

Рівненський державний гуманітарний університет
Факультет математики та інформатики
Кафедра вищої математики

Кваліфікаційна робота
магістерського рівня
на тему

Класична теорія оптимізації в задачах нелінійного моделювання

Виконала: студентка II курсу
магістратури, групи М-М-21
спеціальності 014 Середня освіта
(Математика)
Тимощук Ірина Миколаївна

Керівник: доктор технічних наук,
професор Бичков О.С.

Рецензент: кандидатка технічних наук,
доцентка кафедри автоматизації,
електротехнічних та комп'ютерно-
інтегрованих технологій НУВГП
Присяжнюк О.В.

Рівне-2022 року

ЗМІСТ

ВСТУП.....	3
1. ТЕОРЕТИЧНІ ОСНОВИ КЛАСИЧНОЇ ТЕОРІЇ ОПТИМІЗАЦІЇ.....	5
1.1. Екстремальні задачі без обмежень. Метод Ньютона-Рафсона.....	5
1.2. Екстремальні задачі з обмеженнями у вигляді рівностей.....	11
1.3. Екстремальні задачі з обмеженнями у вигляді нерівностей.....	18
2. МАТЕМАТИЧНІ ОСНОВИ ОПУКЛОГО ПРОГРАМУВАННЯ.....	23
2.1. Похідна за напрямком та градієнт. Опуклі та вгнуті функції.....	23
2.2. Задача опуклого програмування.....	32
2.3. Метод множників Лагранжа.....	33
2.4. Градієнтні методи.....	36
2.5. Теорема Куна-Таккера.....	44
3. ПРИКЛАДИ РОЗВ'ЯЗУВАННЯ ЗАДАЧ НЕЛІНІЙНОГО ПРОГРАМУВАННЯ.....	47
ВИСНОВКИ.....	54
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ.....	55

ВСТУП

Актуальність. Дослідження операцій орієнтоване на розв'язування практичних задач, які можна розв'язати за допомогою математичних моделей.

У моделях дослідження операцій поняття оптимальності розв'язків визначається з урахуванням відповідності цього розв'язку множині обмежень. Це означає, що якість прийнятого рішення, зробленого на основі розв'язування задачі, залежить від адекватності представлення моделлю реальної ситуації, яку вона формально описує за допомогою обмежень. Чим краще модель відображає реальну ситуацію, тим ближчим буде знайдений розв'язок до оптимального.

У дослідженнях операцій немає єдиного загального методу розв'язування усіх математичних моделей, які зустрічаються на практиці. Замість цього, вибір методу розв'язування диктують тип і складність математичної моделі, що вивчається. Найбільш відомими та ефективними методами дослідження операцій є методи лінійного програмування, коли цільова функція і усі обмеження є лінійними функціями. Для розв'язування математичних моделей інших типів призначені методи цілочислового, динамічного та нелінійного програмування (коли цільова функція та/чи обмеження є нелінійними функціями).

Майже усі методи нелінійного програмування не дозволяють отримати розв'язок у замкненій формі. Навпаки, вони породжують обчислювальні алгоритми, які є ітераційними за своєю природою та вимагає великої кількості однотипних обчислень. Унаслідок цього ці алгоритми, в основному, розробляються для реалізації за допомогою обчислювальної техніки.

Мета дослідження: систематизувати відомості про нелінійні задачі математичного програмування та методи їх розв'язування.

Об'єкт дослідження: теорія оптимізації.

Предмет дослідження: задачі нелінійного моделювання.

Завдання:

1. Проаналізувати літературу з предмету дослідження.
2. Розкрити зміст поняття екстремуму функції та задач, пов'язаних із цим поняттям.
3. Систематизувати відомості про задачі опуклого програмування, точні та чисельні методи їх розв'язування.
4. Підібрати приклади розв'язування задач та навести їх повні розв'язання.

Методи дослідження. Під час дослідження у роботі використовувалися методи: аналіз, синтез, математичне моделювання, порівняння результатів дослідження.

Робота містить вступ, три розділи, висновки та список використаних джерел.

РОЗДІЛ 1. ТЕОРЕТИЧНІ ОСНОВИ КЛАСИЧНОЇ ТЕОРІЇ ОПТИМІЗАЦІЇ

1.1. Екстремальні задачі без обмежень. Метод Ньютона-Рафсона

Під математичною оптимізацією мають на увазі вибір найкращого (оптимального) елемента із наявної множини альтернатив. Фактично, теорія оптимізації – це розділ математики, що займається вивченням задач про пошук екстремуму даної функції на множині допустимих значень.

У класичній теорії оптимізації для пошуку точок максимуму та мінімуму функцій за умов як наявності, так відсутності обмежень, що накладаються на змінні, широко використовується апарат диференціального числення. Отримувані при цьому методи не завжди є зручними для їх чисельної реалізації. Однак, відповідні теоретичні результати лежать в основі більшості алгоритмів розв’язування задач нелінійного програмування.

Стандартна задача теорії оптимізації передбачає пошук такої точки $x_0 = (x_1^0, \dots, x_n^0)$, у якій дана (цільова) функція $f(x)$ набуває свого екстремального значення.

Екстремальна точка функції $f(x)$ визначає або її максимальне, або мінімальне значення. За означенням, точка x_0 є точкою максимуму, якщо

$$f(x_0 + h) \leq f(x_0),$$

де $h = (h_1, \dots, h_j, \dots, h_n)$ такий, що $|h_j|$ – досить малі при всіх j . Іншими словами, точка x_0 є точкою максимуму, якщо в її околі значення функції $f(x)$ не перевищують $f(x_0)$. Аналогічно x_0 є точкою мінімуму, якщо для всіх точок її околу виконується нерівність [4]

$$f(x_0 + h) \geq f(x_0).$$

На рисунку 1.1 зображено точки максимуму і мінімуму функції однієї змінної $f(x)$ на відрізку $[a; b]$. Точки x_1, x_2, x_3, x_4 і x_6 складають

множину її екстремальних точок. Тут точки x_1, x_3 і x_6 є точками максимуму, а точки x_2 та x_4 – точками мінімуму функції $f(x)$. Оскільки

$$f(x_6) = \max\{f(x_1), f(x_3), f(x_6)\},$$

значення $f(x_6)$ називають глобальним або абсолютним максимумом (найбільшим значенням на відрізку), а значення $f(x_1)$ і $f(x_3)$ – локальними або відносними максимумами. Аналогічно $f(x_4)$ є локальним, а $f(x_2)$ – глобальним мінімумом функції $f(x)$.

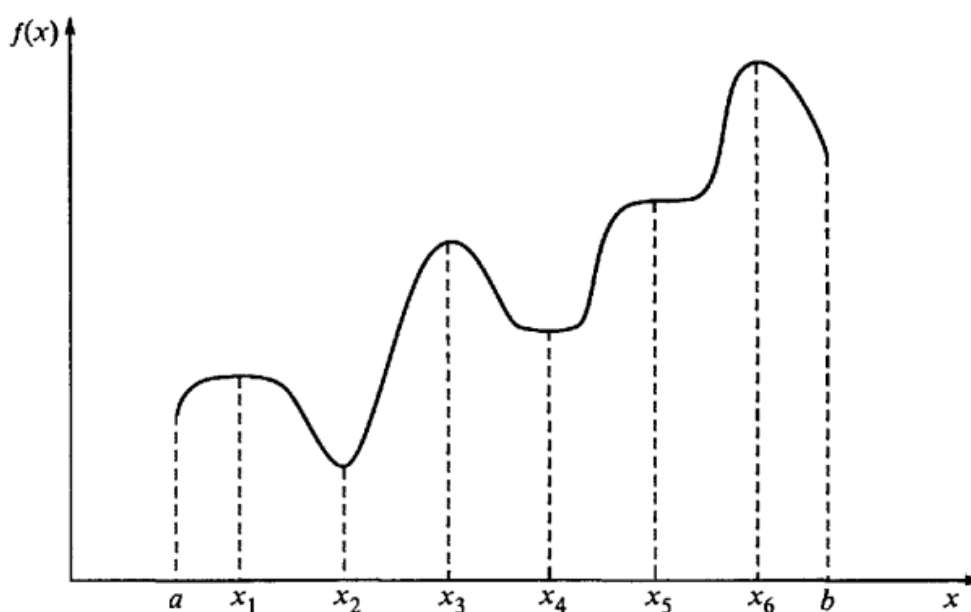


Рис. 1.1

Зауважимо, що хоча точка x_1 є точкою максимуму функції $f(x)$, вона відрізняється від інших локальних максимумів функції $f(x)$ тим, що, принаймні у одній точці її околу, значення функції $f(x)$ співпадає з $f(x_1)$. Точка x_1 із цієї причини називається нестрогим (слабким) максимумом функції $f(x)$, на відміну, наприклад, від точки x_3 , яка є строгим максимумом $f(x)$. Нестрогий максимум передбачає наявність різних точок, яким відповідає одне і те ж максимальне значення функції. Аналогічні результати мають місце у точці x_4 , де функція $f(x)$ має нестрогий мінімум.

У загальному випадку x_0 буде точкою строгого максимуму, якщо для усіх точок її околу виконується нерівність

$$f(x_0 + h) < f(x_0).$$

Розглянемо далі теореми, які дають необхідні і достатні умови існування у точці екстремуму функції $f(x) = f(x_1, \dots, x_n)$ кількох змінних. При цьому будемо вважати функцію $f(x)$ неперервною разом зі своїми частинними похідними до другого порядку включно [1, 18].

Теорема 1.1. Необхідною умовою того, що точка x_0 є екстремальною точкою функції $f(x)$ є рівність

$$\nabla f(x_0) = 0.$$

Доведення

Із теореми Тейлора слідує, що при $0 < \theta < 1$ має місце розклад функції $f(x)$

$$f(x_0 + h) - f(x_0) = \nabla f(x_0)h + \frac{1}{2} h^T Hh \Big|_{x_0 + \theta h},$$

де h – вектор, означений вище.

Для досить малих значень $|h_j|$ залишковий член $\frac{1}{2} h^T Hh$ є величиною порядку h_j^2 . Отже,

$$f(x_0 + h) - f(x_0) = \nabla f(x_0)h + O(h_j^2) \approx \nabla f(x_0)h.$$

Нехай x_0 – точка мінімуму функції $f(x)$. Доведемо, міркуючи від супротивного, що градієнт $\nabla f(x_0)$ функції $f(x)$ в точці мінімуму x_0 рівний нулю. Нехай ця умова не виконується, тоді для деякого j має виконуватись умова

$$\frac{\partial f(x_0)}{\partial x_j} < 0 \quad \text{або} \quad \frac{\partial f(x_0)}{\partial x_j} > 0.$$

Знак h_j завжди можна вибрати таким чином, щоб

$$h_j \frac{\partial f(x_0)}{\partial x_j} < 0.$$

Покладаючи інші h_j рівними нулю, із розкладу Тейлора отримуємо нерівність

$$f(x_0 + h) < f(x_0).$$

А цей результат суперечить припущенню, що x_0 – точка мінімуму. Отже, величина $\nabla f(x_0)$ має дорівнювати нулю. Доведення для точки максимуму проводиться аналогічно.

Так як необхідна умова виконується також у точках перегину і сідлових точках, точки, що задовольняють рівність $\nabla f(x_0) = 0$, називаються стаціонарними [7].

Наступна теорема дає достатні умови того, що стаціонарна точка є екстремальною.

Теорема 1.2. Для того, щоб стаціонарна точка x_0 була екстремальною, достатньо, щоб матриця Гессе H в точці x_0 була:

- а) додатно визначеною (тоді x_0 – точка мінімуму);
- б) від'ємно визначеною (тоді x_0 – точка максимуму).

Доведення

За формулою Тейлора при $0 < \theta < 1$ маємо

$$f(x_0 + h) - f(x_0) = \nabla f(x_0)h + \frac{1}{2} h^T Hh \Big|_{x_0 + \theta h},$$

Оскільки x_0 – стаціонарна точка, то за теоремою 1.1 $\nabla f(x_0) = 0$. Таким чином,

$$f(x_0 + h) - f(x_0) = \frac{1}{2} h^T Hh \Big|_{x_0 + \theta h}.$$

Якщо x_0 – точка мінімуму, то

$$f(x_0 + h) > f(x_0)$$

для усіх ненульових векторів h . Отже, в точці мінімуму x_0 має виконуватись нерівність

$$\frac{1}{2} h^T H h \Big|_{x_0 + \theta h} > 0.$$

Неперервність других частинних похідних функції $f(x)$ гарантує, що величина $\frac{1}{2} h^T H h$ має один і той же знак як у точці x_0 , так і в точці $x_0 + \theta h$.

Так як $\frac{1}{2} h^T H h \Big|_{x_0}$ являє собою квадратичну форму, то її значення (і, отже,

$\frac{1}{2} h^T H h \Big|_{x_0 + \theta h}$) додатне тоді і тільки тоді, коли $H \Big|_{x_0}$ – додатно визначена

матриця. Це означає, що додатна визначеність матриці Гессе у стаціонарній точці x_0 є достатньою умовою існування у цій точці мінімуму. Аналогічними міркуваннями доводиться, що стаціонарна точка є точкою максимуму, якщо матриця Гессе у цій точці є від'ємно визначеною.

У загальному випадку, коли матриця $H \Big|_{x_0}$ є невизначеною, точка x_0 має бути сідловою. Якщо ж матриця $H \Big|_{x_0}$ є напіввизначеною, то відповідна точка x_0 може як бути, так і не бути точкою екстремуму. При цьому формулювання достатньої умови існування екстремуму значно ускладнюється, оскільки для цього необхідно враховувати члени більш високих порядків у розкладі Тейлора.

