

Одеський національний університет імені І. І. Мечникова
Інститут математики, економіки і механіки

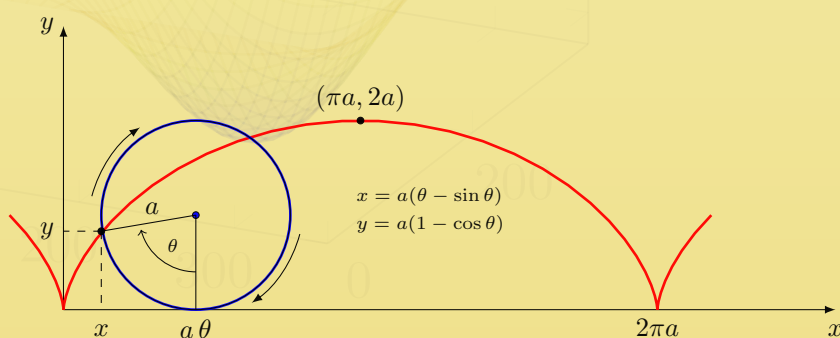
А. Т. Яровий, Є. М. Страхов

МЕТОДИ ОПТИМІЗАЦІЇ

та

ВАРІАЦІЙНЕ ЧИСЛЕННЯ

$$\frac{\partial F(x, y, y')}{\partial y} - \frac{d}{dx} \frac{\partial F(x, y, y')}{\partial y'} = 0$$



$$\dot{x}_1 = x_2$$

$$\dot{x}_2 = u$$

А. Т. Яровий, Є. М. Страхов

**МЕТОДИ ОПТИМІЗАЦІЇ
ТА
ВАРІАЦІЙНЕ ЧИСЛЕННЯ**

**Навчально-методичний посібник
для студентів спеціальності «Математика»**

**Освіта України
2017
Одеса**

УДК 517.97, 519.85

ББК 22.18

Я76

Рекомендовано до друку рішенням Науково-методичної ради ОНУ імені І. І. Мечникова (протокол № 2 від 20.04.2017).

Методи оптимізації та варіаційне числення: навчально-методичний посібник

Автори:

Яровий Анатолій Трохимович, кандидат фіз.-мат. наук, доцент;

Страхов Євген Михайлович, кандидат фіз.-мат. наук.

Рецензенти:

Молчанюк І. В., кандидат фіз.-мат. наук, доцент, завідувач кафедри прикладної, обчислювальної математики і САПР Одеської державної академії будівництва та архітектури;

Трофимов Б. Ф., кандидат технічних наук, доцент кафедри інформаційних систем Одеського національного політехнічного університету;

Таїрова М. С., кандидат фіз.-мат. наук, доцент кафедри обчислювальної математики Одеського національного університету імені І. І. Мечникова.

ISBN 978-617-7366-31-6

Посібник складено на основі курсу лекцій, які читаються студентам спеціальності «Математика» в Одеському національному університеті імені І. І. Мечникова. Розглядаються аналітичні і чисельні методи розв'язування задач нелінійного програмування, теорія опуклого програмування, методи розв'язування найпростіших задач варіаційного числення та оптимального керування.

Посібник буде корисний студентам математичних напрямів, які вивчають дисципліни «Методи оптимізації», «Варіаційне числення», «Теорія керування», а також усім, хто цікавиться застосуванням математичних методів оптимізації у різноманітних галузях людської діяльності.

© Яровий А. Т., Страхов Є. М.

Зміст

Вступ	5
МІНІМІЗАЦІЯ ФУНКЦІЙ	7
1 Мінімізація функцій однієї змінної	8
1.1 Необхідні та достатні умови екстремуму	8
1.2 Чисельні методи одновимірної мінімізації	10
1.2.1 Симетричні методи	11
1.3 Завдання для самостійної роботи	15
1.3.1 Чисельні методи одновимірної мінімізації	15
2 Мінімізація функцій багатьох змінних	17
2.1 Допоміжні означення та теореми	17
2.2 Постановка задачі математичного програмування	24
2.2.1 Збіжність в екстремальних задачах. Існування екстремумів	26
2.3 Мінімізація функцій без обмежень	27
2.4 Задача на умовний мінімум	29
2.4.1 Мінімізація функцій при обмеженнях типу рівностей	29
2.4.2 Мінімізація функцій при обмеженнях типу нерівностей	33
2.5 Мінімізація опуклих функцій	36
2.5.1 Задача лінійного програмування	40
2.6 Завдання для самостійної роботи	42
2.6.1 Задачі нелінійного програмування. Аналітичні методи	42
2.6.2 Лінійне програмування. Симплекс-метод	53
2.6.3 Розв'язування транспортних задач	69
3 Чисельні методи мінімізації функцій багатьох змінних	73
3.1 Методи безумовної мінімізації	73

3.1.1	Метод нульового порядку Нелдера — Міда	73
3.1.2	Методи першого порядку	76
3.1.3	Методи другого порядку	86
3.1.4	Квазіньютонівські методи (методи змінної метрики)	92
3.2	Мінімізація функцій з обмеженнями	96
3.2.1	Метод проекції градієнта	96
3.2.2	Метод умовного градієнта	98
3.2.3	Методи штрафних функцій	100
3.3	Завдання для самостійної роботи	104
3.3.1	Мінімізація функцій багатьох змінних без обмежень	104

МІНІМІЗАЦІЯ ФУНКЦІОНАЛІВ 107

4	Варіаційне числення	108
4.1	Основна задача варіаційного числення. Рівняння Ейлера	108
4.1.1	Часткові випадки рівняння Ейлера	113
4.2	Умови мінімуму в класі кусково-гладких функцій	116
4.3	Дослідження другої варіації. Умови Лежандра та Якобі	117
4.4	Ізопериметрична задача	121
4.5	Задача з рухомими межами	122
4.6	Завдання для самостійної роботи	124
4.6.1	Основна задача варіаційного числення	124
4.6.2	Ізопериметрична задача	126
4.6.3	Варіаційна задача з рухомими межами	128
5	Задача оптимального керування	131
5.1	Постановка основних задач оптимального керування	131
5.2	Приріст критерію якості на припустимих керуваннях	132
5.2.1	Оцінка приросту припустимої траєкторії на голчатій варіації	134
5.3	Принцип максимуму Понтрягіна	135
5.4	Оптимізація лінійних систем. Керованість системи	138
5.5	Достатня умова оптимальності	140
5.6	Синтез лінійних систем, оптимальних по швидкодії	142
5.7	Завдання для самостійної роботи	144
5.7.1	Побудова крайової задачі принципу максимуму	144
5.7.2	Задача синтезу оптимального керування	149
	Рекомендована література	152

Вступ

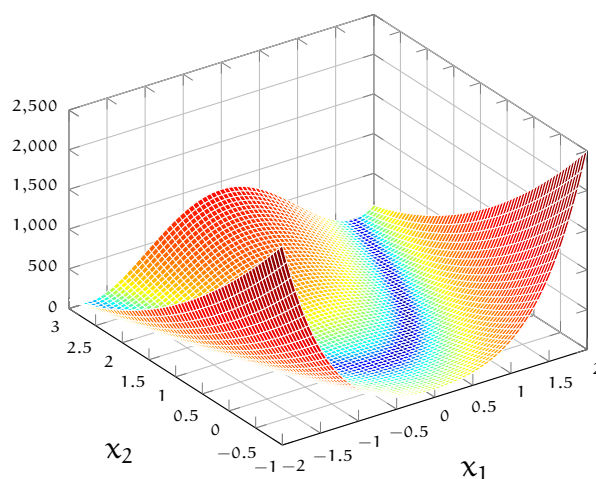
Темпи розвитку економіки, розв’язання багатьох соціальних проблем залежать від інтенсивності впровадження досягнень науково-технічного прогресу в галузях народного господарства. У свою чергу, цю проблему неможливо розв’язати без швидкого розвитку і впровадження в усі сфери людської діяльності сучасних засобів обчислювальної техніки і прикладної математики.

Одним з розділів прикладної математики, до якого інженерно-технічні працівники і економісти проявляють підвищений інтерес, є мінімізація функцій і функціоналів. Велика кількість різноманітних задач і методів їх розв’язання обумовлює мати посібники, в яких у стислій формі було б викладено алгоритми найвідоміших методів і методику їх застосування. До цього спонукають також нові форми навчання студентів.

Основне завдання посібника — допомога студентам в опануванні базових алгоритмів розв’язування задач мінімізації функцій багатьох змінних, задач варіаційного числення і оптимального керування.

Частина I

МІНІМІЗАЦІЯ ФУНКЦІЙ



Мінімізація функцій однієї змінної

1.1 Необхідні та достатні умови екстремуму

Нехай на числовій прямій E^1 задана скалярна функція $\varphi(x)$. Розглянемо задачу пошуку точок, в яких функція досягає свого мінімального або максимального значення. Точки мінімуму або максимуму функції будемо називати екстремальними точками.

Сформулюємо означення, що відносяться до теорії мінімізації функцій.

Означення 1.1. Якщо для всіх $x \in E^1$ виконується умова $\varphi(x^*) \leq \varphi(x)$, то точка x^* називається точкою **глобального (абсолютного) мінімуму** функції $\varphi(x)$.

Означення 1.2. Якщо для достатньо малого $\varepsilon > 0$ виконується нерівність $\varphi(x^*) \leq \varphi(x)$ для всіх $x \in E^1$ таких, що $|x - x^*| \leq \varepsilon$, то точка x^* називається точкою **локального (відносного) мінімуму** функції $\varphi(x)$.

Означення 1.3. Точка x^* називається точкою **строгого мінімуму** (у локальному або глобальному сенсі), якщо відповідні нерівності в означеннях точок локального і глобального мінімумів виконуються як строги (при $x \neq x^*$).

Означення точок локального і глобального максимумів вводяться аналогічно.

Зазначимо, що точки глобального мінімуму є точками локального мінімуму. І тому далі розглядатимемо тільки точки локального мінімуму.

Теорема 1.1 (необхідна умова екстремуму першого порядку). Нехай функція $\varphi(x)$ визначена і диференційовна на E^1 . Якщо x^* — точка локального мінімуму (максимуму), то у ній перша похідна функції дорівнює нулю:

$$\frac{d\varphi(x^*)}{dx} = 0. \quad (1.1)$$

Означення 1.4. Точки, що задовольняють умові (1.1), називаються **стаціонарними**.

Приклад 1.1. Знайти стаціонарні точки функції $\varphi(x) = \frac{1}{3}x^3 - 5x^2 + 24x + 6$.

Знайдемо нулі першої похідної: $\frac{d\varphi(x)}{dx} = x^2 - 10x + 24 = 0$. Маємо $x^1 = 4$, $x^2 = 6$.

Отже, у точках x^1 і x^2 функція може досягати екстремальних значень. 

Теорема 1.2 (необхідна умова екстремуму другого порядку). Нехай функція $\varphi(x)$ визначена і двічі диференційовна на E^1 . Тоді у точці локального мінімуму (максимуму) друга похідна функції невід'ємна (недодатна):

$$\frac{d^2\varphi(x^*)}{dx^2} \geq 0 \quad (\leq 0).$$

Приклад 1.2. Розглянемо ту ж саму функцію $\varphi(x) = \frac{1}{3}x^3 - 5x^2 + 24x + 6$. Друга похідна має вигляд $\frac{d^2\varphi(x)}{dx^2} = 2x - 10$. У стаціонарній точці $x^1 = 4$ друга похідна дорівнює (-2) , тобто від'ємна. У цій точці може досягатися максимум функції. У точці $x^2 = 6$ друга похідна додатна, а значить у ній може досягатися мінімум функції.



Теорема 1.3 (достатня умова екстремуму). Нехай функція $\varphi(x)$ визначена, двічі диференційовна на E^1 . Якщо у стаціонарній точці x^* виконується умова $\frac{d^2\varphi(x^*)}{dx^2} > 0$ (< 0), то точка x^* — точка локального мінімуму (максимуму) функції $\varphi(x)$.

Приклад 1.3. Розглядається функція $\varphi(x) = \frac{1}{3}x^3 - 5x^2 + 24x + 6$. У стаціонарній точці $x^1 = 4$ друга похідна дорівнює (-2) . Отже, у цій точці досягається максимум функції, а у точці $x^2 = 6$ друга похідна дорівнює 2, тобто у ній досягається мінімум.



Якщо у стаціонарній точці x^* друга похідна дорівнює нулю, то питання про мінімум чи максимум у цій точці залишається відкритим.

Теорема 1.4 (загальна достатня умова екстремуму). Нехай функція $\varphi(x)$ визначена на E^1 і має неперервні похідні до k -го порядку включно. Якщо у точці x^* похідні до $(k-1)$ -го порядку дорівнюють нулю:

$$\frac{d\varphi(x^*)}{dx} = 0, \dots, \frac{d^{k-1}\varphi(x^*)}{dx^{k-1}} = 0, \text{ але } \frac{d^k\varphi(x^*)}{dx^k} \neq 0, \text{ то}$$

- 1) x^* є точкою локального мінімуму, якщо k — парне число і $\frac{d^k \varphi(x^*)}{dx^k} > 0$;
- 2) x^* є точкою локального максимуму, якщо k — парне число і $\frac{d^k \varphi(x^*)}{dx^k} < 0$;
- 3) x^* не є екстремальною точкою, якщо k — непарне число.

Далі сформулюємо необхідну умову мінімуму функції на відрізку $[a; b]$.

Теорема 1.5. Якщо точка $x^* = a$ є точкою мінімуму функції $\varphi(x)$ на відрізку $[a; b]$, то $\frac{d\varphi(x^*)}{dx} \geq 0$, а якщо $x^* = b$ — точка мінімуму, то $\frac{d\varphi(x^*)}{dx} \leq 0$.

1.2 Чисельні методи одновимірної мінімізації

Вище показано, що локальний мінімум функції $\varphi(x)$ знаходиться серед коренів рівняння $\frac{d\varphi(x)}{dx} = 0$. Тільки в окремих випадках розв'язок цього рівняння вдається знайти в явному вигляді. Як правило, задача пошуку коренів рівняння $\frac{d\varphi(x)}{dx} = 0$ приблизно так само складна, як і задача мінімізації функції $\varphi(x)$. Обидві ці задачі розв'язуються чисельно. При цьому для розв'язування задач мінімізації можна використовувати чисельні методи знаходження коренів рівняння, але методи, що розроблені спеціально для задач мінімізації, є більш ефективними.

Перерахуємо причини окремого розгляду чисельних методів пошуку екстремуму функції однієї змінної.

По-перше, ці алгоритми використовуються у багатьох алгоритмах пошуку екстремуму функцій багатьох змінних.

По-друге, класи функцій однієї змінної служать зручною моделлю для теоретичного дослідження ефективності методів оптимізації.

По-третє, іноді вдається, використовуючи ті чи інші прийоми, безпосередньо за допомогою алгоритмів одновимірної оптимізації одержати розв'язок багатовимірних задач.

Означення 1.5. Функція $\varphi(x)$ називається **унімодальною** на $R = [a; b]$, якщо існує така точка $x_0 \in [a; b]$, що

$$\varphi(x_1) > \varphi(x_2), \text{ якщо } x_1 < x_2 < x_0, \quad x_1, x_2 \in R,$$

$$\varphi(x_1) < \varphi(x_2), \text{ якщо } x_0 < x_1 < x_2, \quad x_1, x_2 \in R.$$

Лема 1.1. Нехай функція є унімодальною на R , $x_1, x_2 \in R$, $x_1 < x_2$. Тоді, якщо $\varphi(x_1) \leq \varphi(x_2)$, то точка мінімуму x^* така, що $x^* \leq x_2$; якщо ж $\varphi(x_1) \geq \varphi(x_2)$, то $x^* \geq x_1$.

Виникає питання, а як же знайти проміжок $R = [a; b]$, на якому знаходиться точка мінімуму? Це можна зробити так. Якщо припустити, що множина R необмежена, то за допомогою леми легко побудувати процес, що дозволяє визначити відрізок, якому належить точка мінімуму. Для цього обчислимо значення функції в точках $x_1, x_2 = x_1 + h$, де $h > 0$.

Нехай, наприклад, $\varphi(x_1) > \varphi(x_2)$. Якщо R обмежена зверху числом C , то точка мінімуму належить $[x_1; C]$. Якщо ж R необмежена зверху, то варто обчислювати значення $\varphi(x_i)$; ($x_i = x_{i-1} + h, i = 3, 4, \dots$) доти, поки не буде знайдена точка x_i , що $\varphi(x_i) \geq \varphi(x_{i-1})$. Згідно з лемою точка мінімуму належить проміжку $[x_{i-2}; x_i]$.

1.2.1 Симетричні методи

Зробимо опис довільного симетричного методу мінімізації унімодальної функції $\varphi(x)$ на $[a; b]$.

На проміжку $[a; b]$ (його ще називають **інтервалом невизначеності**) обираємо дві симетричні відносно середини відрізка точки x_1 і x_2 , обчислюємо $\varphi(x_1)$ і $\varphi(x_2)$. Якщо $\varphi(x_1) \leq \varphi(x_2)$, то згідно з лемою точка мінімуму належить $[a; x_2]$, інакше вона належить $[x_1; b]$. Ми отримали задачу мінімізації унімодальної функції на більш вузькому відрізку $[a; x_2]$, або $[x_1; b]$. На новому відрізку знову знаходимо за тим же алгоритмом дві симетричні точки і т. д. Нехай ми зменшили наш відрізок до $[a_n; b_n]$. Якщо $b_n - a_n \leq 2\varepsilon$, де ε — точність знаходження точки мінімуму, то за точку мінімуму приймаємо точку $\frac{a_n + b_n}{2}$.

Зрозуміло, що способів вибору двох симетричних точок на $[a; b]$ є нескінченна кількість. Розглянемо деякі з них.

Метод поділу відрізка навпіл

Симетричні точки x' та x'' обираються у такий спосіб:

$$x' = a + \frac{b-a}{4}, \quad x'' = b - \frac{b-a}{4}.$$

Точки x' та x'' разом із серединою відрізка $[a; b]$ ділять його на чотири рівні частини. Таким чином, на кожному кроці довжина інтервалу невизначеності зменшується вдвічі.

Метод дихотомії

У методі дихотомії точки x' та x'' обираються симетрично до середини відрізка $[a; b]$ на відстані $\frac{\delta}{2}$, де δ — мале додатне число ($\delta < \varepsilon$, де ε — точність пошуку мінімуму). Роз-

рахункові формули мають вигляд

$$x' = \frac{a + b - \delta}{2}, \quad x'' = \frac{a + b + \delta}{2}.$$

При малих δ ефективність методів дихотомії та поділу навпіл можна вважати однаковою.

Більш ефективними за два попередні методи є метод Фібоначчі та метод золотого перетину.

Метод Фібоначчі

У цьому методі використовуються числа Фібоначчі, що визначаються у такий спосіб:

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

Параметр n — кількість кроків, що використовуються у методі, знаходиться з нерівності

$$F_{n+1} < \frac{b-a}{\varepsilon} \leq F_{n+2}, \quad (1.2)$$

де ε — точність знаходження точки мінімуму. Якщо $n = 1$, то візьмемо $x'_1 = x''_1 = \frac{a+b}{2}$.

Тоді найближче значення точки мінімуму дорівнюватиме $x^* = \frac{a+b}{2}$.

Якщо $n \geq 2$, то обираємо дві точки

$$x'_1 = a + \frac{F_n}{F_{n+2}}(b-a), \quad x''_1 = a + \frac{F_{n+1}}{F_{n+2}}(b-a) = a + b - x'_1,$$

розташованих на $[a; b]$ симетрично. Подальші дії описані вище. Якщо знайдено відрізок $[a_k; b_k]$ довжини

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

то обираємо на ньому дві точки

$$x'_k = a_k + \frac{F_{n-k+1}}{F_{n+2}}(b-a) \quad \text{і} \quad x''_k = a_k + \frac{F_{n-k+2}}{F_{n+2}}(b-a).$$

Далі за описаним вище алгоритмом звужуємо відрізок $[a_k; b_k]$ доти, поки k не буде дорівнювати n . Якщо $k = n$, то довжина відрізка $[a_k; b_k]$ дорівнює

$$b_n - a_n = \frac{F_3}{F_{n+2}}(b-a) = \frac{2(b-a)}{F_{n+2}},$$

а точки

$$\begin{aligned}x'_n &= a_n + \frac{F_1}{F_{n+2}}(b - a) = a_n + \frac{b - a}{F_{n+2}}, \\x''_n &= a_n + \frac{F_2}{F_{n+2}}(b - a) = a_n + \frac{b - a}{F_{n+2}}\end{aligned}$$

співпадають і ділять відрізок $[a_n; b_n]$ навпіл. Приймаючи $x^* = x'_n = x''_n$ за точку мінімуму ми припускаємося похибки не більше ніж ε незалежно від вигляду функції $\varphi(x)$.

Приклад 1.4. Методом Фібоначчі знайти мінімум функції $\varphi(x) = x^2$ на відрізку $[1; 2]$ з точністю $\varepsilon = 0.2$.

Розв'язування. Спочатку визначимо кількість кроків n . Вона визначається з умови (1.2). Випишемо декілька чисел Фібоначчі: 1, 1, 2, 3, 5, 8, 13, ...

Так як $\frac{b-a}{\varepsilon} = 5$, то з нерівності $F_{n+1} < 5 \leq F_{n+2}$ отримаємо, що $n = 3$. Покладемо $a_1 = a$, $b_1 = b$.

Беремо $k = 1$. Тоді

$$x'_1 = 1 + \frac{F_3}{F_5} = 1.4, \quad x''_1 = 1 + \frac{F_4}{F_5} = 1.6.$$

Так як $\varphi(1.4) < \varphi(1.6)$, то $[a_2; b_2] = [1; 1.6]$.

При $k = 2$

$$x'_2 = 1 + \frac{F_2}{F_5} = 1.2, \quad x''_2 = 1 + \frac{F_3}{F_5} = 1.4.$$

(Зазначимо, що $x''_2 = x'_1$.) І так як $\varphi(1.2) < \varphi(1.4)$, то $[a_3; b_3] = [1; 1.4]$.

Тепер $k = n = 3$, тоді

$$x'_3 = 1 + \frac{F_1}{F_5} = 1.2, \quad x''_3 = 1 + \frac{F_2}{F_5} = 1.2.$$

За точку мінімуму приймаємо точку $x = 1.2$. Із заданою похибкою точку мінімуму знайдено. 

Метод золотого перетину

У цьому методі точки x' та x'' ділять відрізок $[a; b]$ у золотому відношенні.

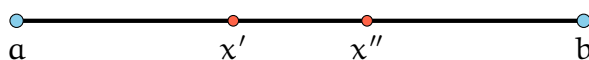
Означення 1.6. Кажуть, що точка ділить відрізок у відношенні **золотого перетину**¹, якщо відношення довжини відрізка до більшої частини дорівнює відношенню більшої частини до меншої.

На відрізку $[a; b]$ є дві симетричні відносно його кінців точки x' і x'' такі, що

$$\frac{b-a}{b-x'} = \frac{b-x'}{x'-a} = \frac{b-a}{x''-a} = \frac{x''-a}{b-x''} = \frac{1+\sqrt{5}}{2} \approx 1.618.$$

Константа $\varphi = \frac{1+\sqrt{5}}{2}$ має назву пропорції золотого перетину.

Зазначимо, що одночасно точка x' утворює золотий перетин відрізка $[a; x'']$, а точка x'' — відрізка $[x'; b]$.



Точки золотого перетину відрізка $[a; b]$

Отже, у методі золотого перетину точки x' , x'' на відрізку $[a; b]$ обираються у такий спосіб:

$$x' = a + \frac{3-\sqrt{5}}{2}(b-a),$$

$$x'' = a + \frac{\sqrt{5}-1}{2}(b-a).$$

На відміну від методу Фібоначчі, на кожному кроці обчислюється величина $b_n - a_n$. Якщо $b_n - a_n < 2\varepsilon$, то за точку мінімуму беремо $x^* = \frac{a_n + b_n}{2}$, і на цьому обчислення закінчуємо.

Приклад 1.5. Методом золотого перетину знайти мінімум функції $\varphi(x) = x^2$ на відрізку $[1; 2]$ з точністю $\varepsilon = 0.2$.

Розв'язування. Так як $b-a = 2-1 = 1 > 2\varepsilon = 0.4$, то розраховуємо дві симетричні

¹Вважається, що цей термін уперше ввів Леонардо да Вінчі. Пропорції золотого перетину пов'язують з естетичним сприйняттям форми. Відомо багато прикладів їх використання у живопису, скульптурі, архітектурі, музиці (див., наприклад, Волошинов А. В. *Математика и искусство*. — М.: Просвещение, 1992. — 335 с.).

точки

$$x' = 1 + \frac{3 - 2.236}{2}(2 - 1) = 1.382,$$
$$x'' = 1 + \frac{2.236 - 1}{2}(2 - 1) = 1.618.$$

Так як $\varphi(1.382) < \varphi(1.618)$, то $a = 1$, $b = 1.618$. Величина $b - a = 1.618 - 1 = 0.618 > 2\varepsilon = 0.4$, тому на новому відрізку $[1; 1.618]$ знаходимо дві симетричні точки

$$x' = 1 + 0.382 \cdot 0.618 = 1.236,$$
$$x'' = 1 + 0.618 \cdot 0.618 = 1.382.$$

Маємо $\varphi(x') < \varphi(x'')$. Тоді $a = 1$, $b = 1.382$. Так як $b - a = 0.382 < 2\varepsilon = 0.4$, то, взявши $x^* = \frac{a + b}{2} = 1.191$, ми отримали точку мінімуму із заданою точністю. ☒

Зауваження 1.1. Як і у методі Фібоначчі, у методі золотого перетину не варто користуватись співвідношенням $x'' = a + b - x'$. Це пов'язано з тим, що методи не стійкі до помилок. Тому на кожному кроці згідно з формулами необхідно знаходити дві симетричні точки.

1.3 Завдання для самостійної роботи

1.3.1 Чисельні методи одновимірної мінімізації

Постановка задачі Виконання завдання включає наступні етапи.

- 1) Побудова графіка функції та дослідження її властивостей (неперервність, гладкість, опуклість, унімодальність).
- 2) Вибір початкового відрізка $[a_0; b_0]$ для пошуку точки мінімуму, а також, за необхідності, інших початкових параметрів. Якщо цільова функція у своїй області визначення має глобальний мінімум, то відрізок $[a_0; b_0]$ бажано взяти таким, який його містить.
- 3) Програмна реалізація запропонованого чисельного методу. У програмі має бути передбачена можливість зміни вихідних даних (функції, відрізка, бажаної точності) при мінімальних змінах коду. Програма має виводити на екран наступне:
 - отриману оптимальну точку та значення функції у цій точці;
 - загальну кількість обчислень значень функції у ході виконання алгоритму;
 - кількість зроблених кроків (ітерацій) алгоритму;
 - візуалізацію процесу розв'язування (не обов'язково).

- 4) Розв'язування тієї ж задачі за допомогою вбудованих методів систем комп'ютерної математики. Порівняння результатів.

Варіанти завдань

	Функція	Метод
Варіант 1	$x \sin \frac{1}{x}$	метод золотого перетину
Варіант 2	$x \operatorname{arctg} x - \frac{1}{2} \ln (1 + x^2)$	метод Фібоначчі
Варіант 3	$(x - 1)^2 \sin x$	метод поділу навпіл
Варіант 4	$(x - 4)^2 + \ln x$	метод поділу навпіл
Варіант 5	$x^4 + e^{-x}$	метод золотого перетину
Варіант 6	$2x + \frac{1}{x}$	метод золотого перетину
Варіант 7	$x^4 + 2x^2 + 4x + 1$	метод Фібоначчі
Варіант 8	$x^5 - 5x^3 + 10x^2 - 5x$	метод дихотомії
Варіант 9	$x^2 + 3x(\ln x - 1)$	метод дихотомії
Варіант 10	$x^2 - 2x - 2 \cos x$	метод поділу навпіл
Варіант 11	$(x + 1)^4 - 2x^2$	метод поділу навпіл
Варіант 12	$3(5 - x)^{\frac{4}{3}} + 2x^2$	метод поділу навпіл
Варіант 13	$-x^3 + 3(1 + x) [\ln(1 + x) - 1]$	метод Фібоначчі
Варіант 14	$x + \frac{1}{x^2}$	метод поділу навпіл
Варіант 15	$x \sin x + 2 \cos x$	метод дихотомії
Варіант 16	$x^4 + 8x^3 - 6x^2 - 72x + 90$	метод Фібоначчі
Варіант 17	$x^6 + 3x^2 + 6x - 1$	метод дихотомії
Варіант 18	$10x \ln x - \frac{x^2}{2}$	метод поділу навпіл
Варіант 19	$-\frac{2x}{\ln 2} + 2x^2$	метод Фібоначчі
Варіант 20	$e^x - \frac{1}{3}x^3 + 2x$	метод золотого перетину
Варіант 21	$x e^{-x}$	метод золотого перетину
Варіант 22	$x^5 - x^2$	метод Фібоначчі
Варіант 23	$x^4 - x$	метод дихотомії
Варіант 24	$2x^2 + 3e^{-x}$	метод дихотомії
Варіант 25	$3x^2 + \frac{12}{x^3} - 5$	метод золотого перетину

Мінімізація функцій багатьох змінних

2.1 Допоміжні означення та теореми

Розглядається евклідовий простір E^n , елементами якого є вектори-стовпчики. Скалярний добуток двох векторів $\mathbf{x}^1$ і $\mathbf{x}^2$ визначається так: $(\mathbf{x}^1, \mathbf{x}^2) = \sum_{i=1}^n x_i^1 x_i^2$. Під довжиною або нормою (евклідовою нормою) вектора $\mathbf{x} \in E^n$ будемо розуміти число

$$\|\mathbf{x}\| = \sqrt{(\mathbf{x}, \mathbf{x})} = \sqrt{\sum_{i=1}^n x_i^2}.$$

Матрицю

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{pmatrix}$$

будемо коротко записувати як $[a_{ij}]_{i,j=1}^{n,m}$.

Нормою матриці (підпорядкованій нормі вектора) називається число

$$\|\mathbf{A}\| = \sup_{\mathbf{x} \neq 0} \frac{\|\mathbf{Ax}\|}{\|\mathbf{x}\|} = \sup_{\|\mathbf{x}\|=1} \|\mathbf{Ax}\|.$$

Означення 2.1. *Відрізком, що з'єднує точки $\mathbf{x}^1$ і $\mathbf{x}^2 \in E^n$, називається множина*

$$[\mathbf{x}^1, \mathbf{x}^2] = \{\mathbf{x} \in E^n \mid \mathbf{x} = \alpha \mathbf{x}^1 + (1 - \alpha) \mathbf{x}^2, 0 \leq \alpha \leq 1\}.$$

Означення 2.2. *Множина $R = \{\mathbf{x} \in E^n \mid (\mathbf{c}, \mathbf{x}) = \alpha, \mathbf{c} \neq 0\}$ називається гіперплощиною. Вектор $\mathbf{c}$ називається нормальним вектором гіперплощини.*

Він ортогональний гіперплощині.

Означення 2.3. Нехай функція $\varphi(\mathbf{x})$ визначена на множині $R \subseteq E^n$. **Лінією (множиною, поверхнею) рівня функції $\varphi(\mathbf{x})$** називається множина

$$\bar{R} = \{\mathbf{x} \in R \mid \varphi(\mathbf{x}) = C = \text{const}\},$$

тобто лінія рівня функції $\varphi(\mathbf{x})$ — це множина точок $\mathbf{x} \in R$, в яких функція $\varphi(\mathbf{x})$ приймає якесь постійне значення.

Приклад 2.1. Розглянемо функцію $\varphi(\mathbf{x}) = x_1^2 + 25x_2^2$. Лінія рівня має вигляд: $x_1^2 + 25x_2^2 = c$. Після перетворень отримаємо: $\frac{x_1^2}{(\sqrt{c})^2} + \frac{x_2^2}{(\sqrt{c}/5)^2} = 1$, а це еліпс. 

Означення 2.4. Множина $R \subseteq E^n$ називається **опуклою**, якщо разом з довільними точками $\mathbf{x}^1$ і $\mathbf{x}^2$ вона містить у собі весь відрізок, що з'єднує їх, тобто якщо точка $\mathbf{x}(\alpha) = \alpha\mathbf{x}^1 + (1 - \alpha)\mathbf{x}^2 \in R$ при довільних $\mathbf{x}^1, \mathbf{x}^2 \in R$ та $\alpha \in [0; 1]$.

Означення 2.5. Симетрична матриця $\mathbf{A}$ виміру $n \times n$ називається **невід'ємною**, якщо відповідна їй квадратична форма є невід'ємно означеною, тобто $\mathcal{I} = \mathbf{x}^T \mathbf{A} \mathbf{x} \geq 0$ для всіх $\mathbf{x} \in E^n$.

Означення 2.6. Симетрична матриця $\mathbf{A}$ виміру $n \times n$ називається **додатною**, якщо відповідна їй квадратична форма є додатно означеною, тобто $\mathcal{I} = \mathbf{x}^T \mathbf{A} \mathbf{x} > 0$ для всіх $\mathbf{x} \in E^n, \mathbf{x} \neq \mathbf{0}$.

Аналогічно вводяться поняття недодатної і від'ємної матриці.

Означення 2.7. **Головним кутовим мінором k -го порядку** деякої квадратної матриці називається визначник матриці, складений з перших k рядків і перших k стовпчиків початкової матриці.

Критерій Сільвестра. Симетрична матриця є:

- а) додатною тоді і тільки тоді, коли всі її головні кутові мінори додатні;
- б) від'ємною, якщо всі її головні кутові мінори непарного порядку від'ємні, а парного — додатні;
- в) невід'ємною тоді і тільки тоді, коли всі мінори, утворені рядками і стовпчиками початкової матриці з однаковими номерами, невід'ємні;
- г) недодатною, якщо всі мінори, утворені рядками і стовпчиками початкової матриці з однаковими номерами непарного порядку недодатні, а парного — невід'ємні.

Теорема 2.1. Перетин довільної кількості опуклих множин теж є опу-

клою множиною.

Важливими теоремами при дослідженні задач на екстремум є теореми про віддільність опуклих множин.

Означення 2.8. Нехай X і Y — множини E^n . Множини **віддільні**, якщо існують вектор $c \neq 0$ і число α такі, що виконуються нерівності

$$c^T x \geq \alpha, \quad c^T y \leq \alpha \quad (2.1)$$

для всіх $x \in X, y \in Y$.

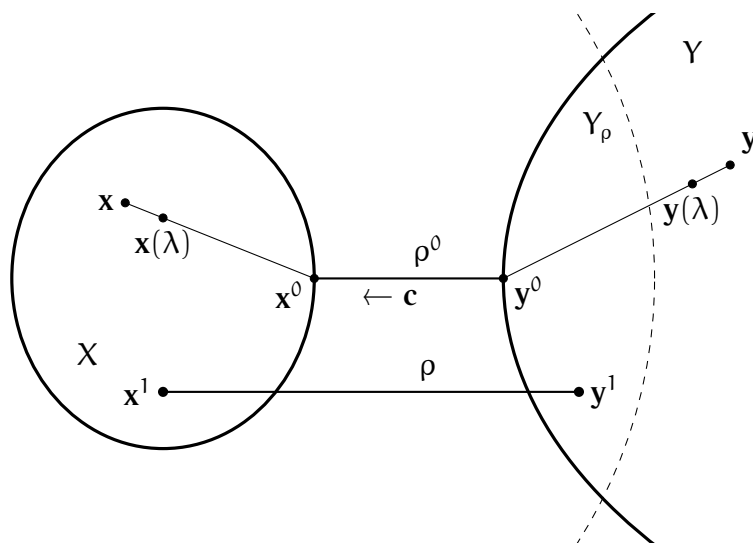
Якщо у (2.1) виконуються строгі нерівності, то множини X і Y **строго віддільні**. Гіперплощина $c^T x = \alpha$ називається **відокремлюючою**.

Теорема 2.2 ([3], про строгу віддільність множин). Якщо множини X та Y опуклі, замкнені, не мають спільних точок і одна з них обмежена, то вони строго віддільні.

Доведення. Нехай, наприклад, обмеженою є множина X . У множинах X і Y візьмемо по одній точці x^1 і y^1 , відстань між якими позначимо через ρ : $\|x^1 - y^1\| = \rho$. Перетин замкненого ρ -околу множини X з множиною Y позначимо через Y_ρ . Множини X і Y_ρ обмежені, замкнені і непусті ($y^1 \in Y_\rho$). Тому неперервна функція $f(x, y) = \|x - y\|$ досягає на (X, Y_ρ) мінімуму

$$\min_{x \in X, y \in Y_\rho} \|x - y\| = \|x^0 - y^0\| = \rho^0. \quad (2.2)$$

Число ρ^0 є мінімальною відстанню між множинами X і Y_ρ та між X і Y . Так як множини X і Y не мають спільних точок, то $\rho^0 > 0$.



Нехай x — довільна точка з X . Так як множина X опукла, то точка $x(\lambda) = \lambda x + (1 -$

$\lambda)x^0$ при довільних λ ($0 \leq \lambda \leq 1$) належить множині X .

За означенням (2.2) точок x^0, y^0 отримаємо, що функція

$$\varphi(\lambda) = \|x(\lambda) - y^0\|^2 = (x(\lambda) - y^0)^\top (x(\lambda) - y^0), \quad 0 \leq \lambda \leq 1$$

у межовій точці $\lambda = 0$ відрізка $[0; 1]$ досягає мінімуму. Тому

$$\left. \frac{d\varphi(\lambda)}{d\lambda} \right|_{\lambda=0} = 2 (x(\lambda) - y^0)^\top \left. \frac{dx(\lambda)}{d\lambda} \right|_{\lambda=0} = 2 (x^0 - y^0)^\top (x - x^0) \geq 0.$$

Нехай

$$c = \frac{x^0 - y^0}{\|x^0 - y^0\|}, \quad (2.3)$$

тоді отримаємо нерівність

$$c^\top x \geq c^\top x^0, \quad (2.4)$$

яка справедлива для всіх $x \in X$.

Візьмемо довільну точку $y \in Y$. Так як Y — опукла множина, то точка $y(\lambda) = \lambda y + (1 - \lambda)y^0$ при $\lambda \in [0; 1]$ теж належить множині X . Тоді функція

$$\psi(\lambda) = \|y(\lambda) - x^0\|^2 = (y(\lambda) - x^0)^\top (y(\lambda) - x^0), \quad 0 \leq \lambda \leq 1$$

досягає мінімуму в точці $\lambda = 0$. Маємо

$$\left. \frac{d\psi(\lambda)}{d\lambda} \right|_{\lambda=0} = 2 (y(\lambda) - x^0)^\top \left. \frac{dy(\lambda)}{d\lambda} \right|_{\lambda=0} = 2 (y^0 - x^0)^\top (y - y^0) \geq 0.$$

Використовуючи (2.3), отримаємо нерівність

$$c^\top y \leq c^\top y^0, \quad y \in Y. \quad (2.5)$$

Покажемо, що гіперплощина

$$c^\top x = \alpha, \quad (2.6)$$

де

$$\alpha = \frac{c^\top (x^0 + y^0)}{2}, \quad (2.7)$$

що проведена через середину відрізка, який з'єднує точки x^0, y^0 , строго розділяє множини X і Y .

Дійсно, з означення (2.3) вектора $\mathbf{c}$ маємо

$$\mathbf{c}^\top (\mathbf{x}^0 - \mathbf{y}^0) = \frac{(\mathbf{x}^0 - \mathbf{y}^0)^\top (\mathbf{x}^0 - \mathbf{y}^0)}{\|\mathbf{x}^0 - \mathbf{y}^0\|} = \|\mathbf{x}^0 - \mathbf{y}^0\| = \rho^0.$$

Тому $\mathbf{c}^\top \mathbf{x}^0 = \mathbf{c}^\top \mathbf{y}^0 + \rho^0$ і $\mathbf{c}^\top \mathbf{y}^0 = \mathbf{c}^\top \mathbf{x}^0 - \rho^0$. Підкладаючи ці вирази в (2.7), отримаємо

$$\alpha = \mathbf{c}^\top \mathbf{x}^0 - \frac{\rho^0}{2} \quad \text{і} \quad \alpha = \mathbf{c}^\top \mathbf{y}^0 + \frac{\rho^0}{2}.$$

Звідси

$$\mathbf{c}^\top \mathbf{x}^0 = \alpha + \frac{\rho^0}{2} > \alpha, \quad (2.8)$$

$$\mathbf{c}^\top \mathbf{y}^0 = \alpha - \frac{\rho^0}{2} < \alpha, \quad (2.9)$$

Використовуючи (2.4) і (2.8), отримаємо

$$\mathbf{c}^\top \mathbf{x} \geq \mathbf{c}^\top \mathbf{x}^0 > \alpha, \quad \mathbf{c}^\top \mathbf{x} > \alpha, \quad \mathbf{x} \in X,$$

а (2.5) і (2.9) дають

$$\mathbf{c}^\top \mathbf{y} \leq \mathbf{c}^\top \mathbf{y}^0 < \alpha, \quad \mathbf{c}^\top \mathbf{y} < \alpha, \quad \mathbf{y} \in Y.$$

Отримали, що $\mathbf{c}^\top \mathbf{x} > \alpha$, $\mathbf{x} \in X$, $\mathbf{c}^\top \mathbf{y} < \alpha$, $\mathbf{y} \in Y$. Множини X і Y строго віддільні. Теорему доведено. 

Означення 2.9. Гіперплощина $\mathbf{c}^\top \mathbf{x} = \alpha$ називається **опорною** до множини R у точці $\mathbf{x}^0$, якщо $\mathbf{c}^\top \mathbf{x}^0 = \alpha$, $\inf \|\mathbf{x} - \mathbf{x}^0\| = 0$ і виконується одна з нерівностей: $\mathbf{c}^\top \mathbf{x} \geq \alpha$, $\mathbf{c}^\top \mathbf{x} \leq \alpha$.

Наслідок 2.1. Опукла множина R у кожній межовій точці має опорну гіперплощину.

Теорема 2.3 ( 3, про нестрогу віддільність множин). Якщо опуклі множини X , Y не мають спільних внутрішніх точок, то їх можна відділити.

Доведення. Побудуємо множину $Z = X - Y = \{\mathbf{x} - \mathbf{y} \mid \mathbf{x} \in X, \mathbf{y} \in Y\}$. Множина Z — опукла. Дійсно, якщо $\mathbf{z}^1 = \mathbf{x}^1 - \mathbf{y}^1 \in Z$ і $\mathbf{z}^2 = \mathbf{x}^2 - \mathbf{y}^2 \in Z$, то для довільного λ ($0 \leq \lambda \leq 1$) точка

$$\begin{aligned} \mathbf{z}(\lambda) &= \lambda \mathbf{z}^1 + (1 - \lambda) \mathbf{z}^2 = \lambda \mathbf{x}^1 - \lambda \mathbf{y}^1 + (1 - \lambda) \mathbf{x}^2 - (1 - \lambda) \mathbf{y}^2 = \\ &= [\lambda \mathbf{x}^1 + (1 - \lambda) \mathbf{x}^2] - [\lambda \mathbf{y}^1 + (1 - \lambda) \mathbf{y}^2] = \mathbf{x}(\lambda) - \mathbf{y}(\lambda) \in Z, \end{aligned}$$

так як $\mathbf{x}(\lambda) = \lambda \mathbf{x}^1 + (1 - \lambda) \mathbf{x}^2 \in X$, $\mathbf{y}(\lambda) = \lambda \mathbf{y}^1 + (1 - \lambda) \mathbf{y}^2 \in Y$ згідно з опуклістю множин X і Y .

Якщо множини X і Y не мають спільних внутрішніх точок, то нульова точка не може бути внутрішньою точкою множини Z .

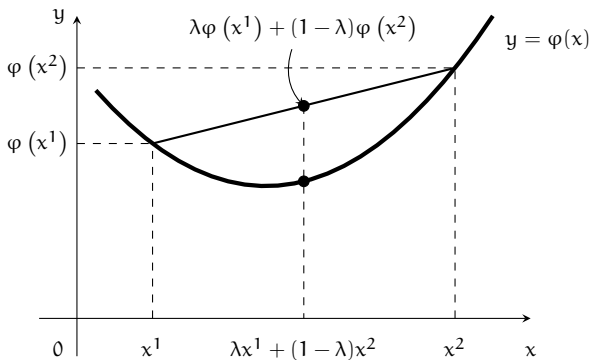
У взаємному розташуванні нульової точки і множини Z можливі два випадки.

- а) Точка $z = 0$ не належить замиканню $\bar{Z}$. У цьому випадку за [теоремою 2.2](#) точка $z = 0$ строго відокремлена від множини Z , тобто $c^T 0 < \alpha$, $c^T z > \alpha$ для всіх $z \in Z$ (за [теоремою 2.2](#)), або $c^T (x - y) > 0$ для $x \in X$ і $y \in Y$. Отримали, що $c^T x > c^T y$, $x \in X$, $y \in Y$. Теорему доведено.
- б) Точка $z = 0$ — межова точка множини Z , тобто $0 \in \bar{Z}$. У цьому випадку за [наслідком 2.1](#) існує гіперплощина, що є опорною до Z у точці $z = 0$. Тоді $c^T 0 = \alpha$ і $c^T z \geq \alpha = 0$ для $z \in Z$. Звідси $c^T z \geq 0$, $c^T (x - y) \geq 0$, $c^T x \geq c^T y$, $x \in X$, $y \in Y$. Теорему доведено.



Означення 2.10. Функція $\varphi(x)$, що визначена на опуклій множині R , називається **опуклою**, якщо для довільних x^1 і $x^2 \in R$ і довільного $\lambda \in [0; 1]$ виконується нерівність

$$\varphi(\lambda x^1 + (1 - \lambda)x^2) \leq \lambda \varphi(x^1) + (1 - \lambda)\varphi(x^2). \quad (2.10)$$



Геометрично це означає, що на проміжку $[x^1, x^2]$ функція завжди знаходиться нижче за хорду, що з'єднує точки $(x^1, \varphi(x^1))$ і $(x^2, \varphi(x^2))$ або співпадає з нею.

Так як у (2.10) присутній знак рівності, то опукла функція може мати прямолінійні відрізки.

Означення 2.11. Функція $\varphi(x)$ називається **строго опуклою** на опуклій множині R , якщо нерівність

$$\varphi(\lambda x^1 + (1 - \lambda)x^2) < \lambda \varphi(x^1) + (1 - \lambda)\varphi(x^2) \quad (2.11)$$

виконується для довільних точок $x^1, x^2 \in R$, $x^1 \neq x^2$ та $\lambda \in (0; 1)$.

Якщо у (2.10) і (2.11) змінити знак нерівності на протилежний, то отримаємо поняття вгнutoї і строго вгнutoї функції.

Теорема 2.4 ([3]). Для того, щоб функція $\varphi(x)$, яка визначена на опуклій множині R , була опуклою, необхідно і достатньо, щоб була опуклою множи-

на $\bar{R} = \{x, y \mid x \in R, y \geq \varphi(x)\}$, тобто надграфік функції був опуклою множиною.

Означення 2.12. Функція $\varphi(x)$, що визначена на опуклій множині R , називається **сильно опуклою**, якщо знайдеться така константа $\gamma > 0$, що для довільних x^1 і $x^2 \in R$ та $\lambda \in [0; 1]$ виконується нерівність

$$\varphi(\lambda x^1 + (1 - \lambda)x^2) \leq \lambda \varphi(x^1) + (1 - \lambda)\varphi(x^2) - \gamma \lambda(1 - \lambda) \|x^1 - x^2\|^2.$$

Приклад 2.2. Довести, що якщо функція $\varphi(x)$ є опуклою, то множина $R = \{x \mid \varphi(x) \leq c\}$ — опукла.

Доведення. Візьмемо дві точки x^1 і x^2 множини R . Так як $x^1 \in R$, то $\varphi(x^1) \leq c$. Аналогічно, $\varphi(x^2) \leq c$. Розглянемо точку $x(\lambda) = \lambda x^1 + (1 - \lambda)x^2$, $\lambda \in [0; 1]$. Далі,

$$\varphi(x(\lambda)) = \varphi(\lambda x^1 + (1 - \lambda)x^2) \leq \lambda \varphi(x^1) + (1 - \lambda)\varphi(x^2) \leq \lambda c + (1 - \lambda)c = c.$$

Отже, взявши дві точки x^1 і $x^2 \in R$, ми довели, що точка $x(\lambda)$ при $\lambda \in [0; 1]$ теж належить множині R . А це означає, що R — опукла множина. 

Приклад 2.3. Якою повинна бути функція $\varphi(x)$, щоб множина

$$R = \{x \mid \varphi(x) \geq c\}$$

була опуклою?

