблиці 10.10.

Таблиця 10.10

Елементи моделі	Економічний Зміст	Математична формалізація
Шукані змінні	Кількість працівників у кожній групі (бригаді)	$x_1, x_2, x_3, x_4, x_5, x_6, x_7$
Цільова функція	Мінімальний тижневий фонд зарплати	$f = 2500 \cdot (x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7) \rightarrow \min$
Основні обмеження моделі	Кількість персоналу у кожний день тижня	$\begin{cases} x_2 + x_3 + x_4 + x_5 + x_6 \geq 22, \\ x_3 + x_4 + x_5 + x_6 + x_7 \geq 17, \\ x_1 + x_4 + x_5 + x_6 + x_7 \geq 13, \\ x_1 + x_2 + x_5 + x_6 + x_7 \geq 14, \\ x_1 + x_2 + x_3 + x_6 + x_7 \geq 15, \\ x_1 + x_2 + x_3 + x_4 + x_7 \geq 18, \\ x_1 + x_2 + x_3 + x_4 + x_5 \geq 24. \end{cases}$
Прямі обмеження	Змінні задачі є цілочисловими	$x_i \geq 0, x_i \in \mathbb{Z}, i = \overline{1,7}.$

Отже, об'єднавши усі співвідношення, зазначені у таблиці 10.10, матимемо таку математичну модель:

$$\left\{ \begin{array}{l} f = 2500 \cdot \sum_{i=1}^7 x_i \rightarrow \min, \\ x_2 + x_3 + x_4 + x_5 + x_6 \geq 22, \\ x_3 + x_4 + x_5 + x_6 + x_7 \geq 17, \\ x_4 + x_5 + x_6 + x_7 \geq 13, \\ x_1 + x_2 + x_5 + x_6 + x_7 \geq 14, \\ x_1 + x_2 + x_3 + x_6 + x_7 \geq 15, \\ x_1 + x_2 + x_3 + x_4 + x_7 \geq 18, \\ x_1 + x_2 + x_3 + x_4 + x_5 \geq 24, \\ x_i \geq 0, x_i \in \mathbb{Z}, i = \overline{1,7}. \end{array} \right. \quad (10.5)$$

Розв'язування задачі у MS Excel

Введемо дані для задачі (10.5) у комірки Excel (рис. 10.17). Зміст використаних комірок такий:

- комірки C2:C8 порожні, зарезервовані для шуканих змінних задачі;
- комірки D2:J8 містять дані про вихідні дні груп працівників;
- у комірку C10 введено розрахункову формулу для цільової функції: =СУММ(C2:C8);
- у комірках D10:J10 записано формули знаходження кількості зайнятого персоналу у кожний з днів тижня. Наприклад, у комірці D10 записана формула: =СУММПРОИЗВ(\$C2:\$C8;D2:D8).
- у комірках D12:J12 записано мінімальні кількості персоналу у кожний з днів тижня.

	A	B	C	D	E	F	G	H	I	J
1	Група	Вихідні дні	Працівники	Нд	Пн	Вт	Ср	Чт	Пт	Сб
2	1	Неділя, понеділок		0	0	1	1	1	1	1
3	2	Понеділок, вівторок		1	0	0	1	1	1	1
4	3	Вівторок, середа		1	1	0	0	1	1	1
5	4	Середа, четвер		1	1	1	0	0	1	1
6	5	Четвер, п'ятниця		1	1	1	1	0	0	1
7	6	П'ятниця, субота		1	1	1	1	1	0	0
8	7	Субота, неділя		0	1	1	1	1	1	0
9										
10			0	0	0	0	0	0	0	0
11										
12				22	17	13	14	15	18	24
13										
14	Тижнева плата працівнику:		3 000 грн.							
15	Загальна тижнева зарплата:		0 грн.							
16										

Рис. 10.17. Вихідні дані для задачі (10.5)

У меню **Дані** вибираємо процедуру **Пошук рішення**. У вікні, що з'явилось, додаємо обмеження (рис. 10.18), вводимо необхідні посилання, фіксуємо галочку для забезпечення виконання невід'ємності кількості спеціалістів та вибираємо метод розв'язування задачі у випадяючому списку.

Параметры поиска решения

Оптимизировать целевую функцию:

До: ☐ Максимум ☒ Минимум ☐ Значения:

Изменяя ячейки переменных:

В соответствии с ограничениями:

☒ Сделать переменные без ограничений неотрицательными

Выберите метод решения:

Метод решения
 Для гладких нелинейных задач используйте поиск решения нелинейных задач методом ОПГ, для линейных задач - поиск решения линейных задач симплекс-методом, а для негладких задач - эволюционный поиск решения.

Справка Найти решение Закрыть

Рис. 10.18. Параметри знаходження розв'язку задачі (10.5)

Після цього у комірках C2:C8 будуть отримані рекомендовані кількості працівників кожної групи, а у комірці C15 – мінімальний тижневий фонд заробітної плати, що складає 62500 грн (рис. 10.19).

	A	B	C	D	E	F	G	H	I	J
1	Група	Вихідні дні	Працівники	Нд	Пн	Вт	Ср	Чт	Пт	Сб
2	1	Неділя, понеділок	3	0	0	1	1	1	1	1
3	2	Понеділок, вівторок	4	1	0	0	1	1	1	1
4	3	Вівторок, середа	7	1	1	0	0	1	1	1
5	4	Середа, четвер	4	1	1	1	0	0	1	1
6	5	Четвер, п'ятниця	6	1	1	1	1	0	0	1
7	6	П'ятниця, субота	1	1	1	1	1	1	0	0
8	7	Субота, неділя	0	0	1	1	1	1	1	0
9										
10			25	22	18	14	14	15	18	24
11										
12				22	17	13	14	15	18	24
13										
14	Тижнева плата працівнику:		2 500 грн.							
15	Загальна тижнева зарплата:		62 500 грн.							