У загальному випадку використання необхідної умови екстремуму для пошуку стаціонарних точок функції може бути пов'язане з певними труднощами, що виникають при чисельному розв'язуванні відповідної системи рівнянь. Метод Ньютона-Рафсона пропонує ітераційну процедуру

розв'язування системи нелінійних рівнянь. Цей метод відноситься до градієнтних методів чисельного пошуку екстремумів функції при відсутності обмежень [13].

Розглянемо систему рівнянь

$$f_i(x) = 0, \quad i = 1, 2, \dots, m. \quad (1.1)$$

Нехай x^k – деяка фіксована точка. Використовуючи розклад Тейлора, маємо

$$f_i(x) \approx f_i(x^k) + \nabla f_i(x^k)(x - x^k), \quad i = 1, 2, \dots, m.$$

Отже, система (1.1) може бути наближено представлена у вигляді

$$f_i(x^k) + \nabla f_i(x^k)(x - x^k) = 0, \quad i = 1, 2, \dots, m. \quad (1.2)$$

Рівняння (1.2) можна записати у матричній формі:

$$A_k + B_k(X - X^k) = 0. \quad (1.3)$$

Припустимо, що вектори $f_i(x)$ – незалежні, тоді матриця B_k є невинродженою. Тоді з (1.3) отримаємо:

$$X = X^k - B_k^{-1} A_k. \quad (1.4)$$

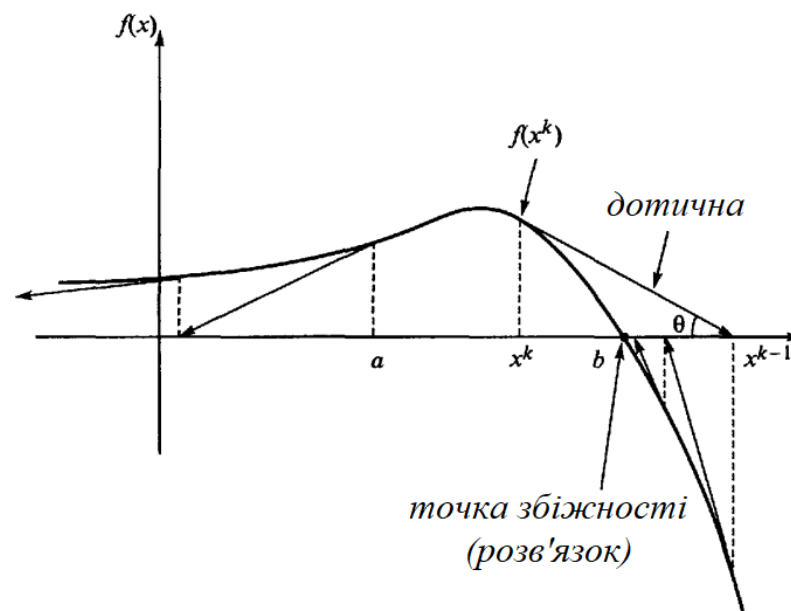


Рис. 1.2

Ідея методу Ньютона-Рафсона полягає у наступному. На першому кроці вибирається початкова точка X^0 . З допомогою (1.4) за відомою точкою X^k можна знайти координати нової точки X^{k+1} . Процес обчислень закінчується

у точці X^m , яка вважається наближеним розв'язком вихідної системи, якщо $X^m \approx X^{m-1}$.

Геометрична інтерпретація даного методу для функції однієї змінної $f_i(x)$ наведена на рисунку 1.2. Зв'язок між точками x^k та x^{k+1} у цьому випадку виражається формулою

$$x^{k+1} = x^k - \frac{f(x^k)}{f'(x^k)}$$

або

$$f'(x^k) = \frac{f(x^k)}{x^k - x^{k+1}}.$$

З рисунка 1.2 помітно, що положення точки x^{k+1} визначається кутом θ нахилу дотичної до графіка функції $f(x)$ у точці x^k , де $\operatorname{tg}\theta = f'(x^k)$.

Одним із недоліків цього методу є те, що для функцій загального вигляду не завжди гарантується його збіжність. На рис. 1.2 видно, що при виборі в якості точки x^0 точки a ітераційний процес розбігається. Простих рецептів вибору хорошого початкового наближення не існує [23].

1.2. Екстремальні задачі з обмеженнями у вигляді рівностей

У даному підрозділі розглянемо два методи розв'язування оптимізаційних задач за наявності обмежень у вигляді рівностей: метод Якобі і метод Лагранжа.

Розглянемо задачу:

$$\text{мінімізувати } z = f(x)$$

при обмеженнях

$$g(x) = 0,$$

де

$$x = (x_1, \dots, x_n),$$

$$g = (g_1, \dots, g_m)^T.$$

Функції $f(x)$ та $g_i(x), i = 1, 2, \dots, m$ вважаються двічі неперервно диференційованими.

Ідея використання зведеного градієнту полягає у тому, щоб знайти замкнений аналітичний вираз для перших частинних похідних функції $f(x)$ у всіх точках, що задовольняють обмеження $g(x) = 0$. Відповідні стаціонарні точки визначаються з умови рівності нулю частинних похідних. Потім можна використовувати достатні умови, сформульовані у підрозділі 1.1, для класифікації знайдених стаціонарних точок.

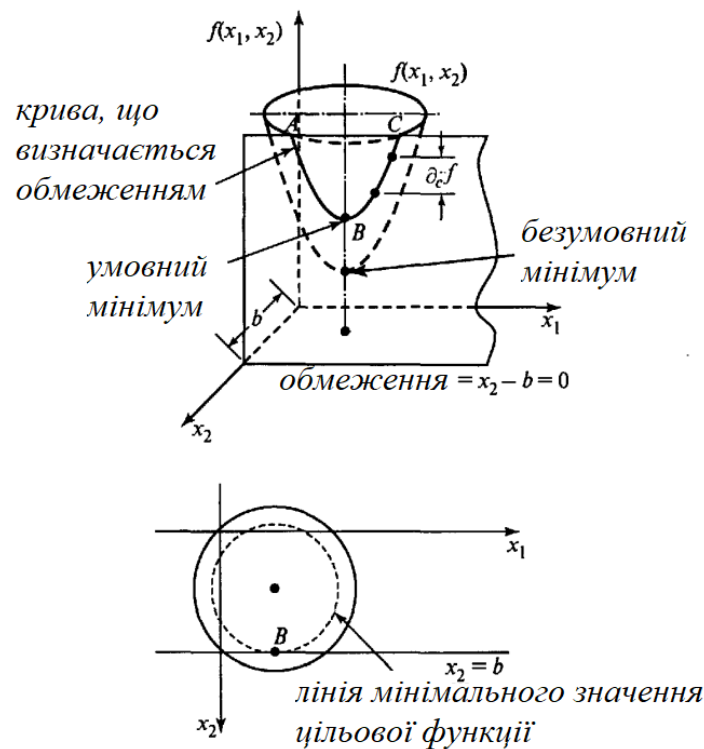


Рис. 1.3

Для пояснення викладеної ідеї, розглянемо функцію $f(x_1, x_2)$, графік якої представлений на рисунку 1.3. Припустимо, що цю функцію необхідно мінімізувати при обмеженні

$$g_1(x_1, x_2) = x_2 - b = 0,$$

де b – деяка константа. На рисунку 1.3 бачимо, що крива, яка проходить через точки A , B і C , складається із значень функції $f(x_1, x_2)$, для яких задане

обмеження виконане. У відповідності до розглядуваного методу, визначаються компоненти зведеного градієнту функції $f(x_1, x_2)$ у кожній точці кривої ABC . Точка B , у якій зведена похідна перетворюється в нуль, є стаціонарною для розглядуваної задачі з обмеженнями [5,16].

Тепер розглянемо загальне математичне формулювання методу. З теореми Тейлора випливає, що для точок $x + \Delta x$ із околу точки x маємо

$$f(x + \Delta x) - f(x) = \nabla f(x)\Delta x + O(\Delta x_j^2),$$

$$g(x + \Delta x) - g(x) = \nabla g(x)\Delta x + O(\Delta x_j^2).$$

При $\Delta x_j \rightarrow 0$ ці рівняння набувають вигляду

$$\partial f(x) = \nabla f(x)\partial x,$$

$$\partial g(x) = \nabla g(x)\partial x.$$

Оскільки $g(x) = 0$, то $\partial g(x) = 0$ у допустимій області. Звідси випливає, що

$$\partial f(x) - \nabla f(x)\partial x = 0,$$

$$\nabla g(x)\partial x = 0.$$

Як видно, задача зводиться до розв'язування $m+1$ рівнянь з $n+1$ невідомими, якими є $\partial f(x)$ і ∂x . Невідому величину $\partial f(x)$ можна визначити, як тільки буде знайдено вектор ∂x . Це означає, що, по суті, є m рівнянь з n невідомими.

При $m > n$ принаймні $m-n$ рівнянь системи є надлишковими. Після усунення надлишковості кількість незалежних рівнянь у системі стає рівною $m (\leq n)$. Якщо $m = n$, розв'язком є $\partial x = 0$. При цьому точка x не має допустимого околу, і, отже, простір розв'язків задачі складається з однієї точки. Така ситуація є тривіальною. Ситуацію, коли $m < n$, розглянемо детально.

Нехай $x = (y, z)$, де

$$y = (y_1, \dots, y_m) \quad \text{і} \quad z = (z_1, \dots, z_{n-m})$$

називаються незалежними і залежними змінними відповідно. Перепишуючи градієнти функцій f і g у нових позначення, отримуємо [11]

$$\nabla f(y, z) = (\nabla_y f, \nabla_z f),$$

$$\nabla g(y, z) = (\nabla_y g, \nabla_z g).$$

Введемо у розгляд матриці

$$J = \nabla_y g = \begin{pmatrix} \nabla_y g_1 \\ \dots \\ \nabla_y g_m \end{pmatrix},$$

$$C = \nabla_z g = \begin{pmatrix} \nabla_z g_1 \\ \dots \\ \nabla_z g_m \end{pmatrix}.$$

Матриця $J_{m \times m}$ називається матрицею Якобі, а $C_{m \times (n-m)}$ – матрицею керування. Матриця Якобі вважається невинродженою. Це завжди можна забезпечити, оскільки розглядувані рівнянь є незалежними за визначенням. Тому компоненти вектора y можна вибрати серед компонентів вектора x таким чином, що матриця J виявиться невинродженою.

Вихідну систему рівнянь з невідомими і можна переписати в наступному вигляді

$$\partial f(y, z) = \nabla_y f \partial y + \nabla_z f \partial z, \quad (1.5)$$

$$J \partial y = -C \partial z. \quad (1.6)$$

Так як матриця J – невинроджена, то існує обернена до неї матриця J^{-1} . Тоді з (1.6)

$$\partial y = -J^{-1} C \partial z. \quad (1.7)$$

Підставляючи (1.7) у (1.5), можна виразити ∂f через ∂z у наступному вигляді

$$\partial f(y, z) = (\nabla_z f - \nabla_y f J^{-1} C) \partial z. \quad (1.8)$$

Із (1.8) отримуємо формулу для похідних функції f по вектору незалежних змінних z :

$$\nabla_c f = \frac{\partial_c f(y, z)}{\partial_c z} = \nabla_z f - \nabla_y f J^{-1} C,$$

де $\nabla_c f$ – представляє вектор зведеного градієнту функції f по z . Тоді вектор має перетворюватись в нуль у стаціонарних точках [22].

Достатні умови екстремуму у стаціонарній точці аналогічні викладеним у підрозділі 1.1. У цьому випадку елементи матриці Гессе будуть відповідати компонентам вектора незалежних змінних z . При цьому елементи матриці Гессе повинні бути зведеними другими похідними.

Метод Якобі можна використовувати для аналізу чутливості оптимального значення цільової функції f до малих змін правих частин обмежень оптимізаційної задачі. Зокрема, можна визначити, як повпливає на оптимальне значення функції f заміна обмеження $g_i(x) = 0$ на $g_i(x) = \delta g_i$. Дослідження такого типу називається аналізом чутливості і, у деякому сенсі, аналогічне відповідній процедурі у лінійному програмуванні. Варто, однак, зауважити, що аналіз чутливості у нелінійному програмуванні приводить до результатів, що справедливі лише у малому околі екстремальної точки. Знайомство з такими процедурами може бути корисним при вивченні методу множників Лагранжа.

Вище було отримано, що

$$\partial f(y, z) = \nabla_y f \partial y + \nabla_z f \partial z, \quad (1.5)$$

$$\partial g = J^{-1} \partial y + C \partial z.$$

Припустимо, що $\partial g \neq 0$, тоді

$$\partial y = J^{-1} \partial g - J^{-1} C \partial z. \quad (1.9)$$

Підставляючи (1.9) у (1.5), отримуємо

$$\partial f(y, z) = \nabla_y f J^{-1} \partial g + \nabla_c f \partial z,$$

де $\nabla_c f = \nabla_z f - \nabla_y f J^{-1} C$, як було визначено раніше. Отриманий вираз для $\partial f(y, z)$ можна використовувати при аналізі змін цільової функції f у малому околі допустимої точки x_0 , обумовлених малими змінами ∂g і ∂z .

У екстремальній точці $x_0 = (y_0, z_0)$ зведений градієнт $\nabla_c f$ має бути рівний нулю. Отже, у точці x_0 маємо

$$\partial f(y_0, z_0) = \nabla_{y_0} f J^{-1} \partial g(y_0, z_0)$$

або

$$\frac{\partial f}{\partial g} = \nabla_{y_0} f J^{-1}, \quad (1.10)$$

обчисленому в точці x_0 . Це означає, що вплив малих змін ∂g на оптимальне значення функції f можна оцінити через швидкість зміни функції f по відношенню до змін g . Ці величини прийнято називати коефіцієнтами чутливості.