Зрозуміло, що функція $\varphi(x)$ повинна бути вігнутою. 

Теорема 2.5. Опукла функція $\varphi(x)$, $x \in R$, неперервна у всіх внутрішніх точках множини R , тобто опукла функція може мати розриви тільки на межі множини R .

Теорема 2.6. Якщо функції $\varphi_i(x)$, ($i = \overline{1, m}$, $x \in R$) опуклі на R , то і функції

$$\psi(x) = \sum_{i=1}^m \alpha_i \varphi_i(x), \quad \alpha_i \geq 0,$$

$$f(x) = \max_{1 \leq i \leq m} \varphi_i(x)$$

теж опуклі.

Теорема 2.7 ( 3). У опуклої функції локальний мінімум буде і глобальним. Строго опукла функція має одну точку мінімуму. Множина точок мінімумів опуклої функції є опуклою.

Теорема 2.8 ( 2). Нехай опукла функція $\varphi(x)$ неперервно диферен-

ційовна на опуклій множині R . Для того, щоб точка $\mathbf{x}^*$ була точкою мінімуму функції $\varphi(\mathbf{x})$, необхідно і достатньо, щоб $\left(\frac{\partial \varphi(\mathbf{x}^*)}{\partial \mathbf{x}}, \mathbf{x} - \mathbf{x}^*\right) \geq 0$ для всіх $\mathbf{x} \in R$.

Якщо $\mathbf{x}^*$ — внутрішня точка множини R , то виконується умова $\frac{\partial \varphi(\mathbf{x}^*)}{\partial \mathbf{x}} = 0$.

Теорема 2.9. Якщо функція $\varphi(\mathbf{x})$ двічі неперервно диференційовна, то для опуклості її на опуклій множині R необхідно і достатньо, щоб матриця $\frac{\partial^2 \varphi(\mathbf{x}^*)}{\partial \mathbf{x}^2}$ була невід'ємною для всіх $\mathbf{x} \in R$. Якщо ж матриця додатна, то функція $\varphi(\mathbf{x})$ — строго опукла.

Зрозуміло, що якщо функція $\varphi(\mathbf{x})$ вігнута на опуклій множині R , то функція $(-\varphi(\mathbf{x}))$ опукла на цій множині, і тому властивості вігнутих функцій можна досить легко отримати з відповідних властивостей опуклих функцій.

2.2 Постановка задачі математичного програмування

Означення 2.13. Точка $\mathbf{x}^* \in R$ називається *точкою глобального мінімуму функції* $\varphi(\mathbf{x})$ на множині R , якщо

$$\varphi(\mathbf{x}^*) \leq \varphi(\mathbf{x}) \quad \text{для всіх } \mathbf{x} \in R.$$

Множину точок глобального мінімуму функції позначимо через X_* .

Означення 2.14. Точка $\mathbf{x}^* \in R$ називається *точкою локального мінімуму* функції $\varphi(\mathbf{x})$ на множині R , якщо знайдеться константа $\varepsilon > 0$ така, що

$$\varphi(\mathbf{x}^*) \leq \varphi(\mathbf{x}) \quad \text{для всіх } \mathbf{x} \in R, \text{ що задовольняють умові } \|\mathbf{x} - \mathbf{x}^*\| \leq \varepsilon.$$

Означення 2.15. Точка $\mathbf{x}^* \in R$ називається *точкою строгого мінімуму* (у локальному чи глобальному сенсі), якщо відповідні нерівності у наведених означеннях виконуються як строги (при $\mathbf{x} \neq \mathbf{x}^*$).

Аналогічно вводяться означення точок локального і глобального максимуму.

Умовне позначення $\varphi(\mathbf{x}) \rightarrow \text{extr}, \mathbf{x} \in R$ застосовується при розгляданні задачі пошуку екстремуму функції $\varphi(\mathbf{x})$ на множині R . Запис

$$\varphi(\mathbf{x}) \rightarrow \min \text{ або } \varphi(\mathbf{x}) \rightarrow \max$$

означає, що досліджується тільки задача мінімізації або максимізації функції $\varphi(\mathbf{x})$.

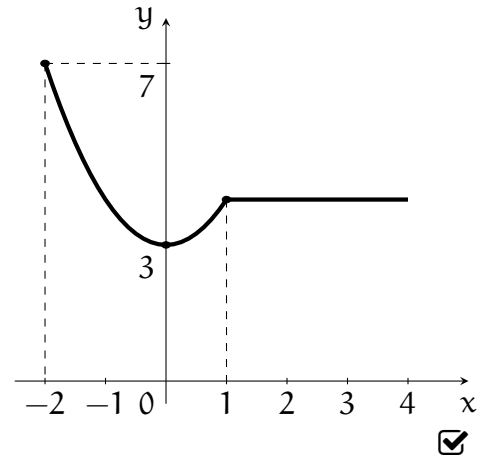
Так як довільна задача максимізації функції $\varphi(\mathbf{x})$ може бути записана у вигляді

задачі мінімізації функції $-\varphi(\mathbf{x})$, то всі теоретичні міркування можна проводити тільки для задачі на мінімум.

Приклад 2.4. Розглянемо функцію

$$\varphi(x) = \begin{cases} x^2 + 3, & x \in [-2; 1], \\ 4, & x \geq 1. \end{cases}$$

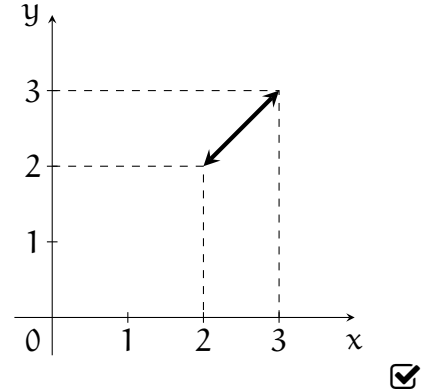
Ця функція на $[-2; +\infty)$ має одну точку строгого глобального мінімуму $x^* = 0$, $\min_{x \in \mathbb{R}} \varphi(x) = \varphi(0) = 3$. Точка $x^* = -2$ є точкою строгого локального максимуму, $\varphi(x^*) = \varphi(-2) = 7$. Промінь $[1; +\infty)$ є множиною нестрогого максимуму, $\max \varphi(x) = 4$.



Приклад 2.5. Функція $\varphi(x) = e^{-x}$, $R = \{x \in E^1 \mid x \geq 1\}$ має одну точку строгого глобального максимуму $x^* = 1$ і не має точок мінімуму. ☒

Приклад 2.6.

Функція $\varphi(x) = x$ не досягає екстремуму в жодній з точок множини $R = \{x \in E^1 \mid 2 < x < 3\}$.



Узагальненням поняття найменшого значення функції є визначення нижньої межі.

Означення 2.16. Нехай функція $\varphi(x)$ обмежена знизу на множині R . Число φ_* називається **нижньою межею (інфімумом)** $\varphi(x)$, якщо воно є найбільшим з нижніх меж функції $\varphi(x)$ на R , тобто

- 1) $\varphi_* \leq \varphi(x), \quad \forall x \in R,$
- 2) $\forall \varepsilon > 0 \quad \exists x_\varepsilon \in R: \varphi(x_\varepsilon) < \varphi_* + \varepsilon.$

Якщо функція $\varphi(x)$ обмежена знизу на R , то існує єдина скінченна нижня межа цієї функції на множині R . Приймаючи у якості інфімуму необмеженої знизу на R функції $\varphi_* = -\infty$, можна вважати, що нижня межа (на відміну від мінімуму) існує завжди.

Аналогічно вводиться поняття верхньої межі (супремуму), як найменшої верхньої межі функції $\varphi(\mathbf{x})$ на R .

2.2.1 Збіжність в екстремальних задачах. Існування екстремумів

Розглянуті приклади показують, що не завжди існує точка, в якій досягається нижня межа цільової функції. Тому краще розглянути узагальнену задачу оптимізації — побудову мінімізуючої послідовності.

Означення 2.17. *Послідовність точок $\{\mathbf{x}^k\}$ з припустимої множини R називається **мінімізуючою** для функції $\varphi(\mathbf{x})$, якщо*

$$\lim_{k \rightarrow \infty} \varphi(\mathbf{x}^k) = \inf_{\mathbf{x} \in R} \varphi(\mathbf{x}) = \varphi_*.$$

Побудова мінімізуючих послідовностей є метою розв’язування задачі мінімізації не тільки у тому випадку, коли точна нижня межа функції не досягається. Більшість методів оптимізації генерують послідовність точок, яка є мінімізуючою.

Розглянемо достатню умову досягнення верхньої і нижньої меж.

Теорема 2.10 (Вейєрштрасса). *Нехай R — обмежена і замкнена множина, функція $\varphi(\mathbf{x})$ — неперервна на R . Тоді $f_* = \inf_{\mathbf{x} \in R} \varphi(\mathbf{x}) > -\infty$, множина точок глобального мінімуму непуста, обмежена і замкнена, а довільна мінімізуюча послідовність збігається до X_* .*

Наслідок 2.2. *Нехай R — непуста замкнута підмножина E^n , функція $\varphi(\mathbf{x})$ неперервна на R і для деякої фіксованої точки $\mathbf{x}^0$ множина Лебега*

$$L(\mathbf{x}^0) = \{\mathbf{x} \in R \mid \varphi(\mathbf{x}) \leq \varphi(\mathbf{x}^0)\}$$

обмежена. Тоді виконуються всі твердження теореми Вейєрштрасса при умові, що елементи мінімізуючої послідовності $\mathbf{x}^k \in L(\mathbf{x}^0)$.

Наслідок 2.3. *Нехай R — непуста замкнута підмножина E^n , функція $\varphi(\mathbf{x})$ неперервна на R і для довільної послідовності $\{\mathbf{x}^k\}$ точок з R , що задовольняють умові $\lim_{k \rightarrow \infty} \|\mathbf{x}^k\| = +\infty$, виконується співвідношення $\lim_{k \rightarrow \infty} \varphi(\mathbf{x}^k) = +\infty$. Тоді виконуються всі твердження теореми Вейєрштрасса.*

Аналогічно формулюється теорема для задачі максимізації.

2.3 Мінімізація функцій без обмежень

Розглянемо задачу

$$\varphi(\mathbf{x}) \rightarrow \text{extr}, \quad \mathbf{x} \in E^n. \quad (2.12)$$

Класичний підхід до пошуку безумовного екстремуму ґрунтується на таких твердженнях.

Т е о р е м а 2.11 (необхідна умова екстремуму першого порядку). Нехай функція $\varphi(\mathbf{x})$ диференційовна у точці $\mathbf{x}^* \in E^n$. Тоді якщо $\mathbf{x}^*$ — локальний розв’язок задачі (2.12), то

$$\frac{\partial \varphi(\mathbf{x}^*)}{\partial \mathbf{x}} = 0. \quad (2.13)$$

Означення 2.18. Розв’язки системи рівнянь (2.13) називаються **стаціонарними точками**.

Т е о р е м а 2.12 (необхідна умова екстремуму другого порядку). Нехай функція $\varphi(\mathbf{x})$ двічі диференційовна у точці $\mathbf{x}^* \in E^n$.

- 1) Якщо $\mathbf{x}^*$ — точка локального мінімуму в задачі (2.12), то матриця $\frac{\partial^2 \varphi(\mathbf{x}^*)}{\partial \mathbf{x}^2}$ невід’ємно визначена, тобто

$$\left(\frac{\partial^2 \varphi(\mathbf{x}^*)}{\partial \mathbf{x}^2} \mathbf{x}, \mathbf{x} \right) \geq 0 \quad \text{для всіх } \mathbf{x} \in E^n.$$

- 2) Якщо $\mathbf{x}^*$ — точка локального максимуму, то матриця $\frac{\partial^2 \varphi(\mathbf{x}^*)}{\partial \mathbf{x}^2}$ недодатно визначена, тобто

$$\left(\frac{\partial^2 \varphi(\mathbf{x}^*)}{\partial \mathbf{x}^2} \mathbf{x}, \mathbf{x} \right) \leq 0 \quad \text{для всіх } \mathbf{x} \in E^n.$$

Т е о р е м а 2.13 (достатня умова екстремуму). Нехай функція $\varphi(\mathbf{x})$ двічі диференційовна у точці $\mathbf{x}^* \in E^n$ і $\frac{\partial \varphi(\mathbf{x}^*)}{\partial \mathbf{x}} = 0$.

- 1) Якщо матриця $\frac{\partial^2 \varphi(\mathbf{x}^*)}{\partial \mathbf{x}^2}$ додатно визначена, то $\mathbf{x}^*$ — точка строгого локального мінімуму функції $\varphi(\mathbf{x})$ на E^n .
- 2) Якщо матриця $\frac{\partial^2 \varphi(\mathbf{x}^*)}{\partial \mathbf{x}^2}$ від’ємно визначена, то $\mathbf{x}^*$ — точка строгого локального максимуму функції $\varphi(\mathbf{x})$ на E^n .

Ці твердження справедливі і для задачі пошуку екстремуму на множині $R \in E^n$, якщо $\mathbf{x}^*$ — внутрішня точка R .

Якщо відомо, що функція має глобальний екстремум, то точкою глобального міні-

муму (максимуму) є та стаціонарна точка, в якій функція приймає найменше (найбільше) значення. Для визначення існування глобального екстремуму інколи можна використувати наслідки з теореми Вейєрштрасса. Якщо ж функція опукла (вгнута) на E^n , то тоді локальні екстремуми точки будуть точками глобального екстремуму.

Можна застосувати таку схему пошуку безумовних екстремумів функцій:

- 1) Складаємо і розв'язуємо систему рівнянь (2.13).
- 2) У стаціонарних точках досліджуємо на знаковизначеність матрицю других похідних; ті точки, в яких матриця додатно визначена, є точками строгого локального мінімуму; стаціонарні точки, в яких матриця від'ємно визначена, є точками строгого локального максимуму.
- 3) Аналізуємо стаціонарні точки, в яких матриця других похідних не є строго знаковизначеною. Якщо матриця невід'ємно (недодатно) визначена, то відповідні точки підозрілі на локальний мінімум (максимум). Інколи вивчення поведінки функції в околі підозрілої точки дає можливість виявити, чи є чи ні точка екстремальною. І якщо у стаціонарній точці матриця других похідних не є знаковизначеною, то така точка не може бути екстремальною.
- 4) Знайдені точки локального екстремуму досліджуємо на глобальний екстремум.

Приклад 2.7. Дослідити на екстремум функцію

$$\varphi(\mathbf{x}) = x_1^4 + x_2^4 - 2(x_1 - x_2)^2.$$

Розраховуємо градієнт функції

$$\begin{aligned}\frac{\partial \varphi(\mathbf{x})}{\partial x_1} &= 4x_1^3 - 4(x_1 - x_2), \\ \frac{\partial \varphi(\mathbf{x})}{\partial x_2} &= 4x_2^3 + 4(x_1 - x_2).\end{aligned}$$

Прирівнявши його до нуля і розв'язавши отриману систему, маємо

$$\mathbf{x}^1 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad \mathbf{x}^2 = \begin{pmatrix} \sqrt{2} \\ -\sqrt{2} \end{pmatrix}, \quad \mathbf{x}^3 = \begin{pmatrix} -\sqrt{2} \\ \sqrt{2} \end{pmatrix}.$$

Матриця Гессе, матриця других похідних, має вигляд

$$\frac{\partial^2 \varphi(\mathbf{x})}{\partial \mathbf{x}^2} = \begin{pmatrix} 12x_1^2 - 4 & 4 \\ 4 & 12x_2^2 - 4 \end{pmatrix}.$$

У точці $\mathbf{x}^1$ матриця має вигляд

$$\frac{\partial^2 \varphi(\mathbf{x}^1)}{\partial \mathbf{x}^2} = \begin{pmatrix} -4 & 4 \\ 4 & -4 \end{pmatrix}.$$

Вона недодатно визначена. Це означає, що точка $\mathbf{x}^1$ підозріла на максимум, $\varphi(\mathbf{x}^1) = \varphi(0, 0) = 0$. Зазначимо, що $\varphi(x_1, x_1) = 2x_1^4 > \varphi(\mathbf{x}^1) = \varphi(0, 0) = 0$ і точки, що лежать на прямій $x_1 = x_2$ можуть бути обрані дуже близько до початку координат. Це означає, що у точці $\mathbf{x}^1$ функція не може досягати максимуму.

Розглянемо матрицю Гессе у точці $\mathbf{x}^2$.

$$\frac{\partial^2 \varphi(\mathbf{x}^2)}{\partial \mathbf{x}^2} = \frac{\partial^2 \varphi(\sqrt{2}, -\sqrt{2})}{\partial \mathbf{x}^2} = \begin{pmatrix} 20 & 4 \\ 4 & 20 \end{pmatrix}.$$

Ця матриця додатно визначена, отже, точка $\mathbf{x}^2$ є точкою строгого локального мінімуму.

У точці $\mathbf{x}^3$ матриця Гессе має той же вигляд, що і в точці $\mathbf{x}^2$. Це означає, що і в точці $\mathbf{x}^3$ функція має строгий локальний мінімум. Так як $\lim_{k \rightarrow \infty} \|\mathbf{x}^k\| = +\infty$, то згідно з другим наслідком теореми Вейерштрасса точки $\mathbf{x}^2$ і $\mathbf{x}^3$ є точками глобального мінімуму функції і $\min_{\mathbf{x} \in E^2} \varphi(\mathbf{x}) = -8$. ☑

2.4 Задача на умовний мінімум

2.4.1 Мінімізація функцій при обмеженнях типу рівностей

Розглядається задача

$$\varphi(\mathbf{x}) \rightarrow \text{extr} \tag{2.14}$$

$$\mathbf{x} \in R = \{\mathbf{x} \in E^n \mid g_i(\mathbf{x}) = 0, i = \overline{1, m}\}. \tag{2.15}$$

Вважаємо, що $m < n$.

У деяких випадках для розв'язування такої задачі можна застосувати метод виключення. Нехай з обмежень-рівностей можна виразити m якихось компонент вектора $\mathbf{x}$ через інші $n - m$ компонент. Наприклад,

$$x_1 = f_1(x_{m+1}, \dots, x_n),$$

$$\dots$$

$$x_m = f_m(x_{m+1}, \dots, x_n).$$

Далі підкладемо ці вирази у функцію $\varphi(\mathbf{x})$ і отримаємо задачу на безумовний екстремум у просторі E^{n-m} . Звичайно, скористатись таким алгоритмом можна не завжди і тому розглянемо більш загальний підхід — **правило множників Лагранжа**.

Для задачі (2.14)–(2.15) розглянемо функцію Лагранжа

$$L(\mathbf{x}, \lambda_0, \boldsymbol{\lambda}) = \lambda_0 \varphi(\mathbf{x}) + \sum_{i=1}^m \lambda_i g_i(\mathbf{x}),$$

де $\lambda_0 \in E^1$, $\boldsymbol{\lambda} \in E^m$.

Теорема 2.14 ([3], правило множників Лагранжа). Нехай $\mathbf{x}^*$ — точка локального екстремуму в задачі (2.14)–(2.15); функції $\varphi(\mathbf{x})$ і $g_i(\mathbf{x})$ ($i = \overline{1, m}$) неперервно диференційовні в околі точки $\mathbf{x}^*$. Тоді існує число λ_0^* і вектор $\boldsymbol{\lambda}^* \in E^m$, які одночасно не дорівнюють нулеві і такі, що

$$\frac{\partial L(\mathbf{x}^*, \lambda_0^*, \boldsymbol{\lambda}^*)}{\partial \mathbf{x}} = 0.$$

Кожній точці локального екстремуму відповідає нескінченна кількість наборів множників Лагранжа (так як невідомих на одну більше, ніж рівнянь).

Означення 2.19. Якщо серед систем множників Лагранжа, що відповідають точці локального екстремуму, немає множників Лагранжа з $\lambda_0^* = 0$, то така точка називається **нормальною точкою екстремуму**.

Теорема 2.15 ([3]). Для нормальної точки локального екстремуму відповідні множники Лагранжа визначаються єдиним способом з $\lambda_0^* = 1$.

Теорема 2.16 ([3]). Точка локального екстремуму $\mathbf{x}^*$ в задачі (2.14)–(2.15) є нормальною точкою екстремуму тоді і тільки тоді, коли вектори $\frac{\partial g_1(\mathbf{x}^*)}{\partial \mathbf{x}}$, ..., $\frac{\partial g_m(\mathbf{x}^*)}{\partial \mathbf{x}}$ лінійно незалежні. Якщо обмеження одне, то тоді $\frac{\partial g(\mathbf{x}^*)}{\partial \mathbf{x}} \neq 0$.

Сформулюємо необхідну умову оптимальності другого порядку.

Теорема 2.17 ([3]). Нехай:

- 1) функції $\varphi(\mathbf{x})$, $g_1(\mathbf{x})$, ..., $g_m(\mathbf{x})$ неперервно диференційовні у деякому околі точки $\mathbf{x}^* \in \mathbb{R}$ і двічі диференційовні у самій цій точці;
- 2) градієнти $\frac{\partial g_1(\mathbf{x}^*)}{\partial \mathbf{x}}$, ..., $\frac{\partial g_m(\mathbf{x}^*)}{\partial \mathbf{x}}$ лінійно незалежні;
- 3) $\mathbf{x}^*$ — точка локального мінімуму (максимуму) задачі (2.14)–(2.15).

Тоді

$$\frac{\partial L(\mathbf{x}^*, \lambda_0^*, \boldsymbol{\lambda}^*)}{\partial \mathbf{x}} = 0, \quad \left(\frac{\partial^2 L(\mathbf{x}^*, \lambda_0^*, \boldsymbol{\lambda}^*)}{\partial \mathbf{x}^2} \mathbf{y}, \mathbf{y} \right) \geq 0 \ (\leq 0)$$

для довільного вектора $y \in E^n$, що задовольняє умові

$$\left(\frac{\partial g_i(x^*)}{\partial x}, y \right) = 0, \quad i = \overline{1, m}.$$

Теорема 2.18 ([ 3], достатня умова оптимальності). Нехай:

- 1) функції $\varphi(x)$, $g_1(x)$, ..., $g_m(x)$ двічі диференційовні у точці $x^* \in E^n$;
- 2) для деякого нетривіального набору $\{\lambda_0^*, \lambda^*\}$ виконується умова

$$\frac{\partial L(x^*, \lambda_0^*, \lambda^*)}{\partial x} = 0;$$

- 3) квадратична форма

$$\mathcal{I} = \left(\frac{\partial^2 L(x^*, \lambda_0^*, \lambda^*)}{\partial x^2} y, y \right) > 0 \quad (< 0)$$

для всіх $y \neq 0$, що задовольняють умові $\left(\frac{\partial g_i(x^*)}{\partial x}, y \right) = 0$, $i = \overline{1, m}$.

Тоді точка x^* є точкою строгого локального мінімуму (максимуму) функції $\varphi(x)$ на R .

При дослідженні на глобальний екстремум досить часто вдається скористатися теоремою Вейерштрасса або її наслідками.

Рекомендується така схема розв'язування задачі з обмеженнями типу рівностей:

- 1) Складається функція Лагранжа.
- 2) Випишується необхідна умова $\frac{\partial L(x, \lambda_0, \lambda)}{\partial x} = 0$, до неї додаються обмеження-рівності з (2.15). З отриманої системи рівнянь знаходяться стаціонарні точки задачі (2.14)–(2.15), для яких існує нетривіальний набір множників Лагранжа. Розглядаються два випадки:
 - а) $\lambda_0 = 0$. Цей вироджений випадок зустрічається на практиці досить рідко. У більшості задач випадок $\lambda_0 = 0$ призводить до несумісності системи рівнянь. У деяких задачах вдається довести, що точки екстремуму нормальні і тоді варіант $\lambda_0 = 0$ розглядати не потрібно;
 - б) розглядається випадок $\lambda_0 = 1$;
- 3) Знайдені стаціонарні точки перевіряють на необхідну умову оптимальності другого порядку. Точки, що не задовольняють умові, відкидаються.
- 4) Ті стаціонарні точки, в яких необхідна умова другого порядку виконується, перевіряють на достатню умову оптимальності.
- 5) Знайдені точки мінімуму і максимуму досліджуються на глобальний екстремум.

Приклад 2.8. Розглянемо задачу

$$\varphi(\mathbf{x}) = 2x_1^2 + x_2^2 \rightarrow \text{extr}, \quad 2x_1 + x_2 = 1.$$

Можна спочатку перевірити нормальність задачі. Для цього скористаємось умовою нормальності. Для цього розрахуємо $\frac{\partial g(\mathbf{x})}{\partial \mathbf{x}} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$. Так як $\frac{\partial g(\mathbf{x})}{\partial \mathbf{x}} \neq 0$ для всіх припустимих точок, то всі точки нормальні і задача нормальна. Тепер можна уже скласти функцію Лагранжа.

А можна піти іншим шляхом. Складаємо узагальнену функцію Лагранжа і розглянемо випадки $\lambda_0 = 0$ і $\lambda_0 = 1$. Маємо

$$L(\mathbf{x}, \lambda_0, \lambda) = \lambda_0 (2x_1^2 + x_2^2) + \lambda(2x_1 + x_2 - 1).$$

Скористаємось правилом множників Лагранжа і отримаємо таку систему рівнянь

$$\begin{aligned} \frac{\partial L(\mathbf{x}, \lambda_0, \lambda)}{\partial x_1} &= 4\lambda_0 x_1 + 2\lambda = 0, \\ \frac{\partial L(\mathbf{x}, \lambda_0, \lambda)}{\partial x_2} &= 2\lambda_0 x_2 + \lambda = 0, \\ 2x_1 + x_2 &= 1. \end{aligned} \tag{2.16}$$

Розглянемо випадок з $\lambda_0 = 0$. Отримаємо, що $\lambda = 0$, а це суперечить твердженню теореми. Отже, $\lambda_0 \neq 0$ і тому розглянемо випадок $\lambda_0 = 1$.

Розв'язавши систему (2.16) з $\lambda_0 = 1$, отримаємо: $x_1^* = \frac{1}{3}$, $x_2^* = \frac{1}{3}$, $\lambda = -\frac{2}{3}$. Розглянемо матрицю других похідних функції Лагранжа:

$$\frac{\partial^2 L(\mathbf{x}^*, 1, -2/3)}{\partial \mathbf{x}^2} = \begin{pmatrix} 4 & 0 \\ 0 & 2 \end{pmatrix}.$$

Скористаємось необхідною умовою другого порядку. Визначимо знак квадратичної форми $\mathcal{I} = \mathbf{y}^T \frac{\partial^2 L(\mathbf{x}^*, \lambda_0, \lambda)}{\partial \mathbf{x}^2} \mathbf{y}$ на гіперплощині $\frac{\partial g^T(\mathbf{x}^*)}{\partial \mathbf{x}} \mathbf{y} = 0$. Отримаємо

$$\mathcal{I} = 4y_1^2 + 2y_2^2, \quad \frac{\partial g^T(\mathbf{x}^*)}{\partial \mathbf{x}} \mathbf{y} = 2y_1 + y_2 = 0.$$

Звідси $y_2 = -2y_1$. Тоді $\mathcal{I} = 4y_1^2 + 2y_2^2 = 12y_1^2 \geq 0$. Необхідна умова мінімуму виконується. Достатня умова теж виконується: $\varphi = 12y_1^2 > 0$ для $\mathbf{y} \neq 0$. Отже, розв'язком задачі є точка $\mathbf{x}^* = \left(\frac{1}{3}, \frac{1}{3}\right)^T$. ☑

2.4.2 Мінімізація функцій при обмеженнях типу нерівностей

Так як умови оптимальності для таких задач досить громіздкі, то розглянемо тільки задачу мінімізації.

Нехай в n -вимірному просторі визначені функції $\varphi(\mathbf{x})$, $g_1(\mathbf{x}), \dots, g_m(\mathbf{x})$. Необхідно знайти точку $\mathbf{x}^* \in E^n$, що задовольняє умовам $g_1(\mathbf{x}^*) \leq 0, \dots, g_m(\mathbf{x}^*) \leq 0$ і таку, що

$$\varphi(\mathbf{x}^*) = \min \varphi(\mathbf{x}), \quad g_i(\mathbf{x}) \leq 0, \quad i = \overline{1, m}.$$

Означення 2.20. Обмеження $g_i(\mathbf{x}) \leq 0$ називається **активним (жорстким)** у точці $\mathbf{x}^0$, якщо $g_i(\mathbf{x}^0) = 0$, і пасивним (нежорстким), якщо $g_i(\mathbf{x}^0) < 0$.

Далі обмеження будемо записувати у векторному вигляді $\mathbf{g}(\mathbf{x}) \leq 0$, де $\mathbf{g}$ — m -вимірний вектор.

Означення 2.21. Припустимо точку $\mathbf{x}^0$ назвемо **звичайною припустимою** при обмеженнях $\mathbf{g}(\mathbf{x}) \leq 0$, якщо вектори

$$\frac{\partial g_{i_1}(\mathbf{x}^0)}{\partial \mathbf{x}}, \dots, \frac{\partial g_{i_k}(\mathbf{x}^0)}{\partial \mathbf{x}}, \quad 1 \leq i_1 \leq \dots \leq i_k \leq m,$$

що відповідають обмеженням $g_{i_1}(\mathbf{x}) \leq 0, \dots, g_{i_k}(\mathbf{x}) \leq 0$, активним у точці $\mathbf{x}^0$, лінійно незалежні.

Сформулюємо необхідні і достатні умови мінімуму для задачі з обмеженнями типу нерівностей.

Теорема 2.19 (необхідна умова мінімуму). Нехай скалярна функція $\varphi(\mathbf{x})$ і m -вимірна функція $\mathbf{g}(\mathbf{x})$ визначені і неперервні разом з двома першими похідними за $\mathbf{x}$. Точка $\mathbf{x}^*$ — точка локального мінімуму — є звичайною припустимою і у цій точці активні обмеження з номерами $i_1, \dots, i_k$. Тоді

- 1) знайдеться невід'ємний множник $\boldsymbol{\lambda}^* = (\lambda_1^*, \dots, \lambda_m^*)^T \geq 0$ такий, що для функції Лагранжа $L(\mathbf{x}, \boldsymbol{\lambda}) = \varphi(\mathbf{x}) + \boldsymbol{\lambda}^T \mathbf{g}(\mathbf{x})$ виконується рівність

$$\frac{\partial L(\mathbf{x}^*, \boldsymbol{\lambda}^*)}{\partial \mathbf{x}} = 0;$$

- 2) виконуються умови доповнюючої нежорсткості

$$\lambda_i^* g_i(\mathbf{x}^*) = 0, \quad i = \overline{1, m};$$

- 3) квадратична форма $\mathcal{I} = \mathbf{y}^T \frac{\partial^2 L(\mathbf{x}^*, \boldsymbol{\lambda}^*)}{\partial \mathbf{x}^2} \mathbf{y} \geq 0$ на гіперплощині, що задається

рівняннями

$$\left(\frac{\partial g_{i_1}(\mathbf{x}^*)}{\partial \mathbf{x}}, \mathbf{y} \right) = 0, \dots, \left(\frac{\partial g_{i_k}(\mathbf{x}^*)}{\partial \mathbf{x}}, \mathbf{y} \right) = 0.$$

Означення 2.22. Припустима точка $\mathbf{x}^*$ називається **умовно-стаціонарною** у сформульованій задачі, якщо для деякого m -вимірного вектора $\boldsymbol{\lambda}^* \geq 0$ виконуються умови

$$\frac{\partial L(\mathbf{x}^*, \boldsymbol{\lambda}^*)}{\partial \mathbf{x}} = 0, \quad \lambda_i^* g_i(\mathbf{x}^*) = 0, \quad i = \overline{1, m}.$$

Теорема 2.20 (достатня умова мінімізації). Для того, щоб умовно-стаціонарна точка $\mathbf{x}^*$ була точкою локального мінімуму достатньо, щоб квадратична форма $\mathcal{I} = \mathbf{y}^\top \frac{\partial^2 L(\mathbf{x}^*, \boldsymbol{\lambda}^*)}{\partial \mathbf{x}^2} \mathbf{y}$ була додатно визначеною на гіперплощині

$$\left(\frac{\partial g_{i_1}(\mathbf{x}^*)}{\partial \mathbf{x}}, \mathbf{y} \right) = 0, \dots, \left(\frac{\partial g_{i_k}(\mathbf{x}^*)}{\partial \mathbf{x}}, \mathbf{y} \right) = 0, \quad \mathbf{y} \neq 0,$$

де $i_1, \dots, i_k$ — номери обмежень, активних у точці $\mathbf{x}^*$.

Приклад 2.9. Необхідно побудувати коло і квадрат, сума периметрів яких дорівнює a , так, щоб сума площ круга і квадрата була максимальною.

Розв'язування. Нехай довжина кола дорівнює x_1 , а периметр квадрата — x_2 . Розрахуємо радіус кола і сторону квадрата. Позначимо їх відповідно через r і y . Маємо $2\pi r = x_1$, $4y = x_2$. Тоді $r = x_1/2\pi$, $y = x_2/4$. Площа круга дорівнює $x_1^2/4\pi$, а квадрата — $x_2^2/16$.

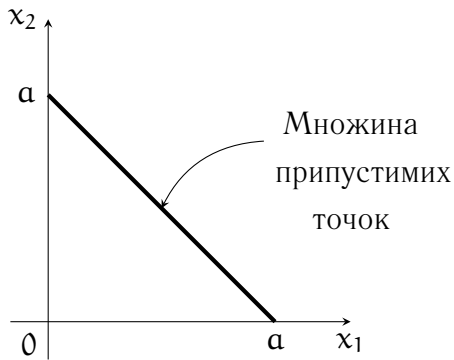
Отже, задача має таку математичну постановку:

$$\begin{aligned} \min & - \left[\frac{x_1^2}{4\pi} + \frac{x_2^2}{16} \right], \\ & x_1 + x_2 = a, \\ & -x_1 \leq 0, \\ & -x_2 \leq 0. \end{aligned} \tag{2.17}$$

Введемо функції $\varphi(\mathbf{x}) = - \left[\frac{x_1^2}{4\pi} + \frac{x_2^2}{16} \right]$, $g_1(\mathbf{x}) \equiv x_1 + x_2 - a$, $g_2(\mathbf{x}) \equiv -x_1$, $g_3(\mathbf{x}) \equiv -x_2$. Тоді задача (2.17) має вигляд

$$\begin{aligned} \min & \varphi(\mathbf{x}), \\ & g_1(\mathbf{x}) = 0, \quad g_2(\mathbf{x}) \leq 0, \quad g_3(\mathbf{x}) \leq 0, \quad \mathbf{x} \in \mathbb{R}^2. \end{aligned} \tag{2.18}$$

У задачі (2.18) присутні дві нерівності і одна чиста рівність.



Так як множина припустимих точок задачі (2.18) обмежена і замкнута, а функція $\varphi(\mathbf{x})$ є неперервною, то за теоремою Вейєрштрасса задача має розв'язок.

Вектори

$$\frac{\partial g_1(\mathbf{x})}{\partial \mathbf{x}} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \frac{\partial g_2(\mathbf{x})}{\partial \mathbf{x}} = \begin{pmatrix} -1 \\ 0 \end{pmatrix}, \quad \frac{\partial g_3(\mathbf{x})}{\partial \mathbf{x}} = \begin{pmatrix} 0 \\ -1 \end{pmatrix}$$

попарно лінійно незалежні, а активними одночасно можуть бути $g_1(\mathbf{x})$, $g_1(\mathbf{x})$ і $g_2(\mathbf{x})$ або $g_1(\mathbf{x})$ і $g_3(\mathbf{x})$. Тому кожна припустима точка є звичайною. А це означає, що усі припустимі точки нормальні. Тому можна скласти нормальну функцію Лагранжа.

$$\begin{aligned} L(\mathbf{x}, \boldsymbol{\lambda}) &= \varphi(\mathbf{x}) + \lambda_1 g_1(\mathbf{x}) + \lambda_2 g_2(\mathbf{x}) + \lambda_3 g_3(\mathbf{x}) = \\ &= -\left(\frac{x_1^2}{4\pi} + \frac{x_2^2}{16}\right) + \lambda_1 (x_1 + x_2 - a) - \lambda_2 x_1 - \lambda_3 x_2. \end{aligned}$$

Згідно з необхідною умовою знайдуться $\lambda_1, \lambda_2, \lambda_3$; $\lambda_i \geq 0$, $i = 1, 2, 3$ такі, що

$$\begin{aligned} \frac{\partial L(\mathbf{x}, \boldsymbol{\lambda})}{\partial x_1} &= -\frac{x_1}{2\pi} + \lambda_1 - \lambda_2 = 0, \\ \frac{\partial L(\mathbf{x}, \boldsymbol{\lambda})}{\partial x_2} &= -\frac{x_2}{8} + \lambda_1 - \lambda_3 = 0, \\ g_1(\mathbf{x}) &\equiv x_1 + x_2 - a = 0, \\ \lambda_2 g_2(\mathbf{x}) &\equiv -\lambda_2 x_1 = 0, \\ \lambda_3 g_3(\mathbf{x}) &\equiv -\lambda_3 x_2 = 0. \end{aligned}$$

Виключимо змінні $x_2, \lambda_2, \lambda_3$ і отримаємо два рівняння

$$\begin{cases} \left(\lambda_1 - \frac{x_1}{2\pi}\right) x_1 = 0, \\ \left(\lambda_1 - \frac{a - x_1}{8}\right) (a - x_1) = 0. \end{cases}$$

Розв'язавши цю систему, отримаємо:

$$\begin{aligned} 1) \quad x_1^* &= \frac{\pi a}{4 + \pi}, \quad x_2^* = \frac{4a}{4 + \pi}, \quad \lambda_1 = \frac{a}{8 + 2\pi}, \quad \lambda_2 = 0, \quad \lambda_3 = 0; \\ 2) \quad x_1^* &= a, \quad x_2^* = 0, \quad \lambda_1 = \frac{a}{2\pi}, \quad \lambda_2 = 0, \quad \lambda_3 = \frac{a}{2\pi}; \end{aligned}$$

$$3) \ x_1^* = 0, x_2^* = a, \lambda_1 = \frac{a}{8}, \lambda_2 = \frac{a}{8}, \lambda_3 = 0.$$

Квадратична форма має вигляд

$$\mathcal{I} = \mathbf{x}^\top \frac{\partial^2 L(\mathbf{x}^*, \boldsymbol{\lambda})}{\partial \mathbf{x}^2} \mathbf{x} = -\frac{1}{2\pi} x_1^2 - \frac{1}{8} x_2^2.$$

Розглянемо кожну з умовно-стаціонарних точок.

- 1) У першій точці активним є тільки обмеження $g_1(\mathbf{x}) = 0$. Гіперплощина $\frac{\partial g_1(\mathbf{x}^*)}{\partial \mathbf{x}} \mathbf{x} = 0$ має вигляд $x_1 + x_2 = 0$. Тоді $x_1 = -x_2$. У першій точці квадратична форма матиме вигляд: $\mathcal{I} = -\left[\frac{1}{2\pi} + \frac{1}{8}\right] x_2^2$ і $\mathcal{I} \leq 0$. Так як необхідна умова мінімуму не виконується, то ця точка не може бути розв'язком задачі.
- 2) Перше і третє обмеження активні у другій точці. Гіперплощина визначається системою

$$\begin{cases} x_1 + x_2 = 0, \\ -x_2 = 0. \end{cases}$$

Так як $x_1 = x_2 = 0$, то $\mathcal{I} = 0$. Необхідна умова мінімуму виконується.

- 3) У третій точці активними є перше і друге обмеження. Тоді отримаємо таку систему

$$\begin{cases} x_1 + x_2 = 0, \\ -x_1 = 0, \end{cases}$$

що визначає гіперплощину. Маємо: $x_1 = x_2 = 0$, а значить $\mathcal{I} = 0$. У третій точці теж виконується необхідна умова мінімуму.

Отримали, що у двох припустимих точках виконується необхідна умова мінімуму, але у жодній із них не виконується достатня. Так як задача має розв'язок і маємо дві точки, в яких може досягатися мінімум, то порівняємо значення функції у цих точках і оберемо розв'язок. У другій точці значення функції дорівнює $-\frac{a^2}{4\pi} \approx -0.0796a^2$, а у третій $-\frac{a^2}{16} \approx -0.0625a^2$. Так як $-0.0796 < -0.0625$, то розв'язком задачі є друга точка. Це означає, що необхідно побудувати коло. 

2.5 Мінімізація опуклих функцій

Розділ нелінійного програмування, в якому задачі мінімізації формуються для опуклих функцій і опуклих множин, називається опуклим програмуванням.

Розглянемо основну теорему опуклого програмування для задачі

$$\begin{aligned} \min \varphi(\mathbf{x}) \\ \mathbf{g}(\mathbf{x}) \leq 0, \quad \mathbf{x} \in \mathbf{R} \subset \mathbb{E}^n, \end{aligned} \quad (2.19)$$

де $\varphi(\mathbf{x})$ і компоненти m -вимірної функції $\mathbf{g}(\mathbf{x})$ є опуклі функції на опуклій множині $\mathbf{R}$.

Означення 2.23. Обмеження задачі (2.19) задовольняють *умові Слейтера*, якщо існує вектор $\mathbf{x}^0 \in \mathbf{R}$ такий, що $\mathbf{g}(\mathbf{x}^0) < 0$.

Розглянемо нормальну функцію Лагранжа для задачі опуклого програмування.

Означення 2.24. Сукупність $(\mathbf{x}^0, \boldsymbol{\lambda}^0)$ векторів $\mathbf{x}^0 \in \mathbf{R}$, $\boldsymbol{\lambda}^0 \geq 0$ називається *сідловою точкою функції Лагранжа*, якщо для всіх $\mathbf{x} \in \mathbf{R}$ і $\boldsymbol{\lambda} \geq 0$ виконується нерівність

$$L(\mathbf{x}^0, \boldsymbol{\lambda}) \leq L(\mathbf{x}^0, \boldsymbol{\lambda}^0) \leq L(\mathbf{x}, \boldsymbol{\lambda}^0).$$

Теорема 2.21 (Куна — Таккера). Нехай обмеження (2.19) задовольняють умові Слейтера. Тоді для того щоб вектор $\mathbf{x}^0 \in \mathbf{R}$ був розв'язком основної задачі опуклого програмування необхідно і достатньо, щоб існував вектор $\boldsymbol{\lambda}^0 \geq 0$ такий, що пара $(\mathbf{x}^0, \boldsymbol{\lambda}^0)$ була сідловою точкою функції Лагранжа $L(\mathbf{x}, \boldsymbol{\lambda}) = \varphi(\mathbf{x}) + \boldsymbol{\lambda}^T \mathbf{g}(\mathbf{x})$.

Доведення.

Необхідність. В $(m+1)$ -вимірному просторі побудуємо множини

$$A = \{\bar{\mathbf{y}} = (y_0, \mathbf{y}) \mid y_0 \geq \varphi(\mathbf{x}), \mathbf{y} \geq \mathbf{g}(\mathbf{x}) \text{ при деякому } \mathbf{x} \in \mathbf{R}\},$$

$$B = \{\bar{\mathbf{y}} = (y_0, \mathbf{y}) \mid y_0 = \varphi(\mathbf{x}), \mathbf{y} = \mathbf{g}(\mathbf{x}) \text{ при деякому } \mathbf{x} \in \mathbf{R}\},$$

$$Z = \{\bar{\mathbf{z}} = (z_0, \mathbf{z}) \mid z_0 < \varphi(\mathbf{x}^0), \mathbf{z} < \mathbf{0}\}.$$

Множини A і Z — опуклі. Дійсно, нехай $(y_0^1, \mathbf{y}^1) \in A$ і $(y_0^2, \mathbf{y}^2) \in A$. Тоді за означенням множини A знайдуться вектори $\mathbf{x}^1 \in \mathbf{R}$ і $\mathbf{x}^2 \in \mathbf{R}$ такі, що

$$y_0^1 \geq \varphi(\mathbf{x}^1), \quad \mathbf{y}^1 \geq \mathbf{g}(\mathbf{x}^1), \quad (2.20)$$

$$y_0^2 \geq \varphi(\mathbf{x}^2), \quad \mathbf{y}^2 \geq \mathbf{g}(\mathbf{x}^2). \quad (2.21)$$

Побудуємо $\lambda y_0^1 + (1 - \lambda)y_0^2$, отримаємо

$$\lambda y_0^1 + (1 - \lambda)y_0^2 \geq \lambda \varphi(\mathbf{x}^1) + (1 - \lambda)\varphi(\mathbf{x}^2) \geq \varphi(\lambda \mathbf{x}^1 + (1 - \lambda)\mathbf{x}^2) = \varphi(\mathbf{x}(\lambda)),$$

так як функція $\varphi(\mathbf{x})$ опукла. Далі,

$$\lambda \mathbf{y}^1 + (1 - \lambda) \mathbf{y}^2 \geq \lambda \mathbf{g}(\mathbf{x}^1) + (1 - \lambda) \mathbf{g}(\mathbf{x}^2) \geq \mathbf{g}(\lambda \mathbf{x}^1 + (1 - \lambda) \mathbf{x}^2) = \mathbf{g}(\mathbf{x}(\lambda))$$

в силу опуклості функції $\mathbf{g}(\mathbf{x})$. Так як $\mathbf{x}(\lambda) = \lambda \mathbf{x}^1 + (1 - \lambda) \mathbf{x}^2 \in \mathbf{R}$, то отримані нерівності доводять, що точка $(\lambda \mathbf{y}_0^1 + (1 - \lambda) \mathbf{y}_0^2, \lambda \mathbf{y}^1 + (1 - \lambda) \mathbf{y}^2)$ належить множині $\mathbf{A}$. Отже, множина $\mathbf{A}$ — опукла.

Розглянемо множину $\mathbf{Z}$. Візьмемо дві точки з цієї множини: $\bar{\mathbf{z}}^1$ і $\bar{\mathbf{z}}^2$. Маємо

$$\bar{\mathbf{z}}^1 = (\mathbf{z}_0^1, \mathbf{z}^1), \quad \bar{\mathbf{z}}^2 = (\mathbf{z}_0^2, \mathbf{z}^2) \quad \text{і} \quad \mathbf{z}_0^1 < \varphi(\mathbf{x}^0), \quad \mathbf{z}^1 < \mathbf{0}, \quad \mathbf{z}_0^2 < \varphi(\mathbf{x}^0), \quad \mathbf{z}^2 < \mathbf{0}.$$

Тоді

$$\lambda \bar{\mathbf{z}}^1 + (1 - \lambda) \bar{\mathbf{z}}^2 = (\lambda \mathbf{z}_0^1 + (1 - \lambda) \mathbf{z}_0^2, \lambda \mathbf{z}^1 + (1 - \lambda) \mathbf{z}^2)$$

і

$$\begin{aligned} \lambda \mathbf{z}_0^1 + (1 - \lambda) \mathbf{z}_0^2 &< \lambda \varphi(\mathbf{x}^0) + (1 - \lambda) \varphi(\mathbf{x}^0) = \varphi(\mathbf{x}^0), \\ \lambda \mathbf{z}^1 + (1 - \lambda) \mathbf{z}^2 &< \lambda \cdot \mathbf{0} + (1 - \lambda) \cdot \mathbf{0} = \mathbf{0}, \end{aligned}$$

а це і означає, що множина $\mathbf{Z}$ опукла.

Множини $\mathbf{A}$ і $\mathbf{Z}$ не мають спільних точок. Дійсно, нехай вони мають спільну точку $\bar{\mathbf{y}} = \bar{\mathbf{z}}$. Тоді якщо $\bar{\mathbf{y}} \in \mathbf{A}$ і $\bar{\mathbf{z}} \in \mathbf{Z}$, то при деякому $\bar{\mathbf{x}} \in \mathbf{R}$ будуть виконуватися нерівності

$$\begin{aligned} \varphi(\bar{\mathbf{x}}) &\leq \mathbf{y}_0 = \mathbf{z}_0 < \varphi(\mathbf{x}^0), \\ \mathbf{g}(\bar{\mathbf{x}}) &\leq \mathbf{y} = \mathbf{z} < \mathbf{0}, \end{aligned}$$

а це означає, що точка $\mathbf{x}^0$ не є точкою мінімуму.