Рис. 10.19. Результати розв'язування задачі (10.5)

Розв'язування задачі за допомогою R language

Знаходження розв'язку задачі (10.5) за допомогою R language здійснюється за алгоритмом:

- 1) підключення потрібної бібліотеки:

```
library(lpSolve)
```

- 2) задання початкових даних задачі:

```
days = c("Нд", "Пн", "Вт", "Ср", "Чт", "Пт", "Сб")
groups = 1:7
demand = c(22, 17, 13, 14, 15, 18, 24)

schedule_matrix = matrix(c(0, 0, 1, 1, 1, 1, 1,
                           1, 0, 0, 1, 1, 1, 1,
                           1, 1, 0, 0, 1, 1, 1,
                           1, 1, 1, 0, 0, 1, 1,
                           1, 1, 1, 1, 0, 0, 1,
                           1, 1, 1, 1, 1, 0, 0,
                           0, 1, 1, 1, 1, 1, 0),
                          ncol = 7, byrow = TRUE)
```

3) розв'язування задачі:

```
result = lp(direction = "min",  
            objective.in = rep(2500, length(groups)),  
            const.mat = schedule_matrix,  
            const.dir = rep(">=", length(days)),  
            const.rhs = demand,  
            all.int = TRUE) # цілочислові змінні
```

4) виведення результатів:

```
recommended_workers <- result$solution  
cat("Рекомендована кількість працівників:",  
    sum(recommended_workers),  
    "\n")  
Рекомендована кількість працівників: 25  
for (i in 1:length(recommended_workers)) {  
  cat(i, "група:", recommended_workers[i], "\n")  
}
```

```
1 група: 0  
2 група: 3  
3 група: 5  
4 група: 6  
5 група: 5  
6 група: 5  
7 група: 1,
```

```
min_weekly_salary <- sum(recommended_workers * 2500)  
cat("Мінімальна сумарна зарплата за тиждень:",  
    min_weekly_salary,  
    "грн")
```

Мінімальна сумарна зарплата за тиждень: 62500 грн.

Розв'язування задачі за допомогою Matlab

Реалізувати алгоритм знаходження розв'язку задачі (10.5) у Matlab потрібно так:

1) ввести початкові дані:

```
f = 2500*ones(7,1);  
intcon = 1:7;  
lb = zeros(7,1);  
A = [0 1 1 1 1 1 0;  
      0 0 1 1 1 1 1;  
      1 0 0 1 1 1 1;  
      1 1 0 0 1 1 1;  
      1 1 1 0 0 1 1;  
      1 1 1 1 0 0 1;  
      1 1 1 1 1 0 0];  
b = [22;17;13;14;15;18;24];
```

2) застосувати функцію:

```
x = intlinprog(f, intcon, -A, -b, [], [], lb);  
disp(x);
```

3) вивести результати обчислень:

Група 1	2
Група 2	5
Група 3	7
Група 4	4
Група 5	6
Група 6	1
Група 7	0

Мінімальна сумарна зарплата за тиждень: 62500 грн.

Зауважимо також, що оптимальні кількості працівників з кожної групи тут інші, що свідчить про те, що задача має не один розв'язок.

10.4. Деякі задачі нелінійного програмування

10.4.1. Задача дробово-лінійного програмування

Приклад 10.4.1. (П.10.4.1). Підприємство використовує три типи технологічного обладнання, необхідного для виробництва двох видів продукції P_1 та P_2 . Витрати часу на обробку одиниці продукції кожного виду та витрати, що пов'язані з виготовленням одиниці продукції кожного виду вказані у таблиці 10.11.

Таблиця 10.11

Тип обладнання	Витрати часу на виробництво продукції:	
	P_1	P_2
Обладнання А	2	8
Обладнання В	1	1
Обладнання С	12	3
Витрати, гр. од	2	3

Обладнання А і С підприємство може використовувати не більше ніж 26 і 39 годин відповідно, а обладнання В доцільно використовувати не менше 4 годин.

Необхідно визначити кількість продукції обох видів, що забезпечать мінімальну собівартість продукції.

Для побудови математичної моделі задачі (П.10.4.1) заповнимо таблицю 10.12.

Таблиця 10.12

№ п/п	Елементи моделі	Економічний зміст	Математична формалізація
1.	Шукані змінні	Кількість продукції Π_1 та Π_2	x_1, x_2
2.	Цільова функція	Мінімальна собівартість виробництва продукції	$\frac{2x_1 + 3x_2}{x_1 + x_2} \rightarrow \min$
3.	Основні обмеження моделі	Максимальний час роботи обладнання А	$2x_1 + 8x_2 \leq 26$
		Мінімальний час роботи обладнання В	$x_1 + x_2 \geq 4$
		Максимальний час роботи обладнання С	$12x_1 + 3x_2 \leq 39$
4.	Прямі обмеження	Невід’ємна кількість продукції Π_1 та Π_2	$x_1 \geq 0, x_2 \geq 0.$

Враховуючи таблицю 10.12, уточнимо математичну модель задачі (П.10.4.1):

$$\left\{ \begin{array}{l} \frac{2x_1 + 3x_2}{x_1 + x_2} \rightarrow \min, \\ 2x_1 + 8x_2 \leq 26, \\ x_1 + x_2 \geq 4, \\ 12x_1 + 3x_2 \leq 39, \\ x_1 \geq 0, x_2 \geq 0. \end{array} \right. \quad (10.6)$$

Розв’язування задачі у MS Excel

Введемо відповідні дані для задачі (10.6) у таблицю Excel (рис. 10.20).

	A	B	C	D	E	F
1	Елементи моделі	Продукція				
2		П1	П2			
3	План випуску продукції			По плану	Умова	Граничні значення
4	Собівартість, гр. од.			#ДЕЛ/0!	min	
5	Обладнання А	2	8	0	<=	26
6	Обладнання В	1	1	0	>=	4
7	Обладнання С	12	3	0	<=	39

Рис. 10.20. Вихідні дані для задачі (10.6)

На рис. 10.20 комірки:

- В3 і С3 відведені для змінних x_1 та x_2 ;
- В5:С7 містять коефіцієнти основних обмежень;
- F5:F7 містять граничні значення обмежень задачі;
- E4:E7 конкретизують мету оптимізації та знаки обмежень;

• D4:D6 містять розрахункову формулу для цільової функції та обмежень задачі (рис. 10.21). Оскільки цільова функція задачі дробово-лінійна, то для уникнення ділення на нуль як початкові значення для x_1, x_2 візьмемо довільні ненульові значення, наприклад, значення, що дорівнюють одиниці.