Коефіцієнти чутливості (1.10) методу Якобі можна використовувати для розв'язування задач з обмеженнями у вигляді рівностей. Нехай

$$\lambda = \nabla_{y_0} J^{-1} = \frac{\partial f}{\partial g}.$$

Тоді

$$\partial f - \lambda \partial g = 0. \quad (1.11)$$

Рівняння (1.11) відповідає необхідним умовам стаціонарності точок, оскільки

формула для отримана $\frac{\partial f}{\partial g}$ з урахуванням того, що $\nabla_c f = 0$. Цю рівність

можна записати у більш зручній формі, якщо перейти до частинних похідних по усіх змінних, що приводить до системи рівнянь

$$\frac{\partial}{\partial x_j} (f - \lambda g) = 0, \quad j = 1, 2, \dots, n. \quad (1.12)$$

Рівняння (1.12) разом з обмеженнями задачі $g(x) = 0$ формують систему, яка дозволяє визначити допустимі вектори x та λ , що задовольняють необхідні умови стаціонарності [6].

Описана процедура є основою методу множників Лагранжа, який дозволяє визначати стаціонарні точки задачі оптимізації з обмеженнями у вигляді рівностей. Формально, схема цього методу може бути подана наступним чином.

Нехай

$$L(x, \lambda) = f(x) - \lambda g(x).$$

Функція L називається функцією Лагранжа, а параметри λ – множниками Лагранжа. Як випливає із означення, ці множники мають той же зміст, що і коефіцієнти чутливості, які фігурують у методі Якобі.

Рівняння

$$\frac{\partial L}{\partial \lambda} = 0 \quad \text{і} \quad \frac{\partial L}{\partial x} = 0$$

дають необхідні умови для визначення стаціонарних точок функції $f(x)$ при обмеженнях $g(x) = 0$. Достатні умови, що використовуються у методі множників Лагранжа, будуть сформульовані без доведення. Визначимо матрицю

$$H^B = \begin{pmatrix} 0 & P \\ P^T & Q \end{pmatrix}_{(m+n) \times (m+n)},$$

де

$$P = \begin{pmatrix} \nabla g_1(x) \\ \dots \\ \nabla g_m(x) \end{pmatrix}_{m \times n} \quad \text{і} \quad Q = \left\| \frac{\partial^2 L(x, \lambda)}{\partial x_i \partial x_j} \right\|_{n \times n} \quad \text{для всіх } i \text{ та } j.$$

Матриця H^B називається матрицею Гессе.

Нехай маємо стаціонарну точку (x_0, λ_0) функції Лагранжа $L(x, \lambda)$, і матриця Гессе H^B обчислена у точці (x_0, λ_0) . Тоді x_0 є:

1) точкою максимуму, якщо, починаючи з кутового мінору порядку $2m+1$, наступні $n-t$ кутових мінорів матриці Гессе H^B утворюють знакозмінний числовий ряд, у якому знак першого члена визначається множником $(-1)^{m+1}$;

2) точкою мінімуму, якщо, починаючи з кутового мінору порядку $2m+1$, наступні $n-t$ кутових мінорів матриці H^B мають знаки, які визначаються множником $(-1)^{m+1}$.

Ці умови є достатніми для визначення екстремальної точки задачі, які є як необхідними, так і достатніми. Проте їх практичне використання часто пов'язане зі значними труднощами.

Визначимо матрицю

$$\Delta = \begin{pmatrix} 0 & P \\ P^T & Q - \mu I \end{pmatrix},$$

обчислену у стаціонарній точці (x_0, λ_0) , де μ – невідомий параметр. Нехай

$|\Delta|$ – визначник матриці Δ , тоді усі $n-t$ дійсних коренів μ многочлена

$|\Delta| = 0$ мають бути:

- 1) від'ємними, якщо x_0 – точка максимуму;
- 2) додатними, якщо x_0 – точка мінімуму [2, 20].

1.3. Екстремальні задачі з обмеженнями у вигляді нерівностей

Розглянемо задачу:

$$\text{максимізувати } z = f(x)$$

при обмеженнях

$$g(x) \leq 0, \quad i = 1, 2, \dots, m.$$

Якщо у розглядуваній задачі є умови невід'ємності змінних $x \geq 0$, то вважають, що вони включені у вказані m обмежень.

Основна ідея узагальненого методу Лагранжа полягає у тому, що, якщо точка безумовного оптимуму функції $f(x)$ не задовольняє усім обмеженням задачі, то оптимальний розв'язок задачі з обмеженнями повинен досягатися у граничній точці області допустимих розв'язків. Отже, одне або кілька з m обмежень задачі мають виконуватись як рівності. Тому обчислювальна процедура узагальненого методу множників Лагранжа містить наступні кроки.

Крок 1. Розв'язується задача без урахування обмежень:

$$\text{максимізувати } z = f(x).$$

Якщо отриманий оптимальний розв'язок задовольняє усі обмеження задачі, то обчислення закінчуються, оскільки усі обмеження є надлишковими. Інакше, слід покласти $k = 1$ і перейти до кроку 2.

Крок 2. Активізуються будь-які k обмежень задачі (тобто перетворюються у рівності), і оптимізується функція $f(x)$ при наявності k обмежень методом множників Лагранжа. Якщо оптимальний розв'язок цієї задачі є допустимим по відношенню до решти обмежень вихідної задачі, то обчислення закінчуються: отримана точка є локальним оптимумом. Інакше, слід зробити активними інші k обмежень вихідної задачі і повторити даний крок. Якщо усі підмножини, що складаються із k активних обмежень, не приводять до допустимого розв'язку, слід перейти до кроку 3.

Крок 3. Якщо $k = m$, обчислення зупиняються: задача не має допустимих розв'язків. Інакше, необхідно покласти $k = k + 1$ і перейти до кроку 2 [5, 22].

У описаній обчислювальній процедурі часто ігнорується та важлива обставина, що вона не гарантує отримання глобального оптимуму навіть у тих випадках, коли задача має єдиний оптимальний розв'язок. Іншим важливим моментом є помилковість інтуїтивного уявлення, що якщо $p < q$,

то оптимальне значення цільової функції $f(x)$ при p обмеженнях-рівностях завжди краще, ніж при q обмеженнях-рівностях. У загальному випадку це має місце лише тоді, коли набір з p обмежень є підмножиною набору з q обмежень.

Розглянемо необхідні умови Куна-Таккера, які дозволяють визначати стаціонарні точки у задачі нелінійного програмування з обмеженнями у вигляді нерівностей. Ці умови є також і достатніми, якщо виконуються певні правила.

Розглянемо задачу:

$$\text{максимізувати } z = f(x)$$

при обмеженнях

$$g(x) \leq 0.$$

Обмеження-нерівності можна перетворити у рівності з допомогою невід'ємних додаткових змінних. Нехай $S_i^2 (\geq 0)$ – додаткова змінна, яка додається до лівої частини i -го обмеження, і нехай

$$S = (S_1, \dots, S_m)^T \quad \text{і} \quad S^2 = (S_1^2, \dots, S_m^2)^T,$$

де m – загальна кількість обмежень-нерівностей. Отже, функція Лагранжа записується у вигляді

$$L(x, S, \lambda) = f(x) - \lambda(g(x) + S^2).$$

При обмеженнях $g(x) \leq 0$ необхідною умовою оптимальності у задачі максимізації (мінімізації) є невід'ємність (недодатність) множників λ . Наведемо обґрунтування цього. Розглянемо задачу максимізації. Так як множники виражають швидкість зміни цільової функції f по відношенню до зміни g , тобто

$$\lambda = \frac{\partial f}{\partial g},$$

то як тільки права частина обмеження $g(x) \leq 0$ збільшується і стає більшою нуля, область допустимих розв'язків задачі розширюється і, отже, оптимальне значення цільової функції f не може зменшитись. Це означає, що $\lambda \geq 0$. Аналогічно у задачі мінімізації при збільшенні правої частини обмежень оптимальне значення функції f не може збільшитись, звідки випливає, що $\lambda \leq 0$. Якщо ж обмеження задачі мають вигляд рівностей, то компоненти вектора λ за знаком не мають обмежень.

Вказані вище обмеження на вектор λ повинні розглядатись як частина необхідних умов Куна-Таккера. Знайдемо тепер інші умови.

Знаходячи частинні похідні функції L по x , S і λ та прирівнюючи їх до нуля, отримуємо:

$$\frac{\partial L}{\partial x} = \nabla f(x) - \lambda \nabla g(x) = 0,$$

$$\frac{\partial L}{\partial S_i} = -2\lambda_i S_i = 0, \quad i = 1, 2, \dots, m, \quad (1.13)$$

$$\frac{\partial L}{\partial \lambda} = -(g(x) + S^2) = 0. \quad (1.14)$$

Із (1.13) випливають наступні результати.

1) Якщо λ_i не дорівнює нулю, то $S_i^2 = 0$. Це означає, що ресурс, який відповідає цьому обмеженню, є дефіцитним і, отже, повністю вичерпаний (обмеження має вигляд рівності).

2) Якщо $S_i^2 > 0$, то $\lambda_i = 0$. Це означає, що i -й ресурс не є дефіцитним і, отже, не впливає на значення цільової функції (тобто).

Із (1.13) та (1.14) випливає, що

$$\lambda_i g_i(x) = 0, \quad i = 1, 2, \dots, m. \quad (1.15)$$

Отримані умови (1.15) фактично підтверджують попередній результат, оскільки якщо $\lambda_i > 0$, то $g_i(x) = 0$ або $S_i^2 = 0$. Аналогічно при $g_i(x) < 0$ $S_i^2 > 0$ і, отже, $\lambda_i = 0$.

Тепер для задачі максимізації можна сформулювати умови Куна-Таккера, необхідні для того, щоб вектори x та λ визначали стаціонарну точку [21]:

$$\begin{aligned} \lambda &\geq 0, \\ \nabla f(x) - \lambda \nabla g(x) &= 0, \\ \lambda_i g_i(x) &= 0, \quad i = 1, 2, \dots, m, \\ g(x) &\leq 0. \end{aligned}$$

Неважко переконатися, що ці умови мають місце також і для задачі мінімізації, за тим лише виключенням, що вектор повинен бути недодатним. При розв'язуванні як задачі максимізації, так і задачі мінімізації множники Лагранжа, що відповідають обмеженням у вигляді рівностей, не мають обмежень за знаком.

Якщо цільова функція і область допустимих розв'язків розглядуваної задачі володіють певними властивостями (табл. 1.1), пов'язаними із опуклістю та увігнутістю, то необхідні умови Куна-Таккера є також достатніми [9].

Таблиця 1.1

Тип оптимізації	Властивості, що вимагаються	
	Цільова функція	Область допустимих розв'язків
Максимізація	Увігнута	Опукла множина
Мінімізація	Опукла	Опукла множина

РОЗДІЛ 2. МАТЕМАТИЧНІ ОСНОВИ ОПУКЛОГО ПРОГРАМУВАННЯ

2.1. Похідна за напрямком та градієнт. Опуклі та вгнуті функції

Нехай дана функція $z=f(x,y)$, визначена в замкнутій області D площини xOy (рис. 2.1). Відомо, що частинні похідні z'_x і z'_y функції $z=f(x,y)$ можуть бути витлумачені як швидкості зміни цієї функції в точці $M_0(x,y)$ у напрямку осей Ox і Oy .

Поняття похідної можна узагальнити, якщо розглядати швидкість зміни функції $z = f(x,y)$ у довільному напрямку, що утворить деякий кут з віссю Ox . Нехай $M_1(x,y,z)$ - точка на поверхні $z = f(x,y)$. Надамо аргументам x и y скінченних приростів Δx і Δy . Тоді функція z набуде приросту Δz . Через M'_1 позначимо точку на поверхні з координатами $(x + \Delta x, y + \Delta y, z + \Delta z)$. Тоді $M_0(x,y)$ і $M_1(x,y)$ – це проекції цих точок на площину xOy . Позначимо через α кут, що утворює M_0M_1 віссю Ox ($0 \leq \alpha \leq 2\pi$), а через ρ — довжину відрізка M_0M_1 .

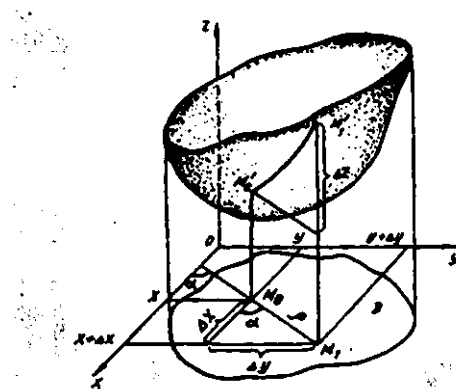


Рис. 2.1

Похідною $\frac{\partial z}{\partial \rho}$ функції $z=f(x,y)$ у точці $M_0(x,y)$ по напрямку α називається границя відношення приросту Δz функції z , що виник при переміщенні точки M з положення M_0 уздовж променя, що складає кут α з

додатним напрямком осі Ox , до величини p цього переміщення, коли p прямує до нуля, тобто

$$\frac{\partial z}{\partial p} = \lim_{p \rightarrow 0} \frac{\Delta z}{p} = \lim_{p \rightarrow 0} \frac{f(x + \Delta x, y + \Delta y) - f(x, y)}{p}. \quad (2.1)$$

Знайдемо формулу для обчислення похідної по напрямку α у точці $M(x, y)$ від функції $z=f(x, y)$, диференційованої у цій точці. З рис. 2.1 видно, що $\Delta x = p \cos \alpha, \Delta y = p \sin \alpha$.