За теоремою про відокремлення опуклих множин (теорема 2.3) множини $\mathbf{A}$ і $\mathbf{Z}$ можна відокремити, тобто знайдеться $(m + 1)$ -вимірний вектор $\bar{\mathbf{c}} = (\mathbf{c}_0, \mathbf{c})$, $\|\bar{\mathbf{c}}\| = 1$ такий, що

$$\bar{\mathbf{c}}^\top \bar{\mathbf{y}} \geq \bar{\mathbf{c}}^\top \bar{\mathbf{z}} \tag{2.22}$$

для всіх $\bar{\mathbf{y}} \in \mathbf{A}$, $\bar{\mathbf{z}} \in \mathbf{Z}$. З (2.22) і визначення множини $\mathbf{Z}$ отримаємо, що вектор $\bar{\mathbf{c}}$ невід'ємний, $\bar{\mathbf{c}} \geq \mathbf{0}$. Якщо припустити, що існує від'ємна координата $\mathbf{c}_i < 0$, то для вектора

$$\bar{\mathbf{z}} = (\varphi(\mathbf{x}^0) + \varepsilon, \varepsilon, \dots, \varepsilon, -\beta_i^2, \varepsilon, \dots, \varepsilon)$$

з фіксованим $\varepsilon < 0$ і великим β отримаємо, що число $\bar{\mathbf{c}}^\top \bar{\mathbf{z}}$ може бути досить великим, що

суперечить (2.22).

Множина B належить множині A , тому нерівність (2.22) виконується і при $\bar{y} \in B$, тобто

$$c_0 \varphi(x) + c^T g(x) \geq \bar{c}^T \bar{z} \quad \text{для всіх } x \in R. \quad (2.23)$$

Так як нерівність (2.23) виконується для всіх точок множини Z , вона буде виконуватись і для граничних точок множини Z , зокрема, для точки $(\varphi(x^0), 0)$:

$$c_0 \varphi(x) + c^T g(x) \geq c_0 \varphi(x^0) \quad \text{для всіх } x \in R. \quad (2.24)$$

Доведемо, що $c_0 > 0$. Якщо $c_0 = 0$, то з $\bar{c} \geq 0$, $\bar{c} \neq 0$ отримаємо, що $c \geq 0$, $c \neq 0$, і тому згідно з (2.24)

$$c^T g(x) \geq 0 \quad \text{для всіх } x \in R. \quad (2.25)$$

За умовою Слейтера знайдеться вектор $x^* \in R$ такий, що $g(x^*) < 0$. Тому $c^T g(x^*) < 0$, що суперечить (2.25).

Покладемо $\lambda^0 = \frac{c}{c_0}$. Ясно, що $\lambda^0 \geq 0$. Тоді з (2.24) отримаємо

$$\varphi(x) + \lambda_0^T g(x) \geq \varphi(x^0) \quad \text{для всіх } x \in R. \quad (2.26)$$

Нехай $x = x^0$, отримаємо $\lambda_0^T g(x^0) \geq 0$. Але з іншого боку $\lambda^0 \geq 0$ і $g(x^0) \leq 0$, тоді $\lambda_0^T g(x^0) \leq 0$. Маємо

$$\lambda_0^T g(x^0) = 0. \quad (2.27)$$

Нерівність (2.26) та рівність (2.27) дають

$$\varphi(x^0) + \lambda_0^T g(x^0) \leq \varphi(x) + \lambda_0^T g(x) \quad \text{для всіх } x \in R. \quad (2.28)$$

Так як $g(x^0) \leq 0$, то для всіх $\lambda \geq 0$ виконується нерівність $\lambda^T g(x^0) \leq 0$. Використовуючи (2.27), отримаємо

$$\varphi(x^0) + \lambda^T g(x^0) \leq \varphi(x^0) = \varphi(x^0) + \lambda_0^T g(x^0) \quad \text{для всіх } \lambda \geq 0. \quad (2.29)$$

Поєднуючи (2.28) і (2.29), маємо, що пара (x^0, λ^0) є сідловою точкою функції Лагранжа.

Достатність. Маємо, що (x^0, λ^0) — сідлова точка функції Лагранжа, тобто

$$\varphi(x^0) + \lambda^T g(x^0) \leq \varphi(x^0) + \lambda_0^T g(x^0) \leq \varphi(x) + \lambda_0^T g(x), \quad x \in R, \quad \lambda \geq 0. \quad (2.30)$$

Ліва частина нерівності дає

$$\lambda^\top g(x^0) \leq \lambda_0^\top g(x^0). \quad (2.31)$$

З (2.31) отримаємо, що $g(x^0) \leq 0$. Дійсно, якщо при деякому i ($1 \leq i \leq m$) $g_i(x^0) > 0$, то покладемо $\lambda_j = 0$, $j \neq i$, $j = \overline{1, m}$, беремо λ_i досить великим і отримаємо що ліва частина (2.31) буде як завгодно великою, що порушує (2.31).

З (2.31) зрозуміло, що виконується

$$\lambda_0^\top g(x^0) = 0. \quad (2.32)$$

Дійсно, якщо припустити, що $\lambda_0^\top g(x^0) < 0$, то візьмемо $\lambda = \frac{1}{2}\lambda_0$ і з (2.31) отримаємо суперечливу нерівність $0 \leq \frac{1}{2}\lambda_0^\top g(x^0) < 0$. З урахуванням (2.32) і правої частини (2.30) отримаємо

$$\varphi(x^0) \leq \varphi(x) + \lambda_0^\top g(x).$$

Звідси маємо, що для всіх $x \in R$, що задовольняють нерівності $g(x) \leq 0$, повинна виконуватись нерівність

$$\varphi(x^0) \leq \varphi(x), \quad x \in R.$$

Отже, x^0 — розв'язок задачі.

Теорему доведено. ✓

Зауваження 2.1. Якщо функція $g(x)$ лінійна, то теорема справедлива і без умови Слейтера.

2.5.1 Задача лінійного програмування

Розділ опуклого програмування, в якому використовуються тільки лінійні функції, називається лінійним програмуванням.

Основною задачею лінійного програмування є задача:

$$\begin{aligned} \min (c, x) \\ x \in R = \{x \in E^n \mid Ax \geq b, x \geq 0\}, \end{aligned} \quad (2.33)$$

де A — $m \times n$ -матриця, b — m -вектор, c — n -вектор.

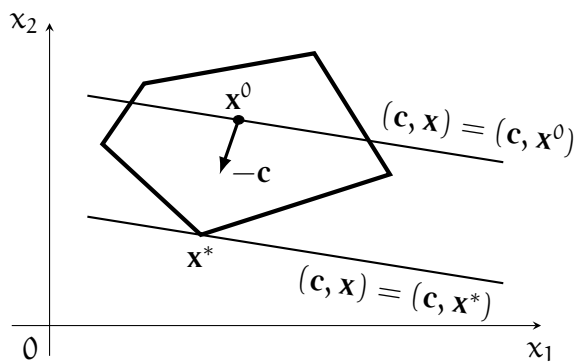
Матриця $A = [a_1, \dots, a_n]$ називається матрицею умов задачі. Її стовпчики a_i , $i = \overline{1, n}$ називаються векторами умов задачі. Вектор b називається вектором обмежень.

Задачею, двоїстою до задачі (2.33), називається така задача:

$$\begin{aligned} & \max (\mathbf{b}, \mathbf{y}) \\ & \mathbf{y} \in Q = \{ \mathbf{y} \in E^m \mid \mathbf{A}^T \mathbf{y} \leq \mathbf{c}, \mathbf{y} \geq 0 \}. \end{aligned} \quad (2.34)$$

Легко довести, що задача (2.33) є двоїстою до задачі (2.34).

У просторі E^n множина R — це перетин напівпросторів $(\mathbf{A}\mathbf{x})_i \geq b_i, i = \overline{1, m}, x_j \geq 0, j = \overline{1, n}$. Перетин напівпросторів і дає нам множину R . Лінії постійного рівня мають вигляд $(\mathbf{c}, \mathbf{x}) = \lambda$.



Розглянемо довільну точку $\mathbf{x}^0 \in R$.

Їй відповідає значення цільової функції $\lambda_0 = (\mathbf{c}, \mathbf{x}^0)$. Якщо зміщувати гіперплощину $(\mathbf{c}, \mathbf{x}) = \lambda$ у напрямку зменшення функції, а це є антиградієнт $(-\mathbf{c})$, до тих пір, поки вона останній раз не торкнеться множини R у точці $\mathbf{x}^*$, то ця точка $\mathbf{x}^*$ і буде розв'язком задачі.

Основні теореми лінійного програмування

Теорема 2.22 (двоїстості). *Пряма і двоїста задачі або обидві мають розв'язок $\mathbf{x}^*$ і $\mathbf{y}^*$, і тоді $(\mathbf{c}, \mathbf{x}^*) = (\mathbf{b}, \mathbf{y}^*)$, або обидві не мають розв'язку.*

Теорема 2.23. *Для довільних припустимих $\mathbf{x} \in R$ і $\mathbf{y} \in Q$ виконується нерівність $(\mathbf{c}, \mathbf{x}) \geq (\mathbf{b}, \mathbf{y})$.*

Теорема 2.24. *Якщо $\mathbf{x}^* \in R$, а $\mathbf{y}^* \in Q$ і $(\mathbf{c}, \mathbf{x}^*) = (\mathbf{b}, \mathbf{y}^*)$, то $\mathbf{x}^*$ і $\mathbf{y}^*$ — оптимальні розв'язки прямої і двоїстої задач.*

Теорема 2.25. *Для існування розв'язків однієї з двоїстих задач (тобто обох) необхідно і достатньо, щоб $R \neq \emptyset, Q \neq \emptyset$.*

Теорема 2.26. *Якщо $(\mathbf{c}, \mathbf{x}^*) = \min (\mathbf{c}, \mathbf{x})$ і якщо існують точки $\mathbf{x}^1, \dots, \mathbf{x}^M$ такі, що $\mathbf{x}^i \in R, i = \overline{1, M}, \mathbf{x}^* = \sum_{i=1}^M \alpha_i \mathbf{x}^i$, де $\sum_{i=1}^M \alpha_i = 1, \alpha_i > 0, i = \overline{1, M}$, то*

$$(\mathbf{c}, \mathbf{x}^*) = (\mathbf{c}, \mathbf{x}^1) = \dots = (\mathbf{c}, \mathbf{x}^M).$$

Означення 2.25. *Точка $\mathbf{x}$ множини R називається **кутовою точкою**, якщо в R не існують точки $\mathbf{x}^1$ і $\mathbf{x}^2, \mathbf{x}^1 \neq \mathbf{x}^2$ такі, що $\mathbf{x} = \alpha \mathbf{x}^1 + (1 - \alpha) \mathbf{x}^2$ для деякого $\alpha \in (0, 1)$.*

Теорема 2.27. Якщо основна задача лінійного програмування має розв'язок $\mathbf{x}^*$, то існує кутова точка $\bar{\mathbf{x}} \in \mathbf{R}$, в якій $(\mathbf{c}, \bar{\mathbf{x}}) = \min (\mathbf{c}, \mathbf{x}) = (\mathbf{c}, \mathbf{x}^*)$.

Отже, остання теорема стверджує, що розв'язок задачі необхідно шукати серед кутових точок. Саме на цій ідеї і побудований симплекс-метод. Це метод пошуку розв'язку основної задачі лінійного програмування.

З алгоритмом симплекс-методу, його модифікаціями, особливими випадками можна ознайомитись у [3], [6], [20].

Наостанок розглянемо основну задачу лінійного програмування у загальному вигляді і для неї напишемо двоїсту задачу.

Введемо позначення: $I = \{i \mid 1, 2, \dots, m\}$, $J = \{j \mid 1, 2, \dots, n\}$. Нехай I_1 означає частину сукупності індексів I , $I_1 \subseteq I$, а $I_2 = I \setminus I_1$. Аналогічно, $J_1 \subseteq J$, $J_2 = J \setminus J_1$.

Нехай основна задача лінійного програмування має вигляд

$$\begin{aligned} \min (\mathbf{c}, \mathbf{x}) \\ (\mathbf{Ax})_i &\geq b_i, \quad i \in I_1, \\ (\mathbf{Ax})_i &= b_i, \quad i \in I_2, \\ x_j &\geq 0, \quad j \in J_1. \end{aligned}$$

Тоді двоїста задача має такий вигляд:

$$\begin{aligned} \max (\mathbf{b}, \mathbf{y}) \\ (\mathbf{A}^T \mathbf{y})_j &\leq c_j, \quad j \in J_1, \\ (\mathbf{A}^T \mathbf{y})_j &= c_j, \quad j \in J_2, \\ y_i &\geq 0, \quad i \in I_1. \end{aligned}$$

2.6 Завдання для самостійної роботи

2.6.1 Задачі нелінійного програмування. Аналітичні методи

Завдання 1. Безумовна мінімізація функцій

Дослідіть задану функцію на екстремум.

Варіант 1 $\varphi(\mathbf{x}) = x_1^3 + x_2^2 + x_3^3 + 12x_1x_2 + 2x_3, \quad \mathbf{x} \in \mathbb{R}^3$

Варіант 2 $\varphi(\mathbf{x}) = x_1^2 + 2x_2^2 + 5x_3^2 - 2x_1x_2 - 4x_2x_3 - 2x_3, \quad \mathbf{x} \in \mathbb{R}^3$

Варіант 3 $\varphi(\mathbf{x}) = (x_1 - x_2 + 1)^2, \quad \mathbf{x} \in \mathbb{R}^2$

Варіант 4 $\varphi(\mathbf{x}) = x_1 x_2 (a - x_1 - x_2), \quad \mathbf{x} \in \mathbb{R}^2, a \in \mathbb{R}$

Варіант 5 $\varphi(\mathbf{x}) = \frac{x_1}{1 + x_1^2 + x_2^2}, \quad \mathbf{x} \in \mathbb{R}^2$

Варіант 6 $\varphi(\mathbf{x}) = e^{2x_1} (x_1 + x_2^2 + 2x_2), \quad \mathbf{x} \in \mathbb{R}^2$

Варіант 7 $\varphi(\mathbf{x}) = \frac{x_1^4}{32} + x_1^2 x_2^2 - x_1 - x_2^2, \quad \mathbf{x} \in \mathbb{R}^2$

Варіант 8 $\varphi(\mathbf{x}) = x_1^2 - x_1 - x_1 x_2 + x_2^2 - x_2 + x_3^4 - 4x_3, \quad \mathbf{x} \in \mathbb{R}^3$

Варіант 9 $\varphi(\mathbf{x}) = x_2^3 - 12x_2 + x_1^2 e^{x_2}, \quad \mathbf{x} \in \mathbb{R}^2$

Варіант 10 $\varphi(\mathbf{x}) = 3x_1^2 - 12x_2 - 2(x_1 - 2x_2)^3, \quad \mathbf{x} \in \mathbb{R}^2$

Варіант 11 $\varphi(\mathbf{x}) = x_2^2 x_1 - 3x_1 x_2 + 2x_1^4, \quad \mathbf{x} \in \mathbb{R}^2$

Варіант 12 $\varphi(\mathbf{x}) = x_1^2 + 4x_2^3 - 2x_1 x_2, \quad \mathbf{x} \in \mathbb{R}^2$

Варіант 13 $\varphi(\mathbf{x}) = x_1^2 - x_2^2 + 2e^{-x_1^2}, \quad \mathbf{x} \in \mathbb{R}^2$

Варіант 14 $\varphi(\mathbf{x}) = \sin x_2 - x_1^2, \quad \mathbf{x} \in \mathbb{R}^2$

Варіант 15 $\varphi(\mathbf{x}) = x_1^4 + x_2^4 - (x_1 + x_2)^4, \quad \mathbf{x} \in \mathbb{R}^2$

Варіант 16 $\varphi(\mathbf{x}) = x_1^2 + x_2^2 + x_3^2 - x_1 x_2 + x_1 - 2x_3, \quad \mathbf{x} \in \mathbb{R}^3$

Варіант 17 $\varphi(\mathbf{x}) = x_1^2 + x_2^2 + 2x_3^2 + x_1 x_2 + 2x_1 x_3 + 3x_2 x_3 - x_1, \quad \mathbf{x} \in \mathbb{R}^3$

Варіант 18 $\varphi(\mathbf{x}) = 3x_1 x_2 - x_1^2 x_2 - x_1 x_2^2, \quad \mathbf{x} \in \mathbb{R}^2$

Варіант 19 $\varphi(\mathbf{x}) = x_1 x_2 + \frac{50}{x_1} + \frac{20}{x_2}, \quad \mathbf{x} \in \mathbb{R}^2$

Варіант 20 $\varphi(\mathbf{x}) = (x_2^2 - x_1)^2, \quad \mathbf{x} \in \mathbb{R}^2$

Варіант 21 $\varphi(\mathbf{x}) = e^{x_1 - x_2} (5 - 2x_1 + x_2), \quad \mathbf{x} \in \mathbb{R}^2$

Варіант 22 $\varphi(\mathbf{x}) = x_1^2 + 4x_2^3 - 2x_1x_2, \quad \mathbf{x} \in \mathbb{R}^2$

Варіант 23 $\varphi(\mathbf{x}) = x_1^4 + x_2^4 - 2x_1^2 - 4x_1x_2 - 2x_2^2, \quad \mathbf{x} \in \mathbb{R}^2$

Варіант 24 $\varphi(\mathbf{x}) = x_1^3 + x_2^2 + x_3^2 + x_2x_3 - 3x_1 + 6x_2 + 2, \quad \mathbf{x} \in \mathbb{R}^3$

Варіант 25 $\varphi(\mathbf{x}) = 2x_1^3 + 4x_1x_2^2 - 10x_1x_2 + x_2^2, \quad \mathbf{x} \in \mathbb{R}^2$

Завдання 2. Умовна мінімізація функцій

Дослідіть задану функцію на екстремум у вказаній області¹ за допомогою методу множників Лагранжа.

Варіант 1

- 1) $\varphi(\mathbf{x}) = x_1x_2$ на кривій $3x_1^2 + x_2^2 = 6$;
- 2) $\varphi(\mathbf{x}) = 2(x_1 - 1)^2 + 3(x_2 - 3)^2$ в області $D = \{\mathbf{x} \in \mathbb{R}^2 \mid x_1 + x_2 = 6, x_1 \geq 0, x_2 \geq 0\}$.

Варіант 2

- 1) $\varphi(\mathbf{x}) = \frac{1}{x_1} + \frac{1}{x_2}$ на кривій $\frac{1}{x_1^2} + \frac{1}{x_2^2} = 4$;
- 2) $\varphi(\mathbf{x}) = x_1^2 - x_2^2$ в області $D = \{\mathbf{x} \in \mathbb{R}^2 \mid x_1^2 + x_2^2 \leq 16, x_1 - x_2 = 4\}$.

Варіант 3

- 1) $\varphi(\mathbf{x}) = x_1 + x_2 + 2x_3$ на поверхні $x_1^2 + x_2^2 + x_3^2 = 3$;
- 2) $\varphi(\mathbf{x}) = (x_1 - 3)^2 + (x_2 - 5)^2$ в області $D = \{\mathbf{x} \in \mathbb{R}^2 \mid x_1^2 + x_2^2 \leq 10, x_2 - 2x_1 = 5\}$.

Варіант 4

- 1) $\varphi(\mathbf{x}) = 3x_1^2 + 2x_2^2 - 3x_1 + 1$ на кривій $x_1^2 + x_2^2 = 4$;
- 2) $\varphi(\mathbf{x}) = (x_1 - 2)^2 + x_2^2$ в області $D = \{\mathbf{x} \in \mathbb{R}^2 \mid x_1^2 + x_2^2 = 1, x_1 \geq 0\}$.

¹Символом $\mathbf{x}$ позначені вектори з простору $\mathbb{R}^n$, символами $x_1, x_2, \dots, x_n$ — компоненти цих векторів.

Варіант 5

- 1) $\varphi(\mathbf{x}) = e^{x_1 x_2}$ на прямій $x_1 + x_2 = 1$;
- 2) $\varphi(\mathbf{x}) = (x_1 - 2)^2 + x_2^2$ в області $D = \{\mathbf{x} \in \mathbb{R}^2 \mid x_1^2 + x_2^2 \leq 1, x_1 \geq 0\}$.

Варіант 6

- 1) $\varphi(\mathbf{x}) = x_1^3 x_2^5$ на прямій $x_1 + x_2 = 8$;
- 2) $\varphi(\mathbf{x}) = 3x_1^2 + 3x_2^2 - 2x_1 - 2x_2 + 2$ в області $D = \{\mathbf{x} \in \mathbb{R}^2 \mid x_1 + x_2 \leq 4, x_1 \geq 0, x_2 \geq 0\}$.

Варіант 7

- 1) $\varphi(\mathbf{x}) = 2x_1 + 3x_2^2 + x_3^2$ на площині $x_1 + x_2 + x_3 = 8$;
- 2) $\varphi(\mathbf{x}) = x_1 x_2^2 + x_1^2 x_2 - 3x_1^2 - 3x_2^2$ в області $D = \{\mathbf{x} \in \mathbb{R}^2 \mid x_1 + x_2 \geq 1, x_1 + x_2 \leq 16\}$.

Варіант 8

- 1) $\varphi(\mathbf{x}) = \frac{1}{2} [(x_1 - 1)^2 + x_2^2]$ на кривій $-x_1 + x_2^2 = 0$;
- 2) $\varphi(\mathbf{x}) = (x_1 + 2)^2 + (x_2 - 2)^2$ в області $D = \{\mathbf{x} \in \mathbb{R}^2 \mid x_1^2 + x_2^2 \leq 1, x_1 \geq 0, x_2 \leq 0\}$.

Варіант 9

- 1) $\varphi(\mathbf{x}) = -4x_1^2 - 4x_1 - x_2^2 + 8x_2 - 5$ на прямій $2x_1 - x_2 = 6$;
- 2) $\varphi(\mathbf{x}) = x_1 + x_2 + x_3$ в області $D = \{\mathbf{x} \in \mathbb{R}^3 \mid x_1^2 + x_2^2 - x_3 \leq 0, x_3 \leq 2\}$.

Варіант 10

- 1) $\varphi(\mathbf{x}) = x_1 + x_2 + x_3$ на поверхні $\frac{1}{x_1} + \frac{1}{x_2} + \frac{1}{x_3} = 1$;
- 2) $\varphi(\mathbf{x}) = 3x_1^2 + 4x_1 x_2 + 5x_2^2$ в області $D = \{\mathbf{x} \in \mathbb{R}^2 \mid x_1 + x_2 = 4, x_1 \geq 0\}$.

Варіант 11

- 1) $\varphi(\mathbf{x}) = x_1^2 - x_2^2$ на поверхні $x_1^2 + 2x_2^2 + 3x_3^2 = 1$;
- 2) $\varphi(\mathbf{x}) = (x_1 + 2)^2 + (x_2 + 2)^2 - 10$ в області $D = \{\mathbf{x} \in \mathbb{R}^2 \mid x_1^2 + x_2^2 = 1, x_1 \geq 0, x_2 \geq 0\}$.

Варіант 12

- 1) $\varphi(\mathbf{x}) = x_1 x_2 x_3$ на поверхні $x_1^2 + 2x_2^2 + 4x_3^2 = 9$;
- 2) $\varphi(\mathbf{x}) = (x_1 + 3)^2 + (x_2 + 3)^2 + 2$ в області $D = \{\mathbf{x} \in \mathbb{R}^2 \mid x_1^2 + x_2^2 = 1, x_2 \geq 0\}$.

Варіант 13

- 1) $\varphi(\mathbf{x}) = x_1^2 + 12x_1 x_2 + x_2^2$ на кривій $4x_1^2 + x_2^2 = 25$;

2) $\varphi(\mathbf{x}) = (x_1 + 3)^2 + (x_2 + 3)^2 + 2$ в області $D = \{\mathbf{x} \in \mathbb{R}^2 \mid x_1^2 + x_2^2 \leq 1, x_1 \leq 0\}$.

Варіант 14

1) $\varphi(\mathbf{x}) = e^{x_1 x_2}$ на кривій $x_1^3 + x_2^3 = 16$;

2) $\varphi(\mathbf{x}) = (x_1 + 4)^2 + (x_2 - 4)^2 + 1$ в області $D = \{\mathbf{x} \in \mathbb{R}^2 \mid 2x_1 - x_2 \leq 2, x_1 \geq 0, x_2 \geq 0\}$.

Варіант 15

1) $\varphi(\mathbf{x}) = x_1^2 + x_2^2 + x_3^2$ на поверхні $x_3 = 1 + 2x_1 x_2$;

2) $\varphi(\mathbf{x}) = x_1 - 2x_2^2 + 4x_2$ в області $D = \{\mathbf{x} \in \mathbb{R}^2 \mid 3x_1 + 2x_2 = -6, x_1 \leq 1\}$.

Варіант 16

1) $\varphi(\mathbf{x}) = x_1^2 + x_2^2$ на кривій $x_1^4 + x_2^4 = 2$;

2) $\varphi(\mathbf{x}) = x_1^2 + (x_2 - 4)^2$ в області $D = \{\mathbf{x} \in \mathbb{R}^2 \mid x_1^2 + x_2^2 \leq 16, x_1^2 + x_2^2 \geq 4\}$.

Варіант 17

1) $\varphi(\mathbf{x}) = x_1^2 + x_2$ на кривій $x_1^2 - x_2^2 = 1$;

2) $\varphi(\mathbf{x}) = x_1 x_2$ в області $D = \{\mathbf{x} \in \mathbb{R}^2 \mid x_1^2 + x_2^2 \leq 16, x_1 \geq 0, x_2 \geq 0\}$.

Варіант 18

1) $\varphi(\mathbf{x}) = x_1^2 + x_2^2 + x_3^2$ на поверхні $x_1 x_2 x_3 = 1$;

2) $\varphi(\mathbf{x}) = x_1 x_2$ в області $D = \{\mathbf{x} \in \mathbb{R}^2 \mid x_1^2 + x_2^2 \leq 16, x_1 \leq 0\}$.

Варіант 19

1) $\varphi(\mathbf{x}) = e^{x_1 x_2}$ на кривій $x_1^2 + x_2^2 = 1$;

2) $\varphi(\mathbf{x}) = x_1 x_2$ в області $D = \{\mathbf{x} \in \mathbb{R}^2 \mid x_1^2 + x_2^2 = 16, x_1 \geq 0, x_2 \geq 0\}$.

Варіант 20

1) $\varphi(\mathbf{x}) = x_1^3 - x_2^3$ на кривій $x_1^2 + x_2^2 = 8$;

2) $\varphi(\mathbf{x}) = e^{-x_1 x_2 / 4}$ в області $D = \{\mathbf{x} \in \mathbb{R}^2 \mid x_1^2 + x_2^2 \leq 1, x_2 \geq 0\}$.

Варіант 21

1) $\varphi(\mathbf{x}) = x_1 + x_2$ на кривій $(x_1 - 4)^2 + (x_2 - 3)^2 = 4$;

2) $\varphi(\mathbf{x}) = x_1^2 - x_2^2$ в області $D = \{\mathbf{x} \in \mathbb{R}^2 \mid x_1^2 + x_2^2 \leq 16, x_1 + x_2 \geq 4\}$.

Варіант 22

- 1) $\varphi(\mathbf{x}) = x_1 + 2x_2 - 2x_3$ на поверхні $x_1^2 + x_2^2 + x_3^2 = 1$;
- 2) $\varphi(\mathbf{x}) = 2x_1^2 + x_2^2$ в області $D = \{\mathbf{x} \in \mathbb{R}^2 \mid x_1^2 + x_2^2 \leq 4, 4x_1^2 + x_2^2 \geq 4\}$.

Варіант 23

- 1) $\varphi(\mathbf{x}) = x_1^2 + x_2^2 + x_3^2$ на поверхні $x_1^2 + (x_2 - 3)^2 + (x_3 - 4)^2 = 1$;
- 2) $\varphi(\mathbf{x}) = x_1^2 + (x_2 - 4)^2$ в області $D = \{\mathbf{x} \in \mathbb{R}^2 \mid x_1^2 + x_2^2 \leq 4, 4x_1^2 + x_2^2 \geq 4\}$.

Варіант 24

- 1) $\varphi(\mathbf{x}) = x_1 x_2$ на кривій $4x_1^2 + x_2^2 = 8$;
- 2) $\varphi(\mathbf{x}) = x_1^2 + x_2^2$ в області $D = \{\mathbf{x} \in \mathbb{R}^2 \mid 2x_1^2 + (x_2 - 4)^2 \leq 1, x_2 \geq 0\}$.

Варіант 25

- 1) $\varphi(\mathbf{x}) = x_1 x_2^2 x_3^3$ на площині $x_1 + x_2 + x_3 = 1$;
- 2) $\varphi(\mathbf{x}) = e^{x_1 - x_2} - x_1 - x_2$ в області $D = \{\mathbf{x} \in \mathbb{R}^2 \mid x_1 + x_2 \leq 1, x_1 \geq 0, x_2 \geq 0\}$.

Завдання 3. Застосування задач на екстремум

Побудуйте математичну модель задачі та знайдіть її розв'язок.

Варіант 1 У задане півколо радіусу R потрібно вписати прямокутник з найбільшим периметром. Якими мають бути сторони такого прямокутника?

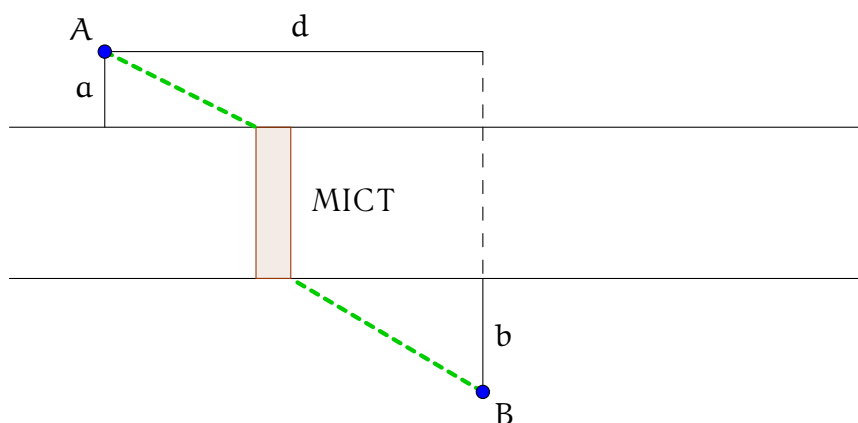
Варіант 2 Містер Баррі планує упорядкувати зберігання олівців та маркерів у своєму кабінеті. У нього є п'ять аркушів картону розміром 9×12 дюймів кожен, з яких він виготовить п'ять однакових коробок з відкритим верхом. Якими повинні бути розміри коробок, щоб їх обсяг був максимальним?

Варіант 3 Поперечний переріз залізничного тунелю має форму прямокутника, завершеного півколом. При якому радіусі півкола площа перерізу буде найбільшою? Периметр P перерізу вважається заданим.

Варіант 4 Ви керуєте магазином електроніки. У період різдвяних свят ви помітили, що одним з найпопулярніших ваших товарів є iPod. Одна з моделей продавалася досить добре за ціною 200 \$: у середньому за тиждень йшло 100 штук. Оптова вартість, за якою ви

купуєте цю модель, є фіксованою: 125 \$ за одиницю. Підприємницька інтуїція підказує вам, що можна поліпшити показники прибутку, продаючи товар зі знижкою. Дослідження ринку показало, що при кожному зниженні ціни на 10 \$ ви будете у середньому за тиждень продавати на 20 одиниць товару більше. За якою ціною потрібно продавати iPod у даний період, щоб максимізувати тижневий прибуток?

Варіант 5 Два села А і В розташовані по різні боки річки з паралельними прямолінійними берегами (див. малюнок). В якому місці слід побудувати міст (перпендикулярний берегам), щоб шлях з одного села до іншого був найменшим? (Параметри a , b і d фіксовані. Вважаємо, що стежки від кожного села до мосту йдуть прямолінійно. Шириною моста можна знехтувати.)



Варіант 6 Джейн та Джим інвестували 20 000 \$ у дизайн і просування нової моделі вудки для риболовлі. Виготовлення однієї одиниці товару обходиться у 2 \$. Потім за допомогою консультантів з маркетингу вони побудували наступну експериментальну модель залежності обсягів продажів S від витрат на рекламу a і вартості вудки p :

$$S(a, p) = 2000 + 4\sqrt{a} - 20p.$$

Грунтуючись на даній моделі, визначте рівень витрат на рекламу і ціну вудки, при яких прибуток Джейн та Джима буде найбільшим.

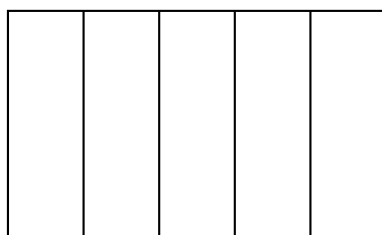
Варіант 7 Компанія «Веселі хлопці» виробляє деякий однорідний товар, який називається «штуковинами», використовуючи два види матеріалів: «аландріт» та «беконіум». Кількість вироблених штуків виражається формулою² $W = 200a^{1/2}b^{1/4}$, де a і b — кількість тонн аландріту та беконіуму відповідно. Припустимо, що вартість аландріту становить 50 \$ за тонну, а вартість беконіуму — 100 \$ за тонну. Яка максимальна кількість

²Така залежність відома в економічній теорії як функція Кобба — Дугласа.

штуковин може бути вироблена «Веселими хлопцями», якщо врахувати, що на матеріали не може бути витрачено більше 150 \$?

Варіант 8 Квіткова клумба має форму прямокутного трикутника. Визначте розміри клумби, що має найбільшу можливу площу, якщо довжина її огорожі дорівнює 1.

Варіант 9 Джон придбав 420 метрів огорожі для будівництва розплідника. Розплідник складається з п'яти однакових суміжних загонів, як показано на малюнку нижче. Знайдіть площу кожного загону, при яких площа розплідника буде найбільшою.



Варіант 10 Пан Штойбен хоче виготовити акваріум у формі прямокутного паралелепіпеда об'ємом 3 кубічних метра, причому довжина основи повинна бути вдвічі більше ширини. Основа і стінки будуть виготовлені зі скла. При яких розмірах акваріуму на його виготовлення піде найменша кількість скла? (Товщиною скла можна знехтувати.)

Варіант 11 Два коридора шириною a і b перетинаються під прямим кутом. Яка максимальна довжина драбини, яку можна пронести (горизонтально) з одного коридору в інший?

Варіант 12 Вантажний контейнер має форму прямокутного паралелепіпеда і повинен мати об'єм 480 кубометрів. Дно контейнера виготовляється з міцнішого матеріалу вартістю 5 \$ за квадратний метр, а бічні стінки і кришка — з менш міцного матеріалу вартістю 3 \$ за квадратний метр. Знайдіть розміри контейнера, що має найменшу вартість виготовлення при заданому об'ємі.

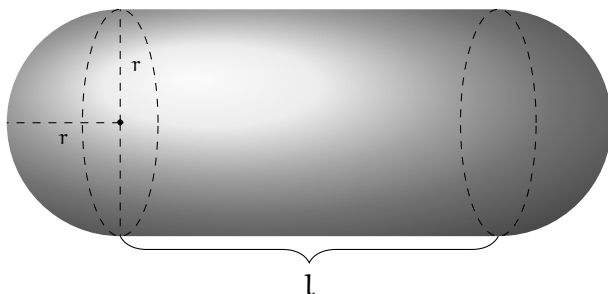
Варіант 13 Місіс Фергюсон купила ділянку землі у місті Вінчестер і планує посадити невеликий яблуневий сад. Відомо, що якщо посадити до 60 дерев на даній ділянці землі, то середній урожай з кожного дерева буде близько 100 кг, однак для кожного додатково посаженого дерева врожайність падає у середньому на 1 кг у результаті «перенаселеності». Пані Фергюсон планує отримати максимальний урожай яблук. Скільки дерев вона має посадити?

Варіант 14 Чарлі щодня вживає тільки яблука (a) і банани (b). Він прагне максимізувати функцію корисності

$$U(a, b) = \sqrt{a} + 4\sqrt{b}.$$

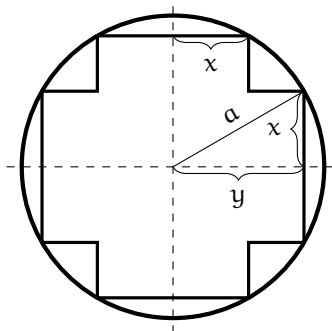
Знаючи, що один банан вдвічі дорожче одного яблука, визначте оптимальний раціон для Чарлі як функцію його денного прибутку m та ціни банана p . При цьому слід врахувати, що доктор прописав Чарлі вживання хоча б одного яблука на день.

Варіант 15 Резервуар для зберігання має форму циліндра з напівсферами на кінцях, як показано на малюнку нижче. Обсяг бака повинен дорівнювати 1000 л. Визначте розміри бака (радіус і довжину), при яких на його виготовлення піде мінімальна кількість матеріалу.



Варіант 16 Консервна банка повинна мати форму циліндра заданого об'єму V . Знайдіть, при яких розмірах цієї банки (діаметрі основи та висоті) на її виготовлення піде мінімальна кількість матеріалу?

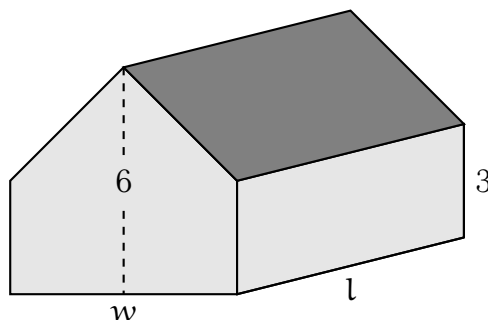
Варіант 17 Трансформатор змінного струму являє собою котушку з хрестоподібним залізним сердечником всередині. При цьому важливо, щоб площа поперечного перерізу сердечника була якомога більшою. Знайдіть відповідні розміри x і y перетину сердечника, якщо радіус котушки дорівнює a (див. малюнок нижче).



Варіант 18 Параболоїд $z = x^2 + 2y^2 + 1$ і площина $x - y + 2z = 4$ перетинаються по кривій C . Знайдіть точку на C , найближчу до початку координат.

Варіант 19 Знайдіть точку на площині $2x+3y+6z = 10$, найближчу до точки $M(-2; 5; 1)$.

Варіант 20 Будова довжиною l і шириною w має висоту 6 метрів посередині і 3 метри по кутах, як показано на малюнку. При яких розмірах будови площа даху і стін буде мінімальною, якщо під будівництво виділена площа 200 квадратних метрів?



Варіант 21 Нехай функція $T(x, y, z) = 100 + x^2 + y^2$ описує температуру в кожній точці з координатами (x, y, z) на поверхні далекої планети Лагранжіан. Сама планета має форму сфери радіусом 50 000 км. В яких точках температура на поверхні Лагранжіана максимальна?

Варіант 22 Форма поверхні планети Земля у хорошому наближенні є сплюснутим еліпсоїдом обертання з великою піввіссю $a = 6\,378\,245$ м (радіус екватора) та малою піввіссю $b = 6\,356\,863$ м (відстань від площини екватора до полюса). Знайдіть координати точки на поверхні Землі, найбільш віддаленої від її центру³.

Варіант 23 Деякі авіакомпанії вимагають, щоб лінійний розмір ручної поклажі (тобто сума довжини, ширини та висоти) не перевищував 45 дюймів, при цьому вона повинна вміщуватися під сидінням. Припустимо, що ручна поклажа має (приблизно) форму прямокутної коробки, причому один з її лінійних розмірів не перевищує половини іншого (наприклад, довжина як мінімум удвічі більше ширини). Якими повинні бути допустимі розміри ручної поклажі, щоб її об'єм був найбільшим?

Варіант 24 Флаєр повинен містити 50 квадратних сантиметрів друкованого тексту і графіки. Поля з кожного боку (зверху і знизу, праворуч і ліворуч) складають по 1.5 см. При яких розмірах флаєра на його виготовлення буде витрачено найменшу кількість паперу?

³Внаслідок опуклості екватора найвіддаленішими від центру точками поверхні Землі є вершина вулкана Чімборасо (6 267 м) в Еквадорі та гора Уаскаран (6 768 м) у Перу. Якби поверхня Землі була ідеальною сферою, найбільш віддаленою точкою від центру була б вершина Евересту.

Варіант 25 Група батьків планує побудувати дитячий майданчик площею 500 квадратних метрів. Для того, щоб заощадити на огороженні, вони хочуть розташувати майданчик так, щоб одна його сторона безпосередньо примикала до будівлі школи, а решту обгородити парканом. Знайдіть розміри майданчика, при яких довжина огорож буде найменшою.

Завдання 4. Застосування задач на екстремум

Зі шматка дроту довжиною L метрів треба сконструювати дві заготовки у формі геометричних фігур так, щоб їх сумарна площа була найбільшою. Види фігур задаються номером варіанта.

- 1) сектор кола із дугою 30° та правильний трикутник;
- 2) коло та рівнобедрений трикутник із бічною стороною, вдвічі більшою за основу;
- 3) коло та рівнобедрений прямокутний трикутник;
- 4) квадрат та рівнобедрений трикутник із бічною стороною, вдвічі більшою за основу;
- 5) квадрат та рівнобедрений прямокутний трикутник;
- 6) півколо та рівнобедрений трикутник із співвідношенням сторін $1:1:1.5$;
- 7) півколо та рівнобедрений прямокутний трикутник;
- 8) півколо та правильний трикутник;
- 9) чверть еліпса із великою піввіссю, вдвічі більшою за малу, та прямокутник із співвідношенням сторін $2:1$;
- 10) чверть кола та квадрат;
- 11) чверть еліпса із співвідношенням півосей $2:1$ та правильний трикутник;
- 12) чверть кола та рівнобедрений трикутник із співвідношенням сторін $2:2:1$;
- 13) чверть кола та рівнобедрений прямокутний трикутник;
- 14) правильний трикутник та прямокутник із співвідношенням сторін $2:1$;
- 15) коло та прямокутник із довжиною, вдвічі більшою за ширину;
- 16) половина еліпса із великою піввіссю, вдвічі більшою за малу, та прямокутник із довжиною, вдвічі більшою за ширину;
- 17) півколо та квадрат;
- 18) прямокутний рівнобедрений трикутник та прямокутник із довжиною, вдвічі більшою за ширину;
- 19) квадрат та еліпс із великою піввіссю, вдвічі більшою за малу;
- 20) правильний трикутник та еліпс із великою піввіссю, вдвічі більшою за малу;
- 21) прямокутний рівнобедрений трикутник та еліпс із співвідношенням півосей $2:1$;
- 22) прямокутник із довжиною, утричі більшою за ширину, та сектор кола із дугою 60° ;
- 23) рівнобедрений трикутник із основою, в 1.5 рази більшою за бічну сторону, та половина еліпса із великою піввіссю, вдвічі більшою за малу;

- 24) квадрат та рівнобедрений прямокутний трикутник;
 25) сектор кола із дугою 45° та квадрат.

2.6.2 Лінійне програмування. Симплекс-метод

Завдання 1. Побудова математичних моделей економічних задач

Сформулюйте математичний опис двох запропонованих практичних ситуацій у вигляді оптимізаційних задач. Побудова моделі повинна включати у себе докладний опис економічного змісту введених змінних, обмежень і цільової функції.

Варіант 1

- М'ясокомбінат виготовляє два типи котлет з індички. Кожен тип котлет містить біле і темне м'ясо. Півкілограма котлет першого типу продається за 40 грн. Даний тип котлет повинен містити, щонайменше, 70 % білого м'яса. Котлети другого типу продаються по 30 грн за 0.5 кг. Вони повинні містити, щонайменше, 60 % білого м'яса. Всього може бути продано не більше 25 кг котлет першого типу і 15 кг котлет другого типу. Для виробництва котлет використовуються дві породи індичок, які купуються на сусідній фермі. Індички першої породи коштують 100 грн за штуку. Кожна з них дає 2.5 кг білого і 1 кг темного м'яса. Кожна індичка другої породи коштує 80 грн. Вона дає 1.5 кг білого і 1.5 кг темного м'яса. Сформулюйте модель лінійного програмування для максимізації прибутку м'ясокомбінату.
- Планове завдання з виготовлення чотирьох видів костюмів необхідно розподілити між трьома швейними фабриками. Виробничі потужності i -ї фабрики дозволяють за розглянутий період часу випустити t_{ij} костюмів j -ї моделі. Костюм j -ї моделі продається по c_j грн. од. за штуку, а собівартість виготовлення j -ї моделі костюму на i -й фабриці становить s_{ij} грн. од. Матриці $\mathbf{R}$, $\mathbf{S}$, а також вектор $\mathbf{c}$ наведені нижче.

$$\mathbf{R} = \begin{pmatrix} 200 & 240 & 300 & 150 \\ 240 & 300 & 200 & 300 \\ 150 & 240 & 300 & 200 \end{pmatrix}, \quad \mathbf{S} = \begin{pmatrix} 400 & 400 & 500 & 200 \\ 250 & 300 & 250 & 400 \\ 400 & 500 & 400 & 300 \end{pmatrix},$$

$$\mathbf{c} = (500 \quad 650 \quad 800 \quad 800).$$

При розподілі планового завдання керівники компанії прагнуть максимізувати прибуток від продажу костюмів. Сформулюйте відповідну оптимізаційну задачу.

Варіант 2

- Компанія ABC поставляє чисті та змішані фруктові соки оптовими партіями у магазини «Овочі-фрукти». Список продуктів, які поставляються компанією, разом з їх складом і отриманим прибутком (у розрахунку на 1 літр продукту) наведено нижче. Компанія ABC прагне, щоб не більше 40 % усього обсягу реалізації у літрах становили чисті соки, а також, принаймні, 30 % яблуневого соку було продано у чистому вигляді. У розпорядженні ABC 778 літрів яблуневого соку, 389 літрів виноградного соку та 584 літри журавлиного соку. Вартість одного літра яблуневого соку — 2 грн, виноградного — 4 грн, журавлиного — 3 грн.

Сік	Відсотковий склад	Прибуток (грн/л)
Яблуневий	100	0.9
Виноградний	100	0.6
Журавлинний	100	0.5
Яблунево-виноградний	70/30	0.7
Яблунево-журавлинний	60/40	0.8
«Все-в-одному»	50/20/30	1.0

Сформулюйте модель лінійного програмування, що дозволяє визначити кількість літрів кожного продукту, яку необхідно виготовити для максимізації прибутку.

- Є дві ґрунтово-кліматичні зони, площі яких відповідно дорівнюють 0.8 і 0.6 млн га. Дані про врожайність зернових культур наведені у таблиці.

Зернові культури	Врожайність, ц/га		Вартість 1 ц, гр. од.
	1-а зона	2-а зона	
Озимі	20	25	8
Ярові	25	20	7

Визначте розміри посівних площ озимих і ярових культур, необхідні для досягнення максимального виходу продукції у вартісному вираженні.

Варіант 3

- Для виробництва двох видів продукції на двох підприємствах може бути використано 480 од. сировини. Норми витрат сировини на один виріб відповідно дорівнюють 4 і 3 од., а прибуток від реалізації одного виробу відповідно дорівнює 5 і 6 гр. од. На кожному з підприємств сировина проходить послідовну обробку, причому використовуються два види технологічного обладнання. Витрати часу при обробці сировини на кожному з типів обладнання у кожному з підприємств наведені у таблиці. У ній же вказано загальний фонд робочого часу кожного з типів обладнання на кожному з підприємств.

Тип обладнання	1-е підприємство			2-е підприємство		
	Час на виготовлення одного виробу		Заг. фонд робочого часу	Час на виготовлення одного виробу		Заг. фонд робочого часу
	Виріб А	Виріб В		Виріб А	Виріб В	
I	2	1	360	2	3	420
II	1	3	420	4	5	340

З урахуванням наявної сировини і можливостей використання підприємствами технологічного обладнання визначте, скільки виробів кожного виду слід виготовити на кожному з підприємств так, щоб прибуток від їх реалізації був максимальним.

- Для перевезення вантажу на трьох лініях можуть бути використані судна трьох типів. Продуктивність суден при використанні їх на різних лініях характеризується даними, наведеними у таблиці. У ній же вказані загальний час, протягом якого судна кожного типу можуть перебувати в експлуатації, і мінімально необхідні обсяги перевезень на кожній з ліній.

Тип судна	Продуктивність суден, млн тонно-миль на добу			Загальний час експлуатації суден, діб
	Лінія А	Лінія В	Лінія С	
I	8	14	11	300
II	6	15	13	300
III	12	12	4	300
Заданий обсяг перевезень, млн тонно-миль	3000	5400	3300	

Потрібно визначити, які судна, на якій лінії і протягом якого часу слід використовувати, щоб забезпечити максимальне завантаження суден з урахуванням можливого часу їх експлуатації.

Варіант 4

1. При складанні добового раціону годівлі тварин можна використовувати свіже сіно (не більше 50 кг) і силос (не більше 85 кг). Раціон повинен мати певну поживність (число кормових одиниць не менше 30) і містити поживні речовини: білок (не менше 1 кг), кальцій (не менше 100 г) і фосфор (не менше 80 г). У наведеній нижче таблиці містяться дані про вміст зазначених компонентів у 1 кг кожного продукту харчування і собівартості (гр. од. за 1 кг) цих продуктів.