	A	B	C	D
1	Елементи моделі	Продукція		
2		П1	П2	
3	План випуску продукції	1	1	По плану
4	Собівартість, гр. од.			=(2*В3+3*С3)/(В3+С3)
5	Обладнання А	2	8	=СУММПРОИЗВ(\$В\$3:\$С\$3;В5:С5)
6	Обладнання В	1	1	=СУММПРОИЗВ(\$В\$3:\$С\$3;В6:С6)
7	Обладнання С	12	3	=СУММПРОИЗВ(\$В\$3:\$С\$3;В7:С7)

Рис. 10.21. Розрахункові формули для цільової функції та обмежень задачі (10.6)

Для знаходження розв'язку оптимізаційної задачі (10.6) у меню **Дані** вибираємо процедуру **Пошук рішення**. У вікні, що з'явилося, додаємо обмеження, вводимо необхідні посилання (рис. 10.22), фіксуємо галочку для забезпечення виконання невід'ємності кількості продукції P_1 і P_2 та вибираємо метод розв'язування задачі у випадваючому списку.

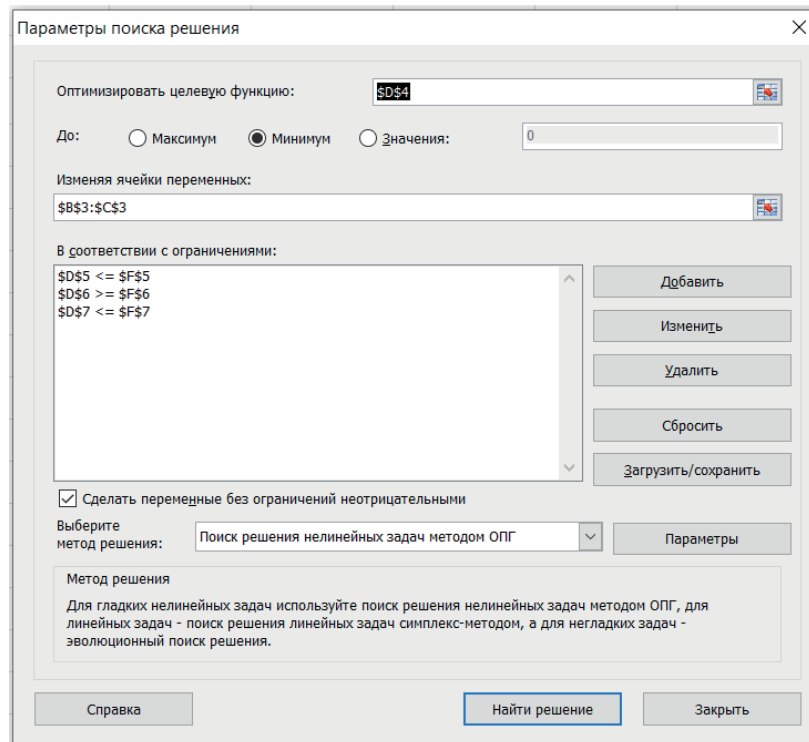


Рис. 10.22. Параметри знаходження розв'язку задачі (10.6)

Для конкретизації розв'язку задачі (10.6) необхідно натиснути кнопку «Знайти розв'язок». При цьому буде отримано повідомлення «Розв'язок знайдено» і після натискання кнопки ОК буде отримано оптимальний розв'язок.

У даному випадку оптимальний розв'язок (оптимальний план виробництва) $x^* = (3; 1)^T$ знаходиться у комірках B3 та C3 і забезпечує мінімальну собівартість виробництва продукції у розмірі 2,25 гр. од., що знаходиться у комірці D4.

Розв'язування задачі за допомогою R language

Алгоритм розв'язування оптимізаційної задачі (10.6) такий:

1) завантаження та підключення потрібної бібліотеки:

```
install.packages("nloptr")  
library(nloptr)
```

2) задання початкових даних:

```
obj_fun <- function(x) {  
  return((2*x[1] + 3*x[2]) / (x[1] + x[2]))  
}  
  
constraint <- function(x) {  
  return(c(  
    2 * x[1] + 8 * x[2] - 26, # 2x1 + 8x2 <= 26  
    4 - x[1] - x[2],         # x1 + x2 >= 4 -> -x1 - x2 <= -4  
    12 * x[1] + 3 * x[2] - 39, # 12x1 + 3x2 <= 39  
    -x[1],                   # x1 >= 0 -> -x1 <= 0  
    -x[2]                    # x2 >= 0 -> -x2 <= 0  
  ))  
}
```

3) задання початкового наближення:

```
x0 <- c(1, 1)
```

4) налаштування параметрів оптимізації (тут використовується алгоритм NLOPT_LN_COBYLA – Constrained Optimization BY Linear Approximations, хоча можна експериментувати також з іншими алгоритмами даного пакету):

```
opts <- list("algorithm" = "NLOPT_LN_COBYLA",  
            "xtol_rel" = 1.0e-8)
```

5) знаходження розв'язку задачі:


```
result <- nloptr( x0 = x0,
                 eval_f = obj_fun,
                 eval_g_ineq = constraint,
                 opts = opts
                )
```

6) виведення результатів:

```
> cat("Оптимальний розв'язок:\n")
Оптимальний розв'язок:
> cat("x1 =", result$solution[1], "\n")
x1 = 3
> cat("x2 =", result$solution[2], "\n")
x2 = 1
> cat("Мінімальне значення цільової функції:",
result$objective, "\n")
Мінімальне значення цільової функції: 2.25.
```

Розв'язування задачі за допомогою Matlab

Розв'язування задачі у Matlab реалізується за таким алгоритмом:

1) введення початкових даних:

```
fun = @(x) (2*x(1)+3*x(2))/(x(1)+x(2));
x0 = [0.1; 0.1];
A = [2 8; -1 -1; 12 3];
b = [26;-4;39];
lb = [0; 0];
```

2) звернення до процедури-функції:

```
x = fmincon(fun,x0,A,b);
f = fun(x);
disp('x = ');
disp(x);
```

disp('f = ');

disp(f);

3) виведення результатів:

x = 3.0000

1.0000

f = 2.2500.

Підкреслимо, що у всіх трьох випадках розв'язок задачі однаковий.

10.4.2. Стандартна задача квадратичного програмування

Приклад 10.4.2. (П.10.4.2). Підприємство використовує два види сировини (A та B) і робочу силу, необхідні для виробництва двох видів продукції Π_1 та Π_2 . Витрати ресурсів на виготовлення однієї тонни кожного виду продукції, прибуток, що отримується підприємством від їх реалізації, а також запаси ресурсів вказані у таблиці 10.13.