Використовуючи еквівалентність повного приросту $\Delta z = f(x + \Delta x, y + \Delta y) - f(x, y)$ і повного диференціала $dz = \frac{\partial z}{\partial x} \Delta x + \frac{\partial z}{\partial y} \Delta y$, з формули (2.1) одержуємо

$$\begin{aligned} \frac{\partial z}{\partial p} &= \lim_{p \rightarrow 0} \frac{\Delta z}{p} = \lim_{p \rightarrow 0} \frac{dz}{p} = \lim_{p \rightarrow 0} \frac{\frac{\partial z}{\partial x} \Delta x + \frac{\partial z}{\partial y} \Delta y}{p} = \\ &= \lim_{p \rightarrow 0} \frac{\frac{\partial z}{\partial x} p \cos \alpha + \frac{\partial z}{\partial y} p \sin \alpha}{p} = \frac{\partial z}{\partial x} \cos \alpha + \frac{\partial z}{\partial y} \sin \alpha, \end{aligned}$$

оскільки значення похідних $\frac{\partial z}{\partial x}, \frac{\partial z}{\partial y}$ і кут α не залежать від p . Отже,

$$\frac{\partial z}{\partial p} = \frac{\partial z}{\partial x} \cos \alpha + \frac{\partial z}{\partial y} \sin \alpha. \quad (2.2)$$

Таким чином, похідна по напрямку α являється лінійною комбінацією частинних похідних [3].

Приклад. Обчислити в точці $M_0(2\sqrt{2}; \sqrt{2})$ похідну функції $z = xy + \ln x$ в напрямку, що складає кут $\alpha = \pi/4$ з віссю Ox .

По формулі (2.1) маємо

$$\begin{aligned} \frac{\partial z}{\partial p} \Big|_{\alpha=\pi/4} &= \left(y + \frac{1}{x} \right)_{M_0} \cos \frac{\pi}{4} + (x)_{M_0} \sin \frac{\pi}{4} = \\ &= \left(\sqrt{2} + \frac{1}{2\sqrt{2}} \right) \frac{\sqrt{2}}{2} + 2\sqrt{2} \cdot \frac{\sqrt{2}}{2} = \frac{13}{4} > 0, \end{aligned}$$

тобто швидкість зміни z виражається додатним числом, отже, функція z зростає в точці M_0 в зазначеному напрямку.

У математичному програмуванні найбільший інтерес представляє питання про напрямок найшвидшого зростання функції в даній точці. Ця задача розв'язується за допомогою вектора, що називається градієнтом функції.

Градієнтом функції $z = f(x, y)$ в даній точці $M_0(x, y)$ називається вектор, розташований у площині xOy і координатами якого є частинні похідні функції z , обчислені в цій точці.

Градієнт позначають $\text{grad } f(x, y), g, \nabla f(x, y)$.

Отже, $\vec{g} = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right)$, а модуль градієнта [14]

$$|\vec{g}| = g = \sqrt{\left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2}.$$

Теорема 2.1. Напрямок градієнта є напрямком найшвидшого зростання функції $z = f(x, y)$, а модуль градієнта дорівнює найбільшій швидкості зростання функції z .

Доведення. На рис. 2.1 напрямок а диференціювання визначається вектором $\overrightarrow{M_0 M_1} = \vec{a}$. Одиничний вектор \vec{a}_0 цього напрямку має координати $\cos \alpha, \sin \alpha$, тому що

$$|\vec{a}_0| = \sqrt{\cos^2 \alpha + \sin^2 \alpha} = 1$$

Знайдемо скалярний добуток векторів \vec{g} і \vec{a}_0 :

$$(\vec{g}, \vec{a}_0) = \frac{\partial f}{\partial x} \cos \alpha + \frac{\partial f}{\partial y} \sin \alpha.$$

Порівнюючи цю рівність з рівністю (2.2), встановлюємо, що

$$\frac{\partial z}{\partial p} = (\vec{g}, \vec{a}_0). \quad (2.3)$$

Позначимо кут між векторами \vec{g} і \vec{a}_0 через φ , тоді

$$(\vec{g}, \vec{a}_0) = |\vec{g}| \cdot |\vec{a}_0| \cos \varphi = g \cos \varphi .$$

З огляду на цю рівність, перепишемо співвідношення (2.3) у виді

$$\frac{\partial z}{\partial p} = g \cos \varphi .$$

Якщо $\varphi = 0$, тобто якщо напрямок вектора \vec{a}_0 (напрямок диференціювання) збігається з напрямком градієнта \vec{g} , то $\cos \varphi$ досягає свого найбільшого значення (рівного 1) і

$$\left(\frac{\partial z}{\partial p} \right)_{\max} = g . \quad (2.4)$$

Таким чином, похідна функції $z = f(x, y)$ досягає найбільшого значення по напрямку градієнта, що і свідчить про те, що в напрямку градієнта функція z зростає з найбільшою швидкістю. З формули (3.4) видно також, що найбільша швидкість зростання функції z дорівнює модулю градієнта [10].

Приклад. Знайти напрямок найшвидшого зростання функції $z = x^2 + xy + 5$ в точці $M_0(1; -1)$ і обчислити значення похідної в цьому напрямку.

Знаходимо координати градієнта даної функції: $dz/dx = 2x + y$, $dz/dy = x$. Отже, $\nabla f(x, y) = (2x + y; x)$. У точці $M_0(1; -1)$ градієнт має координати $\nabla f(1; -1) = (1; 1)$. По координатах градієнта видно, що шуканий напрямок диференціювання складає кут 45° з віссю Ox . Значення похідної в цьому напрямку

$$\left(\frac{\partial z}{\partial p} \right)_{\max} = |\nabla f(1; -1)| = \sqrt{1^2 + 1^2} = \sqrt{2} .$$

Теорема 2.2. Градієнт функції $z = f(x, y)$ у кожній точці $M_0(x; y)$ спрямований пб нормалі до лінії рівня поверхні $z = f(x, y)$, що проходить через цю точку.

Доведення. Нехай рівняння лінії рівня I (рис. 2.2), що проходить через точку M_0 , має вид

$$f(x, y) = C \quad (2.5)$$

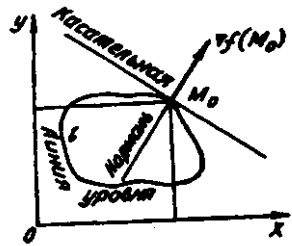


Рис. 2.2

Щоб знайти кутовий коефіцієнт k дотичної в точці M_0 до цієї лінії, продиференціюємо рівність (2.5) як неявну функцію:

$$\frac{\partial f}{\partial x} \frac{dx}{dx} + \frac{\partial f}{\partial y} \frac{dy}{dx} = 0,$$

звідки

$$\frac{dy}{dx} = k = -\frac{\frac{\partial f}{\partial x}}{\frac{\partial f}{\partial y}} \quad (2.6)$$

Знайдемо кутовий коефіцієнт k_1 прямої, на якій лежить градієнт $\nabla f(M_0) = \left(\frac{\partial f}{\partial x}; \frac{\partial f}{\partial y} \right)$. Очевидно, що

$$k_1 = \frac{\frac{\partial f}{\partial y}}{\frac{\partial f}{\partial x}} \quad (2.7)$$

Порівнюючи рівності (2.6) і (2.7), бачимо, що $kk_1 = -1$, тобто градієнт і дотична взаємно перпендикулярні. Інакше кажучи: градієнт у даній точці спрямований по нормалі до лінії рівня поверхні.

Узагальнюючи поняття градієнта на випадок n -вимірного простору, приймемо наступне означення.

Якщо функція $f(x) = f(x_1, x_2, \dots, x_n)$ диференційована у точці $x_0 = (x_1^0, x_2^0, \dots, x_n^0)$, то градієнтом $f(x)$ у точці x_0 називається n -вимірний вектор, координати якого рівні частинним похідним функції $f(x)$, обчисленим у точці x_0 , тобто

$$\overrightarrow{\nabla f(x_0)} = \left(\frac{\partial f(x_0)}{\partial x_1}, \dots, \frac{\partial f(x_0)}{\partial x_n} \right).$$

Усі властивості градієнта, встановлені в тривимірному просторі, переносяться на n -вимірний простір. |

Вектор $-\overrightarrow{\nabla f(x_0)}$, протилежний градієнту, називається, антиградієнтом. Він указує напрямок найшвидшого спадання функції $f(x)$.

На чудовій властивості градієнта вказувати у кожній точці, у якій він існує, напрямок найшвидшого зростання функції f , засновані градієнтні методи розв'язування задач математичного програмування, зокрема опуклого програмування.

Важливий клас задач математичного програмування складають задачі, моделі яких містять опуклі (увігнуті) функції.

Функція $f(x)$, визначена на опуклій множині X , називається опуклою, якщо для будь-яких точок x' і x'' з цієї множини і будь-якого $0 \leq \lambda \leq 1$ справедлива нерівність [11]

$$f(\lambda x' + (1 - \lambda)x'') \leq \lambda f(x') + (1 - \lambda)f(x''). \quad (2.8)$$

Якщо в співвідношенні (2.8) при $0 < \lambda < 1$ і будь-яких $x', x'' \in X$ ($x' \neq x''$) має місце строга нерівність, то $f(x)$ називається строго опуклою.

Функція $f(x)$, визначена на опуклій множині X , називається увігнутою, якщо для будь-яких точок x' і x'' з цієї множини і будь-якого $0 \leq \lambda \leq 1$ справедлива нерівність

$$f(\lambda x' + (1 - \lambda)x'') \geq \lambda f(x') + (1 - \lambda)f(x''). \quad (2.9)$$

Якщо в співвідношенні (2.9) при $0 < \lambda < 1$ і будь-яких $x', x'' \in X$ ($x' \equiv x''$) має місце строга нерівність, то $f(x)$ називається строго увігнутою.

Геометрична ілюстрація даних визначень для двовимірного простору ясна з рис. 2.3: опукла (рис. 2.3, а) функція $f(x)$ на відрізку $[x'; x'']$ не може приймати більших значень, ніж лінійна функція, що інтерполює значення $f(x')$ і $f(x'')$. У свою чергу, увігнута функція $f(x)$ (мал. 7.3, б) не може приймати менших значень, ніж лінійна функція, що інтерполює значення $f(x')$ і $f(x'')$.

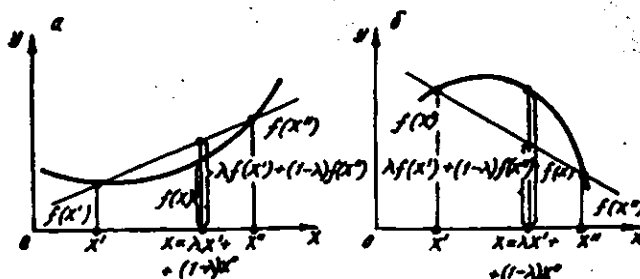


Рис. 2.3

З визначень опуклої та увігнутої функції видно, що властивість опуклості (увігнутості) функцій допускає опуклість множини, на якій задана функція.

Можна довести, що якщо функції $f_i(x)$ ($i = \overline{1, k}$) є опуклими на деякій опуклій множині X , то опуклою на X буде і невід'ємна лінійна комбінація цих функцій, тобто функція $f = \sum_{i=1}^n \lambda_i f_i(x)$ при $\lambda_i \geq 0$, зокрема опуклою буде і сума опуклих функцій.

Теорема 2.3. Якщо $\varphi(x)$ – опукла функція при усіх $x \geq 0$, то буде опуклим і множина розв'язків системи $\varphi(x) \leq b$, $x \geq 0$.

Доведення. Переконаємося, що разом з розв'язками x_1 і x_2 системи $\varphi(x) \leq b$, $x \geq 0$, множині розв'язків цієї системи належить і будь-який вектор

$x_0 = \lambda x_1 + (1 - \lambda)x_2$, де $0 \leq \lambda \leq 1$. Ясно, що $x_0 \geq 0$. Залишається показати, що $\varphi(x_0) \leq b$. По властивості опуклості функції $\varphi(x)$:

$$\varphi(x_0) = \varphi(\lambda x_1 + (1 - \lambda)x_2) \leq \lambda \varphi(x_1) + (1 - \lambda)\varphi(x_2) . \quad (2.10)$$

Але $\varphi(x_1) \leq b$, $\varphi(x_2) \leq b$, тому для кожного $0 \leq \lambda \leq 1$ справедливі $\lambda \varphi(x_1) \leq \lambda b$ и $(1 - \lambda)\varphi(x_2) \leq (1 - \lambda)b$. Додаючи ці нерівності, одержуємо $\lambda \varphi(x_1) + (1 - \lambda)\varphi(x_2) \leq \lambda b + (1 - \lambda)b = b$. З урахуванням співвідношення (2.10) приходимо до нерівності $\varphi(x_0) \leq b$.

Аналогічна теорема доводиться і для увігнутих функцій.

Можна показати, що множина розв'язків системи нерівностей $x \leq 0$, $\varphi_i(x) \{ \leq; \geq \} b_i$ ($i = \overline{1, m}$) є опуклою, якщо в нерівностях зі знаком $\leq \varphi_i$ – опуклі функції, а в нерівностях зі знаком $\geq \varphi_i$ – увігнуті функції для $x \geq 0$.

Доводиться, що опукла функція $f(x)$, визначена на опуклій множині X , неперервна в будь-якій внутрішній точці цієї множини [19].