Продукти	Кількість кормових одиниць	Білок, г/кг	Кальцій, г/кг	Фосфор, г/кг	Собівартість, гр. од./кг
Сіно свіже	0.5	40	1.25	2	1.2
Силос	0.5	10	2.5	1	1.8

Потрібно визначити оптимальний раціон з умови мінімуму собівартості.

2. Підприємству задана місячна програма на виготовлення чотирьох типів виробів у кількостях відповідно 5000, 2000, 3000 і 1800 штук. На підприємстві є три групи верстатів із різною продуктивністю. Сумарний допустимий час для кожної групи верстатів становить відповідно 800, 1000, 1500 годин. Дані про технологічний процес вказані у таблиці.

Групи верстатів	Норми часу на виготовлення одного виробу, год				Витрати на виготовлення одного виробу, гр. од.			
	I	II	III	IV	I	II	III	IV
1	0.5	0.15	0.4	0.6	0.12	0.2	0.3	0.25
2	0.4	0.12	0.2	0.5	0.16	0.14	0.35	0.2
3	0.42	0.14	0.35	0.45	0.17	0.25	0.4	0.3

Розподіліть вироби по верстатах так, щоб місячна програма була виконана при найменших витратах.

Варіант 5

1. Компанія використовує картон для виробництва штучних квітів. Для цього вона поштучно закупає аркуші картону площею 1 м². Аркуш може бути розкритий трьома різними способами. Перший спосіб розкриття дозволяє виготовити 70 троянд і 20 гвоздик. Другий спосіб — 60 троянд і 60 гвоздик. І нарешті, третій спосіб дозволяє виготовити 23 троянди і

117 гвоздик. Єдиними змінними витратами, пов'язаними з процесом розкрою, є витрати на придбання картону, які становлять 10 грн за м².

Компанія продає два типи квіткових наборів. У перший набір входять 8 троянд і 4 гвоздики. У другий — 3 троянди і 12 гвоздик. Набори продаються по 70 і 50 грн відповідно. Компанія повинна самостійно зробити всі квіти, які використовуються для виробництва наборів. Поставки картону обмежені 1000 аркушами. На плановий період компанія домовилася продати, щонайменше, 40 наборів першого типу. Відомо також, що вона не може продати більше 200 одиниць кожного з квіткових наборів. Компанія також вирішила слідувати політиці, при якій виробництво наборів першого типу не повинно перевищувати 40 % від виробництва наборів двох типів.

Сформулюйте модель лінійного програмування для максимізації загального прибутку.

2. Деякій компанії належать три ферми, де вирощують овочі, призначені для подальшої обробки на двох холодильних заводах компанії. Одним із вирощуваних овочів є боби, які холодильні заводи продають по 200 гр. од. за 1 т. У таблицях наведені витрати виробництва для кожної ферми і кожного холодильного заводу, максимальні значення врожаю для кожної ферми, прогнозні значення попиту на наступний сезон для кожного заводу, а також вартість транспортування бобів.

		Витрати виробництва, гр. од./т	Максимальний врожай, т			
Ферми	1	90	2000	Вартість транспортування		
	2	95	3000	Завод 1 Завод 2		
	3	87	1500	Ферма 1	10	15
		Витрати виробництва, гр. од./т	Прогнозний попит, т	Ферма 2	12	12
				Ферма 3	18	9
Заводи	1	20	2750			
	2	23	3250			

Побудуйте транспортну модель, яка дозволяє знайти виробничий план на наступний сезон для ферм і холодильних заводів, що гарантує максимальний прибуток.

Варіант 6

1. Підприємство повинно випускати два види продукції А і В, використовуючи при цьому послідовно чотири верстати. Дані про технологічний процес наведено у таблиці.

Верстат	Витрати часу на 1 од. продукції, год		Фонд часу, год
	Продукція А	Продукція В	
1	3	3	15
2	2	6	18
3	4	—	16
4	1	2	8
Прибуток на од. продукції, гр. од.	2	3	

Складіть план випуску продукції, що забезпечує підприємству найбільший прибуток.

2. Будівельний комбінат виробляє 2 типи облицювальних плит. Для виробництва використовується кольоровий цемент, мармурова та гранітна крихта. Споживання компонентів задано пропорціями 10:5:9:4, 17:10:4:5. Останній член у пропорції задає споживання інших матеріалів (барвники, вода та ін.). Комбінат має матеріали у кількості 50 000 т кольорового цементу, 22 000 т мармурової крихти, 30 000 т гранітної крихти, інші матеріали без обмежень. Потрібно визначити оптимальний план виробництва облицювальних плит, який дозволяє провести обробку максимальної площі, якщо відомо, що на облицювання одного квадратного метра необхідно 0.07 т плит першого типу або 0.1 т другого типу. Складіть математичну модель задачі.

Варіант 7

1. Підприємство випускає два види продукції. На виготовлення продукції витрачається два види ресурсів. Запаси ресурсів 1-го виду становлять 160 од., 2-го виду — 210 од. Норми витрат 1-го ресурсу, що йде на виготовлення одиниці продукції, дорівнюють 2 од. для продукції 1-го виду і 2.67 од. — для продукції 2-го виду; норми витрат 2-го ресурсу становлять 3 од. для продукції 1-го виду та 2 од. — для продукції 2-го виду. Сумарний обсяг випуску повинен бути не менше 40 од. Витрати на виготовлення одиниці продукції відповідно становлять 100 і 140 гр. од.
Потрібно знайти обсяги виробництва продукції 1-го і 2-го виду, при яких сумарні видатки виробництва мінімальні.
2. Компанія виробляє паперову продукцію. На одній з виробничих ліній рулони паперу шириною 80 і 120 см розрізають на рулони меншої ширини, але тієї ж довжини. На поточний тиждень компанія отримала замовлення на 40, 60, 45 і 70 рулонів паперу шириною 50, 45, 32 і 22 см відповідно. Мета — розкроїти 80- і 120-сантиметрові рулони відповідно до потреб і з мінімумом відходів. Вихідний матеріал шириною 80 см наявний у необмеженій кількості, у той час як шириною 120 см — обмежений 50 рулонами.
Сформулюйте модель лінійного програмування для мінімізації загальних витрат: невикористаних залишків і відходів від перерозкрою.

Варіант 8

1. Фірма виробляє і продає два види товарів — панчохи та шкарпетки. Фірма отримує прибуток у розмірі 12000 гр. од. від виробництва і продажу кожної одиниці товару 1 і 4000 гр. од. від виробництва і продажу кожної одиниці товару 2. Фірма складається з трьох підрозділів. Трудові витрати (людино-дні) на виробництво цих товарів у кожному підрозділі подано у таблиці.

Підрозділ	Трудовитрати, люд-днів на 1 од.	
	Товар 1	Товар 2
1	1	2
2	1	3
3	2	3

- Керівництво розрахувало, що у наступному місяці фірма буде мати у своєму розпорядженні наступні можливості забезпечення виробництва: 800 люд-днів у підрозділі 1, 600 — у підрозділі 2 і 2000 — у підрозділі 3. Сформулюйте оптимізаційну задачу, виходячи з максимізації прибутку фірми.
2. З чотирьох видів сировини необхідно скласти суміш, до складу якої має входити не менше 26 од. хімічної речовини А, 30 од. — речовини В і 24 од. — речовини С. Кількість одиниць хімічної речовини, що міститься у 1 кг сировини кожного виду, зазначено у таблиці. У ній

же наведена ціна 1 кг сировини кожного виду.

Речовина	Кількість од. речовини в 1 кг сировини			
	Сировина 1	Сировина 2	Сировина 3	Сировина 4
A	1	1	—	4
B	2	—	3	5
C	1	2	4	6
Ціна 1 кг сировини, гр. од.	5	6	7	4

Потрібно скласти суміш, яка містить не менше потрібної кількості речовин даного виду і має мінімальну вартість.

Варіант 9

1. Цех меблевого комбінату випускає трельяжі, трюмо і тумбочки під телевізори. Норми витрат матеріалу у розрахунку на один виріб, планова собівартість, оптова ціна підприємства, плановий (місячний) асортимент і трудомісткість одиниці продукції наведені у таблиці. Запаси деревостружкових плит, дощок ялинових і березових становлять 90, 30 і 314 м³ відповідно. Плановий фонд робочого часу 16 800 людино-годин.

Показники	Трельяжі	Трюмо	Тумбочки
Норми витрат матеріалу, м ³ :			
деревостружкові плити	0.032	0.031	0.038
дощки ялинові	0.020	0.020	0.006
дощки березові	0.005	0.005	0.006
Трудомісткість, люд-год	10.2	7.5	5.8
Планова собівартість, гр. од.	88.81	63.98	29.60
Оптова ціна, гр. од.	93.00	67.00	30.00
Плановий асортимент, шт.	350	290	1200

Виходячи з необхідності виконання плану за асортиментом і можливості його перевиконання за окремими (або навіть всіма) показниками побудуйте модель, на основі якої можна отримати план виробництва меблів, який максимізує прибуток.

2. З відходів виробництва підприємство може організувати випуск чотирьох видів продукції. Для цього воно планує використовувати два типи взаємозамінного обладнання. Кількість виробів кожного виду, яка може бути виготовлена на відповідному обладнанні протягом 1 год, а також витрати, пов'язані з виробництвом одного виробу, наведені у наступній таблиці.

Тип обладнання	Кількість вироблених на протязі 1 год виробів				Витрати (гр. од.), пов'язані з виробництвом на протязі 1 год			
	Пр. 1	Пр. 2	Пр. 3	Пр. 4	Пр. 1	Пр. 2	Пр. 3	Пр. 4
I	8	7	4	5	2.7	2.6	2.7	2.4
II	6	8	6	4	2.6	2.7	2.6	2.5

Устаткування I типу підприємство може використовувати не більше 80 год, а обладнання II типу — не більше 60 год.

З огляду на те, що підприємству слід виготовити виробів кожного виду відповідно не менше 240, 160, 150 і 220 од., визначте, протягом якого часу і на якому обладнанні слід виготовляти кожен з виробів так, щоб отримати не менше потрібної кількості виробів при мінімальних витратах на виготовлення.

Варіант 10

1. Кондитерська фабрика для виробництва трьох видів карамелі А, В і С використовує три види сировини: цукровий пісок, патоку і фруктове пюре. Норми витрати сировини на виробництво 1 кг карамелі задані у таблиці.

Найменування сировини	Норми витрат		
	А	В	С
Цукровий пісок	0.6	0.5	0.6
Патока	0.4	0.4	0.3
Фруктове пюре	0.1	0.2	0.2

Запаси сировини на складі відповідно дорівнюють 800, 600 і 120 кг. Прибуток від реалізації 1 кг продукції кожного виду визначається значеннями 1.08, 1.12 і 1.28 гр. од. Потрібно знайти план виробництва карамелі, що забезпечує максимальний прибуток.

2. Фірма «Shashlyk Limited» напередодні травневих свят відчуває незвичайний ріст попиту на два типи мангалів для приготування шашлику на відкритому повітрі — газові та вугільні. Фірма уклала контракт на поставку 20 тис. вугільних і 10 тис. газових мангалів. Виробництво мангалів обмежується можливостями трудовитрат на обробку у трьох цехах: виробництва, збірки та упаковки. У таблиці вказано, скільки людино-годин витрачається у кожному цеху на виробництво кожної одиниці продукції, а також скільки людино-годин є у наявності.

Цех	Трудовитрати на		Кількість люд-год у наявності, на місяць
	один мангал, люд-год/шт. вугільний	газовий	
виробництва	0.10	0.20	2500
збірки	0.15	0.25	4000
упаковки	0.05	0.05	2000

Через обмеженість трудових ресурсів фірма не може повністю задовольнити попит своїми силами. Тому вона провела переговори з іншим виробником, який на даний час має у своєму розпорядженні надлишкові потужності. Цей виробник погодився поставляти фірмі у будь-якому співвідношенні вугільні мангали по 50 гр. од. за штуку і газові мангали по 100 гр. од. за штуку. Ці ціни перевищують собівартість мангалів на заводі фірми «Shashlyk Limited» на 10 гр. од. за кожен вугільний і на 15 гр. од. за кожен газовий мангал. Задача фірми полягає у тому, щоб знайти таке співвідношення закуповуваних і вироблюваних мангалів, яке призвело б до мінімальних видатків при задоволенні попиту. Сформулюйте відповідну задачу лінійного програмування.

Варіант 11

1. Завод випускає ексклюзивні складні ножі двох моделей: А і В. Завод повинен випускати на місяць не менше 400 і 600 ножів кожної моделі відповідно. Кожна модель приносить прибуток у розмірі 40 і 60 гр. од. відповідно. Процес виробництва ножів складається із трьох стадій: виготовлення деталей, збірки виробу та його упаковки. Нормативи часу на ці стадії задані таблицею.

Модель	Норми часу (год/шт.)		
	Стадія 1	Стадія 2	Стадія 3
A	0.3	0.4	0.1
B	0.35	0.5	0.15

Виробничі потужності заводу дозволяють витратити на місяць не більше 600 годин на випуск деталей, 500 годин на збірку ножів і 200 годин на упаковку. Складіть план випуску ножів, що забезпечує заводу максимальний прибуток.

2. У трьох пунктах відправлення A, B, C зосереджений однорідний товар. Цей товар необхідно перевезти у чотири пункти призначення P, Q, R, S. Запаси товару у пунктах відправлення і потреби пунктів призначення показані у таблиці.

Запаси, т			Потреби, т			
A	B	C	P	Q	R	S
42	36	40	26	41	25	31

Вартості перевезень 1 т вантажу з i -го пункту відправлення в j -й пункт призначення задані матрицею.

	A	B	C
P	200	400	300
Q	700	500	800
R	600	900	700
S	550	430	650

Потрібно знайти план перевезень вантажу, що забезпечує транспортному підприємству мінімальну вартість перевезень. Врахувати, що у разі нестачі запасів всі потреби не можуть бути задоволені.

Варіант 12

1. Є два види продукції П1 і П2, що містять поживні речовини S1, S2, S3, S4 (жири, білки, вуглеводи, вітаміни). Вміст поживних речовин в одиниці кожного виду продукції і необхідний мінімум поживних речовин наведені у таблиці.

Поживні речовини	Вміст		Необхідний мінімум поживних речовин
	поживних речовин в одиниці продукції		
	Продукт 1	Продукт 2	
S1	1	2	10
S2	3	2	8
S3	2	1	9
S4	2	2	11

Вартість одиниці продукції П1 і П2 відповідно дорівнює 3 і 4 гр. од. Потрібно скласти найбільш дешевий раціон, при якому потреби у поживних речовинах будуть задоволені.

2. На кожній із чотирьох філій виробничого об'єднання можуть виготовлятися вироби чотирьох видів. З огляду на необхідність поглиблення спеціалізації, на філіях вирішено зосередити випуск тільки по одному виду виробів. Собівартість кожного із виробів на кожній філії

різна і визначається матрицею

$$C = \begin{pmatrix} 9 & 8 & 9 & 7 \\ 4 & 6 & 3 & 2 \\ 7 & 2 & 1 & 4 \\ 7 & 8 & 5 & 6 \end{pmatrix}.$$

Знайдіть такий розподіл випуску продукції між філіями, щоб загальна собівартість продукції була мінімальною.

Варіант 13

1. У пекарні випікають два типи тортів: шоколадний і ванільний. Кожен шоколадний торт продається по 100 грн, а ванільний — по 50. Для виготовлення шоколадного торта необхідно 20 хвилин часу і 4 яйця. Ванільний торт вимагає 40 хвилин часу і 1 яйце. У розпорядженні пекарні є 8 годин часу і 30 яєць.

Сформулюйте модель лінійного програмування для максимізації доходу пекарні.

2. Є три ділянки землі, на яких можуть бути засіяні кукурудза, пшениця, ячмінь і просо. Площа кожної з ділянок відповідно дорівнює 600, 180 та 220 га. З урахуванням наявності насіння, кукурудзою, пшеницею, ячменем і просом слід засіяти відповідно 290, 180, 110 і 420 га. Врожайність кожної з культур для кожної ділянки різна і задається матрицею

$$C = \begin{pmatrix} 40 & 45 & 50 \\ 30 & 28 & 22 \\ 18 & 22 & 14 \\ 19 & 15 & 16 \end{pmatrix}.$$

Потрібно визначити, скільки гектарів кожної культури на кожній ділянці слід засіяти, щоб загальний збір зерна був максимальним.

Варіант 14

1. Сталеві пруті довжиною 110 см необхідно розрізати на задані заготовки довжиною 45, 35 і 50 см. Необхідні кількості заготовок кожного виду становлять відповідно 40, 30 і 20 шт. Визначити, скільки прутів по кожному з можливих способів слід розрізати, щоб отримати не менше потрібної кількості заготовок кожного виду при мінімальних відходах.
2. На чотирьох станціях є надлишок порожніх вагонів у розмірі відповідно 100, 120, 150 і 50 вагонів. Необхідно розподілити дані вагони по семи станціям з браком порожняка (відповідно 70, 50, 60, 30, 50, 70 і 90 вагонів). Відстань між кожною станцією відправлення (надлишку вагонів) і кожною станцією призначення (нестачі порожняка) представлено у вигляді матриці

$$C = \begin{pmatrix} 15 & 23 & 28 & 13 & 12 & 16 & 21 \\ 24 & 17 & 24 & 11 & 11 & 12 & 18 \\ 8 & 19 & 16 & 18 & 9 & 17 & 15 \\ 12 & 28 & 18 & 16 & 21 & 20 & 25 \end{pmatrix}$$

Необхідно скласти план розподілу вагонів між зазначеними станціями з мінімальним сумарним пробігом порожніх вагонів.

Варіант 15

1. Фірма має можливість рекламувати свою продукцію, використовуючи радіо- і телевізійну мережу. Витрати на рекламу у бюджеті фірми обмежені сумою 1000 гр. од. на місяць. Кожна хвилина радіореклами обходиться у 5 гр. од., а кожна хвилина телереклами — у 100 гр. од.

на місяць. Фірма хотіла б використовувати радіомережу принаймні у два рази частіше, ніж мережу телебачення. Досвід минулих років показав, що обсяг збуту, який забезпечує кожна хвилина телереклами, у 25 разів більше збуту, що забезпечується однією хвилиною радіореклами.

Визначте оптимальний розподіл фінансових коштів, які щомісяця відпускаються на рекламу, між радіо- і телерекламою.

2. На виробництво надійшла партія стержнів довжиною 250 і 190 см. Необхідно отримати не менше 470 відрізків по 45 см і не менше 450 відрізків по 80 см. Як розрізати наявні стержні, щоб скоротити до мінімуму відходи? Побудуйте математичну модель задачі.

Варіант 16

1. Піротехнічна компанія «Великий Переполох» виробляє три основних типи шумовідтворюючих іграшок: «Шум», «Бабах» і «Тарарам». Для виготовлення «Шума» необхідно 30 хв часу і одне перо. «Бабах» вимагає 20 хв часу, два пера і 0.5 г блискучої пудри. «Тарарам» вимагає 30 хв часу, три пера і 1 г блискучої пудри. Внесок у прибуток одиниці кожного виду іграшок складає: 0.45 гр. од. для «Шуму», 0.55 гр. од. для «Бабаху» і 0.70 гр. од. для «Тарараму». У розпорядженні «Великого переполоху» є наступні ресурси: 4800 хв робочого часу, 90 г блискучої пудри і 360 пір'їв.

Компанія впевнена у тому, що максимально можлива кількість іграшок «Шум», «Бабах» і «Тарарам», яку вона зможе продати, становить 200, 500 і 300 відповідно.

Побудуйте модель для визначення оптимального плану виробництва продуктів, що забезпечує отримання максимального прибутку в умовах наявних обмежень.

2. На трьох хлібокомбінатах щодня виробляється 110, 190 і 90 т борошна. Це борошно споживається чотирма хлібозаводами, щоденні потреби яких дорівнюють відповідно 80, 60, 170 і 80 т. Тарифи перевезень 1 т борошна від комбінатів до заводів задані матрицею

$$C = \begin{pmatrix} 8 & 1 & 9 & 7 \\ 4 & 6 & 2 & 12 \\ 3 & 5 & 8 & 9 \end{pmatrix}.$$

Необхідно скласти такий план доставки борошна, при якому загальна вартість перевезень буде мінімальною.

Варіант 17

1. На промисловому підприємстві виготовляють два типи продукції. Вони виробляються за допомогою обладнання U1, U2, U3, яке протягом дня може працювати відповідно 24 000, 40 000, 27 000 секунд. Норми часу, необхідного для виробництва одиниці продукції за допомогою відповідного обладнання, наводяться у таблиці.

Типи продуктів	Обладнання		
	U1	U2	U3
1	3	8	9
2	6	4	3

Прибуток від реалізації продукту 1 типу становить 9 гр. од., а від продукту 2 типу — 6 гр. од. Розрахуйте, в якому обсязі слід планувати виробництво, щоб отримати максимальний прибуток.

2. Виготовлена на п'яти заводах цегла надходить на шість будівельних майданчиків. Щоденне виробництво цегли і потреби у ній вказані у таблиці нижче. У ній же вказана вартість перевезення 1000 шт. цегли з кожного із заводів до кожного з об'єктів.

Заводи	Об'єкти						Виробництво цегли, тис. шт.
	1	2	3	4	5	6	
I	8	7	5	10	12	8	240
II	13	8	10	6	7	13	360
III	12	4	11	9	10	11	180
IV	14	6	12	13	7	14	120
V	9	12	14	15	8	13	150
Потреби у цеглі, тис. шт.	230	220	130	170	190	110	

Складіть план перевезень, згідно з яким забезпечуються потреби у цеглі на кожному з об'єктів при мінімальній загальній вартості перевезень.

Варіант 18

- При виробництві чотирьох видів кабелю виконується п'ять груп технологічних операцій. Норми витрат на 1 км кабелю даного виду на кожній з груп операцій, прибутки від реалізації 1 км кожного виду кабелю, а також загальний фонд робочого часу, протягом якого можуть виконуватися ці операції, вказані у таблиці.

Технологічна операція	Норми витрат часу (год) на обробку 1 км кабелю				Загальний фонд робочого часу (год)
	Тип 1	Тип 2	Тип 3	Тип 4	
Волочіння	1.2	1.8	1.6	2.4	7200
Накладення ізоляцій	1.0	0.4	0.8	0.7	5600
Скручування елементів у кабель	6.4	5.6	6.0	8.0	11176
Покриття свинцем	3.0	—	1.8	2.4	3600
Випробування і контроль	2.1	1.5	0.8	3.0	4200
Прибуток від реалізації 1 км кабелю (гр. од.)	1.2	0.8	1.0	1.3	

Потрібно визначити такий план випуску кабелю, при якому загальний прибуток від реалізації продукції є максимальним.

- З двох сортів бензину утворюються дві суміші — А і В. Суміш А містить 60 % бензину першого сорту і 40 % бензину другого сорту. Суміш В — 80 % першого сорту і 20 % другого сорту. Ціна 1 кг суміші А — 10 гр. од., а суміші В — 12 гр. од. Складіть план утворення сумішей, при якому буде отримано максимальний прибуток, якщо у наявності є 50 т бензину першого сорту і 30 т бензину другого сорту.

Варіант 19

- Підприємство може виробляти три різних види продукції у пляшках — А, В і С. При цьому використовуються чотири види ресурсів в обмеженій кількості: робочий час першої ділянки, робочий час другої ділянки, сировина і площа для зберігання продукції у процесі виробництва на другій ділянці.

Кожен вид продукції вимагає обробки на обох ділянках. На першій ділянці можуть бути використані 320 годин робочого часу на тиждень. На другій ділянці працюють п'ятеро робітників по 8 годин на день 5 днів на тиждень.

Максимальна кількість сировини, доступна протягом тижня, — 1167 літрів. Майданчик для тимчасового (у процесі виробництва) зберігання продукції може вмістити 400 пляшок

продукції А, якщо ніякі інші види не виробляються.

Продукція А приносить прибуток у розмірі 12 грн за 10 пляшок. Витрати часу на виробництво однієї пляшки на першій ділянці — 1 год. На другій ділянці виробляється дві пляшки даного виду продукції на годину. Для цього необхідний 1 літр сировини.

Продукція В приносить прибуток 2 грн на пляшку. Для виробництва однієї пляшки на першій ділянці потрібно 2 години часу; на другій ділянці виробляються 3 пляшки на годину. Витрати сировини на виробництво продукції виду В — 250 мл у розрахунку на пляшку. Займана площа зберігання у процесі виробництва у два рази більше, ніж для продукції А (у розрахунку на пляшку).

Продукція С приносить прибуток у розмірі 1.5 грн за пляшку. На першій ділянці потрібно $\frac{1}{4}$ години для виробництва однієї пляшки продукції даного виду. На другій ділянці за 1 год може бути вироблено 4 пляшки продукції С. Для її виробництва необхідні 125 мл сировини. Площа, яку займає продукція С у процесі виробництва (в розрахунку на одну пляшку), дорівнює $\frac{1}{3}$ від площі, необхідної для продукції А.

Сформулюйте модель лінійного програмування, яка максимізує прибуток від реалізації продукції.

- Потужності заводу з виробництва добрив дозволяють у поточному місяці виробити 1000 т нітратів, 1800 т фосфатів і 1200 т поташу. В результаті змішування цих активних інгредієнтів з інертними, запаси яких не обмежені, можуть бути отримані три типи добрив. У таблиці зазначено вміст активних інгредієнтів у суміші.

Тип добрива	Вміст активних інгредієнтів, %			Ціна, тис. грн. од./т
	нітрати	фосфати	поташ	
1	5	10	5	40
2	5	10	10	50
3	10	10	10	60
Ціна, тис. грн. од./т	160	40	100	

Ціна інертних інгредієнтів складає 5 тис. грн. од./т. Витрати змішування, упаковки та продажу складають 15 тис. грн. од./т для кожного типу добрива. Існує угода про постачання 6000 т добрив типу 1. Завод прагне максимізувати місячний прибуток від продажу добрив. Побудуйте математичну модель задачі.

Варіант 20

- На меблевій фабриці зі стандартних листів фанери необхідно вирізати заготовки трьох видів у кількостях відповідно 24, 31 і 81 шт. Кожен лист фанери може бути розрізаний на заготовки двома способами. Кількість отриманих заготовок при даному способі розкрою наведено у таблиці. У ній же вказана величина відходів, які утворюються при цьому способі розкрою одного листа фанери.

Вид заготовки	Кількість заготовок	
	Спосіб 1	Спосіб 2
I	2	6
II	5	4
III	2	3
Величина відходів, см ²	12	16

Визначте, скільки листів фанери і за яким способом слід розкрити так, щоб було отримано

не менше потрібної кількості заготовок при мінімальних відходах.

- Для виконання чотирьох типів землерийних робіт можуть бути використані екскаватори чотирьох типів. Продуктивність екскаватора i -го типу при виконанні j -ї роботи задається матрицею

$$C = \begin{pmatrix} 0.9 & 0.6 & 0.7 & 0.9 \\ 0.7 & 0.8 & 0.9 & 0.8 \\ 0.8 & 0.6 & 0.8 & 0.9 \\ 0.8 & 0.7 & 0.8 & 0.7 \end{pmatrix}.$$

З огляду на те, що на кожній з робіт може бути зайнятий тільки один екскаватор і що всі екскаватори повинні бути задіяні, знайдіть такий розподіл екскаваторів між роботами, який забезпечить максимальну продуктивність.

Варіант 21

- Компанія «Суперскло» — виробник високоякісної продукції, у тому числі рам для вікон і дверей. У складі компанії три заводи. На першому виготовляють алюмінієві рами, на другому — дерев'яні. Третій завод спеціалізується на збірці готової продукції (у рами вставляється скло). У зв'язку зі зниженням прибутків головний інженер компанії вирішив перепрофілювати виробничу лінію. Кілька неприбуткових видів продукції знімаються з виробництва, що вивільняє виробничі потужності для виготовлення одного або двох потенційно нових виробів, на які є попит.

Один з цих видів продукції — скляні двері з алюмінієвою рамою, другий вид — вікно з дерев'яною рамою. Служба маркетингу, провівши дослідження, зробила висновок, що компанія зможе продати будь-який з цих двох видів продукції у такій кількості, в якій зможе зробити на наявних виробничих потужностях. У той же час, оскільки обидва види продукції можуть бути вироблені на одних і тих же потужностях третього заводу, не зовсім зрозуміло, яке поєднання їх випусків буде найбільш прибутковим.

Кожна одиниця продукції першого виду вимагає однієї години витрат виробничого часу і трьох годин часу на збірку. Даний продукт приносить прибуток 30 грн на одиницю. Кожна одиниця другого продукту вимагає двох годин виробничого часу, двох годин на збірку і приносить прибуток 50 грн. Загальна доступна тижнева потужність заводів 1, 2 і 3 становить 4, 12 і 18 годин відповідно.

Сформулюйте математичну модель задачі.

- Тканина трьох артикулів виробляється на ткацьких верстатах двох типів з різною продуктивністю. Сировиною для виготовлення тканин є пряжа і барвники. Основні параметри виробництва, його собівартість і ціни реалізації готового продукту наведені у таблиці.

Вид ресурсу	Об'єм ресурсу	Продуктивність і норми витрат		
		Артикул 1	Артикул 2	Артикул 3
Верстати I типу	29000 год	20 м/год	10 м/год	25 м/год
Верстати II типу	30000 год	8 м/год	20 м/год	10 м/год
Пряжа (кг)	30000 кг	0.12 кг/м	0.18 кг/м	0.21 кг/м
Барвники (кг)	3000 кг	0.01 кг/м	0.005 кг/м	0.008 кг/м
Ціна 1 м тканини, грн. од.		45	45	60
Собівартість 1 м тканини, грн. од.		24	20	40

Складіть план виробництва тканин, що забезпечує максимальний прибуток.

Варіант 22

- Цех випускає два типи продукції, використовуючи два види напівфабрикатів. Продукція використовується при комплектуванні виробів, при цьому на кожну одиницю продукції

першого виду потрібно не більше двох одиниць продукції другого виду. Норми витрат напівфабрикатів кожного виду на одиницю продукції, загальні обсяги напівфабрикатів і прибуток від одиниці кожної продукції представлені у таблиці.

Напів- фабрикати	Норми витрат на одиницю продукції		Обсяг напівфабрикатів
	Продукція 1	Продукція 2	
1	1	2	800
2	6	2	2400
Прибуток	10	35	

Потрібно визначити план виробництва, що забезпечує максимальний прибуток.

- Три плодівих господарства поставляють апельсини в ящиках чотирьом оптовим покупцям. Щоденний попит цих покупців становить 20, 110, 40 і 110 ящиків відповідно. Плодові господарства можуть щодня поставляти 60, 120 і 100 ящиків апельсинів відповідно. Транспортні витрати на один ящик апельсинів наведені у таблиці.

Господарства	Покупці			
	A	B	C	D
I	1	2	5	3
II	1	6	5	2
III	6	3	7	4

Знайдіть оптимальний план поставок.

Варіант 23

- Молочний комбінат освоїв випуск двох нових сортів сиру: «Ніжний» і «Петровський». Очікуваний попит на них може скласти відповідно не більше 15 і 12 т на місяць. Оскільки комбінат випускає також традиційні види продукції, кожен з чотирьох цехів може виділити на виробництво сирів обмежений місячний ресурс часу. Кошти, виділені ліміти часу, а також витрати часу роботи кожного цеху на здійснення технологічного процесу при виробленні 1 т сиру кожного сорту наведені у таблиці, де також представлені оптові ціни сирів.

№ цеху	Витрати часу на вироблення 1 т сиру, год		Місячні ліміти часу, год
	«Ніжний»	«Петровський»	
1	2	7	66
2	3	5	45
3	2	4	58
4	1	6	72
Оптова ціна, тис. гр. од./т	156	168	

Ставиться мета: визначити оптимальний обсяг випуску сиру кожного сорту, що забезпечує максимальну виручку від продажу.

- Фермерський кооператив володіє 130 акрами землі (акр ≈ 0.4 га), на яких він виробляє три основні зернові культури: соєві боби, пшеницю і кукурудзу. Продукція, вироблена кооперативом, використовується як для власного споживання, так і для продажу. Кооператив організований таким чином, що потреби його членів у вищевказаних продуктах повинні бути задоволені у першу чергу, тобто до того, як будь-який з них буде проданий. Всі надлишки продукції продаються за ринковими цінами.

У таблиці, наведеній нижче, для кожного продукту показана: врожайність з акра, кількість бушелів (бушель = 36.3 л), необхідних членам кооперативу, максимальна потреба ринку (в бушелях) і очікуваний прибуток.

Культура	Врожайність (бушелів з акра)	Потреби коо- перативу (бушелів)	Потреби рин- ку (бушелів)	Прибуток (\$ за бушель)
Соеві боби	420	2000	10000	1.50
Пшениця	200	5000	8000	1.80
Кукурудза	70	1000	3000	2.50

Сформулюйте модель лінійного програмування, яка дозволить кооперативу визначити кількість акрів землі, що відводяться під кожну культуру, для забезпечення максимального прибутку.

Варіант 24

- Фірма виробляє і продає два продукти: Альфа і Гамма, ціни реалізації яких складають 5 \$ і 4 \$ за одиницю відповідно. Обидва продукти містять два інгредієнти — Д і К. Альфа містить 80 % інгредієнта Д і 20 % інгредієнта К, у той час як Гамма — 40 і 60 % відповідно. Поточні запаси складають 16000 галонів інгредієнта Д і 6000 галонів інгредієнта К. Єдина компанія, яка виробляє ці інгредієнти і поставляє їх фірмі, страйкує. Фірма хотіла б знати кількість галонів продуктів Альфа і Гамма, яке вона повинна буде виробити з наявних запасів для максимізації сумарного прибутку.
- Для обігріву приміщень використовують чотири агрегати, кожен з яких може працювати на будь-якому з п'яти сортів палива, наявному у кількостях 90, 110, 70, 80 і 150 т. Потреба у паливі кожного з агрегатів відповідно дорівнює 80, 120, 140 і 160 т. Теплотворна здатність i -го сорту палива при використанні його на j -му агрегаті задається матрицею

$$C = \begin{pmatrix} 8 & 7 & 9 & 11 & 8 \\ 6 & 5 & 8 & 7 & 6 \\ 7 & 11 & 5 & 8 & 7 \\ 9 & 8 & 7 & 9 & 11 \end{pmatrix}.$$

Знайдіть такий розподіл палива між агрегатами, при якому виходить максимальна кількість теплоти від використання всього палива.

Варіант 25

- Автомобільний завод виробляє легкові та вантажні автомобілі. Завод включає три цехи, виробничі потужності яких за кожним видом автомобілів у розрахунку на місяць подані у таблиці.

Цех	Виробничі потужності у розрахунку на місяць	
	Легкові авто	Вантажні авто
Кузовний	100	200
Моторний	120	144
Збірки	110	120

Прибуток від випуску одиниці легкового автомобіля становить 11 000 гр. од., вантажного — 16 000 гр. од. Розрахуйте обсяги випуску легкових і вантажних автомобілів, що забезпечують максимальний прибуток.

2. М'ясокомбінат має у своєму складі чотири заводи, на кожному з яких може виготовлятися три види ковбасних виробів. Потужності кожного із заводів відповідно 320, 280, 270 і 350 тонн на добу. Щоденні потреби у ковбасних výroбах кожного виду також відомі: відповідно 450, 370 і 400 т. Знаючи собівартість 1 т кожного виду виробів на кожному заводі, яка визначається матрицею

$$C = \begin{pmatrix} 2 & 3 & 4 \\ 1 & 5 & 3 \\ 6 & 4 & 2 \\ 7 & 8 & 5 \end{pmatrix},$$

знайдіть такий розподіл випуску ковбасних виробів між заводами, при якому собівартість виготовленої продукції є мінімальною.

Завдання 2. Графічне розв'язування задач лінійного програмування

1. Знайдіть графічно оптимальний розв'язок задачі 1 із завдання 1. За необхідності потрібно скоротити число змінних до двох (наприклад, вважати, що всі види продукції, окрім першого і другого, зняті з виробництва) і переформулювати модель.
2. Спробуйте використати для цієї задачі інтерактивний симплекс-метод у Sage.

Завдання 3. Симплекс-метод для задачі ЛП у стандартній формі

Розв'яжіть задачу лінійного програмування

$$z = (i + 1) \cdot x_1 + (j + 1) \cdot x_2 \rightarrow \max$$

при обмеженнях

$$\begin{cases} (i + 4) \cdot x_1 + i \cdot x_2 \leq (j + 3) \\ x_1 + (i + 1) \cdot x_2 \leq (j + 5) \\ (i + 2) \cdot x_1 + 5 \cdot x_2 \leq 10 \\ x_k \geq 0, \quad k = 1, 2, \end{cases}$$

де параметри i, j визначаються у залежності від номеру варіанту N :

$$i = N \bmod 4, \quad j = N \bmod 5.$$

Завдання 4. М-метод для задачі ЛП у загальній формі

Розв'яжіть задачу лінійного програмування

$$z = (i + 1) \cdot x_1 + x_2 \rightarrow \min$$

при обмеженнях

$$\begin{cases} (i+5) \cdot x_1 + x_2 = (j+2) \\ (i+2) \cdot x_1 + x_2 \geq (j+1) \\ (i+1) \cdot x_1 + x_2 \leq (j+1) \\ x_k \geq 0, \quad k = 1, 2, \end{cases}$$

де параметри i, j визначаються у залежності від номеру варіанту N :

$$i = N \bmod 4, \quad j = N \bmod 5.$$

Завдання 5. Застосування пакетів прикладних програм для розв'язування ЗЛП

Знайдіть оптимальні розв'язки задач 1 і 2 (див. завдання 1), використовуючи

- надбудову Solver (Пошук Розв'язку) для електронних таблиць Google Sheets, LibreOffice Calc, OpenOffice.org Calc або Excel;
- розв'язувачі `linear_program`, `MixedIntegerLinearProgram` у Sage або інші.

2.6.3 Розв'язування транспортних задач

Основне завдання

Постановка задачі Три плодівих господарства A_1, A_2, A_3 поставляють апельсини в ящиках чотирьом оптовим покупцям B_1, B_2, B_3, B_4 . Відомі щоденні потреби цих покупців, а також максимально можливі обсяги поставок апельсинів кожним плодівим господарством. Крім того, задані транспортні витрати у розрахунку на один ящик апельсинів при перевезенні від постачальників до споживачів.

Знайдіть оптимальний план поставок апельсинів з точки зору мінімізації сумарних витрат на їх транспортування.

1. Побудуйте математичну модель задачі. Які припущення вводяться у даному випадку для спрощення моделі?
2. Знайдіть початковий план перевезень методами північно-західного кута і найменшої вартості. Який з двох методів дає кращий початковий розв'язок і чому?
3. Знайдіть оптимальний план перевезень методом потенціалів.
4. Перевірте правильність отриманого розв'язку, використовуючи один з пакетів прикладних програм.
5. Модифікуйте програму знаходження оптимального плану перевезень з урахуванням наступних додаткових обмежень:

- перевезення від деяких постачальників до деяких споживачів заборонені (конкретне обмеження задайте самостійно);
- існують укладені контракти на обов'язкові поставки у зазначених обсягах (конкретне обмеження задайте самостійно);
- пропускна здатність шляхів між деякими постачальниками і споживачами обмежена (конкретне обмеження задайте самостійно).

Варіанти початкових даних

Варіант №	Запаси (ящ.)			Попит (ящ.)				Витрати на перевезення одиниці вантажу (гр. од.)											
								$A_1 \Rightarrow B_j$				$A_2 \Rightarrow B_j$				$A_3 \Rightarrow B_j$			
	A_1	A_2	A_3	B_1	B_2	B_3	B_4	B_1	B_2	B_3	B_4	B_1	B_2	B_3	B_4	B_1	B_2	B_3	B_4
1	35	45	25	15	20	40	25	4	2	1	3	6	1	2	5	4	3	7	1
2	45	40	75	20	35	40	25	4	1	3	5	6	3	5	7	4	9	6	3
3	25	25	35	20	40	55	20	3	2	3	6	4	2	4	5	8	7	5	4
4	45	25	50	45	20	60	10	1	7	6	2	1	8	3	4	3	5	1	9
5	45	70	40	30	30	25	35	3	6	1	6	3	5	3	2	5	4	6	9
6	20	35	45	75	20	30	10	2	1	4	4	7	1	2	5	8	3	9	2
7	50	65	30	25	35	25	60	1	3	2	7	6	5	8	2	7	4	3	2
8	30	70	35	40	45	55	25	2	7	5	1	3	5	4	3	5	2	1	5
9	35	45	70	20	25	30	20	2	3	1	6	5	3	1	6	7	8	2	5
10	35	45	25	15	20	40	25	4	2	1	3	6	1	2	5	4	3	7	1
11	35	55	25	30	60	10	25	3	2	1	4	5	3	1	6	2	3	6	2
12	50	85	20	75	15	25	70	1	9	6	7	3	5	4	2	3	5	4	9
13	10	25	45	65	40	10	20	7	6	5	6	2	4	3	1	5	3	6	8
14	30	10	25	45	15	60	10	6	9	4	3	1	5	6	8	5	7	6	4
15	25	40	65	10	35	40	15	4	2	3	5	4	3	5	7	4	3	6	2
16	65	15	20	35	40	75	10	5	8	6	9	1	2	4	7	3	6	8	5
17	20	85	45	15	35	65	55	4	6	9	4	5	7	6	8	4	9	3	2
18	50	40	25	10	75	45	40	2	5	3	6	8	5	9	6	4	5	2	1
19	35	40	15	20	35	30	10	7	5	3	1	5	9	8	6	1	4	5	7
20	25	50	70	85	30	40	85	2	5	8	4	6	5	7	5	3	1	5	9
21	55	45	25	65	15	25	40	8	2	4	6	5	3	1	5	7	5	9	5
22	25	30	10	75	10	50	20	2	4	2	3	6	8	8	4	4	2	2	6
23	85	65	50	80	25	10	45	9	5	6	4	7	5	1	5	2	3	4	6
24	20	30	45	15	25	50	35	1	3	7	2	4	1	3	5	2	3	6	2
25	40	50	35	20	30	45	25	5	2	1	4	3	7	2	2	5	1	2	6

Додаткове завдання

Задача для непарних варіантів Автозавод виробляє автомобілі у пунктах A_1 і A_2 та має склад у пункті D . Поставки здійснюються у пункти B_1 і B_2 . Питомі витрати (у дол.) На перевезення між пунктами наведені у таблиці.

З пункту	У пункт				
	A_1	A_2	D	B_1	B_2
A_1	—	140	100	90	225
A_2	145	—	111	110	119
D	105	115	—	113	78
B_1	—	—	121	—	—
B_2	—	—	82	—	—

У пункті A_1 може бути вироблено до 1100 автомобілів, у пункті A_2 — до 2900. Попит у пункті B_1 становить 2400 автомобілів, у пункті B_2 — 1500.

Визначте план перевезень, що враховує можливість транспортування автомобілів між пунктами виробництва, а також через перевалочний пункт (склад), при якому сумарні транспортні витрати будуть найменшими.

- Побудуйте математичну модель задачі. Як звести дану постановку до класичного вигляду?
- Знайдіть оптимальний розв'язок, використовуючи пакет прикладних програм. Зобразіть отриманий план перевезень графічно.
- Модифікуйте відповідь за умови, що поставки між пунктами A_1 і A_2 неможливі.
- Модифікуйте відповідь, якщо поставки між пунктами B_1 і B_2 можливі за ціною 5 дол. у розрахунку на один автомобіль.

Задача для парних варіантів Нафтова компанія видобуває нафту з двох свердловин A_1, A_2 . Перша свердловина дає до 150 000 барелів нафти за добу, друга — до 200 000. Поставки здійснюються у пункти B_1 і B_2 . Можливі як прямі поставки, так і через проміжні пункти D_1 і D_2 . Попит пункту B_1 становить 160 000 барелів нафти на добу, пункту B_2 — 140 000 барелів на добу. Вартість доставки (в дол.) 1000 барелів нафти між пунктами наведена у таблиці.

З пункту	У пункт					
	A_1	A_2	D_1	D_2	B_1	B_2
A_1	—	—	10	13	25	28
A_2	—	—	15	12	26	25
D_1	—	—	—	6	16	17
D_2	—	—	6	—	14	16
B_1	—	—	—	—	—	15
B_2	—	—	—	—	15	—

Знайдіть план поставок нафти з урахуванням можливостей її видобутку і попиту споживачів, при якому сумарні транспортні витрати будуть мінімальними.

- Побудуйте математичну модель задачі. Як звести дану постановку до класичного вигляду?
- Знайдіть оптимальний розв'язок, використовуючи пакет прикладних програм. Зобразіть отриманий план перевезень графічно.
- Припустимо, що перед тим, як нафта буде доставлена у пункти B_1 і B_2 , вона повинна бути очищена у пунктах D_1 і D_2 . Очищення 1000 барелів нафти у пункті D_1 коштує 12 дол., а у пункті D_2 — 10 дол. Припускаючи, що пункти D_1 і D_2 мають необмежені потужності з очищення нафти, сформулюйте задачу з проміжними пунктами для мінімізації добових витрат на перевезення і очищення нафти. Знайдіть оптимальний розв'язок.
- Модифікуйте задачу за умови, що потужності устаткування для очищення нафти обмежені: у пункті D_1 — до 150 000 барелів на добу, у пункті D_2 — до 180 000 барелів на добу. Знайдіть оптимальний розв'язок для цієї ситуації.

Чисельні методи мінімізації функцій багатьох змінних

3.1 Методи безумовної мінімізації

Методи безумовної мінімізації різняться або вибором напрямку спуску, або способом руху вздовж напрямку спуску. Методи спуску мають таку процедуру побудови послідовності $\{\mathbf{x}^k\}$. За початкову точку обирається, взагалі кажучи, довільна точка $\mathbf{x}^0 \in E^n$. Послідовні наближення $\mathbf{x}^1, \mathbf{x}^2, \dots$ знаходимо за схемою

$$\mathbf{x}^{k+1} = \mathbf{x}^k + \beta_k \mathbf{s}^k,$$

де $\mathbf{s}^k$ — напрямок спуску, β_k — довжина кроку руху вздовж $\mathbf{s}^k$.

У залежності від максимального порядку похідної, що бере участь в утворенні $\mathbf{s}^k$, методи бувають нульового, першого і другого порядків.

3.1.1 Метод нульового порядку Нелдера — Міда

У просторі E^n задається $(n + 1)$ точка, які є вершинами регулярного многогранника. Ці точки є пробними у нашому методі. З геометрії відомо, що координати вершин регулярного многогранника визначаються матрицею $\mathbf{D}$, стовпчики якої є вершинами:

$$\mathbf{D} = \begin{pmatrix} 0 & d_1 & d_2 & \dots & d_2 & d_2 \\ 0 & d_2 & d_1 & \dots & d_2 & d_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & d_2 & d_2 & \dots & d_2 & d_1 \end{pmatrix},$$

де $d_1 = \frac{t}{n\sqrt{2}} (\sqrt{n+1} + n - 1)$, $d_2 = \frac{1}{n\sqrt{2}} (\sqrt{n+1} - 1)$, t — відстань між двома вершинами.

Наприклад, для $n = 2$ і $t = 1$ трикутник має такі координати:

вершини	x_1	x_2
1	0	0
2	0.965	0.259
3	0.259	0.965

Розглянемо алгоритм методу. Нехай $\mathbf{x}_i^k = [x_{i1}^k, \dots, x_{in}^k]^\top$, $i = \overline{1, n+1}$ є вершиною в $\mathbb{R}^n$ на k -му етапі пошуку. Відзначимо ті вершини многогранника, в яких $\varphi(\mathbf{x})$ досягає максимального і мінімального значення:

$$\begin{aligned}\varphi(\mathbf{x}_n^k) &= \max [\varphi(\mathbf{x}_1^k), \dots, \varphi(\mathbf{x}_{n+1}^k)], \\ \varphi(\mathbf{x}_l^k) &= \min [\varphi(\mathbf{x}_1^k), \dots, \varphi(\mathbf{x}_{n+1}^k)].\end{aligned}$$

Нехай $\mathbf{x}_{n+2}^k$ буде центром ваги всіх вершин, за виключенням $\mathbf{x}_n$. Координати цього центру визначаються за формулою

$$x_{n+2,j}^k = \frac{1}{n} \left[\sum_{i=1}^{n+1} x_{ij} - x_{nj}^k \right], \quad j = \overline{1, n},$$

де j — номер координати точки.

Крок 1. Відбиття Проектуємо $\mathbf{x}_n^k$ через центр ваги відповідно із співвідношенням

$$\mathbf{x}_{n+3}^k = \mathbf{x}_{n+2}^k + \alpha (\mathbf{x}_{n+2}^k - \mathbf{x}_n^k),$$

де $\alpha > 0$ є коефіцієнт відбиття.