Таблиця 10.13

Сировина	Витрати сировини на виробництво тонни		Запас сировини
	Π_1	Π_2	
Сировина A	3	5	120
Сировина B	4	6	150
Витрати праці	14	12	400
Прибуток від реалізації тони продукції, гр. од	72	103	

Вартість однієї тони кожного виду сировини визначається такими залежностями: $(9 + 0,0088r_1)$ тис. гр. од. для сировини A та $(5 - 0,0086r_2)$ тис. гр. од. для сировини B, де r_1 і r_2 – витрати сировини на виробництво продукції. Вартість однієї години

витрат праці визначається залежністю $(1 - 0,0002r_3)$, де r_3 – витрати часу на виробництво продукції.

Необхідно визначити кількість продукції обох видів, що забезпечить максимальний прибуток.

Розв'язування задачі у MS Excel

Як і у випадку попередніх прикладів алгоритм розв'язування сформульованої у цьому прикладі задачі складається із трьох етапів (формалізації математичної моделі задачі (П10.4.2), побудови її комп'ютерного варіанту та знаходження розв'язку).

Формалізація елементів моделі та їх зміст відображені у таблиці 10.14.

Таблиця 10.14

№ п/п	Елементи моделі	Економічний зміст	Математична формалізація
1.	Шукані змінні	Кількість продукції Π_1 та Π_2	x_1, x_2
2.	Цільова функція	Максимальний прибуток від реалізації продукції	$f = 11x_1 + 16x_2 + 0,0976x_1^2 + 0,1184x_2^2 + 0,216x_1x_2 \rightarrow \max$
3.	Основні обмеження моделі	Максимальні витрати сировини А – 120 од.	$3x_1 + 5x_2 \leq 120$
		Максимальні витрати сировини В – 150 од.	$4x_1 + 6x_2 \leq 150$
		Максимальні витрати праці – 400 год.	$14x_1 + 12x_2 \leq 400$

Продовження таблиці 10.14

4.	Прямі обмеження	Невід'ємна кількість продукції Π_1 та Π_2	$x_1 \geq 0, x_2 \geq 0.$
----	-----------------	---	---------------------------

Стосовно функції прибутку, який є різницею між доходом і витратами, зазначимо, що вона формується так:

$$\begin{aligned}
 f &= 72x_1 + 103x_2 - (3x_1 + 5x_2) \left[9 + 0,0088(3x_1 + 5x_2) \right] - \\
 &\quad - (4x_1 + 6x_2) \left[5 - 0,0086(4x_1 + 6x_2) \right] - \\
 &\quad - (14x_1 + 12x_2) \left[1 - 0,0002(14x_1 + 12x_2) \right] = \\
 &= 11x_1 + 16x_2 + 0,0976x_1^2 + 0,1184x_2^2 + 0,216x_1x_2.
 \end{aligned}$$

Отже, математична формалізація моделі така:

$$\begin{cases}
 11x_1 + 16x_2 + 0,0976x_1^2 + \\
 + 0,1184x_2^2 + 0,216x_1x_2 \rightarrow \max, \\
 3x_1 + 5x_2 \leq 120, \\
 4x_1 + 6x_2 \leq 150, \\
 14x_1 + 12x_2 \leq 400, \\
 x_1 \geq 0, x_2 \geq 0.
 \end{cases} \quad (10.7)$$

Тепер введемо вихідні дані для задачі (10.7) у таблицю Excel (рис. 10.23).

	A	B	C	D	E	F
1	Елементи моделі	Продукція				
2		П1	П2			
3	План випуску продукції			По плану	Умова	Граничні значення
4	Прибуток, гр. од.			0,00	мах	
5	Сировина А	3	5	0	<=	120
6	Сировина В	4	6	0	<=	150
7	Витрати праці	14	12	0	<=	400

Рис. 10.23. Вихідні дані для задачі (10.7)

Відповідні комірки (рис. 10.23) мають такий зміст:

- B3 і C3 відведені для змінних x_1 та x_2 ;
- B5:C7 містять коефіцієнти основних обмежень;
- F5:F7 містять граничні значення ресурсів;
- E4:E7 відображають мету оптимізації та знаки обмежень, що у розрахунках не використовуються, проте дозволяють краще орієнтуватися у моделі;
- D4:D7 містять розрахункову формулу для цільової функції та обмежень задачі (рис. 10.24). Оскільки вони містять посилання на порожні комірки B3:C3, то їх результатом є число нуль.

	A	B	C	D
1	Елементи моделі	Продукція		
2		П1	П2	
3	План випуску продукції			По плану
4	Прибуток, гр. од.			=11*B3+16*C3+0,1*B3^2+0,12*C3^2+0,22*B3*C3
5	Сировина А	3	5	=СУММПРОИЗВ(\$B\$3:\$C\$3;B5:C5)
6	Сировина В	4	6	=СУММПРОИЗВ(\$B\$3:\$C\$3;B6:C6)
7	Витрати праці	14	12	=СУММПРОИЗВ(\$B\$3:\$C\$3;B7:C7)

Рис. 10.24. Розрахункові формули для задачі (10.7)

Для знаходження розв'язку задачі (10.7) у меню **Дані** вибираємо процедуру **Пошук рішення**. У вікні, що з'явилося, додаємо обмеження, вводимо необхідні посилання (рис. 10.25), фіксуємо галочку для забезпечення виконання невід'ємності кількості продукції P_1 і P_2 та вибираємо метод розв'язування задачі у випадяючому списку.