Теорема 2.4. Якщо опукла функція $f(x)$ диференційована у внутрішніх точках множини X , то для будь-яких внутрішніх точках x_1 і x_2 має місце нерівність

$$f(x_2) - f(x_1) \geq \nabla f(x_1)(x_2 - x_1) . \quad (2.11)$$

Доведення. Оскільки $f(x)$ — опукла функція, то

$$f((1 - \lambda)x_1 + \lambda x_2) \leq (1 - \lambda)f(x_1) + \lambda f(x_2) ,$$

звідки

$$(f(x_1 + \lambda(x_2 - x_1)) - f(x_1)) / \lambda \leq f(x_2) - f(x_1) . \quad (2.12)$$

Відомо, що якщо функція $f(x)$ однієї змінної x в проміжку від x_0 до $x_0 + \Delta x$ неперервна і має неперервну похідну, то справедлива наступна формула Тейлора:

$$f(x_0 + \Delta x) = f(x_0) + \Delta x \frac{df(x_0 + \theta \Delta x)}{dx} \quad (0 < \theta < 1).$$

У нашому випадку мова йде про функцію $f(x)$, що залежить від n змінних, тому формула Тейлора у векторному записі приймає наступний вид:

$$f(x_0 + \Delta x) = f(x_0) + \Delta x \nabla f(x_0 + \theta \Delta x) \quad (0 < \theta < 1) .$$

Використовуючи її при $x_0 = x_1$, $\Delta x = \lambda(x_2 - x_1)$, будемо мати

$$f(x_1 + \lambda(x_2 - x_1)) = f(x_1) + \lambda(x_2 - x_1) \nabla f(x_1 + \theta \lambda(x_2 - x_1)) . \quad (2.13)$$

Підставляючи (2.13) у нерівність (2.12), знаходимо

$$(x_2 - x_1) \nabla f(x_1 + \theta \lambda(x_2 - x_1)) \leq f(x_2) - f(x_1) .$$

Переходячи до границі при $\lambda \rightarrow 0$ в цій нерівності, одержуємо співвідношення (2.11).

Теорема 2.5. Опукла функція $f(x)$, визначена на опуклій множині X , досягає свого глобального мінімуму в кожній точці x , у якій градієнт функції перетворюється в нуль.

Доведення. Нехай, наприклад, у точці x_0 градієнт $\nabla f(x_0) = 0$. Покажемо, що в цій точці функція $f(x)$ має глобальний мінімум, тобто що для будь-якої точки $x \in X$ справедлива нерівність $f(x) \geq f(x_0)$. Відомо, що для опуклої множини X функції $f(x)$ справедлива нерівність (2.11), де x_1 і x_2 — будь-які внутрішні точки множини X . Покладаючи в нерівності (2.11) $x_1 = x_0$ і $x_2 = x$, одержуємо $f(x) - f(x_0) \geq \nabla f(x_0)(x - x_0)$, звідки $f(x) \geq f(x_0)$, що і доводить теорему.

Теорема 2.6. Локальний мінімум опуклої функції $f(x)$, визначеної на опуклій множині X , збігається з її глобальним мінімумом на цій множині.

Доведення. Нехай функція $f(x)$ в точці x_0 має локальний мінімум, тобто для всіх точок x з деякого ε -околу точки x_0 справедлива нерівність

$$f(x_0) \leq f(x) . \quad (2.14)$$

Доведемо, що нерівність (2.14) справедлива для будь-якої точки області X . Припустимо супротивне, тобто що в області X існує точка x' , така, що

$$f(x') < f(x_0) , \quad (2.15)$$

і, отже, x_0 не є точкою глобального мінімуму.

Розглянемо відрізок, що з'єднує точки x' і x_0 . Його рівняння має вид $x = \lambda x' + (1 - \lambda)x_0$ ($0 \leq \lambda \leq 1$). Через те що область X опукла, відрізок $x'x_0$ цілком лежить у ній. Візьмемо на ньому точку x'' , що лежить у ε -околі точки x_0 , але яка не співпадає з точкою x_0 . Оскільки точка x'' розташована на відрізку $x'x_0$, їй відповідає деяка визначена величина $\lambda = \lambda''$ і виконується умова $x'' = \lambda'' x' + (1 - \lambda'')x_0$ ($0 < \lambda'' < 1$). У силу опуклості функції $f(x)$ для точок x' і x_0 має місце $f(\lambda'' x' + (1 - \lambda'')x_0) \leq \lambda'' f(x') + (1 - \lambda'')f(x_0)$ чи нерівність

$$f(x'') \leq f(x_0) + \lambda''(f(x') - f(x_0)) . \quad (2.16)$$

Але по припущенню виконується нерівність (2.15), а тому в співвідношенні (2.16) $\lambda''(f(x') - f(x_0)) < 0$. Відкидаючи цей доданок, ми можемо тільки підсилити нерівність (2.16), отже одержуємо

$$f(x'') < f(x_0) . \quad (2.17)$$

Однак нерівність (2.17) суперечить умові (2.14) для точок ε -околу точки x_0 . Виходить, у області X немає точок типу x' , для яких виконується умова (2.15). Отже, точка x_0 дійсно є точкою глобального мінімуму.

Для практичної перевірки функції на опуклість може виявитися корисним наступна властивість опуклих функцій: функція $f(x)$ буде опуклою, якщо її другі частинні похідні утворять матрицю, у якій усі головні мінори невід'ємні [3, 22].

2.2. Задача опуклого програмування

Задача математичного програмування

$$\max(\min) z = f(x) ; \quad (2.18)$$

$$\left. \begin{array}{l} \varphi_i(x) \{ \leq, =, \geq \} b_i \quad (i = \overline{1, m}), \\ x \geq 0, \quad x = (x_1, \dots, x_n), \end{array} \right\} \quad (2.19)$$

у який або цільова функція (2.18), або обмеження (2.19), або і те й інше - нелінійні, називається нелінійною.

Нелінійні задачі складають широкий клас настільки складних задач, що дотепер не вдається розробити загальні методи, подібні симплекс-методу в лінійному програмуванні, що дозволяли б розв'язувати будь-які нелінійні задачі. Але, незважаючи на відсутність універсальних методів, розроблені способи розв'язування окремих спеціальних класів задач, і насамперед задач з опуклими (увігнутими) функціями $f(x)$ і $\varphi_i(x)$.

У теорії опуклого програмування, в якості основної, зазвичай розглядається задача мінімізації опуклої функції n змінних $z = f(x)$ при обмеженнях $\varphi_i(x) \leq 0$ ($i = \overline{1, m}$), $x \geq 0$, де функції $\varphi_i(x)$ передбачаються опуклими. Опуклість множини допустимих розв'язків задачі впливає з теореми 2.3.

Якщо $f(x)$ і $\varphi_i(x)$ є увігнутими функціями, то маємо задачу максимізації $f(x)$ при обмеженнях $\varphi_i(x) \geq 0$ ($i = \overline{1, m}$), $x \geq 0$ [8, 15].

2.3. Метод множників Лагранжа

Метод множників Лагранжа є класичним методом розв'язування задач математичного програмування (зокрема опуклого). На жаль, при практичному застосуванні методу можуть зустрітися значні обчислювальні труднощі, що звужують області його використання.

Ми розглядаємо тут метод Лагранжа головним чином тому, що він є апаратом, активно використовуваним для обґрунтування різних сучасних чисельних методів, широко застосовуваних на практиці. Що ж стосується функції Лагранжа і множників Лагранжа, то вони відіграють самостійну і винятково важливу роль у теорії і додатках не тільки математичного програмування. Розглянемо класичну задачу оптимізації

$$\max(\min) z = f(x); \quad (2.20)$$

$$\varphi_i(x) = b_i \quad (i = \overline{1, m}), \quad x = (x_1, \dots, x_n). \quad (2.21)$$

Ця задача відрізняється від задачі (2.18), (2.19) тим, що серед обмежень (2.21) немає нерівностей, немає умов незаперечності змінних їхньої дискретності, $m < n$ і функції $f(x)$ і $\varphi_i(x)$ неперервні та мають частинні похідні принаймні другого порядку.

Класичний підхід до розв'язування задачі (2.20), (2.21) дає систему рівнянь (необхідні умови), яким повинна задовольняти точка x^* , у якій функція $f(x)$ набуває локального екстремуму на множині точок, що задовольняють обмеженням (2.21) (для задачі опуклого програмування знайдена точка x^* відповідно до теореми 2.6 буде одночасно і точкою глобального екстремуму).

Припустимо, що в точці x^* функція (2.20) має локальний умовний екстремум і ранг матриці $[d\varphi_i/dx_i]_{m \times n}$ дорівнює m . Тоді необхідні умови запишуться у виді:

$$\left. \begin{aligned} \frac{\partial L}{\partial x_j} &= \frac{\partial f}{\partial x_j} - \sum_{i=1}^m \lambda_i \frac{\partial \varphi_i}{\partial x_j} = 0 \quad (j = \overline{1, n}), \\ \frac{\partial L}{\partial \lambda_i} &= b_i - \varphi_i = 0 \quad (i = \overline{1, m}) \end{aligned} \right\}$$

де

$$L(x_1, \dots, x_n, \lambda_1, \dots, \lambda_m) = f(x_1, \dots, x_n) + \sum_{i=1}^m \lambda_i (b_i - \varphi_i(x_1, \dots, x_n)) \quad (2.23)$$

є функція Лагранжа; $\lambda_1, \dots, \lambda_m$ — множники Лагранжа [4, 11].

Існують також і достатні умови, при виконанні яких розв'язок системи рівнянь (2.22) визначає точку екстремуму функції $f(x)$. Це питання зважається на підставі дослідження знаку другого диференціала функції Лагранжа. Однак достатні умови представляють головним чином теоретичний інтерес.

Можна вказати наступний порядок розв'язування задачі (2.20), (2.21) методом множників Лагранжа:

1) скласти функцію Лагранжа (2.23);

2) знайти частинні похідні функції Лагранжа по всіх змінних $x_1, \dots, x_n, \lambda_1, \dots, \lambda_m$ і прирівняти їх до нуля. Тим самим буде отримана система (2.22), що складається з $n+m$ рівнянь. Розв'язати отриману систему (якщо це виявиться можливим) і знайти в такий спосіб усі стаціонарні точки функції Лагранжа;

3) зі стаціонарних точок, узятих без координат $\lambda_1, \dots, \lambda_m$, вибрати точки, у яких функція $f(x)$ має умовні локальні екстремум при наявності обмежень (7.21). Цей вибір здійснюється, наприклад, із застосуванням достатніх умов локального екстремуму. Часто дослідження спрощується, якщо використовувати конкретні умови задачі.

Ми розглянули метод Лагранжа стосовно до класичної задачі оптимізації. Можна узагальнити цей метод на випадок, коли змінні невід'ємні і деякі обмеження задані у формі нерівностей. Однак це узагальнення має переважно теоретичне значення і не приводить до конкретних обчислювальних алгоритмів.

На закінчення дамо множникам Лагранжа економічну інтерпретацію. Для цього звернемося до найпростішої класичної задачі оптимізації

$$\max(\min) z = f(x_1, x_2); \quad (2.24)$$

$$\varphi(x_1, x_2) = b \quad (2.25)$$

Припустимо, що умовний екстремум досягається в точці $x^* = (x_1^*, x_2^*)$. Відповідне екстремальне значення функції $f(x)$

$$f^* = f(x_1^*, x_2^*).$$

Припустимо, що в обмеженнях (2.25) величина b може мінятися. Тоді координати x_1^* і x_2^* точки екстремуму, а отже, і екстремальне значення f^*

функції $f(x)$ стануть величинами, що залежать від b , тобто $x_1^* = x_1^*(b)$, $x_2^* = x_2^*(b)$, $f^* = f(x_1^*(b), x_2^*(b))$. Тому похідна функції (2.24)

$$df^* / db = \partial f / \partial x_1 \cdot dx_1^* / db + \partial f / \partial x_2 \cdot dx_2^* / db . \quad (2.26)$$

З іншого боку, у силу рівності (2.25) $\varphi(x_1^*(b), x_2^*(b)) = b$, звідки після диференціювання маємо

$$\partial \varphi / \partial x_1 \cdot dx_1^* / db + \partial \varphi / \partial x_2 \cdot dx_2^* / db = 1 . \quad (2.27)$$

Крім того, у точці екстремуму x^* виконуються необхідні умови (2.22). З цих рівностей для $n=2$ й $m=1$ одержуємо:

$$\partial f / \partial x_1 = \lambda \partial \varphi / \partial x_1, \quad \partial f / \partial x_2 = \lambda \partial \varphi / \partial x_2 . \quad (2.28)$$

Підставляючи (2.28) у рівність (2.26) і з огляду на співвідношення (2.27),

$$df^* / db = \lambda \partial \varphi / \partial x_1 \cdot dx_1^* / db + \lambda \partial \varphi / \partial x_2 \cdot dx_2^* / db = \lambda \cdot 1$$

чи

$$df^* / db = \lambda .$$

Для задачі (2.20), (2.21) аналогічно одержуємо

$$df^* / db = \lambda_i \quad (i = \overline{1, m}).$$

Якщо f інтерпретувати як дохід чи вартість, а b_i — як обсяги деяких ресурсів, то множники Лагранжа λ_i показують, як зміниться максимальний дохід (чи мінімальна вартість), якщо кількість ресурсу i -го виду збільшиться на одиницю [22, 23].

2.4. Градієнтні методи

Градiєнтним методом можна розв'язувати, узагалі кажучи, будь-яку нелінійну задачу. Однак при цьому знаходиться лише локальний екстремум.

Тому доцільніше застосовувати цей метод при розв'язуванні задач опуклого програмування, у яких будь-який локальний екстремум є одночасно і глобальним.

Будемо розглядати задачу максимізації нелінійної диференційованої функції $f(x)$. Суть градієнтного пошуку точки максимуму x^* дуже проста: треба взяти довільну точку x_0 і за допомогою градієнта $\nabla f(x_0)$, обчисленого в цій точці, визначити напрямок, у якому $f(x)$ зростає з найбільшою швидкістю (рис. 2.4), а потім, зробивши невеликий крок у знайденому напрямку, перейти в нову точку x_1 . Потім знову визначити найкращий напрямок $\nabla f(x_1)$ для переходу в чергову точку x_2 і т.д. На рис. 2.4 пошукова траєкторія являє собою ламану $x_0x_1x_2 \dots$. Таким чином, треба побудувати послідовність точок $x_0, x_1, x_2, \dots, x_k, \dots$ так, щоб вона збігалася до точки максимуму x^* , тобто для точок послідовності виконувалися умови $f_0 < f_1 < f_2 < \dots < f_k < \dots$

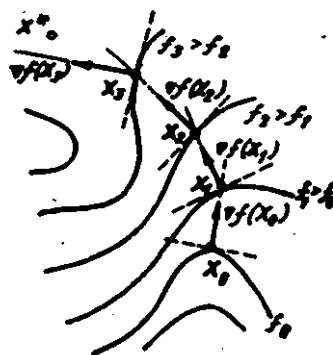


Рис. 2.4

Градієнтні методи, як правило, дозволяють одержувати точний розв'язок за нескінченне число кроків і тільки в деяких випадках — за скінченне. У зв'язку з цим градієнтні методи відносять до наближених методів розв'язування [15].