Крок 2. Розтягування Якщо $\varphi(\mathbf{x}_{n+3}^k) \leq \varphi(\mathbf{x}_l^k)$, то вектор $(\mathbf{x}_{n+3}^k - \mathbf{x}_{n+2}^k)$ розтягується відповідно зі співвідношенням

$$\mathbf{x}_{n+4}^k = \mathbf{x}_{n+2}^k + \gamma (\mathbf{x}_{n+3}^k - \mathbf{x}_{n+2}^k),$$

де $\gamma > 1$ — коефіцієнт розтягування. Якщо $\varphi(\mathbf{x}_{n+4}^k) < \varphi(\mathbf{x}_l^k)$, то $\mathbf{x}_n^k$ замінимо на $\mathbf{x}_{n+4}^k$ і процедура починається з [кроку 1](#) при $k = k + 1$. В іншому випадку $\mathbf{x}_n^k$ замінимо на $\mathbf{x}_{n+3}^k$ і переходимо до [кроку 1](#) при $k = k + 1$.

Крок 3. Стиснення Якщо $\varphi(\mathbf{x}_{n+3}^k) > \varphi(\mathbf{x}_i^k)$ для всіх $i \neq h$, то вектор $(\mathbf{x}_h^k - \mathbf{x}_{n+2}^k)$ стискується у відповідності з формулою

$$\mathbf{x}_{n+5}^k = \mathbf{x}_{n+2}^k + \beta (\mathbf{x}_h^k - \mathbf{x}_{n+2}^k),$$

де $0 < \beta < 1$ — коефіцієнт стиснення. Потім $\mathbf{x}_h^k$ замінюємо на $\mathbf{x}_{n+5}^k$ і повертаємося до [кроку 1](#) при $k = k + 1$.

Крок 4. Редукція Якщо $\varphi(\mathbf{x}_{n+3}^k) > \varphi(\mathbf{x}_h^k)$, то всі вектори $(\mathbf{x}_i^k - \mathbf{x}_l^k)$, $i = \overline{1, n+1}$ зменшуються у два рази відповідно з формулою

$$\mathbf{x}_i^k = \mathbf{x}_l^k + 0.5 (\mathbf{x}_i^k - \mathbf{x}_l^k), \quad i = \overline{1, n+1}.$$

Потім повертаємося до [кроку 1](#) при $k = k + 1$.

Критерій зупинки має вигляд

$$\left\{ \frac{1}{n+1} \sum_{i=1}^{n+1} [\varphi(\mathbf{x}_i^k) - \varphi(\mathbf{x}_{n+2}^k)]^2 \right\}^{\frac{1}{2}} \leq \varepsilon,$$

де ε — мале число, а $\mathbf{x}_{n+2}^k$ — центр ваги на k -му кроці.

Коефіцієнт відбиття α використовується для проектування вершини з найбільшим значенням $\varphi(\mathbf{x})$ через центр ваги многогранника.

Коефіцієнт γ вводиться для розтягування вектора пошуку у випадку, коли відбиття дає вершину зі значенням $\varphi(\mathbf{x})$ меншим, ніж найменше значення $\varphi(\mathbf{x})$, отримане до відбиття.

Коефіцієнт стиснення β використовується для зменшення вектора пошуку, якщо операція відбиття не привела вершини зі значенням $\varphi(\mathbf{x})$ меншим, ніж друге по величині (після найменшого) значення $\varphi(\mathbf{x})$, отримане до відбиття.

Таким чином, за допомогою операції розтягування або стиснення розміри і форма многогранника масштабуються так, щоб вони задовольняли топології задачі.

Яким же чином обирати α , β і γ ? Рекомендації такі.

Нелдер і Мід довели, що при розв'язуванні задачі з $\alpha = 1$ використовується менша кількість обчислень функції, ніж при $\alpha < 1$. З іншого боку, α не повинне бути набагато більшим 1, тому що:

- 1) многогранник краще адаптується до топології задачі при менших значеннях α , особливо коли необхідно змінювати напрямок пошуку, зіткнувшись зі зігнутою впади-

ною;

- 2) в околі локального мінімуму розміри многогранника повинні зменшуватися і велике α у цих випадках уповільнює збіжність. Отже, обирають $\alpha = 1$.

Нелдер і Мід рекомендували брати $\alpha = 1$, $\beta = 0.5$, $\gamma = 2$. Розміри і орієнтація початкового многогранника впливали на час розв'язування задачі, але значення α , β і γ мали більший вплив.

Павіані зазначає, що важко визначити правило вибору α , β і γ , але вплив вибору β на ефективність пошуку більший, ніж вибір γ . Він рекомендує такі діапазони вибору параметрів: $0.4 \leq \beta \leq 0.6$, $2.8 \leq \gamma \leq 3.0$. При $0 < \beta < 0.4$ існує ймовірність того, що через сплюснення многогранника буде мати місце передчасне закінчення процесу. При $\beta > 0.6$ може збільшитись кількість кроків для досягнення точки мінімуму.

Зазначимо, що існує багато методів нульового порядку мінімізації функцій: різні варіанти покоординатного спуску, Розенброка і Девіса, Свенна, Кемпі, Пауелла [13].

3.1.2 Методи першого порядку

Для того щоб у чисельному методі мінімізації функції отримати послідовність точок, що збігається при деяких умовах до точки мінімуму, необхідно знати напрямок спуску і крок руху по напрямку спуску. Відомо, що напрямок $\mathbf{s}^k$ буде напрямком спуску в точці $\mathbf{x}^k$, якщо виконується умова

$$\left(\frac{\partial \varphi(\mathbf{x}^k)}{\partial \mathbf{x}}, \mathbf{s}^k \right) < 0.$$

За допомогою цього співвідношення можна визначити, чи є напрямок $\mathbf{s}^k$ напрямком спуску.

Також необхідно знати, як обирати крок β_k у формулі для отримання послідовності точок

$$\mathbf{x}^{k+1} = \mathbf{x}^k + \beta_k \mathbf{s}^k.$$

Розглянемо декілька способів вибору кроку β_k .

- 1) $\beta_k : \min_{\beta \geq 0} \varphi(\mathbf{x}^k + \beta \mathbf{s}^k)$. Тобто необхідно скористатися одновимірною мінімізацією функції $\varphi(\mathbf{x}^k + \beta \mathbf{s}^k)$ за параметром β . Для цього користуємося методами одновимірної мінімізації, що розглядалися вище.
- 2) $\beta_k = \text{const}$ і перевіряється умова монотонності $\varphi(\mathbf{x}^{k+1}) < \varphi(\mathbf{x}^k)$. Якщо вона порушується, то β зменшуємо до тих пір, поки не відновиться монотонність. Час від часу β збільшують зі збереженням монотонності.
- 3) Задаємо деякі числа β , ε , $0 < \varepsilon < 1$, $0 < \beta < 1$. Далі перевіряємо на кожному кроці

нерівність

$$\varphi(\mathbf{x}^k + \beta \mathbf{s}^k) \leq \varphi(\mathbf{x}^k) - \varepsilon \beta \left\| \frac{\partial \varphi(\mathbf{x}^k)}{\partial \mathbf{x}} \right\|^2.$$

Якщо нерівність не виконується, то замінюємо β на $\beta \varepsilon$ і перевірка повторюється.

- 4) Інколи величину β_k обирають до початку обчислень. Можна скористатися такою процедурою вибору β_k :

$$\beta_k > 0, \quad k = 0, 1, \dots, \quad \sum_{k=0}^{\infty} \beta_k = +\infty, \quad \sum_{k=0}^{\infty} \beta_k^2 < +\infty$$

(наприклад, $\beta_k = \frac{C}{k+1}$). Умова збіжності ряду $\sum_{k=0}^{\infty} \beta_k^2$ дає досить швидку збіжність послідовності β_k до нуля з метою забезпечення збіжності методу в околі точки мінімуму. Умова розбіжності $\sum_{k=0}^{\infty} \beta_k$ забезпечує досягнення точки мінімуму при невдалому виборі точки $\mathbf{x}^0$.

Зазначимо, що у кожному із методів може бути і своє правило вибору β_k .

Критерії зупинки

Майже всі оптимізаційні методи дають розв'язок лише у граничному розумінні. Тому у програмах, що їх реалізують, доводиться передбачати спеціальні правила переривання розрахунку — критерії зупинки.

Оптимізаційний процес бажано перервати, якщо:

- 1) досягнута потрібна точність розв'язку;
- 2) добре наближення ще не знайдено, але швидкість руху до оптимуму досить мала і немає сенсу рахувати далі;
- 3) метод почав розбігатися або зациквився.

Якщо задача погано обумовлена, тобто $q = \frac{l}{L} \ll 1$, де l і L — мінімальне і максимальне власні числа матриці $\frac{\partial^2 \varphi(\mathbf{x}^*)}{\partial \mathbf{x}^2}$, то чекати доброго наближення $\mathbf{x}^k$ до $\mathbf{x}^*$ даремно. Тому регулювати точність будемо параметром τ — кількістю правильних розрядів $\varphi(\mathbf{x}^k)$, що не перевищують за модулем одиниці, старші розряди після десяткової точки повинні враховуватись, навіть якщо у них знаходяться нулі. Наприклад, якщо $\tau = 10^{-5}$, то це означає, що при $|\varphi(\mathbf{x}^*)| \geq 1$ бажано, щоб у $\varphi(\mathbf{x}^k)$ і $\varphi(\mathbf{x}^*)$ співпадали п'ять перших значущих цифр, а при $|\varphi(\mathbf{x}^*)| \leq 1$ — щоб $\varphi(\mathbf{x}^k)$ і $\varphi(\mathbf{x}^*)$ різнилися не більше, ніж на 10^{-5} .

При розв'язуванні задач безумовної мінімізації гладких функцій критерій зупинки

має вигляд:

$$\begin{aligned} A_1: \quad & \varphi(\mathbf{x}^{k-1}) - \varphi(\mathbf{x}^k) < \tau(1 + |\varphi(\mathbf{x}^k)|), \\ A_2: \quad & \|\mathbf{x}^{k-1} - \mathbf{x}^k\| < \sqrt{\tau}(1 + \|\mathbf{x}^k\|), \\ A_3: \quad & \left\| \frac{\partial \varphi(\mathbf{x}^k)}{\partial \mathbf{x}} \right\| \leq \sqrt[3]{\tau}(1 + |\varphi(\mathbf{x}^k)|). \end{aligned}$$

Ще раз згадаємо

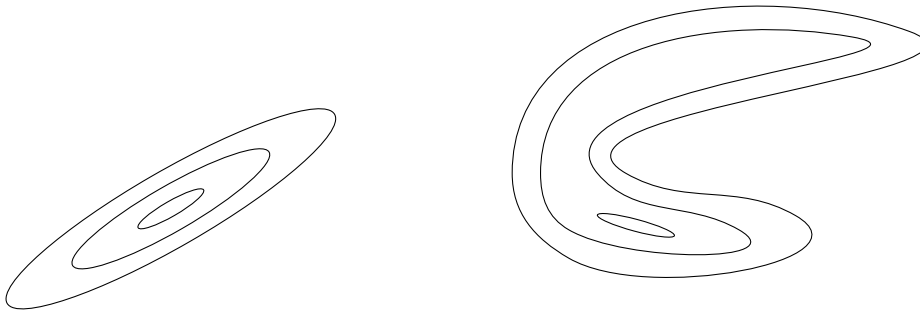
Означення 3.1. *Лініями постійного рівня функції $\varphi(\mathbf{x})$ називається множина точок $\mathbf{x} \in \mathbb{R}^n$, для яких виконується рівність $\varphi(\mathbf{x}) = \text{const}$.*

Приклад 3.1. Розглянемо функцію $\varphi(\mathbf{x}) = x_1^2 + 25x_2^2$. Знайдемо лінії постійного рівня цієї функції. Маємо $x_1^2 + 25x_2^2 = c$; далі,

$$\frac{x_1^2}{(\sqrt{c})^2} + \frac{x_2^2}{\left(\sqrt{\frac{c}{25}}\right)^2} = 1.$$

Отже, лініями постійного рівня функції $\varphi(\mathbf{x}) = x_1^2 + 25x_2^2$ є еліпси.

Означення 3.2. *Функція $y = \varphi(\mathbf{x})$ називається **яристою**, якщо її лінії постійного рівня витягнуті або вигнуті.*



Це геометричне тлумачення яристої функції.

Означення 3.3. *Функція $y = \varphi(\mathbf{x})$ називається **яристою**, якщо її матриця Гессе, що розрахована у точці мінімуму, погано обумовлена, тобто $g = \frac{l}{L} \ll 1$, де l і L — мінімальне і максимальне власні числа цієї матриці.*

3.1.2.1 Градієнтний метод

Розглянемо дуже важливий в ідейному відношенні метод безумовної мінімізації — градієнтний метод. Цей метод рідко використовується на практиці. Він служить моделлю

для побудови більш ефективних методів.

Нехай функція $\varphi(\mathbf{x})$ — неперервно диференційовна. Алгоритм градієнтного методу має вигляд

$$\mathbf{x}^{k+1} = \mathbf{x}^k - \beta_k \frac{\partial \varphi(\mathbf{x}^k)}{\partial \mathbf{x}}.$$

Будемо вважати, що послідовність точок $\mathbf{x}^k$ така, що $\frac{\partial \varphi(\mathbf{x}^k)}{\partial \mathbf{x}} \neq 0$. Тобто у жодній із точок $\mathbf{x}^k$ не виконується необхідна умова мінімуму.

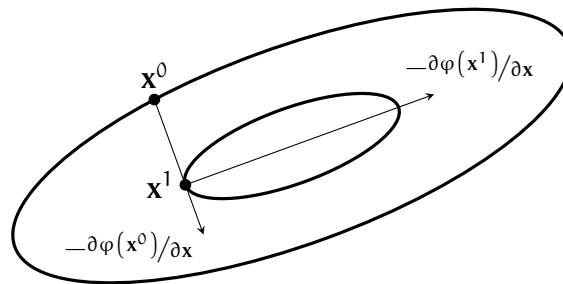
Зазначимо, що

$$\left(\frac{\partial \varphi(\mathbf{x}^k)}{\partial \mathbf{x}}, \mathbf{s}^k \right) = \left(\frac{\partial \varphi(\mathbf{x}^k)}{\partial \mathbf{x}}, -\frac{\partial \varphi(\mathbf{x}^k)}{\partial \mathbf{x}} \right) = - \left\| \frac{\partial \varphi(\mathbf{x}^k)}{\partial \mathbf{x}} \right\| < 0,$$

а це означає, що $\mathbf{s}^k = -\frac{\partial \varphi(\mathbf{x}^k)}{\partial \mathbf{x}}$ є напрямок спуску (це відомо студентам ще на першому курсі при вивченні дисципліни «Математичний аналіз»).

Початкова точка обирається довільно. Якщо β_k обирати за першим способом: $\beta_k : \min_{\beta \geq 0} \varphi \left(\mathbf{x}^k - \beta \frac{\partial \varphi(\mathbf{x}^k)}{\partial \mathbf{x}} \right)$, то метод називається методом найшвидшого спуску.

У методі найшвидшого спуску напрямок руху від точки $\mathbf{x}^k$ дотикається лінії постійного рівня функції у точці $\mathbf{x}^{k+1}$.

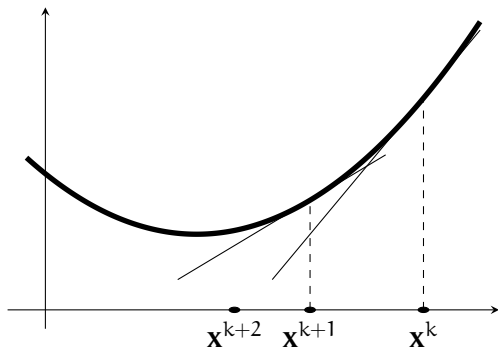


Послідовність точок $\mathbf{x}^0, \mathbf{x}^1, \dots, \mathbf{x}^k, \dots$ зигзагоподібно наближається до точки мінімуму $\mathbf{x}^*$, і при цьому ланки цього зигзагу ортогональні між собою. Дійсно, β обирається з умови мінімуму функції $\psi(\beta) = \varphi \left(\mathbf{x}^k - \beta \frac{\partial \varphi(\mathbf{x}^k)}{\partial \mathbf{x}} \right)$ і тому

$$\frac{d\psi}{d\beta} = -\frac{\partial \varphi(\mathbf{x}^{k+1})}{\partial \mathbf{x}} \frac{\partial \varphi(\mathbf{x}^k)}{\partial \mathbf{x}} = 0.$$

Таким чином, напрямки спуску на двох послідовних ітераціях взаємно ортогональні.

Геометрична інтерпретація градієнтного методу



У градієнтному методі функцію $\varphi(x)$ апроксимуємо в околі точки x^k лінійною функцією, тобто проводимо дотичну до функції $\varphi(x)$ у точці $(x^k, \varphi(x^k))$. Далі, замість функції $\varphi(x)$ мінімізуємо лінійну функцію. А так як вона не має мінімуму $(-\infty)$, то знаходимо її мінімум у невеликому околі точки x^k . Ця точка мінімуму лінійної функції є точкою x^{k+1} і т. д.

Розглянемо теореми збіжності.

Теорема 3.1. Нехай функція $\varphi(x)$ диференційовна, обмежена знизу, а градієнт задовольняє умові Ліпшиця. Тоді послідовність $\{x^k\}$ така, що

$$\lim_{k \rightarrow \infty} \frac{\partial \varphi(x^k)}{\partial x} = 0.$$

Для сильно опуклих функцій можна отримати більш точні оцінки.

Теорема 3.2. Нехай функція $\varphi(x)$ сильно опукла, тобто

$$m\|y\|^2 \leq \left(\frac{\partial^2 \varphi(x)}{\partial x^2} y, y \right) \leq M\|y\|^2, \quad \text{де } M \geq m > 0,$$

а послідовність $\{x^k\}$ будується методом найшвидшого спуску. Тоді $\{x^k\}$ збігається до точки мінімуму зі швидкістю геометричної прогресії зі знаменником $q = \frac{M-m}{M+m}$, тобто

$$\|x^{k+1} - x^*\| \leq \frac{M-m}{M+m} \|x^k - x^*\|.$$

Зауваження 3.1. За величини M і m можна брати максимальне і мінімальне власні числа матриці $\frac{\partial^2 \varphi(x^*)}{\partial x^2}$.

Розглянемо квадратичну функцію $\varphi(x) = \frac{1}{2}x^T A x - b^T x + c$, $A > 0$. Якщо користуватися методом найшвидшого спуску, то крок β_k можна записати у явному вигляді:

$$\beta_k = \frac{\left\| \frac{\partial^2 \varphi(x^*)}{\partial x^2} \right\|^2}{\left(A \frac{\partial^2 \varphi(x^*)}{\partial x^2}, \frac{\partial^2 \varphi(x^*)}{\partial x^2} \right)}.$$

Теорема 3.3. Для квадратичної функції мають місце оцінки:

$$\begin{aligned}\varphi(\mathbf{x}^k) - \varphi(\mathbf{x}^*) &\leq (\varphi(\mathbf{x}^0) - \varphi(\mathbf{x}^*)) \left(\frac{L-l}{L+l} \right)^{2k}, \\ \|\mathbf{x}^k - \mathbf{x}^*\| &\leq q^k \sqrt{2l^{-1}(\varphi(\mathbf{x}^0) - \varphi(\mathbf{x}^*))}, \quad q = \frac{L-l}{L+l},\end{aligned}$$

де l — найменше, а L — найбільше власні числа матриці $\mathbf{A}$.

Зауважимо, що перша оцінка у теоремі є точною.

Якщо порівнювати метод найшвидшого спуску з іншими варіантами градієнтного методу, то можна прийти до висновку, що метод найшвидшого спуску для квадратичної функції збігається не швидше, ніж інші варіанти градієнтного методу. Цей висновок справедливий і для неквадратичних функцій.

Однак, не слід вважати, що неможливо збільшити швидкість градієнтного методу шляхом вибору довжини кроку. Наприклад, якщо для мінімізації квадратичної функції застосувати градієнтний метод з $\beta_k = \frac{1}{\lambda_{k+1}}$, $k = 0, 1, \dots, n-1$, де λ_k — власні числа матриці $\mathbf{A}$ і $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$, то такий варіант методу буде скінченним і $\mathbf{x}^n = \mathbf{x}^*$.

На практиці краще всього користуватися другим способом вибору кроку.

Якщо лінії постійного рівня є гіперсфери, то метод найшвидшого спуску збігається за один крок.

Отже, якщо функції гладкі і опуклі, то варіанти градієнтного методу збігаються зі швидкістю геометричної прогресії. Величина q для сильно опуклих функцій залежить від найбільшого і найменшого власних чисел матриці $\frac{\partial^2 \varphi(\mathbf{x}^*)}{\partial \mathbf{x}^2}$. Знаменник буде малим, якщо ці числа мало відрізняються одне від одного. Якщо ж вони сильно відрізняються (на порядок), то це означає, що функція яриста.

Мала швидкість збігу градієнтних методів або зациклювання не дозволяє розв'язувати за їх допомогою складні задачі мінімізації. Градієнтні методи використовуються у комбінації з більш ефективними методами на початковій стадії мінімізації, коли ще точки $\mathbf{x}^k$ досить далекі від точки мінімуму.

Приклад 3.2. Знайти точку мінімуму функції $\varphi(\mathbf{x}) = x_1^2 + 25x_2^2$ з точністю $\tau = 0.01$ методом найшвидшого спуску.

Розв'язування. За початкову точку візьмемо точку $\mathbf{x}^0 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$. Тоді

$$\mathbf{x}^1 = \mathbf{x}^0 - \beta_0 \frac{\partial \varphi(\mathbf{x}^0)}{\partial \mathbf{x}} = \begin{pmatrix} x_1^0 \\ x_2^0 \end{pmatrix} - \beta_0 \begin{pmatrix} 2x_1^0 \\ 50x_2^0 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} - \beta_0 \begin{pmatrix} 2 \\ 50 \end{pmatrix} = \begin{pmatrix} 1 - 2\beta_0 \\ 1 - 50\beta_0 \end{pmatrix}.$$

Крок β_0 знаходимо з умови

$$\begin{aligned}\beta_0 : \min_{\beta \geq 0} \varphi \left(\mathbf{x}^0 - \beta \frac{\partial \varphi(\mathbf{x}^0)}{\partial \mathbf{x}} \right) &= \min_{\beta \geq 0} \varphi \begin{pmatrix} 1 - 2\beta \\ 1 - 50\beta \end{pmatrix} = \\ &= \min_{\beta \geq 0} [(1 - 2\beta)^2 + 25(1 - 50\beta)^2].\end{aligned}$$

Отримаємо $\beta_0 \approx \frac{1}{50}$. Тоді $\mathbf{x}^1 = \begin{pmatrix} 1 - 2 \cdot \frac{1}{50} \\ 1 - 1 \end{pmatrix} = \begin{pmatrix} 0.96 \\ 0 \end{pmatrix}$.

Перевіримо критерій зупинки. Розглянемо пункт A_1 :

$$\varphi(\mathbf{x}^0) - \varphi(\mathbf{x}^1) < \tau(1 + |\varphi(\mathbf{x}^1)|).$$

Маємо

$$\begin{aligned}26 - 1 &<? 0.01(1 + 1), \\ 25 &\not< 0.02.\end{aligned}$$

Перший пункт критерію зупинки не виконується. Тоді два інші пункти недоцільно перевіряти.

Знайдемо друге наближення.

$$\mathbf{x}^2 = \mathbf{x}^1 - \beta_1 \frac{\partial \varphi(\mathbf{x}^1)}{\partial \mathbf{x}} = \begin{pmatrix} 0.96 \\ 0 \end{pmatrix} - \beta_1 \begin{pmatrix} 2 \cdot 0.96 \\ 50 \cdot 0 \end{pmatrix} = \begin{pmatrix} 0.96 - 1.92\beta \\ 0 \end{pmatrix}.$$

$$\begin{aligned}\beta_1 : \min_{\beta \geq 0} \varphi \left(\mathbf{x}^1 - \beta \frac{\partial \varphi(\mathbf{x}^1)}{\partial \mathbf{x}} \right) &= \min_{\beta \geq 0} \varphi \begin{pmatrix} 0.96 - 1.92\beta \\ 0 \end{pmatrix} = \\ &= \min_{\beta \geq 0} [(0.96 - 1.92\beta)^2 + 25 \cdot 0^2].\end{aligned}$$

Отримаємо $\beta_1 = 0.5$. Отже, $\mathbf{x}^2 = \begin{pmatrix} 0.96 - 1.92 \cdot 0.5 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$.

Перевіримо виконання критерію зупинки. Наприклад, розглянемо умову A_2 :

$$\|\mathbf{x}^{k-1} - \mathbf{x}^k\| < \sqrt{\tau}(1 + \|\mathbf{x}^k\|).$$

Отримали

$$\left\| \begin{pmatrix} 0 \\ 0 \end{pmatrix} - \begin{pmatrix} 0.96 \\ 0 \end{pmatrix} \right\| <? \sqrt{0.01} \left(1 + \left\| \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\| \right),$$

$$0.96 \not\leq 0.33.$$

Так як нерівність не виконується, то будемо знаходити третє наближення (хоча зрозуміло, що точка $\mathbf{x}^2$ є точкою мінімуму).

$$\mathbf{x}^3 = \mathbf{x}^2 - \beta_2 \frac{\partial \varphi(\mathbf{x}^2)}{\partial \mathbf{x}} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} - \beta_2 \begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Точки $\mathbf{x}^2$ і $\mathbf{x}^3$ співпали. Очевидно, що всі три умови критерію зупинки виконуються.

Тепер доведемо, що точка $\bar{\mathbf{x}} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ є точкою мінімуму функції $\varphi(\mathbf{x}) = x_1^2 + 25x_2^2$.

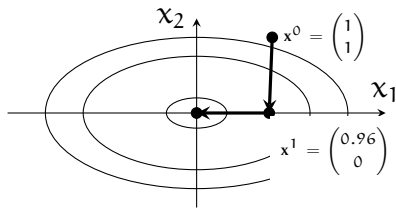
Спочатку доведемо, що функція є опуклою. Дійсно,

$$\mathcal{I} = \mathbf{x}^\top \frac{\partial^2 \varphi(\bar{\mathbf{x}})}{\partial \mathbf{x}^2} \mathbf{x} = 2x_1^2 + 50x_2^2 > 0, \quad \mathbf{x} \neq 0.$$

Так як $\frac{\partial^2 \varphi(\bar{\mathbf{x}})}{\partial \mathbf{x}^2} > 0$, то функція $\varphi(\mathbf{x}) = x_1^2 + 25x_2^2$ є опуклою. Посилаємось на теорему: якщо для опуклої функції $\varphi(\mathbf{x})$ виконується умова $\frac{\partial \varphi(\bar{\mathbf{x}})}{\partial \mathbf{x}} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$, то точка $\bar{\mathbf{x}}$ — точка мінімуму функції. Маємо $\frac{\partial \varphi(\bar{\mathbf{x}})}{\partial \mathbf{x}} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$. Отже, точка $\mathbf{x} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ є точкою мінімуму функції $\varphi(\mathbf{x}) = x_1^2 + 25x_2^2$. $\checkmark$

Зауваження 3.2. Функція $\varphi(\mathbf{x}) = x_1^2 + 25x_2^2$ є яристою. Дійсно, матриця Гессе має вигляд: $\frac{\partial^2 \varphi(\bar{\mathbf{x}})}{\partial \mathbf{x}^2} = \begin{pmatrix} 2 & 0 \\ 0 & 50 \end{pmatrix}$. Власні числа матриці 2 і 50. Вони відрізняються на порядок.

Якщо взяти кроки $\frac{1}{50}$ і $\frac{1}{2}$, то мінімум квадратичної функції знаходиться за два кроки, що ми і отримали.



На малюнку представлені лінії постійного рівня — це еліпси. Дно яру направлене по осі x_1 , а стінки яру — по осі x_2 . Спочатку спуск робиться по стінці яру, а потім по дну яру до точки мінімуму.

3.1.2.2 Метод спряжених градієнтів

Розглянемо двокроковий метод, тобто метод, у якому для знаходження $(k+1)$ -го наближення використовуються $(k-1)$ -е і k -е наближення:

$$\mathbf{x}^{k+1} = \mathbf{x}^k - \alpha_k \frac{\partial \varphi(\mathbf{x}^k)}{\partial \mathbf{x}} + \beta_k (\mathbf{x}^k - \mathbf{x}^{k-1}), \quad (3.1)$$

$$\{\alpha_k, \beta_k\}: \min_{\alpha, \beta} \varphi \left(\mathbf{x}^k - \alpha \frac{\partial \varphi(\mathbf{x}^k)}{\partial \mathbf{x}} + \beta (\mathbf{x}^k - \mathbf{x}^{k-1}) \right). \quad (3.2)$$

Для випадку квадратичної функції $\varphi(\mathbf{x}) = \frac{1}{2} \mathbf{x}^\top \mathbf{A} \mathbf{x} - \mathbf{b}^\top \mathbf{x} + c$, $\mathbf{A} > 0$ задача (3.2) має такі розв'язки:

$$\alpha_k = \frac{\|\mathbf{r}^k\|^2 (\mathbf{A} \mathbf{p}^k, \mathbf{p}^k) - (\mathbf{r}^k, \mathbf{p}^k) (\mathbf{A} \mathbf{r}^k, \mathbf{p}^k)}{(\mathbf{A} \mathbf{r}^k, \mathbf{r}^k) (\mathbf{A} \mathbf{p}^k, \mathbf{p}^k) - (\mathbf{A} \mathbf{r}^k, \mathbf{p}^k)^2},$$

$$\beta_k = \frac{\|\mathbf{r}^k\|^2 (\mathbf{A} \mathbf{r}^k, \mathbf{p}^k) - (\mathbf{r}^k, \mathbf{p}^k) (\mathbf{A} \mathbf{r}^k, \mathbf{r}^k)}{(\mathbf{A} \mathbf{r}^k, \mathbf{r}^k) (\mathbf{A} \mathbf{p}^k, \mathbf{p}^k) - (\mathbf{A} \mathbf{r}^k, \mathbf{p}^k)^2},$$

де $\mathbf{r}^k = \frac{\partial \varphi(\mathbf{x}^k)}{\partial \mathbf{x}} = \mathbf{A} \mathbf{x}^k - \mathbf{b}$, $\mathbf{p}^k = \mathbf{x}^k - \mathbf{x}^{k-1}$.

У квадратичному випадку метод (3.1)–(3.2) є скінченним, тобто дає точний мінімум квадратичної функції за скінченну кількість ітерацій, що дорівнює n .

Для неквадратичних функцій метод (3.1)–(3.2) збігається глобально.

Методу спряжених градієнтів можна надати і іншу форму:

$$\mathbf{x}^{k+1} = \mathbf{x}^k + \alpha_k \mathbf{s}^k,$$

$$\alpha_k: \min_{\alpha} \varphi(\mathbf{x}^k + \alpha \mathbf{s}^k), \quad (3.3)$$

$$\mathbf{s}^k = -\frac{\partial \varphi(\mathbf{x}^k)}{\partial \mathbf{x}} + \beta_k \mathbf{s}^{k-1}, \quad \mathbf{s}^0 = -\frac{\partial \varphi(\mathbf{x}^0)}{\partial \mathbf{x}}, \quad \beta_0 = 0.$$

Метод називається методом спряжених градієнтів, тому що має місце така властивість: вектори $\frac{\partial \varphi(\mathbf{x}^i)}{\partial \mathbf{x}}$, $i = 0, n-1$ взаємно ортогональні, а $\mathbf{s}^0, \dots, \mathbf{s}^{n-1}$ — взаємно спряжені $((\mathbf{A} \mathbf{s}^i, \mathbf{s}^j) = 0)$.

Існує декілька способів вибору параметра β_k :

$$1) \beta_k = \frac{\left\| \frac{\partial \varphi(\mathbf{x}^k)}{\partial \mathbf{x}} \right\|^2}{\left\| \frac{\partial \varphi(\mathbf{x}^{k-1})}{\partial \mathbf{x}} \right\|^2};$$

$$2) \beta_k = \frac{\left(\frac{\partial \varphi(\mathbf{x}^k)}{\partial \mathbf{x}}, \frac{\partial \varphi(\mathbf{x}^k)}{\partial \mathbf{x}} - \frac{\partial \varphi(\mathbf{x}^{k-1})}{\partial \mathbf{x}} \right)}{\left\| \frac{\partial \varphi(\mathbf{x}^{k-1})}{\partial \mathbf{x}} \right\|^2};$$

$$3) \beta_k = \frac{\left(\frac{\partial^2 \varphi(\mathbf{x}^k)}{\partial \mathbf{x}^2} \mathbf{s}^{k-1}, \frac{\partial \varphi(\mathbf{x}^k)}{\partial \mathbf{x}} \right)}{\left(\frac{\partial^2 \varphi(\mathbf{x}^k)}{\partial \mathbf{x}^2} \mathbf{s}^{k-1}, \mathbf{s}^{k-1} \right)}.$$

Для випадку квадратичної функції методи (3.1), (3.2) і (3.3) за першим способом вибору параметру β_k при однаковому $\mathbf{x}^0$ визначають одну і ту саму послідовність точок $\mathbf{x}^k$.

Зазвичай для неквадратичних задач метод спряжених градієнтів застосовується в інший формі. До нього додається процедура оновлення — час від часу крок здійснюється вздовж градієнта. При цьому метод перестає бути скінченним, і напрямки спуску не будуть взаємно спряженими відносно якої-небудь матриці. Оновлення роблять через кількість ітерацій, що дорівнює виміру простору n :

$$\begin{aligned} \mathbf{x}^{k+1} &= \mathbf{x}^k + \alpha_k \mathbf{s}^k, \\ \alpha_k &: \min_{\alpha} \varphi(\mathbf{x}^k + \alpha \mathbf{s}^k), \\ \mathbf{s}^k &= -\frac{\partial \varphi(\mathbf{x}^k)}{\partial \mathbf{x}} + \bar{\beta}_k \mathbf{s}^{k-1}, \quad \mathbf{s}^0 = -\frac{\partial \varphi(\mathbf{x}^0)}{\partial \mathbf{x}}, \\ \bar{\beta}_k &= \begin{cases} 0, & k = 0, n, 2n, \dots, \\ \beta_k, & k \neq 0, n, 2n, \dots, \end{cases} \end{aligned}$$

а β_k обирається одним із вказаних вище способів.

Метод спряжених градієнтів з оновленням збігається глобально; в околі точки мінімуму швидкість збіжності є квадратичною.

Послідовності точок, отриманих за першим і другим способами вибору параметру β_k , співпадають.

Для задач великого виміру ($k < n$) можна гарантувати лише збіжність зі швидкістю геометричної прогресії. При $k < n$ для квадратичних функцій можна гарантувати збіжність зі швидкістю геометричної прогресії зі знаменником $g = \frac{\sqrt{L} - \sqrt{l}}{\sqrt{L} + \sqrt{l}}$, де L і l — найбільше і найменше власні числа матриці $\frac{\partial^2 \varphi(\mathbf{x}^*)}{\partial \mathbf{x}^2}$.

Метод спряжених градієнтів є оптимальним по швидкості збіжності у класі методів першого порядку.

3.1.3 Методи другого порядку

3.1.3.1 Метод Ньютона

У градієнтному методі для визначення напрямку руху використовується лише лінійний член з розвинення функції у ряд Тейлора, тобто використовується найбільш груба апроксимація функції. Тому для функцій з погано обумовленими матрицями Гессе метод повільно збігається. Якщо використовувати другі частинні похідні, то отримані методи не будуть «реагувати» на яристість функцій.

Нехай функція $\varphi(\mathbf{x})$ двічі диференційовна, скористаємося квадратичною апроксимацією функції у точці $\mathbf{x}^k$:

$$\varphi_k(\mathbf{x}) = \varphi(\mathbf{x}^k) + \left(\frac{\partial \varphi(\mathbf{x}^k)}{\partial \mathbf{x}}, \mathbf{x} - \mathbf{x}^k \right) + \frac{1}{2} \left(\frac{\partial^2 \varphi(\mathbf{x}^k)}{\partial \mathbf{x}^2} (\mathbf{x} - \mathbf{x}^k), \mathbf{x} - \mathbf{x}^k \right).$$

У градієнтному методі наступне наближення $\mathbf{x}^{k+1}$ знаходилося з умови мінімуму лінійної апроксимації при додаткових обмеженнях на близькість до точки $\mathbf{x}^k$ (так як лінійна функція не досягає мінімуму на $\mathbb{R}^n$). Для квадратичної функції $\varphi_k(\mathbf{x})$ можна не накладати таких обмежень, так як при $\frac{\partial^2 \varphi(\mathbf{x})}{\partial \mathbf{x}^2} > 0$ функція $\varphi_k(\mathbf{x})$ має безумовний мінімум. Обираємо точку мінімуму $\varphi_k(\mathbf{x})$ за нове наближення. Отже, отримуємо такий алгоритм

$$\mathbf{x}^{k+1} = \mathbf{x}^k - \left[\frac{\partial^2 \varphi(\mathbf{x}^k)}{\partial \mathbf{x}^2} \right]^{-1} \frac{\partial \varphi(\mathbf{x}^k)}{\partial \mathbf{x}}. \quad (3.4)$$

До цього методу можна прийти і іншим шляхом. Точка мінімуму повинна бути розв'язком системи n рівнянь з n невідомими, необхідна умова безумовного мінімуму першого порядку, $\frac{\partial \varphi(\mathbf{x})}{\partial \mathbf{x}} = 0$.

Одним з методів розв'язання таких задач є метод Ньютона або дотичних, а саме: лінеаризуємо систему у точці $\mathbf{x}^k$ і розв'язуємо лінеаризовану систему, яка має вигляд

$$\frac{\partial \varphi(\mathbf{x}^k)}{\partial \mathbf{x}} + \frac{\partial^2 \varphi(\mathbf{x}^k)}{\partial \mathbf{x}^2} (\mathbf{x} - \mathbf{x}^k) = 0,$$

звідки

$$\mathbf{x}^{k+1} = \mathbf{x}^k - \left[\frac{\partial^2 \varphi(\mathbf{x}^k)}{\partial \mathbf{x}^2} \right]^{-1} \frac{\partial \varphi(\mathbf{x}^k)}{\partial \mathbf{x}}.$$

Розглянемо теорему збіжності.

Теорема 3.4. Нехай $\varphi(\mathbf{x})$ двічі диференційовна функція, $\frac{\partial^2 \varphi(\mathbf{x})}{\partial \mathbf{x}^2}$ задоволь-

няє умові Ліпшиця з константою M , $\varphi(\mathbf{x})$ — сильно опукла з константою l і початкове наближення задовольняє умові

$$q = \frac{M}{2l^2} \left\| \frac{\partial \varphi(\mathbf{x}^0)}{\partial \mathbf{x}} \right\| < 1. \quad (3.5)$$

Тоді метод збігається до точки глобального мінімуму $\mathbf{x}^*$ з квадратичною швидкістю:

$$\|\mathbf{x}^k - \mathbf{x}^*\| \leq \frac{2l}{M} q^{2^k}.$$

Зауваження 3.3. Усі умови теореми є суттєвими. Сильна опуклість забезпечує існування $\left[\frac{\partial^2 \varphi(\mathbf{x})}{\partial \mathbf{x}^2} \right]^{-1}$, а невиконання (3.5) може призвести до розбіжності методу. Якщо ж відмовитися від умови Ліпшиця для $\frac{\partial^2 \varphi(\mathbf{x})}{\partial \mathbf{x}^2}$, то збіжність не буде квадратичною (може збігатися зі швидкістю геометричної прогресії).

Зауваження 3.4. У методі Ньютона ніколи не обертають матрицю $\frac{\partial^2 \varphi(\mathbf{x}^k)}{\partial \mathbf{x}^2}$. Зрозуміло, що напрямок спуску у методі є таким:

$$\mathbf{s}^k = - \left[\frac{\partial^2 \varphi(\mathbf{x}^k)}{\partial \mathbf{x}^2} \right]^{-1} \frac{\partial \varphi(\mathbf{x}^k)}{\partial \mathbf{x}}.$$

Тоді (3.4) можна переписати у вигляді:

$$\mathbf{x}^{k+1} = \mathbf{x}^k + \mathbf{s}^k.$$

Вектор $\mathbf{s}^k$ знаходимо з системи рівнянь

$$\frac{\partial^2 \varphi(\mathbf{x}^k)}{\partial \mathbf{x}^2} \mathbf{s}^k = - \frac{\partial \varphi(\mathbf{x}^k)}{\partial \mathbf{x}}.$$

А цю систему можна досить ефективно розв'язати і при $\left\| \frac{\partial^2 \varphi(\mathbf{x}^k)}{\partial \mathbf{x}^2} \right\| \sim \varepsilon$, де $\varepsilon > 0$ — мале число.

Умови теореми можна послабити тільки в одному напрямку — глобальні вимоги до функції $\varphi(\mathbf{x})$ замінити на локальні.

Для квадратичної функції $\varphi(\mathbf{x}) = \frac{1}{2} \mathbf{x}^\top \mathbf{A} \mathbf{x} - \mathbf{b}^\top \mathbf{x} + \mathbf{c}$, $\mathbf{A} > 0$ метод Ньютона збігається за один крок при довільній початковій точці $\mathbf{x}^0$.

Чим ближча функція до квадратичної, тим швидше збігається метод Ньютона. Формально — чим менше M , то згідно з теоремою буде більша область збіжності і тим швид-

ша збіжність, що визначається величиною q .

Отже, вдалий вибір початковою наближення $\mathbf{x}^0$ гарантує збіжність методу Ньютона. Однак пошук вдалого початкового наближення — це нетривіальна задача. Тому необхідно якось змінити метод Ньютона, щоб досягти збіжності процесу незалежно від початкового наближення. Для цього необхідно крім вибору напрямку руху $-\left[\frac{\partial^2 \varphi(\mathbf{x}^k)}{\partial \mathbf{x}^2}\right]^{-1} \frac{\partial \varphi(\mathbf{x}^k)}{\partial \mathbf{x}}$ вибирати і довжину кроку руху. Такий алгоритм називається демпфованим методом Ньютона — Рафсона або узагальненим методом Ньютона:

$$\mathbf{x}^{k+1} = \mathbf{x}^k - \beta_k \left[\frac{\partial^2 \varphi(\mathbf{x}^k)}{\partial \mathbf{x}^2} \right]^{-1} \frac{\partial \varphi(\mathbf{x}^k)}{\partial \mathbf{x}}.$$

Як і у градієнтному методі, у методі Ньютона — Рафсона крок обирається так, щоб забезпечити зменшення цільової функції на кожній ітерації. Наприклад, використовувати перші два способи вибору кроку або такий: крок β_k обирається так, щоб виконувалася нерівність

$$\begin{aligned} \varphi \left(\mathbf{x}^k - \beta_k \left[\frac{\partial^2 \varphi(\mathbf{x}^k)}{\partial \mathbf{x}^2} \right]^{-1} \frac{\partial \varphi(\mathbf{x}^k)}{\partial \mathbf{x}} \right) &\leq \\ &\leq \varphi(\mathbf{x}^k) - \varepsilon \beta_k \left(\frac{\partial \varphi(\mathbf{x}^k)}{\partial \mathbf{x}}, \left[\frac{\partial^2 \varphi(\mathbf{x}^k)}{\partial \mathbf{x}^2} \right]^{-1} \frac{\partial \varphi(\mathbf{x}^k)}{\partial \mathbf{x}} \right), \end{aligned}$$

де $0 < \varepsilon < 1/2$ — деяке число.

Покладемо $\beta_k = 1$, і якщо ця нерівність виконується, то $\beta_k = 1$, а якщо ні, то зменшуємо β_k до тих пір, поки не виконається нерівність.

Метод Ньютона може застосовуватися для мінімізації таких функцій, у яких існує обмежена обернена матриця других похідних. Такі властивості мають сильно опуклі двічі неперервно диференційовні функції, і виконується умова

$$m \|\mathbf{y}\|^2 \leq \left(\frac{\partial^2 \varphi(\mathbf{x})}{\partial \mathbf{x}^2} \mathbf{y}, \mathbf{y} \right) \leq M \|\mathbf{y}\|^2, \quad m > 0,$$

для довільних $\mathbf{x}$ і $\mathbf{y} \in \mathbb{R}^n$. Тоді функції обмеженні і у них існує одна точка мінімуму.

Теорема 3.5. Якщо функція $\varphi(\mathbf{x})$ така, що виконується умова (3.5) і матриця $\frac{\partial^2 \varphi(\mathbf{x})}{\partial \mathbf{x}^2}$ задовольняє умові Ліпшиця з константою L , крок обирається з умови (3.4), то послідовність точок, отриманих методом Ньютона — Рафсона незалежно від вибору початкових точки $\mathbf{x}^0$ збігається до точки мінімуму з

квадратичною швидкістю, тобто

$$\|\mathbf{x}^{k+1} - \mathbf{x}^*\| \leq \frac{L}{m} \|\mathbf{x}^k - \mathbf{x}^*\|.$$

3.1.3.2 Модифікації методу Ньютона — Рафсона

Усі труднощі, що виникають при практичній реалізації методу Ньютона, умовно можна розбити на дві групи. Перші пов'язані з необхідністю обчислювати матрицю $\frac{\partial^2 \varphi(\mathbf{x})}{\partial \mathbf{x}^2}$. Розглянемо дві модифікації методу Ньютона, які використовують не точні значення, а деякі наближені аналоги матриці других похідних. У результаті зменшується трудомісткість методу, але погіршується їх збіжність.

До другої групи відносяться всі ускладнення, що виникають у зв'язку з порушенням у процесі обчислень додатної визначеності матриці других похідних.

Перша модифікація Алгоритм методу має вигляд

$$\mathbf{x}^{k+1} = \mathbf{x}^k - \beta_k \left[\frac{\partial^2 \varphi(\mathbf{x}^0)}{\partial \mathbf{x}^2} \right]^{-1} \frac{\partial \varphi(\mathbf{x}^k)}{\partial \mathbf{x}}, \quad \beta_k \geq 0.$$

У цій схемі використовується один раз розрахована матриця $\frac{\partial^2 \varphi(\mathbf{x}^0)}{\partial \mathbf{x}^2}$.

Друга модифікація Вона пов'язана з оновленням матриці других похідних через певну кількість кроків. Ітераційний процес має вигляд

$$\mathbf{x}^{kt+i+1} = \mathbf{x}^{kt+i} - \beta_{kt+i} \left[\frac{\partial^2 \varphi(\mathbf{x}^{kt})}{\partial \mathbf{x}^2} \right]^{-1} \frac{\partial \varphi(\mathbf{x}^{kt+i})}{\partial \mathbf{x}},$$

$$\beta_{kt+i} \geq 0; k = 0, 1, \dots; i = 0, 1, \dots, t-1, t; t > 0.$$

Якщо виконуються умови [теорема 3.5](#), то справедлива оцінка

$$\|\mathbf{x}^{(k+1)t} - \mathbf{x}^*\| \leq c \|\mathbf{x}^{kt} - \mathbf{x}^*\|^{t+1}, \quad c \geq 0.$$

Ця оцінка означає, що послідовність $\{\mathbf{x}^k\}$ збігається до розв'язку зі швидкістю порядку $t+1$.

Третя модифікація Нехай матриця Гессе функції $\varphi(\mathbf{x})$ не є додатно визначеною. У цьому випадку послідовність точок методу Ньютона буде розбіжною. Левенберг і Марквардт

запропонували додавати до матриці других похідних на кожному кроці величину $\alpha_k \mathbf{I}$, де α_k — деяке число, а $\mathbf{I}$ — одинична матриця. Тоді ітераційний процес матиме вигляд

$$\mathbf{x}^{k+1} = \mathbf{x}^k - \left[\frac{\partial^2 \varphi(\mathbf{x}^k)}{\partial \mathbf{x}^2} + \alpha_k \mathbf{I} \right]^{-1} \frac{\partial \varphi(\mathbf{x}^k)}{\partial \mathbf{x}}.$$

У цьому алгоритмі довжина кроку дорівнює 1, а α_k обирають так, щоб виконувались умови

$$\cos \left(\left(\frac{\partial \varphi(\mathbf{x}^k)}{\partial \mathbf{x}^2} + \alpha_k \mathbf{I} \right)^{-1} \frac{\partial \varphi(\mathbf{x}^k)}{\partial \mathbf{x}}, \frac{\partial \varphi(\mathbf{x}^k)}{\partial \mathbf{x}} \right) \geq \varepsilon_1 > 0;$$

$$\varphi(\mathbf{x}^{k+1}) \leq \varphi(\mathbf{x}^k) - \varepsilon_2 \left(\left[\frac{\partial^2 \varphi(\mathbf{x}^k)}{\partial \mathbf{x}^2} + \alpha_k \mathbf{I} \right]^{-1} \frac{\partial \varphi(\mathbf{x}^k)}{\partial \mathbf{x}}, \frac{\partial \varphi(\mathbf{x}^k)}{\partial \mathbf{x}} \right), \quad 0 < \varepsilon_2 < \frac{1}{2}.$$

Перша нерівність означає, що кут між напрямком спуску антиградієнта у точці $\mathbf{x}^k$ повинен бути гострим, а виконання другої умови гарантує зменшення функції на кожній ітерації.