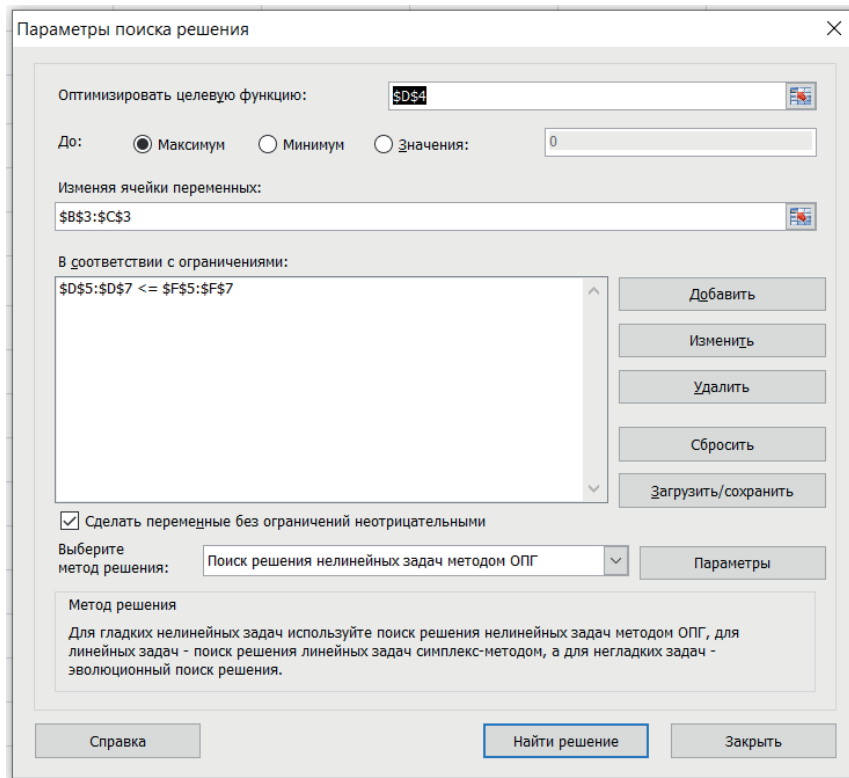


Рис. 10.25. Параметри знаходження розв'язку задачі (10.7)

Потім необхідно натиснути кнопку «Знайти розв'язок». Після отримання повідомлення «Розв'язок знайдено» натискаємо кнопку ОК для зберігання шуканого оптимального розв'язку.

У даному випадку цей розв'язок (оптимальний план виробництва) $x^* = (16,67; 13,89)^T$ знаходиться у комірках B3 та C3 і забезпечує максимальний прибуток у розмірі 507,41 гр. од., що знаходиться у комірці D4.

На основі даних з робочого листа можна також зробити висновок, що всі ресурси є дефіцитними, оскільки їх запаси та витрати збігаються.

Розв'язування задачі за допомогою R language

Для того, щоб знайти розв'язок задачі (10.7) за допомогою R, необхідно виконати такі дії:

1) встановити та підключити потрібну бібліотеку:

```
install.packages("nloptr")  
library(nloptr)
```

2) задати початкові дані задачі та відповідне початкове наближення:

```
obj_fun <- function(x) {  
  return(-(11 * x[1] + 16 * x[2] + 0.0976 * x[1]^2 + 0.1184 *  
x[2]^2 + 0.216 * x[1] * x[2]))  
}  
  
constraints <- function(x) {  
  return(c(  
    3 * x[1] + 5 * x[2] - 120,    # 3x1 + 5x2 <= 120  
    4 * x[1] + 6 * x[2] - 150,    # 4x1 + 6x2 <= 150  
    14 * x[1] + 12 * x[2] - 400, # 14x1 + 12x2 <= 400  
    -x[1],                      # x1 >= 0 -> -x1 <= 0  
    -x[2]                        # x2 >= 0 -> -x2 <= 0  
  ))  
}  
  
x0 <- c(1, 1)
```

3) налаштувати параметри оптимізації:

```
opts <- list("algorithm" = "NLOPT_LN_COBYLA",  
  "xtol_rel" = 1.0e-8)
```

4) знайти розв'язок задачі:

```
result <- nloptr( x0 = x0,  
  eval_f = obj_fun,  
  eval_g_ineq = constraint,  
  opts = opts  
)
```

5) вивести шукані результати:

```
> cat("Оптимальний розв'язок:\n")
Оптимальний розв'язок:
> cat("x1 =", result$solution[1], "\n")
x1 = 16.66667
> cat("x2 =", result$solution[2], "\n")
x2 = 13.88889
> cat("Максимальне значення цільової функції:", -
result$objective, "\n")
Максимальне значення цільової функції: 505.5062.
```

Розв'язування задачі за допомогою Matlab

Для знаходження розв'язку задачі (10.7) послідовне виконання дій у Matlab таке:

1) задання початкових даних:

```
f = [-11; -16];
```

```
A = [3 5; 4 6; 14 12];
```

```
b = [120 150 400];
```

```
lb = [0; 0];
```

```
ub = [];
```

2) застосування до функції пошуку розв'язку:

```
[x, fval, exitflag, output] = linprog(f, A, b, [], [], lb, ub);
```

3) виведення результатів:

```
disp('Оптимальне значення цільової функції:');
```

```
disp(-fval);
```

```
disp('Оптимальні значення змінних:');
```

```
disp(x(1:2));
```

```
disp('Вихідний код:');
```

```
disp(exitflag);
```

```
disp('Інформація виведення:')
```

```
disp(output)
```


$$\text{function_value} = 11 \cdot x(1) + 16 \cdot x(2) + 0.0976 \cdot x(1)^2 + 0.1184 \cdot x(2)^2 + 0.216 \cdot x(1) \cdot x(2)$$

16.6667

13.8889

Максимальне значення цільової функції: 505.5062.

Отже, з урахуванням відповідних наближень можна стверджувати, що у всіх трьох випадках розв'язки збігаються.

10.5. Задача динамічного програмування

У попередньому розділі 9 йшлося про те, що динамічне програмування (ДП) – це метод розв'язування певного класу задач, характерні особливості яких власне кажучи там уже були зазначені. Однак у літературі нерідко зустрічається термінологія «задача ДП» як конкретизація того, що відповідна задача може бути розв'язана методом ДП. Саме у такому сенсі слід розуміти назву цього пункту. Для ілюстрації вищесказаного розглянемо приклад однієї із задач цього класу, яка є частинним випадком уже проаналізованої у контексті ДП задачі, описаної у прикладі 9.1 і яка може бути розв'язана не тільки методом ДП, а й іншими методами.

Приклад 10.5. (П.10.5). *Для збільшення потужності виробництва продукції, яка користується підвищеним попитом, трьом підприємствам виділяються капітальні вклади (КВ) загальним обсягом 700 тис. грн. Використання j – м підприємством $u^{(i)}$ од. КВ дає йому прибуток, вказаний у таблиці 10.15.*

Таблиця 10.15

Обсяг КВ ($u^{(i)}$ тис. грн)	Прибуток підприємства залежно від обсягу КВ (тис. грн. од.)		
	$k^{(1)}$	$k^{(2)}$	$k^{(3)}$
0	0	0	0
100	30	50	40
200	50	80	50
300	90	90	110
400	110	150	120
500	170	190	180
600	180	210	220
700	210	220	240

Необхідно знайти такий розподіл КВ між підприємствами, при якому їх сумарний прибуток буде максимальним.