Рух із точки x_k в нову точку x_{k+1} здійснюється по прямій, що проходить через точку x_k і має рівняння

$$x = x_k + \lambda_k \nabla f(x_k) , \quad (2.29)$$

де λ_k — числовий параметр, від якого залежить величина кроку. Як тільки значення параметра в рівнянні (2.29) обрано: $\lambda_k = \lambda_k^0$, стає визначеною чергова точка $x_{k+1} = x_k + \lambda_k^0 \nabla f(x_k)$ на пошуковій ламаній.

Градiєнтні методи відрізняються один від одного способом вибору величини кроку — значення λ_k^0 параметра λ_k . Можна, наприклад, рухатися з точки в точку з постійним кроком $\lambda_k = \lambda$, тобто при будь-якому k

$$x_{k+1} = x_k + \lambda \nabla f(x_k) .$$

Якщо при цьому виявиться, що $f(x_{k+1}) < f(x_k)$, то варто повернутися в точку x_k і зменшити значення параметра, наприклад до $\lambda/2$.

Іноді величина кроку береться пропорційною модулю градієнта.

Якщо шукається наближений розв'язок, то пошук можна припинити, ґрунтуючись на наступних міркуваннях. Після кожної серії з визначеного числа кроків порівнюють досягнуті значення цільової функції $f(x)$. Якщо після чергової серії зміна $f(x)$ не перевищує якогось наперед заданого малого числа δ , пошук припиняють і досягнуте значення $f(x)$ розглядають як шуканий наближений максимум, а відповідне йому x приймають за x^* .

Якщо цільова функція $f(x)$ увігнута (опукла), то необхідною і достатньою умовою оптимальності точки x^* є рівність нулю градієнта функції в цій точці.

Розповсюдженим є варіант градієнтного пошуку, який називають методом найшвидшого підйому. Суть його в наступному. Після визначення градієнта $\nabla f(x_k)$ в точці x_k рух уздовж прямої $x = x_k + \lambda_k \nabla f(x_k)$ виробляється до точки x_{k+1} , у якій досягається максимальне значення функції $f(x)$ в напрямку градієнта $\nabla f(x_k)$. Потім у цій точці знову визначається градієнт, і рух відбувається по прямій $x = x_{k+1} + \lambda_{k+1} \nabla f(x_{k+1})$ у напрямку нового градієнта $\nabla f(x_{k+1})$ до точки x_{k+2} в якій досягається максимальне в цьому напрямку

значення $f(x)$. Рух продовжується доти, поки не буде досягнута точка x^* , що відповідає найбільшому значенню цільової функції $f(x)$. На рис. 2.5 наведена схема руху до оптимальної точки x^* методом найшвидшого підйому. У даному випадку напрямок градієнта $\nabla f(x_k)$ в точці x_k є дотичним до лінії рівня поверхні $f(x)$ в точці x_{k+1} , отже, градієнт $\nabla f(x_{k+1})$ у точці x_{k+1} ортогональний градієнту $\nabla f(x_k)$ [4, 17].

Переміщення з точки x_k в точку $x_{k+1} = x_k + \lambda_k \nabla f(x_k)$ супроводжується зростанням функції $f(x)$ на величину

$$\Delta f(x) = f(x_{k+1}) - f(x_k) = f(x_{k+1,1}; \dots; x_{k+1,n}) - f(x_{k1}; \dots; x_{kn}) = f(x_{k1} + \lambda_k \partial f(x_k) / \partial x_1; \dots; x_{kn} + \lambda_k \partial f(x_k) / \partial x_n) - f(x_{k1}; \dots; x_{kn}). \quad (2.30)$$

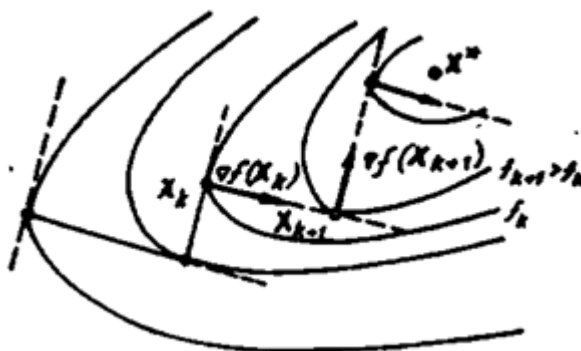


Рис. 2.5

З виразу (2.30) видно, що збільшення Δf є функцією змінної λ_k , тобто $\Delta f = \Delta f(\lambda_k)$. При перебуванні максимуму функції $f(x)$ в напрямку градієнта $\nabla f(x_k)$ необхідно вибирати крок переміщення (множник λ_k), що забезпечує найбільше зростання збільшенню функції, а саме $\Delta f(\lambda_k)$. Величина λ_k , при якій досягається найбільше значення $\Delta f(\lambda_k)$, може бути визначена з необхідної умови екстремуму функції $\Delta f(\lambda_k)$:

$$d(\Delta f(\lambda_k)) / d\lambda = 0. \quad (2.31)$$

Знайдемо вираз для похідної, диференціюючи рівність (2.30) по λ_k як складену функцію:

$$\begin{aligned} d(\Delta f(\lambda_k)) / d\lambda_k &= \partial f(x_{k+1}) / \partial x_1 \cdot \partial f(x_k) / \partial x_1 + \dots + \partial f(x_{k+1}) / \partial x_n \cdot \partial f(x_k) / \partial x_n = \\ &= \nabla f(x_{k+1}) \cdot \nabla f(x_k). \end{aligned}$$

Підставляючи цей результат у рівність (2.31), одержуємо

$$d(\Delta f(\lambda_k)) / d\lambda_k = \nabla f(x_{k+1}) \cdot \nabla f(x_k) = 0. \quad (2.32)$$

Ця рівність має просте геометричне тлумачення: градієнт у наступній точці x_{k+1} ортогональний градієнту в попередній точці x_k .

Відразу ж відзначимо, що якщо цільова функція $f(x)$ в задачі з обмеженнями має єдиний екстремум і він знаходиться усередині області, то для пошуку екстремальної точки x^* застосовується викладена вище методика без яких-небудь змін.

Розглянемо задачу опуклого програмування з лінійними обмеженнями:

$$\max z = f(x), \quad (2.33)$$

$$a_i x \leq a_{i0} \quad (i = \overline{1, m}), \quad (2.34)$$

$$x \geq 0, \quad (2.35)$$

де $x = (x_1; \dots; x_n)$; $a_i = (a_{i1}; \dots; a_{in})$.

Передбачається, що $f(x)$ є увігнутою функцією і має неперервні часткові похідні в кожній точці області.

Почнемо з геометричної ілюстрації процесу розв'язування задачі (рис. 2.6). Нехай початкова точка x_0 розташована усередині області. З точки x_0 можна рухатися в напрямку градієнта $\nabla f(x_0)$, поки $f(x)$ не досягне максимуму. У нашому випадку $f(x)$ увесь час зростає, тому зупинитися треба в точці x_1 на граничній прямій.

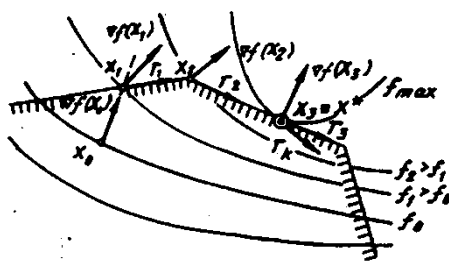


Рис. 2.6

Як видно з рисунка 2.6, далі рухатися в напрямку градієнта $\nabla f(x_1)$ не можна, тому що вийдемо з області. Тому треба знайти інший напрямок переміщення, що, з одного боку, не виводить із припустимої області, а з іншого боку - забезпечує найбільше зростання $f(x)$. Такий напрямок визначить вектор r_1 , що складає з вектором $\nabla f(x_1)$ найменший кут у порівнянні з будь-яким іншим вектором, що виходить із точки x_1 і лежить у області. Аналітично такий вектор знайдеться з умови максимізації скалярного добутку $\nabla f(x_1) \cdot r_1 > 0$. У даному випадку вектор r_1 , що вказує найвигідніший напрямок, збігається з граничною прямою [6].

Таким чином, на наступному кроці рухатися треба по граничній прямій доти, поки зростає $f(x)$; у нашому випадку — до точки x_2 . З рисунка видно, що далі варто переміщатися в напрямку вектора r_2 , що знаходиться з умови максимізації скалярного добутку $\nabla f(x_2) \cdot r_2 > 0$, тобто по граничній прямій. Рух закінчується в точці x_3 , оскільки в цій точці завершується оптимізаційний пошук, тому що в ній функція $f(x)$ має локальний максимум. Через увігнутість у цій точці $f(x)$ досягає також глобального максимуму в області. Градієнт у точці максимуму $x_3 = x^*$ складає тупий кут з будь-яким вектором r_k із області, що проходить через x_3 , тому скалярний добуток $\nabla f(x_3) \cdot r_k$ буде від'ємним для будь-якого припустимого r_k , крім r_3 , спрямованого по граничній прямій. Для нього скалярний добуток $\nabla f(x_3) \cdot r_3 = 0$, тому що $\nabla f(x_3)$ і r_3 взаємно перпендикулярні (гранична пряма стосується лінії рівня поверхні $f(x)$, що проходить через точку максимуму x^*). Ця рівність і служить аналітичною ознакою того, що в точці x_3 функція $f(x)$ досягла максимуму.

Розглянемо тепер аналітичний розв'язок задачі (2.33) — (2.35). Якщо оптимізаційний пошук починається з точки, що лежить у області (всі обмеження задачі виконуються як строгі нерівності), то переміщатися слід по напрямку градієнта так, як встановлено вище. Однак тепер вибір λ_k у рівнянні (2.29) ускладнюється вимогою, щоб чергова точка $x_{k+1} = x_k + \lambda_k \nabla f(x_k)$

залишалася в області. Це означає, що її координати повинні задовольняти обмеженням (2.34), (2.35), тобто повинні виконуватися нерівності:

$$\left. \begin{aligned} a_i(x_k + \lambda_k \nabla f(x_k)) &\leq a_{i0} \quad (i = \overline{1, m}) \\ x_k + \lambda_k \nabla f(x_k) &\geq 0 \end{aligned} \right\}$$

Розв'язуючи систему лінійних нерівностей (2.36), знаходимо відрізок $[\lambda'_k; \lambda''_k]$ припустимих значень параметра λ_k , при яких точка x_{k+1} буде належати області.

Значення λ_k^* , обумовлене в результаті розв'язування рівняння (2.32): $\nabla f(x_k + \lambda_k \nabla f(x_k)) \cdot \nabla f(x_k) = 0$, при якому $f(x)$ має локальний максимум по λ_k в напрямку $\nabla f(x_k)$, повинне належати відрітку $[\lambda'_k; \lambda''_k]$. Якщо ж знайдене значення λ_k виходить за межі зазначеного відрізка, то в якості λ_k^* приймається λ''_k . У цьому випадку чергова точка шуканої траєкторії виявляється на граничній гіперплощині, що відповідає тій нерівності системи (2.36), по якій при розв'язуванні системи отримана права кінцева точка λ''_k відрізка припустимих значень параметра λ_k .

Якщо оптимізаційний пошук початий із точки, що лежить на граничній гіперплощині, чи чергова точка пошукової траєкторії виявилася на граничній гіперплощині, то для продовження руху до точки максимуму насамперед необхідно знайти найкращий напрямок руху. З цією метою варто розв'язати допоміжну задачу математичного програмування, а саме: максимізувати функцію

$$T_k = \nabla f(x_k) \cdot r_k \quad (2.37)$$

при обмеженнях

$$a_i \cdot r_k \leq 0 \quad (2.38)$$

для тих i , при яких

$$a_i \cdot x_k = a_{i0}, \quad |r_k| = 1 \quad (2.39)$$

де $r_k = (r_{k1}; \dots; r_{kn})$; $|r_k| = \sqrt{\sum_{j=1}^n r_{kj}^2}$.

У результаті розв'язування задачі (2.37) — (2.40) буде знайдений вектор r_k , що утворює з градієнтом найменший гострий кут.

Умова (2.39) говорить про те, що точка x_k належить межі області, а умова (2.38) означає, що переміщення з x_k по векторі r_k буде спрямовано усередину області чи по її межі. Умова нормалізації (2.40) необхідна для обмеження величини r_k , тому що в протилежному випадку значення цільової функції (2.37) можна зробити як завгодно великим. Відомі різні форми умов нормалізації, і в залежності з від цього задача (2.37) — (2.40) може бути лінійною чи нелінійною [11, 15].

Після визначення напрямку r_k знаходиться значення λ_k^* для наступної точки $x_{k+1} = x_k + \lambda_k^* r_k$ шуканої траєкторії. При цьому використовується необхідна умова екстремуму у формі, аналогічній рівнянню (2.32), але з заміною $\nabla f(x_k)$ на вектор r_k , тобто

$$\nabla f(x_{k+1}) \cdot r_k = 0. \quad (2.41)$$

Оптимізаційний пошук припиняється, коли досягнута точка x_k^* , у якій $\max T_k = \nabla f(x_k^*) \cdot r_k = 0$.