При досить великому значенні α_k метод вдалині від точки мінімуму веде себе як градієнтний і початкову точку можна обирати довільно, а при наближенні α_k до нуля метод переходить у звичайний метод Ньютона.

3.1.3.3 Обговорення властивостей методу Ньютона

Встановлено, що метод Ньютона з регулюванням кроку збігається до розв'язку незалежно від початкової точки $\mathbf{x}^0$ і має або зверхлінійну швидкість збіжності, або ж квадратичну в залежності від вимог до функції $\varphi(\mathbf{x})$.

Збіжність з довільної початкової точки є суттєвою перевагою методу Ньютона — Рафсона у порівнянні зі звичайним методом Ньютона, в якому збіжність гарантується лише при наявності досить хорошого початкового наближення. Крім того, перевірити умови, які гарантують, що дане початкове наближення забезпечить збіжність процесу, у методі Ньютона практично важко, тому що вимагається знання таких відомостей про функцію, які звичайно невідомі.

Якщо порівнювати метод Ньютона і градієнтні методи стосовно до розв'язування задач мінімізації опуклих функцій, то метод Ньютона забезпечує більш високу швидкість збігу послідовності точок до розв'язку. Однак більш точна суть поняття ефективності методу є в оцінці загальної кількості обчислень для розв'язування задачі з заданою точністю. Отже, про ефективність того чи іншого алгоритму необхідно судити по кількості

ітерацій, необхідних для розв'язування задачі, і по кількості обчислень на кожній ітерації.

Кількість обчислень на ітерації методу Ньютона, як правило, значно більша, ніж у градієнтних методах, за рахунок обчислення і обертання матриці других похідних. Але на отримання розв'язку з досить високим ступенем точності за допомогою методу Ньютона необхідно у десятки, а то і у сотні раз менше ітерацій, ніж при використанні градієнтних методів. А у деяких задачах трудомісткість ітерації методу Ньютона може бути непомірно великою за рахунок обчислення матриці других похідних (а не її обертання). У цих випадках для розв'язування задач можна використовувати модифікації методу Ньютона. В одній з таких модифікацій обчислювати і обертати матрицю Гессе потрібно лише один раз, а у другій це робиться через скінченну кількість ітерацій. При цьому, якщо початкове наближення вдало обране, то швидкість збігу буде досить високою.

Однак використання модифікації методу Ньютона не є кардинальним розв'язком питання про скорочення трудомісткості розв'язування задачі. Тому виникає питання про побудову методів мінімізації, які б по швидкості збігу були близькі до методу Ньютона, але для своєї реалізації вимагали значно меншу кількість обчислень на кожній ітерації.

Приклад 3.3. Знайти мінімум функції $\varphi(\mathbf{x}) = x_1^2 + 25x_2^2$ методом Ньютона з точністю $\varepsilon = 0.1$.

Розв'язування. За початкову точку візьмемо точку $\mathbf{x}^0 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

Тоді $\mathbf{x}^1 = \mathbf{x}^0 - \left[\frac{\partial^2 \varphi(\mathbf{x}^0)}{\partial \mathbf{x}^2} \right]^{-1} \frac{\partial \varphi(\mathbf{x}^0)}{\partial \mathbf{x}}$ або $\mathbf{x}^1 = \mathbf{x}^0 + \mathbf{s}^0$,

де $\mathbf{s}^0 = - \left[\frac{\partial^2 \varphi(\mathbf{x}^0)}{\partial \mathbf{x}^2} \right]^{-1} \frac{\partial \varphi(\mathbf{x}^0)}{\partial \mathbf{x}}$. Отримаємо систему для визначення $\mathbf{s}^0$:

$$\frac{\partial^2 \varphi(\mathbf{x}^0)}{\partial \mathbf{x}^2} \mathbf{s}^0 = - \frac{\partial \varphi(\mathbf{x}^0)}{\partial \mathbf{x}}.$$

Маємо

$$\frac{\partial \varphi(\mathbf{x})}{\partial \mathbf{x}} = \begin{pmatrix} 2x_1 \\ 50x_2 \end{pmatrix}, \quad \frac{\partial^2 \varphi(\mathbf{x})}{\partial \mathbf{x}^2} = \begin{pmatrix} 2 & 0 \\ 0 & 50 \end{pmatrix}.$$

Система має вигляд

$$\begin{pmatrix} 2 & 0 \\ 0 & 50 \end{pmatrix} \begin{pmatrix} s_1^0 \\ s_2^0 \end{pmatrix} = - \begin{pmatrix} 2 \\ 50 \end{pmatrix}, \quad \text{або} \quad \begin{cases} 2s_1^0 = -2, \\ 50s_2^0 = -50. \end{cases}$$

Розв'язавши її, отримаємо $s_1^0 = -1$, $s_2^0 = -1$. Тоді $\mathbf{x}^1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \begin{pmatrix} -1 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$.

Критерій зупинки не виконується, наприклад, пункт A_1 , і тому знаходимо друге наближення

$$\mathbf{x}^2 = \mathbf{x}^1 + \mathbf{s}^1, \quad \frac{\partial^2 \varphi(\mathbf{x}^1)}{\partial \mathbf{x}^2} \mathbf{s}^1 = -\frac{\partial \varphi(\mathbf{x}^1)}{\partial \mathbf{x}},$$

$$\begin{pmatrix} 2 & 0 \\ 0 & 50 \end{pmatrix} \begin{pmatrix} s_1^1 \\ s_2^1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad \begin{cases} 2s_1^1 = 0, \\ 50s_2^1 = 0, \end{cases} \quad x_1^1 = 0, \quad x_2^1 = 0.$$

Усі пункти критерію зупинки виконуються. Так як функція $\varphi(\mathbf{x}) = x_1^2 + 25x_2^2$ є опуклою (див. [приклад 3.2](#)), то точка $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ є розв'язком нашої задачі. ☑

Метод Ньютона є методом другого порядку.

3.1.4 Квазіньютонівські методи (методи змінної метрики)

В основі цих методів лежить ідея відновлення квадратичної апроксимації функції по значенням її градієнтів у ряді точок. Таким чином методи об'єднують переваги градієнтного методу і методу Ньютона. Ці методи мають таку загальну структуру:

$$\mathbf{x}^{k+1} = \mathbf{x}^k - \beta_k \mathbf{H}_k \frac{\partial \varphi(\mathbf{x}^k)}{\partial \mathbf{x}}, \quad (3.6)$$

де матриця $\mathbf{H}_k$ перераховується рекурентним способом на основі інформації, отриманої на k -й ітерації так, що

$$\mathbf{H}_k - \left[\frac{\partial^2 \varphi(\mathbf{x}^k)}{\partial \mathbf{x}^2} \right]^{-1} \rightarrow 0 \quad \text{при} \quad k \rightarrow +\infty.$$

Таким чином, методи переходять у ньютонівські, і тому мають таку назву.

Т е о р е м а 3.6. Нехай $\mathbf{x}^*$ — невироджена точка мінімуму¹ функції $\varphi(\mathbf{x})$, $\varphi(\mathbf{x})$ — двічі неперервно диференційовна в околі точки $\mathbf{x}^*$ і $\left\| \mathbf{H}_k - \left[\frac{\partial^2 \varphi(\mathbf{x}^k)}{\partial \mathbf{x}^2} \right]^{-1} \right\| \rightarrow 0$ при $k \rightarrow \infty$. Тоді метод (3.6) з $\beta_k = 1$ локально збігається до $\mathbf{x}^*$ швидше довільної геометричної прогресії.

Метод при додатно визначених $\mathbf{H}_k$ має глобальну збіжність.

Перейдемо до побудови матриць $\mathbf{H}_k$. Їх будемо будувати так, щоб вони апроксиму-

¹ Точка $\mathbf{x}^*$ є невиродженою, якщо $\left\| \frac{\partial^2 \varphi(\mathbf{x}^*)}{\partial \mathbf{x}^2} \right\| \neq 0$.

вали $\left[\frac{\partial^2 \varphi(\mathbf{x}^k)}{\partial \mathbf{x}^2} \right]^{-1}$. Має місце співвідношення

$$\frac{\partial \varphi(\mathbf{x}^k)}{\partial \mathbf{x}} - \frac{\partial \varphi(\mathbf{x}^{k+1})}{\partial \mathbf{x}} = \frac{\partial^2 \varphi(\mathbf{x}^{k+1})}{\partial \mathbf{x}^2} (\mathbf{x}^k - \mathbf{x}^{k+1}) + o(\|\mathbf{x}^k - \mathbf{x}^{k+1}\|).$$

Якщо матриця $\frac{\partial^2 \varphi(\mathbf{x}^{k+1})}{\partial \mathbf{x}^2}$ не вироджена, то з точністю до членів більш високого порядку малості у порівнянні з $\|\mathbf{x}^k - \mathbf{x}^{k+1}\|$ маємо

$$\left[\frac{\partial^2 \varphi(\mathbf{x}^{k+1})}{\partial \mathbf{x}^2} \right]^{-1} \left(\frac{\partial \varphi(\mathbf{x}^{k+1})}{\partial \mathbf{x}} - \frac{\partial \varphi(\mathbf{x}^k)}{\partial \mathbf{x}} \right) \approx \mathbf{x}^{k+1} - \mathbf{x}^k.$$

Якщо функція $\varphi(\mathbf{x})$ є квадратичною, тобто $\varphi(\mathbf{x}) = \frac{1}{2}(\mathbf{A}\mathbf{x}, \mathbf{x}) - (\mathbf{b}, \mathbf{x})$, то наближена рівність перетворюється у точну

$$\left[\frac{\partial^2 \varphi(\mathbf{x}^{k+1})}{\partial \mathbf{x}^2} \right]^{-1} \Delta \mathbf{y}^k = \Delta \mathbf{x}^k,$$

де $\Delta \mathbf{y}^k = \frac{\partial \varphi(\mathbf{x}^{k+1})}{\partial \mathbf{x}} - \frac{\partial \varphi(\mathbf{x}^k)}{\partial \mathbf{x}}$, $\Delta \mathbf{x}^k = \mathbf{x}^{k+1} - \mathbf{x}^k$. Тому природно вимагати, щоб для матриці $\mathbf{H}_k$, що наближає $\left[\frac{\partial^2 \varphi(\mathbf{x}^{k+1})}{\partial \mathbf{x}^2} \right]^{-1}$, виконувалась умова

$$\mathbf{H}_{k+1} \Delta \mathbf{y}^k = \Delta \mathbf{x}^k. \quad (3.7)$$

Ця умова має назву квазіньютонівської.

Нехай наближення до $\left[\frac{\partial^2 \varphi(\mathbf{x}^{k+1})}{\partial \mathbf{x}^2} \right]^{-1}$ перераховується за формулою

$$\mathbf{H}_{k+1} = \mathbf{H}_k + \Delta \mathbf{H}_k.$$

Укажемо будь-яку матрицю $\Delta \mathbf{H}_k$, що забезпечує (3.7). Для цього перепишемо (3.7) таким чином:

$$\begin{aligned} (\mathbf{H}_k + \Delta \mathbf{H}_k) \Delta \mathbf{y}^k &= \Delta \mathbf{x}^k, \\ \Delta \mathbf{H}_k \Delta \mathbf{y}^k &= \Delta \mathbf{x}^k - \mathbf{H}_k \Delta \mathbf{y}^k. \end{aligned}$$

Цій рівності задовольняє, наприклад, така матриця:

$$\Delta \mathbf{H}_k = \frac{1}{(\mathbf{z}^k, \Delta \mathbf{y}^k)} (\Delta \mathbf{x}^k - \mathbf{H}_k \Delta \mathbf{y}^k) \mathbf{z}^k,$$

де $\mathbf{z}^k$ — довільний вектор такий, що $(\mathbf{z}^k, \Delta \mathbf{y}^k) \neq 0$.

Обираємо $\mathbf{z}^k = \Delta \mathbf{x}^k - \mathbf{H}_k \Delta \mathbf{y}^k$, і тоді отримаємо:

$$\mathbf{H}_{k+1} = \mathbf{H}_k + \frac{(\Delta \mathbf{x}^k - \mathbf{H}_k \Delta \mathbf{y}^k)^\top (\Delta \mathbf{x}^k - \mathbf{H}_k \Delta \mathbf{y}^k)}{(\Delta \mathbf{x}^k - \mathbf{H}_k \Delta \mathbf{y}^k, \Delta \mathbf{y}^k)}. \quad (3.8)$$

При такому виборі перерахунку $\mathbf{H}_k$ отримуємо метод Бroyдена.

Якщо β_k обрати за способом 1 (тобто мінімізуємо по одновимірному напрямку), то цей метод дає спряжені напрямки. Якщо функція $\varphi(\mathbf{x})$ є квадратичною, то за n кроків отримаємо, що $\mathbf{H}_n = \mathbf{A}^{-1}$.

Цікава особливість цього методу: β_k не обов'язково повинно давати мінімум по одновимірному напрямку, воно має бути довільним параметром, поки не виникла сингулярність матриці $\mathbf{H}_k$ ² або знаменник (3.8) не перетворився в нуль.

У випадку, коли функція $\varphi(\mathbf{x})$ не є квадратичною, при застосуванні методу Бroyдена на можуть виникнути такі небажані явища:

- 1) Матриця $\mathbf{H}_k$ може перестати бути додатно визначеною. Для цього існують методи, що перетворюють матрицю $\mathbf{H}_k$ у додатно визначену (див. Д. Хіммельблау. Прикладное нелинейное программирование, роз. 3.2).
- 2) Величина $\Delta \mathbf{H}_k$ може стати необмеженою (інколи і для квадратичних функцій внаслідок похибок округлення).
- 3) Якщо на k -му кроці напрямок спуску співпадає з напрямком спуску на $(k-1)$ -му кроці, то матриця $\mathbf{H}_{k+1}$ може бути сингулярною.

Якщо $\mathbf{H}_{k+1}$ перерахувати за формулою

$$\mathbf{H}_{k+1} = \mathbf{H}_k + \frac{(\Delta \mathbf{x}^k)^\top \Delta \mathbf{x}^k}{(\Delta \mathbf{x}^k, \Delta \mathbf{y}^k)} - \frac{\mathbf{H}_k \left((\Delta \mathbf{y}^k)^\top \Delta \mathbf{y}^k \right) \mathbf{H}_k}{(\mathbf{H}_k \Delta \mathbf{y}^k, \Delta \mathbf{y}^k)}, \quad (3.8')$$

то отримаємо метод Девідона — Флетчера — Пауелла (ДФП).

Якщо $\mathbf{H}_0 = \mathbf{I}$, то при оптимізації маємо поступовий перехід від градієнтного напрямку до ньютонівського.

При використанні (3.8') похибка обчислень $\frac{\partial \varphi(\mathbf{x}^k)}{\partial \mathbf{x}}$ повинна бути малою, і тоді $\mathbf{H}_k$

²Тобто $\mathbf{H}_k = \mathbf{0}$.

не стане «поганою». Другий доданок у формулі (3.8') забезпечує виконання умови: $\mathbf{H}_k \rightarrow \left[\frac{\partial^2 \varphi(\mathbf{x}^*)}{\partial \mathbf{x}^2} \right]^{-1}$, а третій доданок забезпечує додатновизначеність $\mathbf{H}^{k+1}$ на всіх етапах.

У випадку квадратичної функції в алгоритмі (3.8') використовуються спряжені напрямки. Якщо функція загального виду, то ефективність методу ДФП є швидше наслідком використання спряжених напрямків, ніж близької апроксимації $\left[\frac{\partial^2 \varphi(\mathbf{x})}{\partial \mathbf{x}^2} \right]^{-1}$ матрицею $\mathbf{H}$.

Практика показала, що у деяких задачах неможливо досягти мінімуму цільової функції за допомогою квазіньютонівських методів, якщо точність одновимірного пошуку недостатня. Тому рекомендується, щоб точність одновимірного пошуку була еквівалентна точності закінчення головного алгоритму.

Метод ДФП можна використовувати і у випадку, коли компоненти градієнта оцінюються за допомогою різницевого співвідношення.

Існують ще квазіньютонівські методи під назвою: Пірсона 1, Пірсона 2, Пірсона 3, Девідона, Флетчера, Мурґата і Сарджента і інші. Досить ефективним є метод Бройдена — Флетчера — Гольдфарба — Шенно (БФГШ).

Вище квазіньютонівські методи були отримані як наближення до методу Ньютона. Проте на них можна поглянути і з іншого боку.

Якщо поряд з евклідовою нормою розглянути таку: $\|\mathbf{x}\|_1 = \sqrt{(\mathbf{A}\mathbf{x}, \mathbf{x})}$, де матриця $\mathbf{A} > 0$, то у цій метриці градієнт матиме вигляд

$$(\text{grad } \varphi(\mathbf{x}))_1 = \mathbf{A}^{-1} \frac{\partial \varphi(\mathbf{x})}{\partial \mathbf{x}}.$$

Тоді у новій метриці градієнтний метод набуває вигляду

$$\mathbf{x}^{k+1} = \mathbf{x}^k - \beta_k \mathbf{A}^{-1} \frac{\partial \varphi(\mathbf{x}^k)}{\partial \mathbf{x}}$$

і відрізняється від алгоритму градієнтного методу наявністю матриці $\mathbf{A}^{-1}$.

Природно намагатися вибрати метрику так, щоб прискорити збіжність методу. Для квадратичної функції покладають $\mathbf{A} = \frac{\partial^2 \varphi(\mathbf{x})}{\partial \mathbf{x}^2}$, тобто отримують (при $\beta_k = 1$) метод Ньютона, і градієнтний метод збігається за один крок.

Для неквадратичної функції метод

$$\mathbf{x}^{k+1} = \mathbf{x}^k - \beta_k \mathbf{H}_k \frac{\partial \varphi(\mathbf{x}^k)}{\partial \mathbf{x}}, \quad \mathbf{H}_k > 0$$

може розглядатися як градієнтний у метриці $(\mathbf{x}, \mathbf{y})_1 = (\mathbf{H}_k^{-1} \mathbf{x}, \mathbf{y})$, і оптимальним вибором

метрики є $\mathbf{H}_k = \left[\frac{\partial \varphi(\mathbf{x}^k)}{\partial \mathbf{x}^2} \right]^{-1}$. Іншими словами, квазіньютонівські методи можуть трактуватися як градієнтні, у яких на кожному кроці обирається нова метрика, за можливістю близька до найкращої. У зв'язку з цим часто вживають термін «методи змінної метрики» як синонім квазіньютонівських методів.

3.2 Мінімізація функцій з обмеженнями

3.2.1 Метод проекції градієнта

Розглядається задача

$$\min \varphi(\mathbf{x}), \quad \mathbf{x} \in \mathbf{R} \subset \mathbb{E}^n.$$

Застосування градієнтного методу може призвести до того, що точка $\mathbf{x}^{k+1}$ при деякому k може не належати множині $\mathbf{R}$.

Однак це явище можна обійти, якщо отриману точку $\mathbf{x}^{k+1}$ спроектувати на множину $\mathbf{R}$. Отже, якщо маємо наближення $\mathbf{x}^k$, то

$$\mathbf{x}^{k+1} = \mathcal{P}_{\mathbf{R}} \left(\mathbf{x}^k - \beta_k \frac{\partial \varphi(\mathbf{x}^k)}{\partial \mathbf{x}} \right), \quad k = 0, 1, \dots \quad (3.9)$$

де $\mathcal{P}_{\mathbf{R}}$ — оператор проектування.

Параметр β_k розраховується як і у градієнтному методі. Якщо в (3.9) на деякій ітерації $\mathbf{x}^{k+1} = \mathbf{x}^k$, то процес припиняють.

На практиці для пришвидшення збіжності досить часто користуються більш загальним варіантом методу проекції градієнта:

$$\mathbf{x}^{k+1} = \mathbf{x}^k + \alpha_k \left(\mathcal{P}_{\mathbf{R}} \left(\mathbf{x}^k - \beta_k \frac{\partial \varphi(\mathbf{x}^k)}{\partial \mathbf{x}} \right) - \mathbf{x}^k \right), \quad 0 < \alpha_k \leq 1. \quad (3.10)$$

Зазначимо, що у методах (3.9) і (3.10) на кожній ітерації необхідно проектувати точку на множину $\mathbf{R}$, тобто розв'язувати задачу мінімізації квадратичної функції на множині $\mathbf{R}$:

$$\min_{\mathbf{x} \in \mathbf{R}} \left\| \mathbf{x} - \left(\mathbf{x}^k - \beta_k \frac{\partial \varphi(\mathbf{x}^k)}{\partial \mathbf{x}} \right) \right\|^2. \quad (3.11)$$

Звичайно, задачу (3.11) не завжди легко розв'язати. Тому методом проекції градієнта ко-

ристуються тільки у тих випадках, якщо проекції точки на множину легко визначаються. Наприклад, якщо множина R — сфера в E^n , паралелепіпед, гіперплощина, напівпростір або додатний октант, то задача проектування точки розв'язується просто і в явному вигляді.

Нехай $\mathbf{a}$ — довільна точка E^n :

1) Якщо $R = \{\mathbf{x} \in E^n \mid \|\mathbf{x} - \mathbf{x}^0\| \leq r\}$ — сфера, то

$$\mathcal{P}_R(\mathbf{a}) = \mathbf{x}^0 + \frac{\mathbf{a} - \mathbf{x}^0}{\|\mathbf{a} - \mathbf{x}^0\|} r;$$

2) Якщо $R = \{\mathbf{x} \in E^n \mid b_i \leq x_i \leq c_i, i = \overline{1, n}\}$ — координатний паралелепіпед, то

$$(\mathcal{P}_R(\mathbf{a}))_j = \begin{cases} b_j, & \text{якщо } a_j < b_j, \\ a_j, & \text{якщо } b_j \leq a_j \leq c_j, \\ c_j, & \text{якщо } a_j > c_j; \end{cases}$$

3) Якщо $R = \{\mathbf{x} \in E^n \mid x_i \geq 0, i = \overline{1, n}\}$ — невід'ємний ортант, то

$$\mathcal{P}_R(\mathbf{a}) = (\max(0, a_1), \dots, \max(0, a_n));$$

4) Якщо $R = \{\mathbf{x} \in E^n \mid (\mathbf{b}, \mathbf{x}) = \beta\}$ — гіперплощина, $\mathbf{b} \neq \mathbf{0}$, то

$$\mathcal{P}_R(\mathbf{a}) = \mathbf{a} + (\beta - (\mathbf{b}, \mathbf{a})) \frac{\mathbf{b}}{\|\mathbf{b}\|^2};$$

5) Якщо $R = \{\mathbf{x} \in E^n \mid (\mathbf{b}, \mathbf{x}) \geq \beta\}$ — напівпростір, $\mathbf{b} \neq \mathbf{0}$, то

$$\mathcal{P}_R(\mathbf{a}) = \mathbf{a} + \max(0, \beta - (\mathbf{b}, \mathbf{a})) \frac{\mathbf{b}}{\|\mathbf{b}\|^2};$$

6) Якщо $R = \{\mathbf{x} \in E^n \mid \mathbf{A}\mathbf{x} = \mathbf{b}\}$, до того ж рядки матриці лінійно незалежні, то

$$\mathcal{P}_R(\mathbf{a}) = \mathbf{a} - \mathbf{A}^\top (\mathbf{A}\mathbf{A}^\top)^{-1} (\mathbf{A}\mathbf{a} - \mathbf{b}).$$

Розглянемо збіжність алгоритму (3.9).

Теорема 3.7. *Нехай множина R опукла і замкнута, функція $\varphi(\mathbf{x})$ сильно опукла з константою $\gamma > 0$ і диференційовна на R , градієнт її задовольняє умові Ліпшиця*

$$\left\| \frac{\partial \varphi(\mathbf{x}^1)}{\partial \mathbf{x}} - \frac{\partial \varphi(\mathbf{x}^2)}{\partial \mathbf{x}} \right\| \leq M \|\mathbf{x}^1 - \mathbf{x}^2\|, \quad \mathbf{x}^1, \mathbf{x}^2 \in R.$$

Тоді послідовність $\{\mathbf{x}^k\}$, що отримана за формулою (3.9), де $\mathbf{x}^0$ — довільна точка з $\mathbb{R}$, а $\beta_k \equiv \beta \in \left(0, \frac{4\gamma}{M^2}\right)$, збігається до розв'язку зі швидкістю геометричної прогресії:

$$\|\mathbf{x}^{k+1} - \mathbf{x}^*\| \leq q \|\mathbf{x}^k - \mathbf{x}^*\|,$$

де $q = \sqrt{1 - 4\gamma\beta + \beta^2 M^2} \in (0, 1)$.

3.2.2 Метод умовного градієнта

Метод умовного градієнта може застосовуватись для розв'язання задач мінімізації нелінійних функцій в області, в якій задача мінімізації лінійної функції легко розв'язується.

Розглядається задача мінімізації диференційовної в обмеженій, замкнутій і опуклій множині $\mathbb{R}$ функції $\varphi(\mathbf{x})$:

$$\min \varphi(\mathbf{x}), \quad \mathbf{x} \in \mathbb{R} \subset \mathbb{E}^n.$$

Опишемо алгоритм методу. Нехай $\mathbf{x}^0 \in \mathbb{R}$ — деяке початкове наближення. Якщо відоме k -е наближення $\mathbf{x}^k$, то розглянемо лінійну частину розвинення Тейлора функції $\varphi(\mathbf{x})$ в околі точки $\mathbf{x}^k$ без константи:

$$\varphi_k(\mathbf{x}) = \left(\frac{\partial \varphi(\mathbf{x}^k)}{\partial \mathbf{x}}, \mathbf{x} - \mathbf{x}^k \right)$$

і знайдемо точку $\bar{\mathbf{x}}^k$ таку, що

$$\min_{\mathbf{x} \in \mathbb{R}} \varphi_k(\mathbf{x}) = \left(\frac{\partial \varphi(\mathbf{x}^k)}{\partial \mathbf{x}}, \bar{\mathbf{x}}^k - \mathbf{x}^k \right).$$

Так як множина $\mathbb{R}$ замкнута і обмежена, а лінійна функція неперервна, то точка $\bar{\mathbf{x}}^k$ завжди існує.

Введемо позначення $\eta_k = \left(\frac{\partial \varphi(\mathbf{x}^k)}{\partial \mathbf{x}}, \bar{\mathbf{x}}^k - \mathbf{x}^k \right)$. Далі отримаємо

$$\eta_k \leq \left(\frac{\partial \varphi(\mathbf{x}^k)}{\partial \mathbf{x}}, \bar{\mathbf{x}}^k - \mathbf{x}^k \right) = 0.$$

Тому можливі два випадки: $\eta_k = 0$ або $\eta_k < 0$.

Якщо $\eta_k = 0$, то $\left(\frac{\partial \varphi(\mathbf{x}^k)}{\partial \mathbf{x}}, \mathbf{x} - \mathbf{x}^k \right) \geq \eta_k = 0$ для всіх $\mathbf{x} \in \mathbb{R}$, тобто точка $\mathbf{x}^k$

задовольняє необхідній і достатній умовам оптимальності для опуклої диференційованої функції на опуклій множині. Якщо $\mathbf{R}$ не є опуклою, то виконується тільки необхідна умова оптимальності. Далі необхідне додаткове дослідження.

Якщо $\eta_k < 0$, то розглядається точка $\mathbf{x}^{k+1} = \mathbf{x}^k + \beta_k (\bar{\mathbf{x}}^k - \mathbf{x}^k) \in \mathbf{R}$. Вектор $\bar{\mathbf{x}}^k - \mathbf{x}^k$ називають умовним градієнтом.

Параметр β_k обирається з $(0, 1]$ так, щоб виконувалась умова $\varphi(\mathbf{x}^{k+1}) < \varphi(\mathbf{x}^k)$. Вибір конкретного β_k можна робити, наприклад, з умови одновимірної мінімізації:

$$\beta_k : \min_{0 \leq \beta \leq 1} \varphi(\mathbf{x}^k + \beta (\bar{\mathbf{x}}^k - \mathbf{x}^k)),$$

$$\text{або } 0 < \beta_k < 1, \sum_{k=0}^{\infty} \beta_k = \infty, \sum_{k=0}^{\infty} \beta_k^2 < +\infty.$$

Існують і інші способи вибору β_k (див. [2], [11], [12]).

Має місце така

Теорема 3.8. Нехай $\mathbf{R}$ — опукла, обмежена і замкнута множина, $\varphi(\mathbf{x})$ — опукла і диференційовна функція на $\mathbf{R}$, а градієнт її задовольняє умові Ліпшиця. Тоді довільна гранична точка послідовності $\{\mathbf{x}^k\}$ є розв'язком задачі і має місце оцінка

$$\varphi(\mathbf{x}^k) - \varphi(\mathbf{x}^*) \leq \frac{c}{k}, \quad c > 0 — \text{константа}.$$

Зрозуміло, що описаний метод доцільно використовувати лише у випадку, якщо задача мінімізації лінійної функції на $\mathbf{R}$ розв'язується досить просто. Наприклад, множина $\mathbf{R}$ утворена лінійними нерівностями та рівностями. У цьому випадку можна застосувати симплекс-метод. Якщо ж множина має складну структуру, то задача мінімізації лінійної функції на $\mathbf{R}$ має таку ж складність, як і початкова.

Все ж розглянутий метод цікавий тим, що у ньому використовується лінійна апроксимація функції, що відіграє важливу роль у чисельних методах оптимізації.

Ідея лінійної апроксимації приводить до ідеї квадратичної апроксимації. У цьому випадку точка $\bar{\mathbf{x}}^k$ знаходиться з умови

$$\bar{\mathbf{x}}^k : \min \left[\left(\frac{\partial \varphi(\mathbf{x}^k)}{\partial \mathbf{x}}, \mathbf{x} - \mathbf{x}^k \right) + \frac{1}{2} \left(\frac{\partial^2 \varphi(\mathbf{x}^k)}{\partial \mathbf{x}^2} (\mathbf{x} - \mathbf{x}^k), \mathbf{x} - \mathbf{x}^k \right) \right].$$

Коефіцієнт β_k обирається, як і раніше.

3.2.3 Методи штрафних функцій

Усі методи цієї групи, не дивлячись на різні варіанти і схеми, мають одну загальну особливість: у цих методах робиться перехід від задачі умовної оптимізації до еквівалентної задачі (або послідовності задач) безумовної оптимізації.

Розглянемо деякі з методів штрафних функцій: методи внутрішніх штрафів, методи зовнішніх штрафів, комбіновані методи.

При використанні методів внутрішніх штрафів поточна точка постійно знаходиться всередині припустимої області за допомогою штрафної функції, яка у цьому випадку називається бар'єрною.

Методи зовнішніх штрафів, навпаки, генерують послідовність точок, що виходять за межі припустимої області, і границя цієї послідовності є припустимий розв'язок.

У комбінованих методах у процесі мінімізації одні з обмежень задовольняються, а інші — ні. Однак при досягненні розв'язку всі умови в межах заданого допуску задовольняються.

3.2.3.1 Метод внутрішніх штрафів

Розглядається задача

$$\begin{aligned} \min \varphi(\mathbf{x}) \\ \mathbf{x} \in R = \{\mathbf{x} \mid g_i(\mathbf{x}) \geq 0, i = \overline{1, p}\}. \end{aligned} \quad (3.12)$$

Від задачі (3.12) перейдемо до послідовності задач безумовної мінімізації

$$\min P(\mathbf{x}, r), \quad r \rightarrow 0,$$

де штрафна функція $P(\mathbf{x}, r)$ має вигляд

$$P(\mathbf{x}, r) = \varphi(\mathbf{x}) + \Phi(\mathbf{x}, r).$$

Функція $\Phi(\mathbf{x}, r)$ називається штрафом і має такі властивості:

$$\Phi(\mathbf{x}, r) = \begin{cases} \infty, & \text{якщо } \mathbf{x} \notin R, \\ \rightarrow 0, & \text{якщо } \mathbf{x} \in R, \mathbf{x} \notin \partial R, r \rightarrow 0, \\ \rightarrow \infty, & \text{якщо } \mathbf{x} \in R, \mathbf{x} \rightarrow \partial R, \end{cases}$$

де ∂R — межа множини R .

У якості штрафу $\Phi(\mathbf{x}, r)$ можна обирати такі функції:

- 1) $\Phi_1(\mathbf{x}, r) = r \sum_{i=1}^p \frac{\omega_i}{g_i(\mathbf{x})},$
- 2) $\Phi_2(\mathbf{x}, r) = r \sum_{i=1}^p \frac{\omega_i}{g_i^2(\mathbf{x})},$
- 3) $\Phi_3(\mathbf{x}, r) = -r \sum_{i=1}^p \omega_i \ln(g_i(\mathbf{x})),$

де ω_i — додатні вагові коефіцієнти.

Побудувавши штрафну функцію і визначивши внутрішню точку, починаємо реалізовувати процедуру мінімізації $P(\mathbf{x}, r)$ при заданому початковому значенні r^0 . Тоді кінцева для першого етапу обчислювальної процедури точка $\mathbf{x}$ буде початковою точкою для мінімізації функції P при зменшеному значенні r і т. д.

Безумовну мінімізацію штрафної функції проводимо довільним з відомих методів безумовної мінімізації.

На ефективність методу суттєво впливають вибір початкового значення r , метод скорочення значення r під час мінімізації і вибір вагових коефіцієнтів ω_i . Якщо в $P(\mathbf{x}, r)$ значення r^0 обирається дуже малим на початковій стадії процесу мінімізації, ми прийдемо до мінімуму функції $\varphi(\mathbf{x})$, який навряд чи буде поблизу умовного мінімуму в точці $\mathbf{x}^*$. Отже, рух вздовж траєкторії, яка повинна привести до $\mathbf{x}^*$, буде пов'язаний зі значними часовими витратами. З іншого боку, якщо значення r^0 дуже велике, то на перших етапах обчислювального процесу поточна точка неминуче виявиться дуже далеко за межами припустимої області і пошук через необхідність повернення до меж припустимої області також буде дуже затяжним. Вважається, що на першому етапі оптимізаційного процесу кут між $\frac{\partial \varphi(\mathbf{x}^0)}{\partial \mathbf{x}}$ і $\frac{\partial P(\mathbf{x}^0, r^0)}{\partial \mathbf{x}}$ повинен бути гострим, а значення r повинно збільшуватись або зменшуватись шляхом множення чи ділення початкового значення цього параметра на деяке число до тих пір, поки кут між $\frac{\partial \varphi(\mathbf{x}^k)}{\partial \mathbf{x}}$ і $\frac{\partial P(\mathbf{x}^k, r^k)}{\partial \mathbf{x}}$ не стане тупим. Крім того, повинна виконуватись умова $r^0 \geq 10^{-4}$. На практиці частіше за все беруть $r^0 = 50$. Рекомендується при переході від одного етапу обчислювального процесу до наступного помножити поточне значення r на 0.1 (або на число, що знаходиться між 0.02 та 0.1).

Відносно ω_i рекомендації зводяться до наступного: значення ω_i повинні обиратися з таким розрахунком, щоб усі доданки в 1, 2 і 3 після відповідного масштабування були по можливості величинами одного порядку.

Знаходження внутрішньої точки Задача полягає у тому, щоб знайти точку, що задовольняє строгим нерівностям $g_i(\mathbf{x}) > 0$, $i = \overline{1, m}$, як вимагає цього алгоритм внутрішньої точки. Нехай задана точка $\mathbf{x}^0$, що не задовольняє всім нерівностям. Розглянемо дві множини $S = \{s \mid g_s(\mathbf{x}) \leq 0, 1 \leq s \leq m\}$ і $T = \{t \mid g_t(\mathbf{x}) > 0, 1 \leq t \leq m\}$. Стратегія полягає

у побудові такої послідовності точок, що значення $\sum_{s \in S} g_s(\mathbf{x})$ збільшується, але жодне з виконаних обмежень не порушується.

Для цього ми шукаємо $\max \sum_{s \in S} g_s(\mathbf{x})$ (або $\min[-\sum_{s \in S} g_s(\mathbf{x})]$), мінімізуючи функцію

$$U(\mathbf{x}, r_k) = - \sum_{s \in S} g_s(\mathbf{x}) + \sum_{t \in T} \Phi(g_t, r_k),$$

де послідовність $\{r_k\}$ строго прямує до нуля.

Це і є допоміжна задача. На кожному кроці процедури з'являється нова точка, що задовольняє одному або більше обмеженням з тих, що раніше не виконувались. Відповідні індекси переходять від множини S до множини T , при цьому змінюється і сама допоміжна задача. Якщо множина S буде пустою, то внутрішня точка знайдена. І якщо у точці розв'язку допоміжної задачі $S \neq \emptyset$, то умови початкової задачі несумісні.

3.2.3.2 Метод зовнішніх штрафів

У методі зовнішніх штрафів від задачі (3.12) переходимо до послідовності задач безумовної мінімізації

$$\min P(\mathbf{x}, r), \quad r \rightarrow +\infty,$$

де штрафна функція $P(\mathbf{x}, r)$ має вигляд $P(\mathbf{x}, r) = \varphi(\mathbf{x}) + \Psi(\mathbf{x}, r)$.

Функція $\Psi(\mathbf{x}, r)$ має таку властивість:

$$\Psi(\mathbf{x}, r) = \begin{cases} 0, & \text{при } \mathbf{x} \in R, \\ \rightarrow +\infty, & \text{при } \mathbf{x} \notin R, r \rightarrow +\infty. \end{cases}$$

У якості функції $\Psi(\mathbf{x}, r)$ можна, наприклад, обирати такі функції:

- 1) $\Psi(\mathbf{x}, r) = r \sum_{i=1}^m g_i(\mathbf{x})$, якщо $R = \{\mathbf{x} \mid \mathbf{g}(\mathbf{x}) = 0\}$;
- 2) $\Psi(\mathbf{x}, r) = r \sum_{i=1}^m \min(0, g_i(\mathbf{x}))$, якщо $R = \{\mathbf{x} \mid \mathbf{g}(\mathbf{x}) \geq 0\}$.

Параметр r_k рекомендується обирати так: $r_0 = 1$, $r_k = 4r_{k-1}$.

Відмінність між методами внутрішньої і зовнішньої точки Методи внутрішньої і зовнішньої точки засновані на суттєво різних принципах. У методах внутрішньої точки штрафний член перешкоджає порушенню обмежень, а у методах зовнішньої точки він потрібен для того, щоб попередити блукання точок досить далеко від припустимої області. У цьому розумінні метод зовнішньої точки більш простий.

У методах зовнішньої точки виникають труднощі, пов'язані з наявністю обмежень у формі нерівностей, багато з яких ніяк не пов'язані з розв'язком. Ці обмеження можна виключити з обчислювального процесу, так як ставши припустимими на деякій ітерації, вони залишаються припустимими і надалі. Тому, якщо обмеження виконалися як строгі нерівності, то не потрібно розраховувати для них першу і другу частинні похідні, необхідні для знаходження безумовних мінімумів штрафних функцій. Для методів внутрішньої точки, навпаки, потрібна інформація про всі обмеження протягом всього обчислювального процесу і не доводиться замислюватися, які з обмежень не впливають на розв'язок.

За цю перевагу алгоритми зовнішньої точки платять пониженням порядку диференційовності (за $\mathbf{x}$) безумовно мінімізуючої функції у довільній межевій точці припустимої області. У методах внутрішньої точки, навпаки, порядок диференційовності той же, що і у функцій задачі, при умові, що $\Phi(\mathbf{x}, \mathbf{r})$ нескінченно диференційовна.

У методах зовнішньої точки не вимагається існування внутрішності припустимої області, що задається нерівностями. Тому з обмеженнями у формі рівностей поведуться так само просто, як і з обмеженнями у формі нерівностей.

Очевидно, що при обмеженнях у формі рівностей методи внутрішньої точки не можуть бути прийнятними. Вважається, що мінімізація безумовних функцій у методах зовнішньої точки буде більш трудомісткою, ніж у методах внутрішньої точки.

3.2.3.3 Комбінований метод

Для деяких задач, наприклад, для задач з обмеженнями у формі рівностей, методи внутрішньої точки не застосовуються за відсутності внутрішності припустимої області. Для таких задач хотілося б «обробляти» кожен підмножину обмежень окремо. З іншої сторони, наприклад, коли виконання деяких обмежень вимагається весь час, доцільно застосувати комбінований алгоритм. Цей алгоритм буде гарантувати виконання деяких обмежень протягом всього обчислювального процесу. Інші обмеження будуть виконуватись при наближенні до оптимальної точки. Розглянемо алгоритм комбінованого методу, що є комбінацією методів внутрішньої і зовнішньої точок. Нехай розв'язується така задача:

$$\begin{aligned} \min \varphi(\mathbf{x}) \\ g_i(\mathbf{x}) \geq 0, \quad i = \overline{1, m}, \\ h_i(\mathbf{x}) = 0, \quad i = \overline{m+1, p}. \end{aligned} \quad (3.13)$$

Задача нелінійного програмування (3.13) перетворюється у послідовність задач без обмежень шляхом побудови, наприклад, таких штрафних функцій:

$$1) P(\mathbf{x}, \mathbf{r}) = \varphi(\mathbf{x}) + \frac{1}{\mathbf{r}} \sum_{i=1}^p h_i^2(\mathbf{x}) - \mathbf{r} \sum_{i=1}^m \ln g_i(\mathbf{x});$$

$$\begin{aligned}
2) \quad P(\mathbf{x}, r) &= \varphi(\mathbf{x}) + \frac{1}{r} \sum_{i=1}^p h_i^2(\mathbf{x}) + r \sum_{i=1}^m \frac{1}{g_i(\mathbf{x})}; \\
3) \quad P(\mathbf{x}, r) &= \varphi(\mathbf{x}) + \frac{1}{\sqrt{r}} \sum_{i=1}^p h_i^2(\mathbf{x}) - r \sum_{i=1}^m \ln g_i(\mathbf{x}); \\
4) \quad P(\mathbf{x}, r) &= \varphi(\mathbf{x}) + \frac{1}{\sqrt{r}} \sum_{i=1}^p h_i^2(\mathbf{x}) + r \sum_{i=1}^m \frac{1}{g_i(\mathbf{x})}.
\end{aligned}$$

Тут параметр r прямує до нуля. При мінімізації цих штрафних функцій можуть виникнути ускладнення: матриця Гессе функції $\Phi(\mathbf{x}, r)$ при наближенні до точки мінімуму веде себе досить «погано», і тому напрямки спуску можуть бути помилковими. По-друге, швидкість збіжності залежить від початкового вибору r_0 і від способу його зменшення. Існує декілька способів вибору r_0 .

При дуже великих значеннях r_0 мінімізаційний процес починається з внутрішніх точок, що лежать далеко від межі припустимої області, тоді як при дуже малих r_0 початковий розв'язок буде досить близьким до межі, що задається обмеженнями задачі. З практичної точки зору найбільш зручним є такий: $r_0 = 1$. Зменшення r відбувається за алгоритмом $r_k = \frac{1}{4} r_{k-1}$.

Момент знаходження $\min P(\mathbf{x}, r_k)$ на кожному етапі обчислювального процесу визначається за допомогою одного з таких критеріїв:

$$\begin{aligned}
1) \quad & \left| \frac{\partial P(\mathbf{x}^k, r_k)}{\partial x_i} \Delta x_i \right| < \varepsilon, \quad i = \overline{1, n}; \\
2) \quad & \left| \frac{\partial P(\mathbf{x}^k, r_k)}{\partial x_i} \Delta x_i \right| < \frac{P(\mathbf{x}^{k-1}, r_{k-1}) - P(\mathbf{x}^k, r_k)}{5}, \quad i = \overline{1, n}; \\
3) \quad & \left| \frac{\partial P(\mathbf{x}^k, r_k)}{\partial x_i} \right| < \varepsilon, \quad i = \overline{1, n}.
\end{aligned}$$

3.3 Завдання для самостійної роботи

3.3.1 Мінімізація функцій багатьох змінних без обмежень

Знайти точку мінімуму заданої функції за допомогою двох чисельних методів. Середовище та мова програмування для реалізації алгоритмів обираються студентом. Тестові функції та методи задаються викладачем.

У програмі, окрім точки мінімуму, прорахувати кількість ітерацій та кількість обчислень значень функції і, за наявності, її градієнта та матриці Гессе.

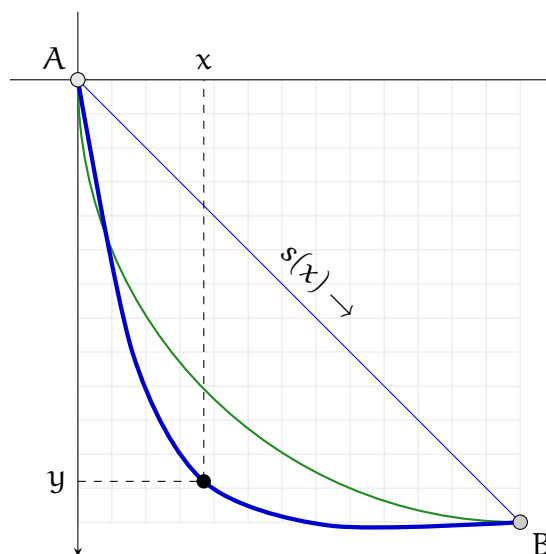
Порівняти результати чисельного експерименту для обох методів та зробити висновки.

Перелік методів:

- метод Нелдера — Міда;
- градієнтний метод із постійним кроком;
- градієнтний метод із регулюванням кроку;
- метод найшвидшого спуску;
- метод спряжених градієнтів із різними способами вибору кроку;
- метод змінної метрики (Девідона — Флетчера — Пауелла);
- метод Ньютона;
- метод Ньютона — Рафсона (три модифікації);
- метод Левенберга — Марквардта.

Частина II

МІНІМІЗАЦІЯ ФУНКЦІОНАЛІВ



Варіаційне числення

4.1 Основна задача варіаційного числення.

Рівняння Ейлера

Багато задач сьогодення зводиться до мінімізації функціоналів. У цих задачах розв'язком буде не скінченновимірний вектор, а функція, на якій досягається екстремум функціоналу.

Розглянемо основну задачу варіаційного числення.

Означення 4.1. Скалярні функції $y = y(x)$, визначені і неперервні на відрізках $[a; b]$ разом з першими похідними $y'(x)$ і що приймають на кінцях відрізка задані значення $y(a) = c$, $y(b) = d$, називаються **припустимими кривими**.

Задача мінімізації функціоналу

$$\mathcal{J}(y) = \int_a^b F(x, y(x), y'(x)) dx \quad (4.1)$$

на множині припустимих кривих $y(x)$, $x \in [a; b]$ називається **основною задачею варіаційного числення**.

Означення 4.2. Припустиму криву $y^0(x)$, $x \in [a; b]$ називають **сильною мінімаллю** функціоналу (4.1), якщо для деякого $\varepsilon > 0$ для всіх припустимих кривих $y(x)$, $x \in [a; b]$ таких, що

$$|y(x) - y^0(x)| \leq \varepsilon, \quad x \in [a; b] \quad (4.2)$$

виконується нерівність

$$\mathcal{J}(y(x)) \geq \mathcal{J}(y^0(x)). \quad (4.3)$$

Якщо нерівність (4.3) виконується тільки для припустимих кривих $y(x)$, $x \in [a; b]$, які крім (4.2) задовольняють ще умові $|y'(x) - y^{0'}(x)| \leq \varepsilon$, $x \in [a; b]$, то кажуть, що $y^0(x)$, $x \in [a; b]$ є **слабка мінімаль** функціоналу (4.1).

Сильна мінімаль є і слабкою, тому всі необхідні умови слабого мінімуму будуть і необхідними умовами сильного мінімуму. Далі будемо розглядати необхідні умови слабого мінімуму.

Вважаємо, що функція $F(x, y, y')$ по сукупності своїх аргументів належить класу $C^2_{[a; b]}$.

Означення 4.3. Нехай $y = y(x)$, $x \in [a; b]$ — припустима крива основної задачі варіаційного числення. Функцію $\delta y(x)$, $x \in [a; b]$ називають **варіацією кривої** $y = y(x)$, якщо функція $\bar{y}(x) = y(x) + \delta y(x)$ є припустимою кривою.