Розв'язування задачі у MS Excel

Сформуємо таблицю 10.16, у якій поставлена задача формалізована окремими блоками (елементами).

Таблиця 10.16

Елементи моделі	Економічний зміст	Математична формалізація
Шукані змінні	Виділення $u^{(i)}$ тис. грн. для j -го підприємства	$x_{ij} = \begin{cases} 1, & \text{виділ. } u^{(i)} j\text{-му підпр.}, \\ 0, & \text{не вид. } u^{(i)} j\text{-му підпр.} \end{cases}$ $i = \overline{1,8}, j = \overline{1,3}.$
Цільова функція	Максимальний прибуток від КВ	$30x_{21} + 50x_{22} + 40x_{23} + 50x_{31} + 80x_{32} + 50x_{33} + 90x_{41} + 90x_{42} + 110x_{43} + 110x_{51} + 150x_{52} + 120x_{53} + 170x_{61} + 190x_{62} + 180x_{63} + 180x_{71} + 210x_{72} + 220x_{73} + 210x_{81} + 220x_{82} + 240x_{83} \rightarrow \max$
Основні обмеження моделі	за кількістю КВ одному підприємству	$\sum_{i=1}^8 x_{i1} \leq 1, \sum_{i=1}^8 x_{i2} \leq 1, \sum_{i=1}^8 x_{i3} \leq 1.$
	за загальною сумою КВ	$\sum_{i=1}^8 (x_{i1} + x_{i2} + x_{i3}) u^{(i)} = 700.$
Прямі обмеження	невід'ємність і цілочисловість виділених КВ підприємствам	$x_{ij} \geq 0, x_{ij} \leq 1, x_{ij} \in \mathbb{Z}, i = \overline{0,7}, j = \overline{1,3}.$

Уточнимо, що x_{ij} – це змінна, яка фіксує факт виділення чи невиділення КВ у розмірі $u^{(i)}$ тис. грн. j –ому підприємству.

Із таблиці 10.16 випливає, що математична модель задачі (П.10.5) має вигляд:

$$\begin{cases}
30x_{21} + 50x_{22} + 40x_{23} + 50x_{31} + 80x_{32} + \\
+ 50x_{33} + 90x_{41} + 90x_{42} + 110x_{43} + 110x_{51} + \\
+ 150x_{52} + 120x_{53} + 170x_{61} + 190x_{62} + \\
+ 180x_{63} + 180x_{71} + 210x_{72} + 220x_{73} + \\
+ 210x_{81} + 220x_{82} + 240x_{83} \rightarrow \max, \\
\sum_{i=1}^8 x_{i1} \leq 1, \sum_{i=1}^8 x_{i2} \leq 1, \sum_{i=1}^8 x_{i3} \leq 1, \\
\sum_{i=1}^8 (x_{i1} + x_{i2} + x_{i3}) u^{(i)} = 700, \\
x_{ij} \geq 0, x_{ij} \leq 1, x_{ij} \in \mathbb{Z}, i = \overline{0,7}, j = \overline{1,3}.
\end{cases} \quad (10.8)$$

Тепер побудуємо комп'ютерної моделі задачі (10.8).

Комірки MS Excel (рис. 10.26), у які вводимо вхідні дані, мають такий зміст:

- комірки B3:D10 порожні, зарезервовані для шуканих змінних задачі;
- комірки I3:K10 містять коефіцієнти цільової функції задачі (прибуток підприємства залежно від обсягу KB);
- у комірку I12 введено розрахункову формулу для цільової функції: =СУММПРОИЗВ(I3:K10;B3:D10);
- комірки B11:D11 містять формули для розрахунку кількості KB, інвестованих у кожне підприємство (наприклад, у комірку B11 введена формула =СУММ(B3:B10));
- комірки E3:E10 містять формули для розрахунку суми KB (наприклад, у комірку E3 введена формула =СУММ(B3:D3)*A3), а у комірку F3 введено загальну суму KB, що визначається формулою=СУММ(E3:E10);

- у комірках B12:D12 введені максимальні кількості КВ у кожне підприємство, а у комірку F12 – максимальну суму КВ, задану умовою задачі.

	A	B	C	D	E	F	G	H	I	J	K
1	Обсяг капіталовкладень (у тис. гр. од.)	Прибуток підприємства залежно від обсягу КВ (тис. гр. од.)			Сума вкладень	Загальна сума вкладень		Обсяг капіталовкладень (у тис. гр. од.)	Прибуток підприємства залежно від обсягу КВ (тис. гр. од.)		
2		$k^{(1)}$	$k^{(2)}$	$k^{(3)}$					$k^{(1)}$	$k^{(2)}$	$k^{(3)}$
3	0				0	0		0	0	0	0
4	100				0			100	30	50	40
5	200				0			200	50	80	50
6	300				0			300	90	90	110
7	400				0			400	110	150	120
8	500				0			500	170	190	180
9	600				0			600	180	210	220
10	700				0			700	210	220	240
11	Кількість вкладень	0	0	0		Максимальна сума вкладень					
12	Максимальна кількість вкладень	1	1	1		700		Прибуток	0	max	

Рис. 10.26. Вихідні дані для задачі (10.8)

З метою знаходження розв'язку задачі (10.8) використаємо **Пошук рішення**. У його робочому вікні вводимо посилання на цільову функцію та шукані змінні, додаємо обмеження, фіксуємо галочку для забезпечення невід'ємності виділених КВ підприємствам та вибираємо метод розв'язування задачі у випадіючому списку (рис. 10.27).

Рис. 10.27. Параметри пошуку розв'язку задачі (10.8)

Після реалізованих операцій у комірках B3:D10 будуть отримані значення оптимального розподілу КВ, які дозволяють зробити висновок, що 100 тис. грн варто виділити другому підприємству, а 600 тис. грн – третьому підприємству. Максимальний прибуток відображений у комірці I12 і дорівнює 270 тис. грн. (рис. 10.28).