Якщо в задачах з лінійними обмеженнями рух по граничних прямих виявляється можливим і навіть доцільним, то при нелінійних обмеженнях, що визначають опуклу область, будь-яке як завгодно мале переміщення з граничної точки може відразу вивести за межі області припустимих рішень, і виникне необхідність у поверненні в припустиму область (рис. 2.7). Подібна ситуація характерна для задач, у яких екстремум функції $f(x)$ досягається на межі області. У зв'язку з цим застосовуються різні способи переміщення, що забезпечують побудову послідовності точок, розташованих поблизу границі й усередині припустимої області, чи зигзагоподібний рух уздовж межі з перетинанням останньої. Як видно з рисунка, повернення з точки x_1 в припустиму область варто здійснювати уздовж градієнта тієї граничної функції $\varphi_k(x)$ яка виявилася порушеною.

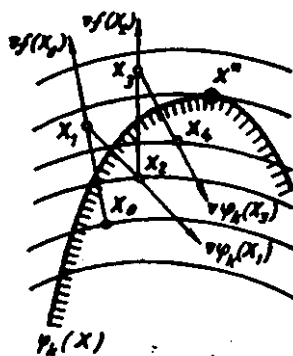


Рис. 2.7

Це забезпечить відхилення чергової точки x_2 у бік точки екстремуму x^* . Ознакою екстремуму в подібному випадку буде колінеарність векторів ∇f і $\nabla \varphi_k$ [1].

2.5. Теорема Куна-Таккера

У теорії нелінійного програмування центральне місце займає теорема Куна — Таккера, що узагальнює класичний метод множників Лагранжа на випадок, коли в нелінійній задачі, крім обмежень-рівностей, містяться також і нерівності. Зокрема, для задачі опуклого програмування: мінімізувати $z = f(x)$ при обмеженнях $\varphi_i(x) \leq 0$ ($i = \overline{1, m}$), $x \geq 0$, де усі функції f і φ_i опуклі, теорема Куна — Таккера встановлює зв'язок між оптимальним розв'язком задачі і сідловою точкою функції Лагранжа для цієї задачі:

$$L(x, \bar{\lambda}) = f(x) + \sum_{i=1}^m \lambda_i \varphi_i(x). \quad (2.42)$$

Точка $(x^*, \bar{\lambda}^*)$ називається сідловою точкою функції (2.42), якщо n -вимірна точка x^* є точкою мінімуму функції $L(x, \bar{\lambda}^*)$, а m -вимірна точка $\bar{\lambda}^*$ — точкою максимуму функції $L(x^*, \bar{\lambda})$, так що для всіх x і $\bar{\lambda}$ виконується нерівність

$$L(x^*, \bar{\lambda}) \leq L(x^*, \bar{\lambda}^*) \leq L(x, \bar{\lambda}^*). \quad (2.43)$$

У цьому зміст теореми Куна — Таккера. Сама ж теорема формулюється в такий спосіб.

Теорема (Куна — Таккера). Припустимо, що існує вектор $x \geq 0$, такий, що $\varphi_i(x) < 0$ ($i = \overline{1, m}$). Тоді необхідною і достатньою умовою оптимальності вектора x^* , що належить припустимій області, є існування такого вектора $\bar{\lambda}^*$, що для усіх $x \geq 0$ і $\bar{\lambda} \geq 0$ має місце нерівність (2.43).

Спочатку ця теорема була доведена тільки для випадку диференційованих функцій $f(x)$ і $\varphi_i(x)$. Узагальнення на випадок опуклих функцій належить Слейтеру.

Якщо функції $f(x)$ і $\varphi_i(x)$ є диференційованими, то нерівності (2.43), де $x^* \geq 0$, $\bar{\lambda}^* \geq 0$, еквівалентні наступним "локальним" умовам Куна — Таккера:

$$\partial L(x^*, \bar{\lambda}^*) / \partial x \geq 0, \quad (2.44)$$

$$x^* \partial L(x^*, \bar{\lambda}^*) / \partial x = 0, \quad (2.45)$$

$$x^* \geq 0, \quad (2.46)$$

$$\partial L(x^*, \bar{\lambda}^*) / \partial \bar{\lambda} \leq 0, \quad (2.47)$$

$$\bar{\lambda}^* \partial L(x^*, \bar{\lambda}^*) / \partial \bar{\lambda} = 0, \quad (2.48)$$

$$\bar{\lambda} \geq 0. \quad (2.49)$$

Покажемо, наприклад, що якщо виконуються умови (2.44) — (2.49), то виконуються й умови (2.43), де $x^* \geq 0$, $\bar{\lambda}^* \geq 0$. При фіксованому $\bar{\lambda}^*$ функція Лагранжа (2.42) $L(x, \bar{\lambda}^*)$ є опуклою по x , а тому відповідно до співвідношення (2.11) має місце нерівність $L(x, \bar{\lambda}^*) - L(x^*, \bar{\lambda}^*) \geq \partial L(x^*, \bar{\lambda}^*) / \partial x \cdot (x - x^*)$, звідки, з урахуванням рівності, (2.45) одержуємо $L(x, \bar{\lambda}^*) \geq L(x^*, \bar{\lambda}^*) + x \cdot \partial L(x^*, \bar{\lambda}^*) / \partial x$.

Оскільки $x \geq 0$, то, опускаючи невід'ємний другий доданок у правій частині, одержуємо лише підсилену нерівність $L(x, \bar{\lambda}^*) \geq L(x^*, \bar{\lambda}^*)$. Таким чином, доведено виконання правої нерівності в співвідношенні (2.43). Аналогічно можна показати, що при $\bar{\lambda} \geq 0$ виконується і ліва нерівність у співвідношенні (2.43).

Так само неважно встановлюється, що з умов (2.43), де $x^* \geq 0$, $\bar{\lambda}^* \geq 0$, випливають умови (2.44) — (2.49) [8, 13, 20].

РОЗДІЛ 3. ПРИКЛАДИ РОЗВ'ЯЗУВАННЯ ЗАДАЧ НЕЛІНІЙНОГО ПРОГРАМУВАННЯ

Розглянемо приклади розв'язування задач нелінійного програмування, що ілюструють теоретичний матеріал, наведений у попередніх розділах.

Приклад 1. Мінімізувати функцію $z = x_1^2 + x_2^2 + x_3^2$ при обмеженні $4x_1 + x_2^2 + 2x_3 - 14 = 0$.

Функція Лагранжа має вигляд

$$L(x, \lambda) = x_1^2 + x_2^2 + x_3^2 - \lambda(4x_1 + x_2^2 + 2x_3 - 14).$$

Звідси отримуємо необхідні умови екстремуму у вигляді системи:

$$\begin{cases} \frac{\partial L}{\partial x_1} = 2x_1 - 4\lambda = 0, \\ \frac{\partial L}{\partial x_2} = 2x_2 - 2\lambda x_2 = 0, \\ \frac{\partial L}{\partial x_3} = 2x_3 - 2\lambda = 0, \\ \frac{\partial L}{\partial \lambda} = -(4x_1 + x_2^2 + 2x_3 - 14) = 0, \end{cases}$$

розв'язками якої є вектори

$$(x_0, \lambda_0)_1 = (2, 2, 1, 1),$$

$$(x_0, \lambda_0)_2 = (2, -2, 1, 1),$$

$$(x_0, \lambda_0)_3 = (2, 8, 0, 1, 4, 1, 4).$$

Використовуючи достатні умови, обчислимо елементи матриці Гессе

$$H^B = \begin{pmatrix} 0 & 4 & 2x_2 & 2 \\ 4 & 2 & 0 & 0 \\ 2x_2 & 0 & 2 - 2\lambda & 0 \\ 2 & 0 & 0 & 2 \end{pmatrix}.$$

Так як $m=1$ і $n=3$, то для того, щоб стаціонарна точка була точкою мінімуму, знак останніх $3-1=2$ кутових мінорів повинен визначатися знаком множника $(-1)^m = -1$. Таким чином, у точці $(x_0, \lambda_0)_1 = (2, 2, 1, 1)$, маємо

$$\begin{vmatrix} 0 & 4 & 4 \\ 4 & 2 & 0 \\ 4 & 0 & 0 \end{vmatrix} = -32 < 0 \quad \text{і} \quad \begin{vmatrix} 0 & 4 & 4 & 2 \\ 4 & 2 & 0 & 0 \\ 4 & 0 & 0 & 0 \\ 2 & 0 & 0 & 2 \end{vmatrix} = -64 < 0.$$

Звідси випливає, що $(x_0)_1$ і $(x_0)_2$ – точки мінімуму. Той факт, що у точці $(x_0)_3$ не задовольняються достатні умови максимуму чи мінімуму, не означає, що ця точка не є екстремальною. Це наслідок того, що умови, які використовуються, є лише достатніми.

Приклад 2. Максимізувати функцію $z = -(2x_1 - 5)^2 - (2x_2 - 1)^2$ при обмеженнях

$$x_1 + 2x_2 \leq 2, \quad (3.1)$$

$$x_1, x_2 \geq 0.$$

Графічне представлення даної задачі зображено на рисунку 3.1.

Як видно, цільова функція задачі є увігнутою, а область допустимих розв'язків – опуклою. Звідси випливає, що ефективний алгоритм повинен гарантувати визначення глобального оптимуму. Проте узагальнений метод множників Лагранжа дозволяє знайти лише точку локального максимуму.

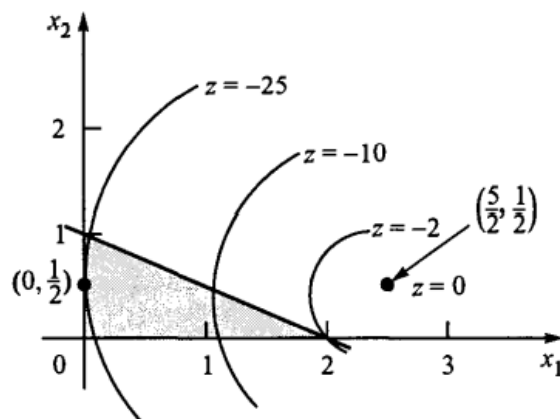


Рис. 3.1

Точку безумовного екстремуму знаходимо як розв'язок системи

$$\begin{cases} \frac{\partial z}{\partial x_1} = -4(2x_1 - 5) = 0, \\ \frac{\partial z}{\partial x_2} = -4(2x_2 - 1) = 0. \end{cases}$$

Звідси $(x_1, x_2) = \left(\frac{5}{2}, \frac{1}{2}\right)$. Так як цей розв'язок не задовольняє умову (3.1), то

обмеження по чергово активізуються. Розглянемо обмеження $x_1 = 0$. Функція Лагранжа у цьому випадку має вигляд

$$L(x_1, x_2, \lambda) = -(2x_1 - 5)^2 - (2x_2 - 1)^2 - \lambda x_1.$$

Звідси

$$\begin{cases} \frac{\partial L}{\partial x_1} = -4(2x_1 - 5) - \lambda = 0, \\ \frac{\partial L}{\partial x_2} = -4(2x_2 - 1) = 0, \\ \frac{\partial L}{\partial \lambda} = -x_1 = 0. \end{cases}$$

Розв'язком цієї системи буде точка $(x_1, x_2) = \left(0, \frac{1}{2}\right)$, яка є точкою максимуму, у чому можна переконатись, використовуючи достатню умову. Оскільки ця точка задовольняє інші обмеження вихідної задачі, обчислювальна процедура закінчується: $(x_1, x_2) = \left(0, \frac{1}{2}\right)$ – точка локального максимуму задачі. Оптимальне значення цільової функції при цьому $z_{\max} = -25$.

Як видно з рисунку 3.1, допустимому розв'язку $(x_1, x_2) = (2, 0)$ розглядуваної задачі, який визначається точкою перетину двох прямих $x_2 = 0$ і $x_1 + 2x_2 = 2$, відповідає значення цільової функції $z = -2$. Воно

краще за значення функції, що отримане з урахуванням одного активного обмеження.

Приклад 3. Знайти оптимальний розподіл обмеженого ресурсу в a одиниць між n споживачами, якщо прибуток, одержуваний при виділенні j -му споживачу x_j одиниць ресурсу, обчислюється по формулі $c_j\sqrt{x_j}$.

Математична модель задачі має наступний вид:

$$\begin{aligned} \max z &= \sum_{j=1}^n c_j \sqrt{x_j}; \\ \sum_{j=1}^n x_j &= a \end{aligned}$$

Складаємо функцію Лагранжа:

$$L(x_1, \dots, x_n, \lambda) = \sum_{j=1}^n c_j \sqrt{x_j} + \lambda(a - \sum_{j=1}^n x_j).$$

Знаходимо частинні похідні функції Лагранжа і прирівнюємо їх до нуля:

$$\left. \begin{aligned} \partial L / \partial x_j &= c_j / 2\sqrt{x_j} - \lambda = 0 \quad (j = \overline{1, n}), \\ \partial L / \partial \lambda &= a - \sum_{j=1}^n x_j = 0 \end{aligned} \right\}$$

Розв'язуючи цю систему рівнянь, отримуємо:

$$x_j^* = ac_j^2 / \sum_{j=1}^n c_j^2 \quad (j = \overline{1, n}), \quad \lambda^* = \sqrt{\sum_{j=1}^n c_j^2 / (4a)}.$$

Таким чином, якщо j -му ($j = \overline{1, n}$) споживачу буде виділено $ac_j^2 / \sum_{j=1}^n c_j^2$ одиниць ресурсу, то сумарний прибуток досягне максимальної величини і складе $\sum_{j=1}^n c_j \sqrt{x_j^*}$ грошових одиниць.

Приклад 4. Визначити максимум функції $f = 4x_1 + 8x_2 - 2x_1^2 - 2x_2^2$, почавши оптимізаційний пошук із точки $x_0 = (5; 10)$.