При дослідженні слабких мінімалей зручно варіацію кривої $\delta y(x)$ розглядати у вигляді

$$\delta y(x) = \varepsilon h(x), \quad x \in [a; b],$$

де ε — дійсне число. Далі функцію $h(x)$ будемо називати варіацією кривої. Зрозуміло, що $h(x)$ задовольняє умовам: $h(x) \in C^1$, $x \in [a; b]$; $h(a) = h(b) = 0$.

Означення 4.4. Нехай $\mathcal{J}(y)$ — функціонал, що визначений на припустимих кривих. Якщо $y(x)$ і $h(x)$ фіксовані, $x \in [a; b]$ і функціонал має розвинення

$$\mathcal{J}(y + \varepsilon h) = \mathcal{J}(y) + \varepsilon \delta \mathcal{J}(y) + \frac{\varepsilon^2}{2} \delta^2 \mathcal{J}(y) + o(\varepsilon^2), \quad (4.4)$$

то коефіцієнт $\delta \mathcal{J}(y)$ при ε називається **першою варіацією функціоналу** $\mathcal{J}(y)$ на кривій $y(x)$ і варіації $h(x)$. Коефіцієнт $\delta^2 \mathcal{J}(y)$ при $\frac{1}{2}\varepsilon^2$ у розвиненні (4.4) називається **другою варіацією функціоналу** $\mathcal{J}(y)$.

Перша і друга варіації функціоналу мають вигляд:

$$\delta \mathcal{J}(y) = \int_a^b \left[\frac{\partial F(x, y, y')}{\partial y} h(x) + \frac{\partial F(x, y, y')}{\partial y'} h'(x) \right] dx,$$

$$\delta^2 \mathcal{J}(y) = \int_a^b \left[\frac{\partial^2 F(x, y, y')}{\partial y^2} h^2(x) + 2 \frac{\partial^2 F(x, y, y')}{\partial y \partial y'} h(x) h'(x) + \frac{\partial^2 F(x, y, y')}{\partial y'^2} h'^2(x) \right] dx.$$

Т е о р е м а 4.1 (необхідні умови слабого мінімуму в термінах варіацій функціоналу). На слабкій мінімалі $y^0(x)$, $x \in [a; b]$ і довільній варіації $h(x)$, $x \in [a; b]$

перша варіація функціоналу дорівнює нулеві (умова стаціонарності):

$$\delta \mathcal{J}(y^0(x)) = 0,$$

а друга варіація — невід’ємна:

$$\delta^2 \mathcal{J}(y^0(x)) \geq 0.$$

Користуватись результатом теореми досить складно, тому що у вирази першої і другої варіацій входить невідома функція $h(x)$. Тому будемо намагатись отримати більш прості для користування необхідні умови слабкого мінімуму.

Лема Лагранжа. Якщо рівність $\int_a^b c(x)k(x) dx = 0$ виконується для неперервної функції $c(x)$, $x \in [a; b]$ і всіх функцій $k(x)$, $x \in [a; b]$ класу C^1 , $k(a) = k(b) = 0$, то $c(x) \equiv 0$, $x \in [a; b]$.

Теорема Ейлера (необхідна умова слабкого мінімуму). Для того, щоб припустима крива $y = y(x)$, $x \in [a; b]$ давала слабкий мінімум функціоналу (4.1) основної задачі варіаційного числення, необхідно, щоб вона задовольняла рівнянню Ейлера

$$\frac{\partial F(x, y, y')}{\partial y} - \frac{d}{dx} \frac{\partial F(x, y, y')}{\partial y'} = 0.$$

Доведення. Скористаємося необхідною умовою слабкого мінімуму в термінах варіацій: $\delta \mathcal{J}(y) = 0$. Маємо

$$\int_a^b \left[\frac{\partial F(x, y, y')}{\partial y} h(x) + \frac{\partial F(x, y, y')}{\partial y'} h'(x) \right] dx = 0.$$

Далі,

$$\begin{aligned} \int_a^b \frac{\partial F(x, y, y')}{\partial y'} h'(x) dx &= \left[\begin{array}{ll} u = \frac{\partial F(x, y, y')}{\partial y'} & dv = h'(x) dx \\ du = \frac{d}{dx} \frac{\partial F(x, y, y')}{\partial y'} & v = h(x) \end{array} \right] = \\ &= \frac{\partial F(x, y, y')}{\partial y'} h(x) \Big|_a^b - \int_a^b \left[\frac{d}{dx} \frac{\partial F(x, y, y')}{\partial y'} \right] h(x) dx = [h(a) = h(b) = 0] = \\ &= - \int_a^b \left[\frac{d}{dx} \frac{\partial F(x, y, y')}{\partial y'} \right] h(x) dx. \end{aligned}$$

Остаточно отримаємо

$$\delta \mathcal{J}(y) = \int_a^b \left[\frac{\partial F(x, y, y')}{\partial y} - \frac{d}{dx} \frac{\partial F(x, y, y')}{\partial y'} \right] h(x) dx = 0.$$

Використавши [лема Лагранжа](#), отримаємо, що розв'язок основної задачі варіаційного числення з необхідністю задовольняє рівнянню

$$\frac{\partial F(x, y, y')}{\partial y} - \frac{d}{dx} \frac{\partial F(x, y, y')}{\partial y'} = 0. \quad (4.5)$$

Теорему доведено. 

Рівняння (4.5) має назву **рівняння Ейлера**.

Припустима крива, що задовольняє рівнянню Ейлера, називається **екстремаллю**.

Рівняння (4.5) можна написати у розгорнутому вигляді:

$$\frac{\partial F(x, y, y')}{\partial y} - \frac{\partial^2 F(x, y, y')}{\partial x \partial y'} - \frac{\partial^2 F(x, y, y')}{\partial y \partial y'} y' - \frac{\partial^2 F(x, y, y')}{\partial y'^2} y'' = 0.$$

Доведення [теорему Ейлера](#) має недоліки, на які вперше вказав Дюбуа-Реймонд. При інтегруванні по частинам неявно вважалось, що функція $\frac{\partial F(x, y, y')}{\partial y'}$ є гладкою. У розгорнутому вигляді $\frac{d}{dx} \frac{\partial F(x, y, y')}{\partial y'}$ має доданок з $y''(x)$, але у постановці задачі цього не передбачалось. Для строгого доведення теореми Ейлера використовується

Лема Дюбуа-Реймонда. Якщо рівність $\int_a^b c(x)k(x) dx = 0$ виконується для неперервної функції $c(x)$, $x \in [a; b]$ і для всіх неперервних функцій $k(x)$, $x \in [a; b]$, що задовольняють умові $\int_a^b k(x) dx = 0$, то необхідно, щоб $c(x) \equiv \text{const}$, $x \in [a; b]$.

Рівняння Ейлера у вигляді (4.5) називається **рівнянням Ейлера у диференціальній формі**. Далі розглянемо рівняння Ейлера в інтегральній формі.

Т е о р е м а 4.2 (рівняння Ейлера в інтегральній формі). Слабка мінімаль $y^0 = y^0(x)$, $x \in [a; b]$ основної задачі задовольняє інтегральному рівнянню Ейлера

$$\frac{\partial F(x, y, y')}{\partial y'} = \int_a^x \frac{\partial F(s, y, y')}{\partial y} ds + \text{const}.$$

Доведення. Знову розглядаємо першу варіацію на кривій, що є розв'язком основної

задачі варіаційного числення:

$$\delta \mathcal{J}(y) = \int_a^b \left[\frac{\partial F(x, y, y')}{\partial y} h(x) + \frac{\partial F(x, y, y')}{\partial y'} h'(x) \right] dx = 0.$$

Розглянемо

$$\begin{aligned} \int_a^b \frac{\partial F(x, y, y')}{\partial y} h(x) dx &= \left[\begin{array}{ll} u = h(x) & dv = \frac{\partial F(x, y, y')}{\partial y} dx \\ du = h'(x) dx & v = \int_a^x \frac{\partial F(s, y, y')}{\partial y} ds \end{array} \right] = \\ &= \left[h(x) \int_a^x \frac{\partial F(s, y, y')}{\partial y} ds \right] \Big|_a^b - \int_a^b \int_a^x \frac{\partial F(s, y, y')}{\partial y} ds h'(x) dx = [h(a) = h(b) = 0] = \\ &= - \int_a^b \int_a^x \frac{\partial F(s, y, y')}{\partial y} ds h'(x) dx. \end{aligned}$$

Отримали

$$\delta \mathcal{J}(y) = \int_a^b \left[- \int_a^x \frac{\partial F(s, y, y')}{\partial y} ds + \frac{\partial F(x, y, y')}{\partial y'} \right] h'(x) dx = 0.$$

Застосувавши [лему Дюбуа-Реймонда](#), отримаємо рівняння Ейлера в інтегральній формі:

$$\frac{\partial F(x, y, y')}{\partial y'} - \int_a^x \frac{\partial F(s, y(s), y'(s))}{\partial y} ds = \text{const.} \quad (4.6)$$

Теорему доведено. 

Рівняння (4.6) напишемо у вигляді

$$\frac{\partial F(x, y, y')}{\partial y'} = \int_a^x \frac{\partial F(s, y(s), y'(s))}{\partial y} ds + \text{const.} \quad (4.7)$$

Інтеграл справа є функція класу $C^1_{[a,b]}$. Тому і ліва частина має неперервну повну похідну за x . Якщо взяти похідну за x від обох частин (4.7), то отримаємо рівняння Ейлера у диференціальній формі.

У першому доведенні вважалось, що $\frac{\partial F(x, y, y')}{\partial y'}$ диференційовна вздовж розв'язку, а у другому доведенні доведено наявність цієї властивості.

Розгорнутий вигляд рівняння Ейлера має другу похідну $y(x)$, тобто $y''(x)$. Але у постановці задачі $y(x) \in C_{[a;b]}^1$. Доведемо, що екстремаль належить класу $C_{[a;b]}^2$.

Означення 4.5. Екстремаль $y = y(x)$ називається **неособливою**, якщо вздовж неї виконується нерівність

$$\frac{\partial^2 F(x, y(x), y'(x))}{\partial y'^2} \neq 0, \quad x \in [a; b].$$

Теорема 4.3 (Гільберта). Неособлива екстремаль належить класу $C_{[a;b]}^2$.

Тепер усі проблеми, що пов'язані з рівнянням Ейлера, повністю розв'язані.

Приклад 4.1. Знайти екстремалі функціоналу

$$\mathcal{J}(y) = \int_0^1 (y'^2 + xy) \, dx, \quad y(0) = 0, \quad y(1) = 1.$$

Розв'язування. Рівняння Ейлера має вигляд

$$x - 2y'' = 0.$$

Загальний розв'язок цього рівняння такий: $y = \frac{1}{12}x^3 + C_1x + C_2$. Використовуючи межові умови, отримаємо:

$$\begin{cases} C_2 = 0, \\ \frac{1}{2} + C_1 + C_2 = 1, \end{cases}$$

звідки $C_1 = \frac{11}{12}$, $C_2 = 0$. Отримали, що мінімум функціоналу може досягатись на кривій $y = \frac{1}{12}x^3 + \frac{11}{12}x$. ☑

4.1.1 Часткові випадки рівняння Ейлера

Розглянемо часткові випадки інтегрування рівняння Ейлера.

1) Підінтегральна функція не залежить від y .

У цьому випадку рівняння Ейлера має вигляд

$$\frac{d}{dx} \frac{\partial F(x, y, y')}{\partial y'} = 0,$$

звідки $\frac{\partial F(x, y, y')}{\partial y'} = C$. Це рівняння першого порядку. Розв'язавши його відносно

y' , отримаємо $y' = f(x, c)$. Звідки $y(x)$ знаходимо квадратурою.

Приклад 4.2. Знайти екстремалі функціоналу

$$\mathcal{J}(y) = \int_1^2 y'^2 x^2 dx, \quad y(1) = 1, \quad y(2) = 3.$$

Розв'язування. Так як функція $F(x, y, y')$ не залежить від y , то рівняння Ейлера має вигляд

$$\frac{d}{dx} \frac{\partial F(x, y, y')}{\partial y'} = 0,$$

тобто $\frac{d}{dx} 2y'x^2 = 0$. Отримаємо, що $2y'x^2 = C_1$ або $y' = \frac{C_1}{2x^2}$. Звідси $y = -\frac{C_1}{2x} + C_2$.

Використовуючи межові умови, отримаємо систему

$$\begin{cases} 1 = -\frac{C_1}{2} + C_2, \\ 3 = -\frac{C_1}{4} + C_2. \end{cases}$$

Остаточно, $C_1 = 8$, $C_2 = 5$; $y = -\frac{4}{x} + 5$.



2) Підінтегральна функція не залежить від x .

Рівняння Ейлера має вигляд

$$\frac{\partial F(y, y')}{\partial y} - \frac{\partial^2 F(y, y')}{\partial y \partial y'} y' - \frac{\partial^2 F(y, y')}{\partial y'^2} y'' = 0.$$

Домножимо вираз лівої частини на y' і отримаємо точну похідну

$$\frac{d}{dx} \left(F(y, y') - y' \frac{\partial F(y, y')}{\partial y'} \right) = 0,$$

тоді отримаємо таке рівняння

$$F(y, y') - y' \frac{\partial F(y, y')}{\partial y'} = C_1.$$

Приклад 4.3. Знайти екстремалі функціоналу

$$\mathcal{J}(y) = \int_0^1 y y'^2 dx, \quad y(0) = 0, \quad y(1) = 1.$$

Розв'язування. Рівняння Ейлера має вигляд $yy'^2 = C_1$. Тоді $\sqrt{y} dy = \sqrt{C_1} dx$, або $y = x^{2/3}$. ☒

3) **Функція F не залежить від y' .**

Тоді рівняння Ейлера матиме вигляд $\frac{\partial F(x, y)}{\partial y} = 0$, воно не диференціальне, тому до розв'язку не входять константи і межові умови можуть не виконуватись.

Приклад 4.4. Знайти екстремаль функціоналу

$$\mathcal{J}(y) = \int_1^2 (2xy - y^2) dx, \quad y(1) = 1, \quad y(2) = 7.$$

Розв'язування. Рівняння Ейлера має вигляд $x - y = 0$, звідки $y = x$. Крива $y = x$ задовольняє умові $y(1) = 1$, але умові $y(2) = 7$ не задовольняє. Отже, варіаційна задача розв'язку не має. ☒

4) **Функція F залежить лише від y' .**

Тоді рівняння Ейлера має вигляд

$$\frac{\partial^2 F(y')}{\partial y'^2} y'' = 0.$$

Якщо $\frac{\partial^2 F(y')}{\partial y'^2} y'' \neq 0$, то екстремальними є прямі $y = C_1 x + C_2$.

Приклад 4.5. Знайти екстремаль функціоналу

$$\mathcal{J}(y) = \int_0^1 \sqrt{1 + y'^2} dx, \quad y(0) = 0, \quad y(1) = 1.$$

Цей функціонал визначає довжину кривої, що з'єднує ці дві точки.

Розв'язування. Так як $\frac{\partial^2 F(y')}{\partial y'^2} = \frac{1}{(1 + y'^2)^{3/2}} \neq 0$, то рівняння Ейлера має вигляд $y'' = 0$ і $y = C_1 x + C_2$. Враховуючи умови, отримуємо $y = x$. ☒

5) **Функція F залежить від y' лінійно,**

тобто

$$F(x, y, y') = M(x, y) + y'N(x, y).$$

У цьому випадку рівняння Ейлера має вигляд

$$\frac{\partial M}{\partial y} - \frac{\partial N}{\partial x} = 0.$$

Отримане рівняння не є диференціальним, тому отримана крива не завжди задовольняє межовим умовам і тоді задача не має розв'язку. Якщо ж у деякій області $\frac{\partial M}{\partial y} - \frac{\partial N}{\partial x} \equiv 0$, то вираз $F(x, y, y') = M(x, y) dx + N(x, y) dy$ є повним диференціалом, і функціонал

$$\mathcal{J}(y) = \int_a^b F(x, y, y') dx = \int_{(a,c)}^{(b,d)} M(x, y) dx + N(x, y) dy$$

не залежить від шляху інтегрування: значення функціоналу $\mathcal{J}(y)$ однакове на всіх припустимих кривих. У цьому випадку варіаційна задача не має сенсу.

Приклад 4.6. Знайти екстремаль функціоналу

$$\mathcal{J}(y) = \int_1^2 (y^2 + 2xyy') dx, \quad y(1) = 2, \quad y(2) = 4.$$

Розв'язування. Функціонал залежить лінійно від y' . Маємо

$$\frac{\partial M}{\partial y} = 2y, \quad \frac{\partial N}{\partial x} = 2y, \quad \text{тоді} \quad \frac{\partial M}{\partial y} - \frac{\partial N}{\partial x} \equiv 0.$$

Отже, підінтегральний вираз є повним диференціалом, і інтеграл не залежить від шляху інтегрування. Варіаційна задача не має сенсу. 

4.2 Умови мінімуму в класі кусково-гладких функцій

Основна задача може не мати розв'язку в класі гладких припустимих кривих. Тому будемо розглядати ширший клас припустимих кривих, а саме — кусково-гладких. Це такі неперервні функції $y(x)$, $x \in [a; b]$, $y(a) = c$, $y(b) = d$, що мають неперервні похідні всюди на $[a; b]$ за виключенням, можливо, скінченної кількості точок, у яких похідні мають розриви першого роду. Має місце

Т е о р е м а 4.4 (Вейєрштрасса — Ердмана). Нехай $y = y(x)$, $x \in [a; b]$ — не-

перервна функція, похідна якої неперервна всюди на $[a; b]$ за виключенням точок $x_1, \dots, x_k$. Якщо на кривій $y = y(x)$ функціонал (4.1) досягає слабкого мінімуму, то необхідно, щоб

- 1) функція $y = y(x)$ в інтервалах між точками $a, x_1, \dots, x_k, b$ задовольняла рівнянню Ейлера;
- 2) У точках $x_1, \dots, x_k$ кривої $y = y(x)$ виконувались умови Вейєрштраса — Ердмана:

$$\left. \frac{\partial F(x, y(x), y'(x))}{\partial y'} \right|_{x=x_i-0} = \left. \frac{\partial F(x, y(x), y'(x))}{\partial y'} \right|_{x=x_i+0}, \quad i = \overline{1, k};$$

$$\begin{aligned} F(x, y(x), y'(x)) - y'(x) \left. \frac{\partial F(x, y(x), y'(x))}{\partial y'} \right|_{x=x_i-0} &= \\ &= F(x, y(x), y'(x)) - y'(x) \left. \frac{\partial F(x, y(x), y'(x))}{\partial y'} \right|_{x=x_i+0}. \end{aligned}$$

4.3 Дослідження другої варіації.

Умови Лежандра та Якобі

Нижче розглядаються дві необхідні умови слабкого мінімуму, що впливають з невід'ємності другої варіації функціоналу. Має місце така

Теорема Лежандра. Якщо на припустимій кривій $y = y(x)$, $x \in [a; b]$ функціонал основної задачі досягає слабкого мінімуму, то вздовж цієї кривої виконується умова Лежандра

$$\frac{\partial^2 F(x, y(x), y'(x))}{\partial y'^2} \geq 0, \quad x \in [a; b].$$

Приклад 4.7. Розглянемо умови [прикладу 4.1](#).

$$\mathcal{J}(y) = \int_0^1 (y'^2 + xy) \, dx, \quad y(0) = 0, \quad y(1) = 1.$$

Отримали, що припустима крива $y = \frac{1}{12}x^3 + \frac{11}{12}x$ може доставляти слабкий мінімум

функціоналу. Розглянемо необхідну умову Лежандра:

$$\left. \frac{\partial^2 F(x, y(x), y'(x))}{\partial y'^2} \right|_{y = \frac{1}{12}x^3 + \frac{11}{12}x} = 2 > 0.$$

Необхідна умова Лежандра виконується. Отже, припустима крива $y = \frac{1}{12}x^3 + \frac{11}{12}x$ може доставляти слабкий мінімум функціоналу. 

Означення 4.6. *Вздовж припустимої кривої $y = y(x)$ виконується посилена умова Лежандра, якщо $\frac{\partial^2 F(x, y(x), y'(x))}{\partial y'^2} > 0$, $x \in [a; b]$.*

Якщо позначити підінтегральний вираз у другій варіації через

$$w(x, h, h') = h^2 \frac{\partial^2 F(x, y(x), y'(x))}{\partial y^2} + 2hh' \frac{\partial^2 F(x, y(x), y'(x))}{\partial y \partial y'} + h'^2 \frac{\partial^2 F(x, y(x), y'(x))}{\partial y'^2},$$

то друга варіація функціоналу основної задачі вздовж припустимої кривої $y = y(x)$, $x \in [a; b]$ матиме вигляд

$$\delta^2 \mathcal{J}(y, h) = \int_a^b w(x, h, h') dx. \quad (4.8)$$

Означення 4.7. *Задача мінімізації другої варіації (4.8) на варіаціях $h(x)$, $x \in [a; b]$ припустимої кривої $y(x)$, $x \in [a; b]$ називається **приєднаною задачею про мінімум**.*

Так як $\delta^2 \mathcal{J}(y^0(x), h) \geq 0$ на слабкій мінімалі при всіх варіаціях $h(x)$, $x \in [a; b]$, то приєднана задача про мінімум вздовж слабкої мінімалі $y^0(x)$ завжди має розв'язок $h^0(x) \equiv 0$, $x \in [a; b]$ і $\delta^2 \mathcal{J}(y^0(x), h^0(x)) = 0$.

Рівняння Ейлера

$$\frac{\partial w(\cdot)}{\partial h} - \frac{d}{dx} \frac{\partial w(\cdot)}{\partial h'} = 0 \quad (4.9)$$

для функціоналу (4.8) називається **рівнянням Якобі основної задачі варіаційного числення**. Його можна записати у вигляді

$$a(x)h'' + b(x)h' + c(x)h = 0,$$

де

$$\begin{aligned} a(x) &= \frac{\partial^2 F(x, y(x), y'(x))}{\partial y'^2}, & b(x) &= \frac{d}{dx} a(x), \\ c(x) &= \frac{d}{dx} \frac{\partial^2 F(x, y(x), y'(x))}{\partial y' \partial y} - \frac{\partial^2 F(x, y(x), y'(x))}{\partial y^2}. \end{aligned}$$

Означення 4.8. Кажуть, що вздовж припустимої кривої $y = y(x)$, $x \in [a; b]$ точка $\bar{x} \in [a; b]$ **спряжена з точкою a** , якщо існує такий нетривіальний розв'язок $h(x) \neq 0$, $x \in [a; b]$ рівняння Якобі (4.9), що $h(a) = h(\bar{x}) = 0$.

Теорема Якобі. Вздовж неособливої мінімалі $y^0(x)$, $x \in [a; b]$ не існує точок $\bar{x} \in (a; b)$, спряжених з точкою a .

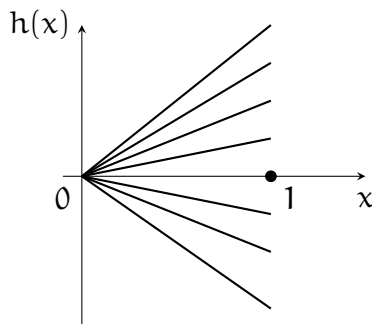
Приклад 4.8. Розглянемо [приклад 4.1](#). Ми довели, що припустима крива

$$y = \frac{1}{12}x^3 + \frac{11}{12}x$$

задовольняє двом необхідним умовам — Ейлера і Лежандра і є неособливою. Отже, вона може доставляти слабкий мінімум функціоналу

$$\mathcal{J}(y) = \int_0^1 (y'^2 + xy) dx, \quad y(0) = 0, \quad y(1) = 1.$$

Складемо рівняння Якобі. Маємо $a(x) = 2$, $b(x) = 0$, $c(x) = 0$, тоді рівняння Якобі має вигляд $h'' = 0$ і $h(x) = C_1x + C_2$. Так як $h(0) = 0$, то $C_2 = 0$, і тоді $h(x) = C_1x$.



На малюнку представлено розв'язок рівняння Якобі. На проміжку $(0; 1)$ $h(x)$ не перетворюється в нуль, тобто на проміжку $(0; 1)$ не існує точок, спряжених з точкою 0. Необхідна умова Якобі виконується. ☒

Приклад 4.9. Досліджується задача

$$\mathcal{J}(y) = \int_0^{2\pi} (y'^2 - y^2) dx \rightarrow \min, \quad y(0) = y(2\pi) = 0.$$

Рівняння Ейлера має вигляд $y'' + y = 0$. Його розв'язок $y(x) = C_1 \sin x + C_2 \cos x$. Використовуючи умови, знаходимо невідомі C_1 і C_2 :

$$y(0) = C_2 = 0, \quad y(2\pi) = C_2 = 0.$$

Екстремаль має вигляд $y = C_1 \sin x$, $x \in [0; 2\pi]$. На ній виконується умова Лежан-

дра

$$\left. \frac{\partial^2 F(x, y(x), y'(x))}{\partial y'^2} \right|_{y=C_1 \sin x} = 2 > 0$$

і вона є неособливою. Рівняння Якобі має вигляд

$$h'' + h = 0, \quad h(0) = 0, \quad h(2\pi) = 0.$$

Розв'язком цього рівняння є функція $h(x) = C \sin x$. Він перетворюється в нуль в точці $x = \pi$. Значить, точка $x = \pi$ є спряженою з точкою нуль. Необхідна умова Якобі не виконується. 

Означення 4.9. Припустима крива $y = y(x)$, $x \in [a; b]$ задовольняє **посиленій умові Якобі**, якщо вздовж неї не існує точок $\bar{x} \neq a$ з $(a; b]$, спряжених з точкою a .

Теорема 4.5 (достатня умова слабкого мінімуму). Для того, щоб припустима крива $y = y(x)$ доставляла слабкий мінімум в основній задачі варіаційного числення, достатньо, щоб

- 1) крива $y = y(x)$, $x \in [a; b]$ задовольняла рівнянню Ейлера;
- 2) вздовж $y = y(x)$ виконувалась посилена умова Лежандра;
- 3) на $(a; b]$ не було точок, спряжених з точкою a , тобто виконувалась посилена умова Якобі.

Приклад 4.10. Розглядаються умови [прикладу 4.1](#):

$$\mathcal{J}(y) = \int_1^2 (y'^2 + 2xy) dx, \quad y(1) = 2, \quad y(2) = 4.$$

Вище отримали, що усі три необхідні умови слабкого мінімуму виконуються. Більше того, виконується посилена умова Лежандра:

$$\left. \frac{\partial^2 F(x, y, y')}{\partial y'^2} \right|_{y=\frac{1}{12}x^3 + \frac{11}{12}x} = 2 > 0$$

і на $(0; 1]$ нетривіальний розв'язок рівняння Якобі не перетворюється в нуль, тобто на $(0; 1]$ відсутні точки, спряжені з точкою 0.

Отже, усі три необхідні умови слабкого мінімуму і достатня умова виконуються. Це означає, що припустима крива $y = \frac{1}{12}x^3 + \frac{11}{12}x$ доставляє слабкий мінімум функціоналу. 

4.4 Ізопериметрична задача

Розглянемо таку задачу: серед кривих $y = y(x)$, $y \in C^1$, $x \in [a; b]$, що задовольняють умовам $y(a) = C_1$, $y(b) = C_2$, на яких функціонал

$$\mathcal{K}(y) = \int_a^b G(x, y, y') dx \quad (4.10)$$

приймає задане значення l , знайти ту, на якій функціонал

$$\mathcal{J}(y) = \int_a^b F(x, y, y') dx \quad (4.11)$$

досягає екстремуму.

Вважається, що функції $G(x, y, y')$ і $F(x, y, y')$ мають неперервні похідні першого і другого порядків для $a \leq x \leq b$ і довільних y і y' . Також розв'язок задачі не є екстремаллю функціонала (4.10).

Алгоритм розв'язання сформульованої задачі дає

Теорема 4.6. Якщо крива $y = y(x)$ дає екстремум функціоналу (4.11), задовольняє умовам

$$\int_a^b G(x, y, y') dx = l, \quad y(a) = C_1, \quad y(b) = C_2$$

і не є екстремаллю функціоналу (4.10), то існує така константа λ , що ця крива $y = y(x)$ є екстремаллю функціоналу

$$\int_a^b [F(x, y, y') + \lambda G(x, y, y')] dx.$$

Приклад 4.11. Знайти екстремаль в ізопериметричній задачі

$$\min \mathcal{J}(y) = \int_0^1 y'^2 dx, \quad \int_0^1 y dx = 2, \quad y(0) = 1, \quad y(1) = 3. \quad (4.12)$$

Розв'язування. Розглянемо функціонал

$$\int_0^1 [y'^2 + \lambda y] dx. \quad (4.13)$$

Згідно з [теоремою 4.6](#), розв'язком задачі є екстремаль функціоналу (4.13). Рівняння Ейлера має вигляд

$$\lambda - 2y'' = 0.$$

Проінтегрувавши його, отримаємо

$$y(x) = \frac{\lambda}{4}x^2 + C_1x + C_2.$$

Постійні C_1 , C_2 і λ знайдемо з умов (4.12): $C_1 = \frac{1}{2}$, $C_2 = 1$, $\lambda = 6$. Отже, шукана екстремаль має вигляд $y(x) = \frac{3}{2}x^2 + \frac{1}{2}x + 1$. Очевидно, що отримана екстремаль не є екстремаллю функціоналу $\int_0^1 y dx$. ☑

4.5 Задача з рухомими межами

Розглянемо задачу знаходження екстремуму функціоналу $\mathcal{J}(y) = \int_a^b F(x, y, y') dx$, визначеного на гладких кривих $y = y(x)$, кінці яких лежать на заданих кривих $y = \varphi(x)$ і $y = \psi(x)$, $\varphi(x)$ і $\psi(x) \in C^1$, $x \in [a; b]$. Точки a і b не задані.

Теорема 4.7. Нехай крива $y = y(x)$ дає екстремум функціоналу серед усіх кривих класу C^1 , що з'єднують криві $\varphi(x)$ і $\psi(x)$. Тоді крива $y = y(x)$ є екстремаллю функціоналу і у точках a і b — точках перетину кривої $y = y(x)$ з кривими $\varphi(x)$ і $\psi(x)$ — виконуються умови трансверсальності:

$$\begin{aligned} \left[F(x, y(x), y'(x)) + \frac{\partial F(x, y(x), y'(x))}{\partial y'} (\varphi'(x) - y'(x)) \right]_{x=a} &= 0, \\ \left[F(x, y(x), y'(x)) + \frac{\partial F(x, y(x), y'(x))}{\partial y'} (\psi'(x) - y'(x)) \right]_{x=b} &= 0. \end{aligned}$$

Отже, алгоритм розв'язування задачі такий:

- 1) Написати і розв'язати відповідне рівняння Ейлера. Отримаємо сукупність кривих $y = f(x, C_1, C_2)$, що залежать від невідомих параметрів C_1, C_2 .

2) З умов трансверсальності і перетину екстремалі з кривими $y = \varphi(x)$ і $y = \psi(x)$

$$f(a, C_1, C_2) = \varphi(a), \quad f(b, C_1, C_2) = \psi(b)$$

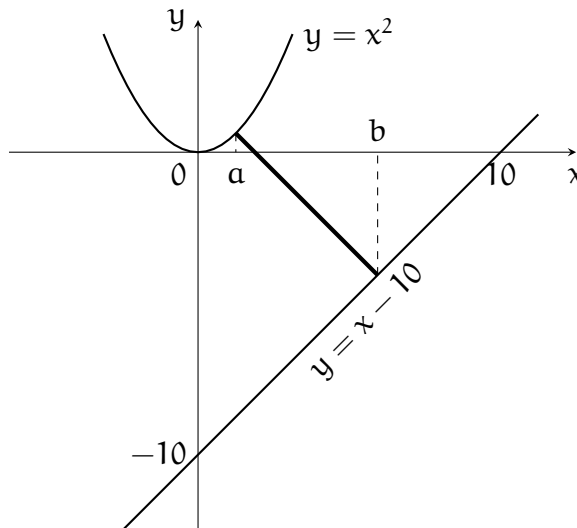
визначимо параметри C_1, C_2, a і b .

Приклад 4.12. Знайти відстань від параболи $y = x^2$ до прямої $y = x - 10$.

Розв'язування. Задача зводиться до мінімізації функціоналу

$$\mathcal{J}(y) = \int_a^b \sqrt{1 + y'^2} dx$$

при умові, що кінці екстремалі рухаються по кривим $\varphi(x) = x^2$ і $\psi(x) = x - 10$.



Так як функція $F(x, y, y')$ залежить тільки від y' , то згідно з частковим випадком екстремаль має вигляд $y = C_1 x + C_2$. Параметри C_1, C_2, a і b знаходяться з умов трансверсальності і перетину екстремалі з кривими $\varphi(x) = x^2$ і $\psi(x) = x - 10$. Отримаємо таку систему рівнянь:

$$\begin{cases} \sqrt{1 + C_1^2} + (2a - C_1) \frac{C_1}{\sqrt{1 + C_1^2}} = 0, \\ \sqrt{1 + C_1^2} + (1 - C_1) \frac{C_1}{\sqrt{1 + C_1^2}} = 0, \\ C_1 a + C_2 = a^2, \\ b - 10 = C_1 b + C_2. \end{cases}$$

Розв'язавши цю систему, отримаємо такі значення параметрів:

$$C_1 = -1, \quad C_2 = \frac{3}{4}, \quad a = \frac{1}{2}, \quad b = \frac{43}{8}.$$

Отже, рівняння екстремалі таке: $y = -x + \frac{3}{4}$. А так як очевидно, що задача має розв'язок, то відстань між параболою $y = x^2$ і прямою $y = x - 10$ буде дорівнювати

$$\mathcal{J}(y) = \int_{1/2}^{48/3} \sqrt{1 + (-1)^2} dx = \frac{39\sqrt{2}}{8}.$$



Існують і інші постановки варіаційних задач. Їх можна знайти у [4], [8].

4.6 Завдання для самостійної роботи

4.6.1 Основна задача варіаційного числення

Дослідіть на екстремум функціонал, використовуючи необхідні та достатні умови слабого екстремуму.

Варіант 1 $\mathcal{J}[y] = \int_0^1 (2y'^2 + 7y' + x) dx, \quad y(0) = 0, \quad y(1) = 1$

Варіант 2 $\mathcal{J}[y] = \int_1^2 (4y'^2 + 4 \cos y') dx, \quad y(1) = 1, \quad y(2) = 3$

Варіант 3 $\mathcal{J}[y] = \int_0^1 (5y'^2 - 2 \cos y' + x^2 - 1) dx, \quad y(0) = 2, \quad y(1) = 5$

Варіант 4 $\mathcal{J}[y] = \int_0^1 (2y'^2 + 4y - y^2) dx, \quad y(0) = 0, \quad y(1) = 1$

Варіант 5 $\mathcal{J}[y] = \int_1^2 (7y'^2 + 7x^3 - x) dx, \quad y(1) = 0, \quad y(2) = 2$

Варіант 6 $\mathcal{J}[y] = \int_2^3 (2y'^2 - xy + x) dx, \quad y(2) = 0, \quad y(3) = 2$

Варіант 7 $\mathcal{J}[y] = \int_1^2 (9y'^2 + 2x^2 - xy) dx, \quad y(1) = 1, \quad y(2) = 8$

Варіант 8 $\mathcal{J}[y] = \int_2^7 (7y'^2 - 4xy + 1) dx, \quad y(2) = 0, \quad y(7) = 2$

Варіант 9 $\mathcal{J}[y] = \int_0^3 (xy'^2 + x^3) dx, \quad y(0) = 1, \quad y(3) = 0$

Варіант 10 $\mathcal{J}[y] = \int_1^2 (4y'^2 - 7x^2 + y) dx, \quad y(1) = 0, \quad y(2) = 3$

Варіант 11 $\mathcal{J}[y] = \int_0^7 (5y'^2 + \cos y' + x^3) dx, \quad y(0) = 1, \quad y(7) = 2$

Варіант 12 $\mathcal{J}[y] = \int_0^1 (\cos y' + \sin y') dx, \quad y(0) = 1, \quad y(1) = 2$

Варіант 13 $\mathcal{J}[y] = \int_1^3 \left(\frac{1}{\cos y'} + x - 1 \right) dx, \quad y(1) = 0, \quad y(3) = 2$

Варіант 14 $\mathcal{J}[y] = \int_0^4 \left(1 + 2x^2 + \frac{1}{\sin y'} \right) dx, \quad y(0) = 0, \quad y(4) = 21$

Варіант 15 $\mathcal{J}[y] = \int_0^1 (\operatorname{tg} y' + x) dx, \quad y(0) = 1, \quad y(1) = 5$

Варіант 16 $\mathcal{J}[y] = \int_2^3 (\operatorname{ctg} y' + x^2 - 1) dx, \quad y(2) = 1, \quad y(3) = 4$

Варіант 17 $\mathcal{J}[y] = \int_1^2 \left(\frac{1}{\cos y'} + \sin y' \right) dx, \quad y(1) = 1, \quad y(2) = 5$

Варіант 18 $\mathcal{J}[y] = \int_0^2 (\sin y' + x) dx, \quad y(0) = 1, \quad y(2) = 3$

$$\text{Варіант 19} \quad \mathcal{J}[y] = \int_1^2 (y'^2 - \sin y') dx, \quad y(1) = 1, \quad y(2) = 1$$

$$\text{Варіант 20} \quad \mathcal{J}[y] = \int_1^3 (2y'^2 + 2 \cos y') dx, \quad y(1) = 0, \quad y(3) = 4$$

$$\text{Варіант 21} \quad \mathcal{J}[y] = \int_2^4 (7y'^2 + \operatorname{tg} y') dx, \quad y(2) = 10, \quad y(4) = 5$$

$$\text{Варіант 22} \quad \mathcal{J}[y] = \int_0^1 (y'^2 - \sin y' - x) dx, \quad y(0) = 1, \quad y(1) = 4$$

$$\text{Варіант 23} \quad \mathcal{J}[y] = \int_1^2 (2y'^2 + \cos y' + x^2) dx, \quad y(1) = 1, \quad y(2) = 5$$

$$\text{Варіант 24} \quad \mathcal{J}[y] = \int_2^7 (17y'^2 + \operatorname{ctg} y' - x) dx, \quad y(2) = 0, \quad y(7) = 2$$

$$\text{Варіант 25} \quad \mathcal{J}[y] = \int_9^{10} (y'^2 + \operatorname{tg} y' + \cos y') dx, \quad y(9) = 1, \quad y(10) = 4$$

4.6.2 Ізопериметрична задача

Знайдіть припустимі екстремалі ізопериметричної задачі. Проаналізуйте поведінку функціоналу на отриманих кривих.

$$\text{Варіант 1} \quad \mathcal{J}[y] = \int_0^1 (y^2 + y'^2) dx, \quad y(0) = 0, \quad y(1) = 4e, \quad \int_0^1 y e^x dx = 1 + e^2$$

$$\text{Варіант 2} \quad \mathcal{J}[y] = \int_1^2 x^3 y'^2 dx, \quad y(1) = 4, \quad y(2) = 1, \quad \int_1^2 y dx = 2$$

$$\text{Варіант 3} \quad \mathcal{J}[y] = \int_0^\pi y'^2 dx, \quad y(0) = 1, \quad y(\pi) = -1, \quad \int_0^\pi y \cos x dx = \frac{\pi}{2}$$

$$\text{Варіант 4} \quad \mathcal{J}[y] = \frac{1}{2} \int_0^1 y'^2 dx, \quad y(0) = 0, \quad y(1) = 0, \quad \int_0^1 x y' dx = 2$$

$$\text{Варіант 5} \quad \mathcal{J}[y] = \int_0^1 y'^2 dx, \quad y(0) = 0, \quad y(1) = 0, \quad \int_0^1 y dx = -1$$

$$\text{Варіант 6} \quad \mathcal{J}[y] = \int_0^1 y'^2 dx, \quad y(0) = 0, \quad y(1) = \frac{1}{4}, \quad \int_0^1 (y - y'^2) dx = \frac{1}{12}$$

$$\text{Варіант 7} \quad \mathcal{J}[y] = \int_1^2 x^2 y'^2 dx, \quad y(1) = 1, \quad y(2) = 2, \quad \int_1^2 xy dx = \frac{7}{3}$$

$$\text{Варіант 8} \quad \mathcal{J}[y] = \int_0^1 y'^2 dx, \quad y(0) = 0, \quad y(1) = e - 3, \quad \int_0^1 ye^x dx = 0$$

$$\text{Варіант 9} \quad \mathcal{J}[y] = \int_0^\pi y \sin x dx, \quad y(0) = 0, \quad y(\pi) = 0, \quad \int_0^\pi y'^2 dx = \frac{\pi}{2}$$

$$\text{Варіант 10} \quad \mathcal{J}[y] = \int_0^1 y'^2 dx, \quad y(0) = 0, \quad y(1) = 2, \quad \int_0^1 xy dx = 1$$

$$\text{Варіант 11} \quad \mathcal{J}[y] = \int_0^1 (y'^2 + y^2) dx, \quad y(0) = 0, \quad y(1) = \frac{1}{e}, \quad \int_0^1 ye^{-x} dx = \frac{1 - 3e^{-2}}{4}$$

$$\text{Варіант 12} \quad \mathcal{J}[y] = \int_0^{\pi/6} (y'^2 - 9y^2) dx, \quad y(0) = 1, \quad y\left(\frac{\pi}{6}\right) = 0, \quad \int_0^{\pi/6} 2y dx = 1$$

$$\text{Варіант 13} \quad \mathcal{J}[y] = \int_0^1 (2xy + y'^2) dx, \quad y(0) = 0, \quad y(1) = 3, \quad \int_0^1 xy dx = 1$$

$$\text{Варіант 14} \quad \mathcal{J}[y] = \int_1^2 xy'^2 dx, \quad y(1) = 0, \quad y(2) = 12, \quad \int_1^2 xy dx = 9$$

$$\text{Варіант 15} \quad \mathcal{J}[y] = \int_0^1 y'^2 dx, \quad y(0) = 0, \quad y(1) = 1, \quad \int_0^1 xy dx = 0$$

$$\text{Варіант 16} \quad \mathcal{J}[y] = \int_0^1 (y'^2 + y^2) dx, \quad y(0) = 0, \quad y(1) = e, \quad \int_0^1 ye^x dx = \frac{e^2 + 1}{4}$$

$$\text{Варіант 17} \quad \mathcal{J}[y] = \int_0^1 y'^2 dx, \quad y(0) = 0, \quad y(1) = 1, \quad \int_0^1 ye^x dx = 0$$

Варіант 18 $\mathcal{J}[y] = \int_0^1 y'^2 dx, \quad y(0) = 2e + 1, \quad y(1) = 2, \quad \int_0^1 y e^{-x} dx = e$

Варіант 19 $\mathcal{J}[y] = \int_0^1 y'^2 dx, \quad y(0) = 1, \quad y(1) = 6, \quad \int_0^1 y dx = 3$

Варіант 20 $\mathcal{J}[y] = \int_0^\pi y \sin x dx, \quad y(0) = 0, \quad y(\pi) = \pi, \quad \int_0^\pi y'^2 dx = \frac{3\pi}{2}$

Варіант 21 $\mathcal{J}[y] = \int_0^1 y'^2 dx, \quad y(0) = 0, \quad y(1) = 2, \quad \int_0^1 y^2 dx = 4$

Варіант 22 $\mathcal{J}[y] = \int_0^\pi y'^2 dx, \quad y(0) = 0, \quad y(\pi) = 1, \quad \int_0^\pi y \sin x dx = 0$

Варіант 23 $\mathcal{J}[y] = \int_0^1 y'^2 dx, \quad y(0) = -4, \quad y(1) = 4, \quad \int_0^1 xy dx = 0$

Варіант 24 $\mathcal{J}[y] = \int_0^1 (x^2 + y'^2) dx, \quad y(0) = 0, \quad y(1) = 0, \quad \int_0^1 y^2 dx = 2$

Варіант 25 $\mathcal{J}[y] = \int_0^\pi y'^2 dx, \quad y(0) = 0, \quad y(\pi) = \pi, \quad \int_0^\pi y \sin x dx = 0$

4.6.3 Варіаційна задача з рухомими межами

Варіант 1 Знайти найкоротшу відстань від точки $(1; 0)$ до кола $(x - 2)^2 + y^2 = 9$.

Варіант 2 Знайти найкоротшу відстань від точки $(-1; 5)$ до кривої $y = \sqrt{x}$.

Варіант 3 Знайти криву, на якій функціонал $\mathcal{J}[y] = \int_0^2 (xy' + y'^2) dx$ може досягати екстремуму, якщо лівий кінець кривої закріплений у точці $(0; 5)$, а правий ковзає по прямій $y = x + 2$.

Варіант 4 Знайти криву, на якій функціонал $\mathcal{J}[y] = \int_0^b (y'^2 + y^2) dx$ може досягати екстремуму, якщо лівий кінець кривої закріплений у початку координат, а правий ковзає

по прямій $y = -x - 1$.

Варіант 5 Знайти криву, на якій функціонал $\mathcal{J}[y] = \int_0^b \frac{\sqrt{1+y'^2}}{y} dx$ може досягати екстремуму, якщо лівий кінець кривої закріплений у точці $(0; 1)$, а правий ковзає по прямій $y = -x - 1$.

Варіант 6 Знайти криву, на якій функціонал $\mathcal{J}[y] = \int_0^1 y'(y' - x) dx$ може досягати екстремуму, якщо лівий та правий кінці кривої ковзають по прямим $y = -x$ і $y = x$ відповідно.

Варіант 7 Знайти криву, на якій функціонал $\mathcal{J}[y] = \int_0^b (y'^2 + y) dx$ може досягати екстремуму, якщо лівий кінець кривої закріплений у початку координат, а правий ковзає по прямій $y = x$.

Варіант 8 Знайти криву, на якій функціонал $\mathcal{J}[y] = \int_0^b \frac{\sqrt{1+y'^2}}{x-1} dx$ може досягати екстремуму, якщо лівий кінець кривої закріплений у початку координат, а правий рухається по колу $(x-5)^2 + (y-1)^2 = 4$.

Варіант 9 Знайти криву, на якій функціонал $\mathcal{J}[y] = \int_0^b y'^3 dx$ може досягати екстремуму, якщо лівий кінець кривої закріплений у початку координат, а правий ковзає по прямій $y = -x + 1$.

Варіант 10 Знайти найкоротшу відстань від точки $(-1; 3)$ до прямої $y = 1 - 3x$.

Варіант 11 Знайти найкоротшу відстань між еліпсом $4x^2 + y^2 = 9$ та прямою $y = x - 5$.

Варіант 12 Знайти найкоротшу відстань між колом $x^2 + y^2 = 1$ та прямою $x + y = 4$.

Варіант 13 Знайти криву, на якій функціонал $\mathcal{J}[y] = \int_0^b \frac{\sqrt{1+y'^2}}{x-3} dx$ може досягати екстремуму, якщо лівий кінець кривої закріплений у початку координат, а правий ковзає по прямій $y = -4x + 4$.

Варіант 14 Знайти найкоротшу відстань між параболою $y = x^2 + 2$ та прямою $y = x$.

Варіант 15 Знайти найкоротшу відстань між колом $x^2 + y^2 = 1$ та прямою $x + y = 4$.

Варіант 16 Знайти криву, на якій функціонал $\mathcal{J}[y] = \int_0^b \frac{\sqrt{1+y'^2}}{y} dx$ може досягати екстремуму, якщо лівий кінець кривої закріплений у точці $(0; 1)$, а правий ковзає по прямій $y = x - 1$.

Варіант 17 Знайти криву, на якій функціонал $\mathcal{J}[y] = \int_0^b y'^2 dx$ може досягати екстремуму, якщо лівий кінець кривої закріплений у початку координат, а правий рухається по кривій, що задана рівнянням $(x - 1)y^2 + 2 = 0$.

Варіант 18 Знайти криву, на якій функціонал $\mathcal{J}[y] = \int_0^b \frac{\sqrt{1+y'^2}}{y} dx$ може досягати екстремуму, якщо лівий кінець кривої закріплений у початку координат, а правий ковзає по прямій $y = x - 10$.

Варіант 19 Знайти криву, на якій функціонал $\mathcal{J}[y] = \int_0^b (y'^2 + y^2) dx$ може досягати екстремуму, якщо лівий кінець кривої закріплений у початку координат, а правий ковзає по прямій $y = 1$.

Варіант 20 Знайти найкоротшу відстань між параболою $y = x^2$ та прямою $y = x - 5$.

Варіант 21 Знайти найкоротшу відстань від точки $(1; 0)$ до еліпса $4x^2 + 9y^2 = 36$.

Варіант 22 Знайти криву, на якій функціонал $\mathcal{J}[y] = \int_0^b y'^2 dx$ може досягати екстремуму, якщо лівий кінець кривої закріплений у початку координат, а правий ковзає по прямій $y = -x - 1$.