	A	B	C	D	E	F	G	H	I	J	K
1	Обсяг капіталовкладень (у тис. гр. од.)	Прибуток підприємства залежно від обсягу КВ (тис. гр. од.)			Сума вкладень	Загальна сума вкладень		Обсяг капіталовкладень (у тис. гр. од.)	Прибуток підприємства залежно від обсягу КВ (тис. гр. од.)		
2		$k^{(1)}$	$k^{(2)}$	$k^{(3)}$					$k^{(1)}$	$k^{(2)}$	$k^{(3)}$
3	0	0	0	0	0	700		0	0	0	0
4	100	0	1	0	100			100	30	50	40
5	200	0	0	0	0			200	50	80	50
6	300	0	0	0	0			300	90	90	110
7	400	0	0	0	0			400	110	150	120
8	500	0	0	0	0			500	170	190	180
9	600	0	0	1	600			600	180	210	220
10	700	0	0	0	0			700	210	220	240
11	Кількість вкладень	0	1	1		Максимальна сума вкладень					
12	Максимальна кількість вкладень	1	1	1		700		Прибуток	270	max	

Рис. 10.28. Результати розв'язування задачі (10.8)

Розв'язування задачі за допомогою R language

Знаходження розв'язку задачі (10.8) у R здійснюється виконанням таких операцій:

1) автоматичне встановлення та підключення потрібного пакету до сесії R:

```
if (!require(lpSolve)) install.packages("lpSolve")
library(lpSolve)
```

2) введення вхідних даних задачі:

```
investments <- seq(0, 700, by = 100)
profits <- matrix(c(
  0, 0, 0,
  30, 50, 40,
  50, 80, 50,
```

```
90, 90, 110,  
110, 150, 120,  
170, 190, 180,  
180, 210, 220,  
210, 220, 240  
) , nrow = 8, ncol = 3, byrow = TRUE)
```

3) введення:

- цільової функції:

```
f.obj <- as.vector(profits)
```

- кількості підприємств і варіантів інвестицій:

```
n_options <- nrow(profits)  
n_enterprises <- ncol(profits)
```

- матриці обмежень:

```
mat <- matrix(0, nrow = n_enterprises + 1, ncol = n_options *  
n_enterprises)  
  
# обмеження на вибір лише одного варіанту для кожного  
підприємства  
for (i in 1:n_enterprises) {  
  mat[i, ((i - 1) * n_options + 1):(i * n_options)] <- 1  
}  
  
# обмеження на загальну суму інвестицій  
mat[n_enterprises + 1, ] <- rep(investments, n_enterprises)
```

- типу обмежень (усі рівності) та їх відповідних значень (кожне підприємство отримує одну інвестицію, а загальна сума інвестицій – 700).

```
dir <- c(rep("==", n_enterprises + 1))  
b <- c(rep(1, n_enterprises), 700)
```

4) розв'язування задачі з бінарними змінними:

```
solution <- lp("max", f.obj, mat, dir, b, all.bin = TRUE)
```

5) виведення статусу розв'язку та максимального прибутку, отриманого у результаті оптимізації:

```
if (solution$status == 0) {  
  cat("Оптимальне рішення знайдено\n")  
  cat("Максимальний прибуток:", solution$objval, "тис.  
грн\n\n")  
} else {  
  cat("Помилка в розрахунках\n")  
  stop("Оптимальне рішення не знайдено.")  
}
```

Оптимальне рішення знайдено
Максимальний прибуток: 270 тис. грн

6) виведення розв'язку задачі (10.8):

```
allocation <- matrix(solution$solution, nrow = n_options, ncol =  
n_enterprises, byrow = FALSE)  
investment_values <- numeric(n_enterprises)  
for (i in 1:n_enterprises) {  
  inv <- investments[which(allocation[, i] == 1)]  
  investment_values[i] <- inv  
  cat("Підприємству", i, "потрібно виділити", inv, "тис. грн\n")  
}
```

Підприємству 1 потрібно виділити 0 тис. грн
Підприємству 2 потрібно виділити 100 тис. грн
Підприємству 3 потрібно виділити 600 тис. грн

7) візуалізація розподілу інвестицій між підприємствами за допомогою стовпчикової діаграми (рис. 10.29).

```
barplot(investment_values,  
names.arg = paste("Підприємство", 1:n_enterprises),  
main = "Розподіл інвестицій між підприємствами",
```



```
ylab = "Інвестиції (тис. грн)",  
col = rainbow(n_enterprises))
```

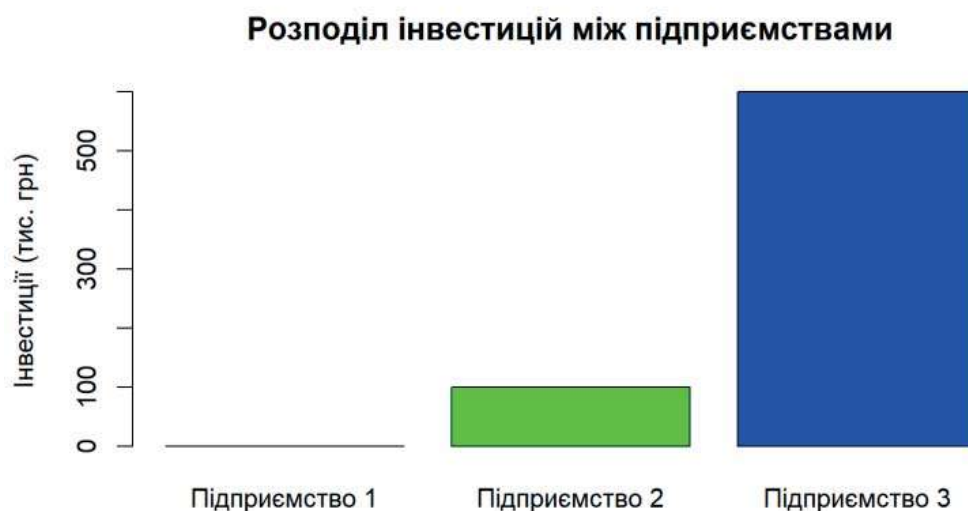


Рис. 10.29. Стовпчикова діаграма для візуалізації розподілу інвестицій між підприємствами