Будемо супроводжувати розв'язування задачі графічною ілюстрацією (рис. 3.1). По даному рівнянню параболоїда обертання побудовані лінії рівня цієї поверхні, для чого рівняння приведенне до виду $(x_1 - 1)^2 + (x_2 - 2)^2 = 5 - 0,5f$, з якого

ясно, що лініями перетину параболоїда з площинами, рівнобіжними площини x_1Ox_2 (лініями рівня), є кола радіусом $\sqrt{5-0,5f}$. При $f = -150, -100, -50$ їхні радіуси рівні відповідно $\sqrt{80}, \sqrt{55}, \sqrt{30}$, а загальний центр знаходиться в точці $(1; 2)$.

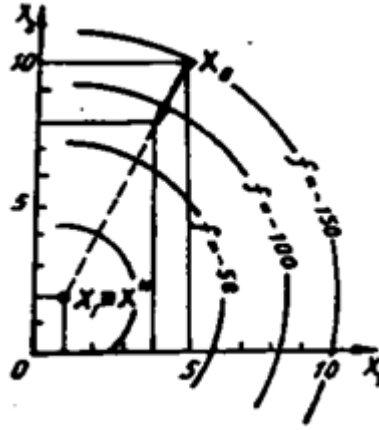


Рис. 3.1

Знаходимо градієнт даної функції:

$$\nabla f(x) = \nabla f(x_1, x_2) = (\partial f / \partial x_1; \partial f / \partial x_2) = (-4x_1 + 4; -4x_2 + 8).$$

1 крок. Обчислюємо:

$$\nabla f(x_0) = \nabla f(5; 10) = (-4 \cdot 5 + 4; -4 \cdot 10 + 8) = (-16; -32).$$

На рис. 3.1 у точці $x_0 = (5; 10)$ побудований вектор $1/16 \nabla f(x_0) = (-1; -2)$, що вказує напрямок найшвидшого зростання функції в точці x_0 . На цьому напрямку розташована наступна точка

$$x_1 = x_0 + \lambda_0 \nabla f(x_0) = (5; 10) + \lambda_0 (-16; -32) = (5 - 16\lambda_0; 10 - 32\lambda_0).$$

У цій точці

$$\nabla f(x_1) = (-4(5 - 16\lambda_0) + 4; -4(10 - 32\lambda_0) + 8) = (-16 + 64\lambda_0; -32 + 128\lambda_0).$$

Використовуючи умову (2.32), одержуємо

$$d(\Delta f(\lambda_0)) / d\lambda_0 = \nabla f(x_1) \cdot \nabla f(x_0) = (-16 + 64\lambda_0; -32 + 128\lambda_0) \cdot (-16; -32) = 0$$

чи $1 - 4\lambda_0 = 0$, звідки $\lambda_0 = 1/4$. Тому що $d^2(\Delta f(\lambda_0)) / d\lambda_0^2 < 0$, то знайдене значення λ_0 є точкою максимуму $\Delta f(x)$. Знаходимо $x_1 = (5 - 16/4; 10 - 32/4) = (1; 2)$.

2 крок. Початкова точка для другого кроку $x_1 = (1; 2)$. Обчислюємо $\nabla f(x_1) = (-4 \cdot 1 + 4; -4 \cdot 2 + 8) = (0; 0)$. Отже, $x_1 = (1; 2)$ є стаціонарною точкою. Але

оскільки дана функція увігнута, то в знайденій точці (1; 2) досягається глобальний максимум $f_{\max} = 10$.

Приклад 5. Максимізувати функцію $f = x_1 + 2x_2 - 0,2x_1^2 - 0,2x_2^2$ при обмеженнях $x_1 + 4x_2 \leq 14$, $7x_1 + 3x_2 \leq 42$, $x_1 \geq 0$, $x_2 \geq 0$.

Для наочного представлення процесу оптимізації будемо супроводжувати його графічною ілюстрацією. На рис. 3.2 зображено кілька ліній рівня даної поверхні і припустима область, у якій варто знайти точку x^* , у якій функція набуває максимального значення.

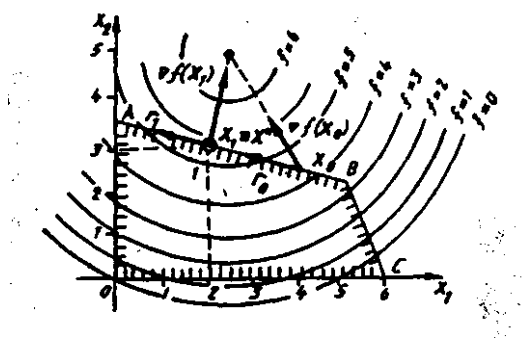


Рис. 3.2

Почнемо оптимізаційний пошук, наприклад, із точки $x_0 = (4; 2,5)$, що лежить на граничній прямій $AB: x_1 + 4x_2 = 14$. При цьому $f(x_0) = 4,55$.

Знайдемо значення градієнта $\nabla f(x) = (1 - 0,4x_1; 2 - 0,4x_2)$ в точці x_0 : $\nabla f(x_0) = (-0,6; 1) \neq (0; 0)$. Крім того, із малюнка видно, що через припустиму область проходять лінії рівня з позначками більш високими, ніж $f(x_0) = 4,55$. Словом, треба шукати напрямок $r_0 = (r_{01}; r_{02})$ переміщення в наступну точку x_1 , більш близьку до оптимальної. З цією метою розв'язуємо задачу (2.37) — (2.40) максимізації функції $T_0 = \nabla f(x_0) \cdot r_0 = (-0,6; 1)(r_{01}; r_{02}) = -0,6r_{01} + r_{02}$ при наступних обмеженнях:

$$a_1 \cdot r_0 = (1; 4)(r_{01}; r_{02}) = r_{01} + 4r_{02} = 0, \quad |r_0| = \sqrt{r_{01}^2 + r_{02}^2} = 1.$$

Оскільки точка x_0 розташовується тільки на одній (першій) граничній прямій ($i=1$): $x_1 + 4x_2 = 14$, то умова (2.38) записується у формі рівності.

Система обмежувальних рівнянь цієї задачі має тільки два розв'язки: $(-0,9700; 0,2425)$ і $(0,9700; -0,2425)$. Безпосередньою підстановкою їх у функцію T_0 встановлюємо, що максимум T_0 відмінний від нуля і досягається при розв'язку

$(0,9700;0,2425)$. Таким чином, переміститися з x_0 потрібно по напрямку вектора $r_0 = (-0,9700;0,2425)$, тобто по граничній прямій ВА.

Два визначення координат наступної точки $x_1 = (x_{11}; x_{12})$:

$$\left. \begin{aligned} x_{11} &= x_{01} + r_{01}\lambda_0 = 4 - 0,9700\lambda_0, \\ x_{12} &= x_{02} + r_{02}\lambda_0 = 2,5 + 0,2425\lambda_0 \end{aligned} \right\}$$

необхідно знайти значення λ_0^* параметра λ_0 , при якому функція $f(x)$ в точці x_1 досягає можливо більшого значення. Але спочатку знайдемо інтервал допустимих значень параметра λ_0 , при яких точка x_1 буде належати припустимій області. Система (2.36) у даному випадку має вид

$$\left. \begin{aligned} 7(4 - 0,9700\lambda_0) + 3(2,5 + 0,2425\lambda_0) &\leq 42, \\ 4 - 0,9700\lambda_0 &\geq 0, \\ 2,5 + 0,2425\lambda_0 &\geq 0. \end{aligned} \right\}$$

Перше обмеження ми опустили, оскільки точка x_1 лежить на прямій АВ. Розв'язуючи систему (2.43), установлюємо, що λ_0 варто шукати в напівінтервалі $(0; 4,1237]$ (помітимо, що нас цікавлять тільки невід'ємні значення параметра λ_0). Знайдемо значення λ_0^* , при якому досягається найбільше зростання приросту Δf функції, викликане переміщенням із точки x_0 в точку x_1 . Відповідно до умови (2.41)

$$\begin{aligned} d(\Delta f)/d\lambda_0 &= \nabla f(x_1) \cdot r_0 = (-0,6 + 0,388\lambda_0; 1 - 0,0970\lambda_0)(-0,9700; 0,2425) = \\ &= -0,3999\lambda_0 + 0,8245 = 0, \end{aligned}$$

звідки $\lambda_0 = 2,0618$. При цьому $d^2(\Delta f)/d\lambda_0^2 = -0,3999 < 0$. Виходить, $\lambda_0^* = 2,0618$. По формулі (2.42) знаходимо координати нової точки: $x_1 = (2; 3)$.

Якщо продовжити оптимізаційний пошук, то при розв'язуванні чергової допоміжної задачі (2.37) — (2.40) буде встановлено, що $T_1 = \nabla f(x_1) \cdot r_1 = 0$, а це говорить про те, що точка x_1 є точкою максимуму x^* цільової функції в припустимій області. Це ж видно і з рис. 3.2: у точці x_1 одна з ліній рівня стосується границі припустимої області. Отже, точка x_1 є точкою максимуму x^* . При цьому $f_{\max} = f(x^*) = 5,4$ [14, 24].

ВИСНОВКИ

У будь-якій галузі діяльності людині повсякчас доводиться стикатись із необхідністю прийняття тих чи інших рішень. На те, щоб розробити математичні методи та засоби обґрунтування таких рішень і спрямована класична теорія оптимізації. Найкраще розроблені та найпростішими у застосуванні є методи лінійного програмування. Однак, розв'язування складних задач приводить до необхідності роботи із нелінійними виразами.

У ході виконання роботи було розглянуто основні поняття класичної теорії оптимізації, формулювання та методи розв'язування нелінійних задач математичного програмування.

Для розв'язування оптимізаційних задач із обмеженнями у вигляді рівностей ефективним є використання методів Якобі та Лагранжа. У роботі також розглянуто узагальнення методу множників Лагранжа на випадок задачі із обмеженнями у вигляді нерівностей.

Велика увага приділена виведенню та використанню умов Куна-Такера, які відіграють важливу роль у нелінійному програмуванні.

Теоретичні основи розглянутих методів розв'язування задач нелінійного моделювання підкріплюється прикладами, наведеними із повним їх розв'язанням у третьому розділі.

Магістерська робота може бути корисна здобувачам вищої освіти та викладачам при вивченні математичного моделювання у ЗВО.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Бакулин М.Г., Варукина Л.А., Крейнделин В.Б. Технология МІМО : принципы и алгоритмы. Москва : Телеком, 2014. 244 с.
2. Бенькович Е.С., Колесов Ю.Б., Сениченков Ю.Б. Практическое моделирование динамических систем. Санкт-Петербург : БХВ-Петербург, 2002. 464 с.
3. Блехман И.И., Мышкис А.Д., Пановко Н.Г., Прикладная математика : С примерами из механики : Учебное пособие. Москва : УРСС, 2006. 376 с.
4. Бублик Б.Н., Кириченко Н.Ф. Основы теории управления. Изд.об. «Высшая школа», 1975. 328с.
5. Бунін С.Г., Войтер А.П., Ільченко М.Е., Романюк В.А. Самоорганізовані мережі. Київ : Наукова думка, 2012. 444 с.
6. Васильев Ф.П. Численные методы решения экстремальных задач. Москва : Наука, 1980. 518 с.
7. Вентцель Е.С. Исследование операций. Москва : Советское радио, 1972. 552 с.
8. Волков И.К., Загоруйко Е.А. Исследование операций : Учеб. для вузов. Москва : Изд-во МГТУ им. Н.Э. Баумана, 2002. 436 с.
9. Галеев Э.М., Тихомиров В.М. Краткий курс теории экстремальных задач. Москва : Изд-во Моск. ун-та, 1989. 204 с.
10. Глоба Л.С. Математичні основи побудови інформаційно-телекомунікаційних систем. Київ : Норіта-плюс, 2007. 360 с.
11. Жалдак М.І., Триус Ю.В. Основы теории і методів оптимізації: Навчальний посібник. Черкаси : Брама-Україна, 2005. 608 с.
12. Зайченко Ю.П. Дослідження операцій : Підручник. 6 вид., перероб. і доп. Київ : Видавничий дім «Слово», 2003. 688 с.
13. Зайченко О.Ю., Зайченко Ю.П. Дослідження операцій. Збірник задач. Київ : Видавничий Дім “Слово”, 2007. 472 с.
14. Згуровський М.З., Панкратова Н.Д. Основы системного аналізу. Київ : Видавнича група ВНУ, 2007. 544с.

15. Ільченко М.Ю., Кравчук С.О. Сучасні телекомунікаційні системи. Київ : НВП «Видавництво «Наукова думка» НАН України», 2008. 328 с.
16. Корн Г., Корн Т. Справочник по математике для научных работников и инженеров. Москва : Наука, 1984. 832с.
17. Костевич Л.С. Математическое программирование : Информ. Технологии оптимальных решений : Учеб. пособие. Минск : Новое знание, 2003. 424 с.
18. Крак Ю.В., Ливошич О.Л. Оптимальне керування: Навч.-метод. посібник для студентів фак.-ту кібернетики спеціальності «Інформатика». Київ : Видавничо-поліграфічний центр «Київський університет», 2003. 106 с
19. Ларіонов Ю.І., Левикін В.М., Хажмурадов М.А. Дослідження операцій в інформаційних системах. Харків : Компанія СМІТ, 2005. 364 с.
20. Самсонов В.В. Алгоритми розв'язання задач оптимізації: Навчальний посібник. Київ : НУХТ, 2014. 300 с.
21. Томашевський В.М. Моделювання систем. Підручник. Київ : Видавнича група ВНУ, 2007. 352 с.
22. Hamdy A. Taha. Operations research: An introduction. New Jersey : Pearson Education, 2003. 903 p.
23. Obretenov A., Dimitrov B. Teoria masowej obsługi. PWN, 1989.
24. Tijms H.C. Stochastic Models - An Algorithmic Approach, Wiley, New York, 1994.