Розв'язування задачі за допомогою Matlab

Алгоритм знаходження розв'язку задачі (10.8) у Matlab конкретизується так:

```
1) введення початкових даних та обмежень задачі:  
u = [0;100;200;300;400;500;600;700];  
f = [0 0 0; 30 50 40; 50 80 50; 90 90 110; 110 150 120; 170  
190 180; 180 210 220; 210 220 240];  
f = reshape(f',[24,1]);  
intcon = 1:24;  
lb = zeros(24,1);  
ub = ones(24,1);  
A = [1 0 0 1 0 0 1 0 0 1 0 0 1 0 0 1 0 0 1 0 0;  
      0 1 0 0 1 0 0 1 0 0 1 0 0 1 0 0 1 0 0 1 0;  
      0 0 1 0 0 1 0 0 1 0 0 1 0 0 1 0 0 1 0 0 1];  
b = [1;1;1];
```

```
Aeq = [1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0;
       0 0 0 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0;
       0 0 0 0 0 0 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0;
       0 0 0 0 0 0 0 0 0 1 1 1 0 0 0 0 0 0 0 0 0 0 0;
       0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 0 0 0 0 0 0 0;
       0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 0 0 0 0;
       0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 0 0;
       0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1];
```

```
Aeq = Aeq * u;
```

```
Aeq = sum(Aeq,1);
```

```
beq = 700;
```

2) звернення до процедури функції:

```
x = intlinprog(-f,intcon,A,b,Aeq,beq,lb,ub);
```

```
profit = sum(f*x);
```

```
x = reshape(x, [3,8]);
```

```
x = x.';
```

```
disp('x = ');
```

```
disp(x);
```

```
disp('profit = ');
```

```
disp(profit);
```

3) виведення результатів:

```
x = 0
```

```
    100
```

```
    600
```

Максимальний прибуток: 270 тис. грн.

Розв'язки збігаються у всіх розглянутих вище варіантах їх знаходження.

СПИСОК ЛІТЕРАТУРИ

1. Вища математика. Курс лекцій у трьох частинах. Частина 3. Математичні методи дослідження операцій : Навчальний посібник / В.П. Лавренчук, Т.І. Готинчан, В.С. Дронь, О.С. Кондур. Чернівці : Рута, 2007. – 296 с.
2. Вітлінський В.В. Економіко-математичні методи та моделі: оптимізація : навч. посібник / Вітлінський В.В., Терещенко Т.О., Савіна С.С. К. : КНЕУ, 2016. 303 с.
3. Вовк В.М., Зомчак Л.М. Оптимізаційні методи та моделі: навч. посібник. Львів: ЛНУ імені Івана Франка, 2014. 360 с.
4. Григорків В.С. Економічна кібернетика : Навчальний посібник. Чернівці : Рута, 2006. 198 с.
5. Григорків В. С. Моделювання економіки: підручник. Чернівці : Чернівецьк. нац. ун-т ім. Ю. Федьковича, 2019. 360 с.
6. Григорків В.С. Оптимальне керування в економіці : Навч. посібник. Чернівці : Чернівецький нац. ун-т, 2011. 200 с.
7. Григорків В.С. Практикум з математичного програмування : навч. посібник для студентів економічних спеціальностей вузів / В.С. Григорків, М.В. Бойчук. Чернівці : Прут, 1995. 244 с.
8. Григорків В.С., Григорків М.В. Економічна кібернетика: практикум: навч. посіб. Чернівці : Чернівецький нац. ун-т, 2024. 200 с.
9. Григорків В.С., Григорків М.В. Моделі прийняття рішень в економіці: навч. посібник / В.С. Григорків,

М.В. Григорків. Чернівці : Чернівецьк. нац. ун-т, 2021. 256 с.

10. Григорків В.С., Григорків М.В. Моделі прийняття рішень в економіці: вибрані завдання для тематичного контролю : навч. посіб. / В.С. Григорків, М.В. Григорків. Чернівці : Чернівецький нац. ун-т, 2021. 48 с.

11. Григорків В.С., Григорків М.В. Моделювання економіки: практикум: навч. посіб. / В.С. Григорків, М.В. Григорків. Чернівці: Чернівецький нац. ун-т, 2023. 208 с.

12. Григорків В.С., Григорків М.В., Ярошенко О.І. Оптимізаційні методи та моделі : підручник / В.С. Григорків, М.В. Григорків, О.І. Ярошенко. Чернівці : Чернівецький нац. ун-т, 2022. 440 с.

13. Григорків М.В. Оптимізаційні методи та моделі: тестові завдання : навч. посібник / М.В. Григорків, В.С. Григорків, О.І. Ярошенко, О.Ю. Вінничук, Л.В. Скращук. Чернівці : Чернівецький нац. ун-т, 2022. 136 с.

14. Заболоцький М.В. Математичний аналіз : Підручник / М.В. Заболоцький, О.Г. Сторож, С.І. Тарасюк. К. : Знання, 2008. 421 с.

15. Забуранна Л.В., Попрозман Н.В., Клименко Н.А., Попрозман О.І., Зарубанний С.В. Оптимізаційні методи та моделі: підручник. Київ: НУБІП, 2014. 372 с.

16. Кігель В.Р. Елементи лінійного цілочислового лінійного, нелінійного програмування : навч. посібник / В.Р. Кігель. К. : ІСДО, 1995. 400 с.

17. Кісельова О.М., Шевельова А.Є. Чисельні методи оптимізації: навч. посіб. Д.: Вид-во ДНУ, 2008. 208 с.

18. Клочко О.В., Клочко В.І., Потапова Н.А. Методи оптимізації в економіці: Навчальний посібник. Вінниця: Вінницька газета, 2013. 456 с.

19. Математичне програмування : Навчальний посібник / І.М. Богаєнко, В.С. Григорків, М.В. Бойчук, М.О. Рюмшин. К. :

Логос, 1996. 266 с.

20. Нагнибіда М.І. Математичний аналіз. Завдання для самостійної роботи. Посібник / М.І. Нагнибіда, П.П. Настасієв. К. : Вища школа, 1981. 224 с.

21. Наконечний С.І., Савіна С.С. Математичне програмування: Навч. посіб. К.: КНЕУ, 2003. 452 с.

22. Оптимізаційні методи та моделі: вибрані завдання для тематичного контролю: навч. посіб. / В.С. Григорків, М.В. Григорків, О.І. Ярошенко, О.Ю. Вінничук, Л.В. Скращук. Чернівці : Чернівецький нац. ун-т, 2022. 168 с.