Варіант 23 Знайти найкоротшу відстань між кривими $x^2 + y^2 = 1$ та $(x - 10)^2 + y^2 = 4$.

Варіант 24 Знайти найкоротшу відстань між параболою $y = x^2$ та прямою $y = x - 1$.

Варіант 25 Знайти найкоротшу відстань від точки $(0; 0)$ до кривої $y = \frac{1}{x^2}$.

Задача оптимального керування

5.1 Постановка основних задач оптимального керування

Нехай рух деякого об'єкту описується системою

$$\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}, \mathbf{u}, t), \quad t \in T = [t_0; t_1], \quad \mathbf{x}(t_0) = \mathbf{x}^0, \quad (5.1)$$

де $\mathbf{x}$ — n -вектор, $\mathbf{u}$ — r -вектор, t — скаляр. Функція $\mathbf{f}(\mathbf{x}, \mathbf{u}, t)$ неперервна разом з першою похідною $\frac{\partial \mathbf{f}(\mathbf{x}, \mathbf{u}, t)}{\partial \mathbf{x}}$.

Означення 5.1. Припустимим керуванням назовемо r -вимірну кусково-неперервну функцію $\mathbf{u}(t)$, $t \in T$, що приймає значення з заданої множини $\mathcal{U} \subset \mathbb{R}^r$.

Кожному припустимому керуванню $\mathbf{u}(t)$, $t \in T$ відповідає розв'язок $\mathbf{x}(t)$, $t \in T$ системи (5.1). Якість процесу $\mathbf{x}(t)$, $t \in T$ оцінюється критерієм якості $\mathcal{J}(\mathbf{u})$. Вектор $\mathbf{x}(t)$ називають фазовим.

Задачу оптимального керування можна сформулювати так: *серед припустимих керувань знайти те, на якому критерій якості досягає мінімального значення*. Це керування називається оптимальним, а відповідна йому траєкторія — оптимальною траєкторією.

У залежності від критерію якості отримуємо різні задачі оптимального керування.

Якщо $\mathcal{J}(\mathbf{u}) = \varphi(\mathbf{x}(t_1))$, то задача називається задачею термінального керування або Майєра, $\mathcal{J}(\mathbf{u}) = \int_{t_0}^{t_1} g(\mathbf{x}, \mathbf{u}, t) dt$ — задачею Лагранжа. Задача Больца — це задача оптимального керування з критерієм якості $\mathcal{J}(\mathbf{u}) = \varphi(\mathbf{x}(t_1)) + \int_{t_0}^{t_1} g(\mathbf{x}, \mathbf{u}, t) dt$.

Якщо ввести змінну

$$x_{n+1}(t) = \int_{t_0}^t g(x, u, t) dt,$$

що задовольняє рівнянню

$$\dot{x}_{n+1}(t) = g(x, u, t), \quad x_{n+1}(t_0) = 0,$$

то задача Лагранжа зводиться до задачі Майєра:

$$\begin{aligned} \min x_{n+1}(t_1), \\ \dot{x}(t) &= f(x, u, t), \\ \dot{x}_{n+1}(t) &= g(x, u, t), \\ x(t_0) &= x^0, \quad x_{n+1}(t_0) = 0, \quad u \in U. \end{aligned}$$

Аналогічним способом задача Больца зводиться до задачі Майєра. Далі будемо розглядати задачу Майєра.

5.2 Приріст критерію якості на припустимих керуваннях

Якщо керування $u = u(t)$, $t \in T$ є оптимальним, а $\bar{u}(t)$ — довільне припустиме керування, то

$$\Delta \mathcal{J}(u) = \mathcal{J}(\bar{u}(t)) - \mathcal{J}(u(t)) \geq 0. \quad (5.2)$$

Нехай

$$\Delta u(t) = \bar{u}(t) - u(t), \quad \Delta x(t) = \bar{x}(t) - x(t). \quad (5.3)$$

Тут $\bar{x}(t)$ — траєкторія, що відповідає припустимому керуванню $\bar{u}(t)$, а $x(t)$ — оптимальна траєкторія. Далі, $\Delta x(t_0) = \bar{x}(t_0) - x(t_0)$.

Використовуючи (5.3) і розвинення Тейлора, отримаємо

$$\begin{aligned} \Delta \mathcal{J}(u) &= \varphi(\bar{x}(t_1)) - \varphi(x(t_1)) = \varphi(x(t_1) + \Delta x(t_1)) - \varphi(x(t_1)) = \\ &= \frac{\partial \varphi^T(x(t_1))}{\partial x} \Delta x(t_1) + o(\|\Delta x(t_1)\|), \end{aligned} \quad (5.4)$$

де $o(\alpha)$ — величина, що $\lim_{\alpha \rightarrow 0} \frac{o(\alpha)}{\alpha} = 0$.

Запишемо очевидну тотожність для функції $\Delta x(t)$, $t \in T$ і довільної кусково-гладкої

n -вимірної функції $\psi(t)$, $t \in T$:

$$\frac{d}{dt} \psi^\top(t) \Delta \mathbf{x}(t) = \dot{\psi}^\top(t) \Delta \mathbf{x}(t) + \psi^\top(t) \Delta \dot{\mathbf{x}}(t). \quad (5.5)$$

Враховуючи вираз (5.5) і те, що $\Delta \mathbf{x}(t_0) = 0$, отримаємо

$$\psi^\top(t_1) \Delta \mathbf{x}(t_1) = \int_{t_0}^{t_1} [\dot{\psi}^\top(t) \Delta \mathbf{x}(t) + \psi^\top(t) \Delta \dot{\mathbf{x}}(t)] dt. \quad (5.6)$$

Функція $\psi(t)$ довільна, тому нехай

$$\psi(t_1) = -\frac{\partial \varphi(\mathbf{x}(t_1))}{\partial \mathbf{x}}. \quad (5.7)$$

Тоді з (5.4), (5.6) і (5.7) отримаємо, що

$$\Delta \mathcal{J}(\mathbf{u}) = - \int_{t_0}^{t_1} [\dot{\psi}^\top(t) \Delta \mathbf{x}(t) + \psi^\top(t) \Delta \dot{\mathbf{x}}(t)] dt + o(\|\Delta \mathbf{x}(t_1)\|). \quad (5.8)$$

З (5.3) маємо

$$\Delta \dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t) + \Delta \mathbf{x}(t), \mathbf{u}(t) + \Delta \mathbf{u}(t), t) - \mathbf{f}(\mathbf{x}(t), \mathbf{u}(t), t), \quad \Delta \mathbf{x}(t_0) = 0. \quad (5.9)$$

Введемо функцію Гамільтона системи (5.1)

$$\mathcal{H}(\mathbf{x}, \psi, \mathbf{u}, t) = \psi^\top(t) \mathbf{f}(\mathbf{x}, \mathbf{u}, t).$$

Тоді, використовуючи розвинення Тейлора, отримаємо

$$\begin{aligned} \psi^\top(t) \Delta \dot{\mathbf{x}}(t) &= \mathcal{H}(\mathbf{x}(t) + \Delta \mathbf{x}(t), \psi(t), \mathbf{u}(t) + \Delta \mathbf{u}(t), t) - \mathcal{H}(\mathbf{x}(t), \psi(t), \mathbf{u}(t), t) = \\ &= \mathcal{H}(\mathbf{x}(t), \psi(t), \mathbf{u}(t) + \Delta \mathbf{u}(t), t) + \frac{\partial \mathcal{H}^\top(\mathbf{x}(t), \psi(t), \mathbf{u}(t) + \Delta \mathbf{u}(t), t)}{\partial \mathbf{x}} \Delta \mathbf{x}(t) + \\ &\quad + o_1(\|\Delta \mathbf{x}(t)\|) - \mathcal{H}(\mathbf{x}(t), \psi(t), \mathbf{u}(t), t) = \\ &= \Delta_{\bar{\mathbf{u}}} \mathcal{H}(\mathbf{x}(t), \psi(t), \mathbf{u}(t), t) + \frac{\partial \mathcal{H}^\top(\mathbf{x}(t), \psi(t), \bar{\mathbf{u}}, t)}{\partial \mathbf{x}} \Delta \mathbf{x}(t) + o_1(\|\Delta \mathbf{x}(t)\|). \end{aligned}$$

Довизначимо функцію $\psi(t)$, $t \in T$ таким чином:

$$\dot{\psi}(t) = -\frac{\partial \mathcal{H}(\mathbf{x}(t), \psi(t), \mathbf{u}(t), t)}{\partial \mathbf{x}} \quad (5.10)$$

і підкладемо значення $\dot{\psi}(t)$ у (5.8). Отримаємо такий вираз приросту критерію якості:

$$\Delta \mathcal{J}(\mathbf{u}) = - \int_{t_0}^{t_1} \Delta_{\bar{\mathbf{u}}} \mathcal{H}(\mathbf{x}(t), \boldsymbol{\psi}(t), \mathbf{u}(t), t) dt + \eta, \quad (5.11)$$

де

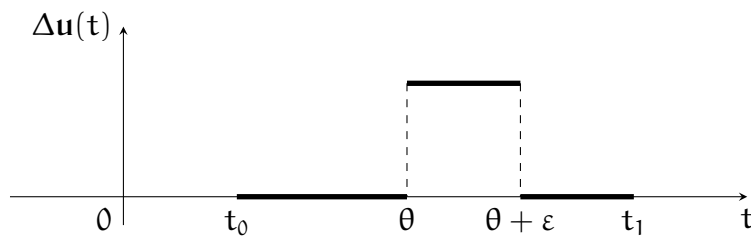
$$\begin{aligned} \eta &= \eta_1 + \eta_2 + \eta_3; \quad \eta_1 = o(\|\Delta \mathbf{x}(t_1)\|), \quad \eta_2 = - \int_{t_0}^{t_1} o_1(\|\Delta \mathbf{x}(t)\|) dt, \\ \eta_3 &= - \int_{t_0}^{t_1} \frac{\partial \Delta_{\bar{\mathbf{u}}} \mathcal{H}^\top(\mathbf{x}(t), \boldsymbol{\psi}(t), \mathbf{u}(t), t)}{\partial \mathbf{x}} \Delta \mathbf{x}(t) dt. \end{aligned} \quad (5.12)$$

Система диференціальних рівнянь (5.10) називається спряженою системою до системи (5.1), вектор $\boldsymbol{\psi}(t)$ — імпульсом або вектором спряжених змінних.

Введемо поняття варіації керування:

$$\Delta \mathbf{u}(t) = \begin{cases} \mathbf{v} - \mathbf{u}(t), & t \in [\theta, \theta + \varepsilon), \\ 0, & t \notin [\theta, \theta + \varepsilon), \end{cases} \quad (5.13)$$

де $\mathbf{v}$ — елемент множини $\mathcal{U}$, θ — точка інтервалу $[t_0; t_1]$, $\varepsilon > 0$ — достатньо мале число таке, що $\theta + \varepsilon < t_1$. Цю варіацію називають голчатою.



Зазначимо, що варіація у варіаційному численні мала на великому відрізку, а у оптимальному керуванні вона велика на малому відрізку.

5.2.1 Оцінка приросту припустимої траєкторії на голчатій варіації

Оцінимо приріст траєкторії $\Delta \mathbf{x}(t)$ на голчатій варіації керування (5.13). Розглянемо три випадки.

- 1) $t \in [t_0, \theta]$. Згідно з рівнянням (5.9) на цьому відрізку $\Delta \mathbf{x}(t) \equiv 0$, так як воно є розв'язком диференціального рівняння

$$\Delta \dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t) + \Delta \mathbf{x}(t), \mathbf{u}(t), t) - \mathbf{f}(\mathbf{x}(t), \mathbf{u}(t), t)$$

з умовою $\Delta \mathbf{x}(t_0) = 0$.

- 2) $t \in [\theta, \theta + \varepsilon]$. На цьому інтервалі функція $\Delta \mathbf{x}(t)$ є розв'язком рівняння

$$\Delta \dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t) + \Delta \mathbf{x}(t), \mathbf{v}(t), t) - \mathbf{f}(\mathbf{x}(t), \mathbf{u}(t), t), \quad \Delta \mathbf{x}(\theta) = 0.$$

За теоремою про інтегральну неперервність розв'язку диференціального рівняння виконується асимптотичне співвідношення $\Delta \mathbf{x}(t) \sim \varepsilon$, $t \in [\theta, \theta + \varepsilon]$.

- 3) $t \in [\theta + \varepsilon, t_1]$. На цьому відрізку $\Delta \mathbf{x}(t)$ задовольняє рівнянню

$$\Delta \dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t) + \Delta \mathbf{x}(t), \mathbf{u}(t), t) - \mathbf{f}(\mathbf{x}(t), \mathbf{u}(t), t), \quad \Delta \mathbf{x}(\theta + \varepsilon) \sim \varepsilon.$$

Використовуючи неперервну залежність розв'язку диференціального рівняння від початкових умов, отримуємо, що $\Delta \mathbf{x}(t) \sim \varepsilon$, $t \in [\theta + \varepsilon, t_1]$.

5.3 Принцип максимуму Понтрягіна

Теорема 5.1 (принцип максимуму Понтрягіна). Нехай $\mathbf{u}^0(t)$, $t \in T$ — оптимальне керування задачі Майєра, $\mathbf{x}^0(t)$, $t \in T$ — оптимальна траєкторія, $\psi^0(t)$, $t \in T$ — розв'язок системи

$$\dot{\psi}(t) = -\frac{\partial \mathcal{H}(\mathbf{x}^0(t), \psi(t), \mathbf{u}^0(t), t)}{\partial \mathbf{x}}, \quad \psi(t_1) = -\frac{\partial \varphi(\mathbf{x}^0(t_1))}{\partial \mathbf{x}}. \quad (5.14)$$

Тоді виконується умова максимуму

$$\mathcal{H}(\mathbf{x}^0(t), \psi^0(t), \mathbf{u}^0(t), t) = \max_{\mathbf{u} \in \mathcal{U}} \mathcal{H}(\mathbf{x}^0(t), \psi^0(t), \mathbf{u}, t), \quad t \in T.$$

Доведення. Оцінимо приріст критерію якості. На голчатій варіації (5.13) приріст визначений формулами (5.11) і (5.12). Згідно з пунктами 1), 2), 3) доданки залишкового члена η мають вигляд

$$\eta_1 = o(\|\Delta \mathbf{x}(t_1)\|) = o_1(\varepsilon), \quad \eta_2 = -\int_{t_0}^{t_1} o_1(\|\Delta \mathbf{x}(t)\|) dt = o_2(\varepsilon),$$

$$\eta_3 = - \int_{\theta}^{\theta+\varepsilon} \frac{\partial \Delta_v \mathcal{H}^\top(\mathbf{x}(t), \psi(t), \mathbf{u}(t), t)}{\partial \mathbf{x}} \Delta \mathbf{x}(t) dt = o_3(\varepsilon).$$

Перший доданок в (5.11), використовуючи формулу Тейлора, можна записати так:

$$\int_{\theta}^{\theta+\varepsilon} \Delta_v \mathcal{H}(\mathbf{x}(t), \psi(t), \mathbf{u}(t), t) dt = \varepsilon \Delta_v \mathcal{H}(\mathbf{x}(\theta), \psi(\theta), \mathbf{u}(\theta), \theta) + \bar{o}(\varepsilon).$$

Остаточно отримаємо

$$\Delta \mathcal{J}(\mathbf{u}) = -\varepsilon \Delta_v \mathcal{H}(\mathbf{x}(\theta), \psi(\theta), \mathbf{u}(\theta), \theta) + \tilde{o}(\varepsilon). \quad (5.15)$$

Нехай $\mathbf{u}(t)$, $t \in T$ — оптимальне керування. Тоді для малих ε таких, що $\theta + \varepsilon < t_1$, на голчатій варіації (5.13) виконується нерівність $\Delta \mathcal{J}(\mathbf{u}) \geq 0$. Це буде виконуватись, враховуючи (5.15), якщо

$$\Delta_v \mathcal{H}(\mathbf{x}(\theta), \psi(\theta), \mathbf{u}(\theta), \theta) \leq 0. \quad (5.16)$$

Дійсно, якщо припустити, що

$$\Delta_v \mathcal{H}(\mathbf{x}(\theta), \psi(\theta), \mathbf{u}(\theta), \theta) = \alpha > 0,$$

то з (5.15) отримаємо

$$\Delta \mathcal{J}(\mathbf{u}) = \varepsilon \left[-\alpha + \frac{\tilde{o}(\varepsilon)}{\varepsilon} \right] < 0$$

при достатньо малому $\varepsilon > 0$, що неможливо.

Нерівність (5.16) отримана для довільного $\mathbf{v}$, $\mathbf{v} \in \mathcal{U}$ і у кожній точці $t = \theta$, $\theta \in [t_0; t_1]$. Тоді нерівність (5.16) можна записати у вигляді

$$\mathcal{H}(\mathbf{x}^0(t), \psi^0(t), \mathbf{u}^0(t), t) = \max_{\mathbf{u} \in \mathcal{U}} \mathcal{H}(\mathbf{x}^0(t), \psi^0(t), \mathbf{u}(t), t), \quad t \in T. \quad (5.17)$$

Теорему доведено. ☑

Означення 5.2. Припустиме керування $\mathbf{u}(t)$, $t \in T$ задовольняє *умові максимуму*, якщо для нього виконується умова (5.17).

Зазначимо, що керування, яке задовольняє умові (5.17), називається екстремальним; воно не обов'язково оптимальне.

Тепер можна сформулювати алгоритм знаходження екстремального керування і екстремальної траєкторії.

- 1) Складаємо функцію Гамільтона $\mathcal{H}(\mathbf{x}, \boldsymbol{\psi}, \mathbf{u}, t) = \boldsymbol{\psi}^\top \mathbf{f}(\mathbf{x}, \mathbf{u}, t)$.
- 2) Екстремальне керування знаходимо з умови максимуму

$$\mathcal{H}(\mathbf{x}(t), \boldsymbol{\psi}(t), \mathbf{u}(t), t) = \max_{\mathbf{u} \in \mathbf{U}} \mathcal{H}(\mathbf{x}(t), \boldsymbol{\psi}(t), \mathbf{u}, t).$$

- 3) Підкладаємо значення екстремального керування у системи (5.1) і (5.14), отримаємо крайову задачу

$$\begin{aligned} \dot{\mathbf{x}} &= \mathbf{f}(\mathbf{x}, \mathbf{u}(\mathbf{x}, \boldsymbol{\psi}, t), t), \quad \mathbf{x}(t_0) = \mathbf{x}^0, \\ \dot{\boldsymbol{\psi}} &= -\frac{\partial \mathbf{f}^\top(\mathbf{x}, \mathbf{u}(\mathbf{x}, \boldsymbol{\psi}, t), t)}{\partial \mathbf{x}} \boldsymbol{\psi}, \quad \boldsymbol{\psi}(t_1) = -\frac{\boldsymbol{\varphi}(\mathbf{x}(t_1))}{\partial \mathbf{x}}. \end{aligned}$$

Якщо ця крайова задача має розв'язок, то оптимальна траєкторія знаходиться серед розв'язків крайової задачі.

Приклад 5.1. Написати крайову задачу принципу максимуму для задачі Майєра

$$\begin{aligned} \min_{|u| \leq 5} (3x_1(4) + 5x_2(4)), \quad t \in [0, 4], \\ \dot{x}_1 = 2x_1 - 3x_2 + 2u, \\ \dot{x}_2 = 5x_1 + 8u, \\ x_1(0) = 6, \quad x_2(0) = 7. \end{aligned}$$

Розв'язування. Складаємо функцію Гамільтона

$$\begin{aligned} \mathcal{H}(\mathbf{x}, \boldsymbol{\psi}, \mathbf{u}, t) &= (2x_1 - 3x_2 + 2u)\psi_1 + (5x_1 + 8u)\psi_2 = \\ &= 2x_1\psi_1 - 3x_2\psi_1 + 5x_1\psi_2 + u(2\psi_1 + 8\psi_2). \end{aligned}$$

Знаходимо екстремальне керування $u^0(t)$:

$$u^0(t) = 5 \operatorname{sign}(2\psi_1(t) + 8\psi_2(t)).$$

Отримаємо таку крайову задачу:

$$\begin{aligned} \dot{x}_1 &= 2x_1 - 3x_2 + 10 \operatorname{sign}(2\psi_1 + 8\psi_2), \\ \dot{x}_2 &= 5x_1 + 40 \operatorname{sign}(2\psi_1 + 8\psi_2), \\ \dot{\psi}_1 &= -2\psi_1 - 5\psi_2, \\ \dot{\psi}_2 &= 3\psi_1, \\ x_1(0) &= 6, \quad x_2(0) = 7, \quad \psi_1(4) = -3, \quad \psi_2(4) = -5. \end{aligned}$$



Розглянемо необхідні умови оптимальності для різних постановок задач оптимального керування.

№ з/п	Постановка задачі					Необхідні умови				
	система	функціо- нал	час	область кінцевих положень	гаміль- тоніан	1	2	3	4	
1	$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{u}, t), \quad \mathbf{x}(t_0) = \mathbf{x}^0$	$\mathcal{J} = \varphi(\mathbf{x}(t_1))$	t_1 фіксоване	E^n	$\mathcal{H}(\mathbf{x}, \psi, \mathbf{u}, t) = \psi^T \mathbf{f}(\mathbf{x}, \mathbf{u}, t)$	$\dot{\mathbf{x}}^0 = \frac{\partial \mathcal{H}}{\partial \psi} \bigg _0, \quad \psi^0 = - \frac{\partial \mathcal{H}}{\partial \mathbf{x}} \bigg _0$	$\mathcal{H}(\mathbf{x}^0, \psi^0, \mathbf{u}^0, t) = \max_{\mathbf{u} \in U} \mathcal{H}(\mathbf{x}^0, \psi^0, \mathbf{u}, t)$		$\psi^0(t_1) = - \frac{\partial \varphi(\mathbf{x}^0(t_1))}{\partial \mathbf{x}}$	
2				$\mathbf{x}^1$					Умови на $\psi(t_1)$ відсутні	
3				$g_i(\mathbf{x}(t_1)) \leq 0, \quad i = \overline{1, m},$ $h_i(\mathbf{x}(t_1)) = 0, \quad i = \overline{1, l}$					$\psi^0(t_1) = - \frac{\partial \varphi(\mathbf{x}^0(t_1))}{\partial \mathbf{x}} -$ $- \sum_{i=1}^m \mu_i \frac{\partial g_i(\mathbf{x}^0(t_1))}{\partial \mathbf{x}} -$ $- \sum_{i=1}^l \gamma_i \frac{\partial h_i(\mathbf{x}^0(t_1))}{\partial \mathbf{x}}$	
4		$\mathcal{J} = \varphi(\mathbf{x}(t_1), t_1)$	t_1 не фіксоване	E^n			$\mathcal{H}(\mathbf{x}^0, \psi^0, \mathbf{u}^0, t) = \max_{\mathbf{u} \in U} \mathcal{H}(\mathbf{x}^0, \psi^0, \mathbf{u}, t)$	$\mathcal{H}(\mathbf{x}^0(t_1), \mathbf{u}^0(t_1), t_1) =$ $= - \frac{\partial \varphi(\mathbf{x}^0(t_1), t_1)}{\partial t}$	$\psi(t_1) = - \frac{\partial \varphi(\mathbf{x}^0(t_1^0), t_1^0)}{\partial \mathbf{x}}$	
5	$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{u}, t), \quad \mathbf{x}(t_0) = \mathbf{x}^0$	$\mathcal{J} = \int_{t_0}^{t_1} L(\mathbf{x}, \mathbf{u}, t) dt$	t_1 фіксоване	E^n	$\mathcal{H}(\mathbf{x}, \psi, \mathbf{u}, t) = \psi^T \mathbf{f}(\mathbf{x}, \mathbf{u}, t) - L(\mathbf{x}, \mathbf{u}, t)$	$\dot{\mathbf{x}}^0 = \frac{\partial \mathcal{H}}{\partial \psi} \bigg _0, \quad \psi^0 = - \frac{\partial \mathcal{H}}{\partial \mathbf{x}} \bigg _0$	$\mathcal{H}(\mathbf{x}^0, \psi^0, \mathbf{u}^0, t) = \max_{\mathbf{u} \in U} \mathcal{H}(\mathbf{x}^0, \psi^0, \mathbf{u}, t)$		$\psi(t_1) = 0$	
6				$\mathbf{x}^1$					Умови на $\psi(t_1)$ відсутні	
7			t_1 не фіксоване	E^n					$\mathcal{H}(\mathbf{x}^0(t_1), \psi^0(t_1),$ $\mathbf{u}^0(t_1), t_1) = 0$	$\psi(t_1) = 0$
8				$\mathbf{x}^1$						

5.4 Оптимізація лінійних систем. Керуваність системи

Далі будемо розглядати лінійну автономну систему $\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u}$, де матриці $\mathbf{A}$ і $\mathbf{B}$ є постійними.

Означення 5.3. Автономний лінійний процес керування

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u}$$

називається **керуванім**, якщо для кожної пари точок $\mathbf{x}^0$ і $\mathbf{x}^1$ з E^n існує припустиме керування $\mathbf{u}(t)$ на деякому скінченному інтервалі $0 \leq t \leq t_1$ таке, що відповідна траєкторія $\mathbf{x}(t)$ переводить систему з точки $\mathbf{x}(0) = \mathbf{x}^0$ у точку $\mathbf{x}(t_1) = \mathbf{x}^1$ за скінченний час.

Сформулюємо критерій керуваності для лінійних автономних систем.

Теорема 5.2. Лінійний автономний процес

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u},$$

де $\mathbf{A}$ — $n \times n$ -матриця, $\mathbf{B}$ — $n \times r$ -матриця, є керуванім тоді і тільки тоді, коли ранг $n \times n$ -матриці

$$[\mathbf{B}, \mathbf{A}\mathbf{B}, \mathbf{A}^2\mathbf{B}, \dots, \mathbf{A}^{n-1}\mathbf{B}]$$

дорівнює n .

Приклад 5.2. Визначити, чи є лінійний автономний процес

$$\dot{x}_1 = x_2,$$

$$\dot{x}_2 = u$$

керуванім.

Розв'язування. Маємо $\mathbf{A} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, $\mathbf{B} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$. Складаємо матрицю

$$[\mathbf{B}, \mathbf{A}\mathbf{B}] = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$$

ранг якої дорівнює $2 = n$. Процес керований.

**Приклад 5.3.** Визначити, чи керована система

$$\dot{x}_1 = x_1 - 2x_2 + 3u_1,$$

$$\dot{x}_2 = 2x_1 + x_2 - u_2.$$

Розв'язування. Маємо $\mathbf{A} = \begin{pmatrix} 1 & -2 \\ 2 & 1 \end{pmatrix}$, $\mathbf{B} = \begin{pmatrix} 3 & 0 \\ 0 & -1 \end{pmatrix}$. Тоді

$$[\mathbf{B}, \mathbf{AB}] = \begin{pmatrix} 3 & 0 & 3 & 2 \\ 0 & -1 & 6 & 1 \end{pmatrix}.$$

Ранг цієї матриці дорівнює 2 і дорівнює n . Система керована. ☑

Існує така властивість керованих систем:

$$\psi^T(t)\mathbf{B} \neq 0, \quad t \geq 0, \quad \text{якщо} \quad \psi(t) \neq 0.$$

5.5 Достатня умова оптимальності

Розглянемо задачу на швидкодію для лінійної автономної системи

$$\begin{aligned} \dot{\mathbf{x}} &= \mathbf{Ax} + \mathbf{Bu}, \\ \mathbf{x}(0) &= \mathbf{x}^0, \quad \mathbf{x}(t_1) = \mathbf{0}, \\ \mathbf{u}(t) &\in \mathcal{U}, \quad t_1 \rightarrow \min, \end{aligned} \tag{5.18}$$

де $\mathbf{A}$ — $n \times n$ -матриця, $\mathbf{B}$ — $n \times r$ -матриця, $\mathbf{x} \in E^n$, $\mathbf{u} \in \mathcal{U} \subset E^r$, $\mathbf{0} \in \text{int } \mathcal{U}$.

Функція Гамільтона має вигляд

$$\mathcal{H}(\mathbf{x}, \psi, \mathbf{u}, t) = \psi^T \mathbf{Ax} + \psi^T \mathbf{Bu}.$$

Спряжена система така:

$$\dot{\psi} = -\mathbf{A}^T \psi. \tag{5.19}$$

Принцип максимуму Понтрягіна стверджує, що знайдеться такий ненульовий розв'язок $\psi^0(t)$ системи (5.19), що

$$\psi^{0T} \mathbf{Bu}^0(t) = \max_{\mathbf{u} \in \mathcal{U}} \psi^{0T} \mathbf{Bu}.$$

Доведемо, що принцип максимуму для лінійних автономних керованих систем є не тільки необхідною, але і достатньою умовою оптимальності. Має місце

Теорема 5.3. Нехай $\mathbf{u}(t)$, $t \in T$ — припустиме керування, вздовж якого виконується умова максимуму

$$\psi^T(t)\mathbf{Bu}(t) = \max_{\mathbf{u} \in \mathcal{U}} \psi(t)\mathbf{Bu}, \quad t \in T,$$

де $\mathbf{x}(t)$, $t \in [0, t_1]$ — розв'язок системи (5.18), що відповідає керуванню $\mathbf{u}(t)$ і задовольняє умові $\mathbf{x}(t_1) = \mathbf{0}$; $\boldsymbol{\psi}(t)$, $t \in T$ — деякий ненульовий розв'язок системи (5.19).

Тоді $\mathbf{u}(t)$, $t \in T$ — оптимальне керування, що переводить траєкторію $\mathbf{x}(t)$ з положення $\mathbf{x}^0$ у початок координат за мінімальний час.

Доведення. Розглянемо тотожність

$$\begin{aligned} \frac{d}{dt} \boldsymbol{\psi}^\top(t) \mathbf{x}(t) &= \dot{\boldsymbol{\psi}}^\top(t) \mathbf{x}(t) + \boldsymbol{\psi}^\top(t) \dot{\mathbf{x}}(t) = \\ &= -\boldsymbol{\psi}^\top(t) \mathbf{A} \mathbf{x}(t) + \boldsymbol{\psi}^\top(t) \mathbf{A} \mathbf{x}(t) + \boldsymbol{\psi}^\top(t) \mathbf{B} \mathbf{u} = \boldsymbol{\psi}^\top(t) \mathbf{B} \mathbf{u}, \end{aligned}$$

яка виконується у всіх точках неперервності керування $\mathbf{u}(t)$, $t \in T$. Далі отримаємо

$$\boldsymbol{\psi}^\top(t) \mathbf{x}(t) - \boldsymbol{\psi}^\top(s) \mathbf{x}(s) = \int_s^t \boldsymbol{\psi}^\top(\tau) \mathbf{B} \mathbf{u}(\tau) d\tau \quad (5.20)$$

для всіх $t \geq s \geq 0$.

Нехай керування $\mathbf{u}(t)$, $t \in [0, t_1]$ не є оптимальним, тобто існує припустиме керування $\bar{\mathbf{u}}(t)$, $t \in [0, \theta]$, яке переводить траєкторію $\bar{\mathbf{x}}(t)$ системи (5.18) з $\mathbf{x}^0$ в $\mathbf{0}$ за час θ , $\theta < t_1$. Тоді, використовуючи (5.20), отримаємо

$$\begin{aligned} \boldsymbol{\psi}^\top(\theta) \mathbf{x}(\theta) &= [\boldsymbol{\psi}^\top(\theta) \mathbf{x}(\theta) - \boldsymbol{\psi}^\top(0) \mathbf{x}(0)] - [\boldsymbol{\psi}^\top(\theta) \bar{\mathbf{x}}(\theta) - \boldsymbol{\psi}^\top(0) \bar{\mathbf{x}}(0)] = \\ &= \int_0^\theta [\boldsymbol{\psi}^\top(\tau) \mathbf{B} \mathbf{u}(\tau) - \boldsymbol{\psi}^\top(\tau) \mathbf{B} \bar{\mathbf{u}}(\tau)] d\tau. \end{aligned} \quad (5.21)$$

Так як керування $\mathbf{u}(t)$, $t \in [0, t_1]$ задовольняє принципу максимуму, то підінтегральний вираз у (5.21) невід'ємний, і тоді

$$\boldsymbol{\psi}^\top(\theta) \mathbf{x}(\theta) \geq 0. \quad (5.22)$$

З іншого боку, так як система керована, то $\boldsymbol{\psi}^\top(\tau) \mathbf{B} \neq 0$ і $\max_{\mathbf{u} \in \mathcal{U}} \boldsymbol{\psi}^\top(\tau) \mathbf{B} \mathbf{u} > 0$ при $\boldsymbol{\psi}^\top(\tau) \mathbf{B} \neq 0$. Тоді отримаємо

$$\boldsymbol{\psi}^\top(\theta) \mathbf{x}(\theta) = \boldsymbol{\psi}^\top(\theta) \mathbf{x}(\theta) - \boldsymbol{\psi}^\top(t_1) \mathbf{x}(t_1) = - \int_\theta^{t_1} \boldsymbol{\psi}^\top(\tau) \mathbf{B} \mathbf{u}(\tau) d\tau < 0.$$

Ця нерівність суперечить (5.22). Теорема доведена. 

5.6 Синтез лінійних систем, оптимальних по швидкодії

Доведена вище теорема дозволяє здійснити синтез оптимальних систем, тобто побудувати оптимальне керування як функцію фазових координат: $\mathbf{u} = \mathbf{u}(\mathbf{x})$.

Така задача досить просто розв'язується для двовимірних систем.

Розглянемо задачу перевodu системи

$$\begin{aligned}\dot{x}_1 &= x_2, \\ \dot{x}_2 &= u\end{aligned}\tag{5.23}$$

з довільної точки $\mathbf{x}^0$ у точку $\mathbf{x}^1 = (1; 1)^\top$ за мінімальний час при умові, що керування задовольняє умові $|u(t)| \leq 1$.

Так як система (5.23) є керованою (див. [приклад 5.2](#)), то принцип максимуму не тільки необхідна, але і достатня умова оптимальності.

Функція Гамільтона має вигляд $\mathcal{H}(\mathbf{x}, \boldsymbol{\psi}, \mathbf{u}, t) = \psi_1 x_2 + \psi_2 u$. Припустимо керування $u = \text{sign } \psi_2$ дає максимум функції Гамільтона. Функції $\psi_1(t)$ і $\psi_2(t)$ є розв'язками системи

$$\dot{\psi}_1 = 0, \quad \dot{\psi}_2 = -\psi_1.$$

Розв'язок системи такий: $\psi_1(t) = C_1$, $\psi_2(t) = -C_1 t + C_2$. Оптимальне керування таке: $u^0(t) = \text{sign } \psi_2 = \text{sign } (-C_1 t + C_2)$.

На проміжку часу, на якому $u^0(t) = +1$, з системи (5.23) отримаємо рівняння

$$\frac{dx_1}{dx_2} = x_2.$$

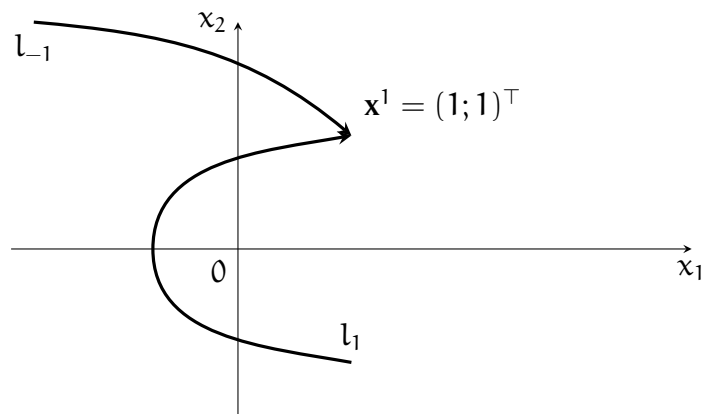
Інтегрування його дає $x_1 = \frac{1}{2}x_2^2 + \alpha_1$.

Отже, при $u^0(t) = +1$ рух точок системи проходить по параболі знизу вгору ($\dot{x}_2 = 1 > 0$).

Якщо проінтегрувати систему при $u^0(t) = -1$, то отримаємо сукупність траєкторій $x_1 = -\frac{1}{2}x_2^2 + \alpha_2$.

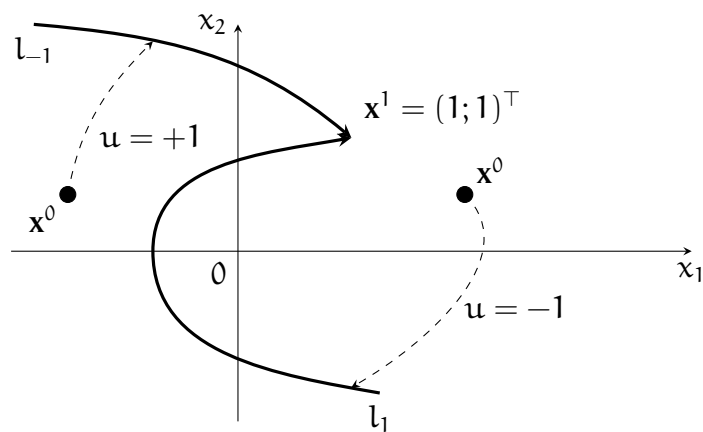
У кожної із сукупностей траєкторій є лише одна траєкторія, що проходить через точку $\mathbf{x}^1 = (1; 1)^\top$. Нас буде цікавити тільки та частина траєкторії, що входить у точку

$\mathbf{x}^1 = (1; 1)^\top$. Ці частини траєкторії позначені l_1 і l_{-1} (мал. 1).



Мал. 1

Якщо початкова точка $\mathbf{x}^0$ знаходиться на l_1 або l_{-1} , то, обираючи керування $+1$ або -1 будемо рухатись у точку $\mathbf{x}^1 = (1; 1)^\top$ за мінімальний час. Крива $l_{-1}\mathbf{x}^1 l_1$ називається лінією перемикання. Якщо початкова точка знаходиться справа від лінії $l_{-1}\mathbf{x}^1 l_1$, то обираємо керування (-1) і рухаємось до лінії l_1 . На ній керування перемикаємо на $(+1)$ і рухаємось по ній до точки $\mathbf{x}^1$ (мал. 2).



Мал. 2

Якщо точка $\mathbf{x}^0$ знаходиться зліва від лінії перемикання, то обираємо керування $(+1)$ і рухаємось до лінії перемикання l_{-1} , перемикаємо керування на (-1) і по кривій рухаємось до точки $\mathbf{x}^1$.

Отже, керування обирається у залежності від положення системи, тобто $\mathbf{u} = \mathbf{u}(\mathbf{x})$.

5.7 Завдання для самостійної роботи

5.7.1 Побудова крайової задачі принципу максимуму

Нехай стан об'єкту описується системою диференціальних рівнянь¹

$$\frac{d\mathbf{x}(t)}{dt} = \mathbf{f}(\mathbf{x}(t), \mathbf{u}(t), t), \quad (5.24)$$

де $t \in [t_0, T]$ — часова змінна, $\mathbf{x}(t) = (x_1(t), x_2(t))^T \in \mathbb{R}^2$ — вектор стану, $\mathbf{u}(t) \in \mathbb{R}^1$ — функція керування.

Розглядається задача оптимального керування заданим об'єктом: знайти мінімум функціоналу

$$\mathcal{J}(\mathbf{u}) \equiv \mathcal{J}(\mathbf{x}(t), \mathbf{u}(t), t) = \varphi(\mathbf{x}(T), T) + \int_{t_0}^T g(\mathbf{x}(t), \mathbf{u}(t), t) dt \quad (5.25)$$

при диференціальних зв'язках (5.24), початкових умовах $\mathbf{x}(0) = \mathbf{x}^0$ та обмеженнях на керування $\mathbf{u}(t) \in \mathcal{U}$, $t \in [t_0, T]$.

Завдання Використовуючи принцип максимуму Понтрягіна, зведіть задану задачу вигляду (5.24)–(5.25) до крайової задачі.

Варіант 1

$$\begin{cases} \dot{x}_1 = u, \\ \dot{x}_2 = -x_1 + 2u; \end{cases} \quad x_1(0) = 2, \quad x_2(0) = 2; \\ |u| \leq 1; \quad t \in [0, 1]; \quad \mathcal{J}(u) = x_2^2(1).$$

Варіант 2

$$\begin{cases} \dot{x}_1 = x_2, \\ \dot{x}_2 = x_1 + u; \end{cases} \quad x_1(0) = 2, \quad x_2(0) = 2; \\ u \in \{-1, 0, 1\}; \quad t \in [0, 4]; \quad \mathcal{J}(u) = \int_0^4 (x_1 + u) dt.$$

¹Похідну за часом будемо далі позначати крапкою: $\frac{d\mathbf{x}(t)}{dt} \equiv \dot{\mathbf{x}}$.

Варіант 3

$$\begin{cases} \dot{x}_1 = x_2, \\ \dot{x}_2 = u; \end{cases} \quad x_1(0) = 0, \quad x_2(0) = 0;$$

$$|u| \leq 1; \quad t \in [0, 1]; \quad \mathcal{J}(u) = \int_0^1 x_1 \, dt.$$

Варіант 4

$$\begin{cases} \dot{x}_1 = 2u, \\ \dot{x}_2 = x_1; \end{cases} \quad x_1(0) = 0, \quad x_2(0) = 2;$$

$$-2 \leq u \leq 1; \quad t \in [0, 2]; \quad \mathcal{J}(u) = \frac{1}{2}x_2^2(2) + \int_0^2 (x_1^2 + u^2) \, dt.$$

Варіант 5

$$\begin{cases} \dot{x}_1 = x_2, \\ \dot{x}_2 = u; \end{cases} \quad x_1(0) = 0, \quad x_2(0) = 0;$$

$$|u| \leq 1; \quad t \in [0, 1]; \quad \mathcal{J}(u) = x_1^2(1).$$

Варіант 6

$$\begin{cases} \dot{x}_1 = x_2, \\ \dot{x}_2 = u; \end{cases} \quad x_1(0) = -1, \quad x_2(0) = -1;$$

$$|u| \leq 1; \quad t \in [0, 1]; \quad \mathcal{J}(u) = x_1^2(1) + x_2^2(1).$$

Варіант 7

$$\begin{cases} \dot{x}_1 = -x_2 + u^2, \\ \dot{x}_2 = x_1 + 2u; \end{cases} \quad x_1(0) = 0, \quad x_2(0) = 1;$$

$$u \in \{-1, 0, 1\}; \quad t \in [0, \pi]; \quad \mathcal{J}(u) = x_2(\pi) + \int_0^\pi u \, dt.$$

Варіант 8

$$\begin{cases} \dot{x}_1 = x_2, \\ \dot{x}_2 = -x_1 + u; \end{cases} \quad x_1(0) = 1, \quad x_2(0) = 1;$$

$$|u| \leq 1; \quad t \in [0, 1]; \quad \mathcal{J}(u) = \int_0^1 u^2 dt.$$

Варіант 9

$$\begin{cases} \dot{x}_1 = -u, \\ \dot{x}_2 = x_1 + u; \end{cases} \quad x_1(0) = 0, \quad x_2(0) = 0;$$

$$0 \leq u \leq 2; \quad t \in [0, 2]; \quad \mathcal{J}(u) = x_2(2).$$

Варіант 10

$$\begin{cases} \dot{x}_1 = x_2, \\ \dot{x}_2 = u; \end{cases} \quad x_1(0) = 1, \quad x_2(0) = 2;$$

$$|u| \leq 1; \quad t \in [0, 1]; \quad \mathcal{J}(u) = x_1^2(1).$$

Варіант 11

$$\begin{cases} \dot{x}_1 = -x_2, \\ \dot{x}_2 = u; \end{cases} \quad x_1(0) = 0, \quad x_2(0) = 0;$$

$$-1 \leq u \leq 4; \quad t \in [0, \pi]; \quad \mathcal{J}(u) = x_1(\pi) + x_2(\pi).$$

Варіант 12

$$\begin{cases} \dot{x}_1 = x_2, \\ \dot{x}_2 = u; \end{cases} \quad x_1(0) = 3, \quad x_2(0) = 0;$$

$$|u| \leq 2; \quad t \in [0, 1]; \quad \mathcal{J}(u) = \int_0^1 (x_1^2 + x_2^2) dt.$$

Варіант 13

$$\begin{cases} \dot{x}_1 = x_2 - u, \\ \dot{x}_2 = x_1 + u; \end{cases} \quad x_1(0) = 2, \quad x_2(0) = 0;$$

$$|u| \leq 2; \quad t \in [0, 3]; \quad \mathcal{J}(u) = -x_2(3) + \int_0^3 (x_1 + x_2 + 2u) dt.$$

Варіант 14

$$\begin{cases} \dot{x}_1 = -2x_2 + u, \\ \dot{x}_2 = x_1; \end{cases} \quad x_1(0) = 1, \quad x_2(0) = 0;$$

$$|u| \leq 1; \quad t \in [0, 1]; \quad \mathcal{J}(u) = \int_0^1 (x_1^2 + x_2^2 + u^2) dt.$$

Варіант 15

$$\begin{cases} \dot{x}_1 = x_2, \\ \dot{x}_2 = x_1 + u^2; \end{cases} \quad x_1(0) = 1, \quad x_2(0) = 0;$$

$$-1 \leq u \leq 2; \quad t \in [0, 1]; \quad \mathcal{J}(u) = x_1(1) + x_2(1).$$

Варіант 16

$$\begin{cases} \dot{x}_1 = x_2, \\ \dot{x}_2 = x_1 + u; \end{cases} \quad x_1(0) = 1, \quad x_2(0) = 1;$$

$$|u| \leq 1; \quad t \in [0, 1]; \quad \mathcal{J}(u) = x_1(1) + x_2(1).$$

Варіант 17

$$\begin{cases} \dot{x}_1 = x_2, \\ \dot{x}_2 = u; \end{cases} \quad x_1(0) = 0, \quad x_2(0) = 0;$$

$$|u| \leq 2; \quad t \in [0, 1]; \quad \mathcal{J}(u) = \int_0^1 x_2 dt.$$

Варіант 18

$$\begin{cases} \dot{x}_1 = x_2 + u, \\ \dot{x}_2 = x_1 + u; \end{cases} \quad x_1(0) = 2, \quad x_2(0) = 2;$$

$$|u| \leq 1; \quad t \in [0, 1]; \quad \mathcal{J}(u) = x_1^2(1) + \int_0^1 u^2 dt.$$

Варіант 19

$$\begin{cases} \dot{x}_1 = x_2, \\ \dot{x}_2 = u; \end{cases} \quad x_1(0) = 0, \quad x_2(0) = 0;$$

$$|u| \leq 2; \quad t \in [0, 1]; \quad \mathcal{J}(u) = \int_0^1 x_1 dt.$$

Варіант 20

$$\begin{cases} \dot{x}_1 = x_2, \\ \dot{x}_2 = x_1 + 2u; \end{cases} \quad x_1(0) = 0, \quad x_2(0) = 1;$$

$$|u| \leq 2; \quad t \in [0, 4]; \quad \mathcal{J}(u) = \int_0^4 (2x_1^2 + u^2) dt.$$

Варіант 21

$$\begin{cases} \dot{x}_1 = x_2 + 2u, \\ \dot{x}_2 = -x_1 + u^2; \end{cases} \quad x_1(0) = 1, \quad x_2(0) = 0;$$

$$u \in \{-1, 0, 2\}; \quad t \in [0, \pi]; \quad \mathcal{J}(u) = x_1(\pi) + \int_0^\pi u dt.$$

Варіант 22

$$\begin{cases} \dot{x}_1 = x_2, \\ \dot{x}_2 = -x_1 + u; \end{cases} \quad x_1(0) = 1, \quad x_2(0) = 1;$$

$$|u| \leq 1; \quad t \in [0, 1]; \quad \mathcal{J}(u) = x_1^2(1) + x_2^2(1).$$

Варіант 23

$$\begin{cases} \dot{x}_1 = -u, \\ \dot{x}_2 = x_1 + 2u; \end{cases} \quad x_1(0) = 0, \quad x_2(0) = 0;$$

$$|u| \leq 1; \quad t \in [0, 1]; \quad \mathcal{J}(u) = x_2(1).$$

Варіант 24

$$\begin{cases} \dot{x}_1 = -u, \\ \dot{x}_2 = x_1 + 3u; \end{cases} \quad x_1(0) = 0, \quad x_2(0) = 0;$$

$$-1 \leq u \leq 2; \quad t \in [0, 2]; \quad \mathcal{J}(u) = x_2(2) + \int_0^2 (4x_1 + u^2) dt.$$

Варіант 25

$$\begin{cases} \dot{x}_1 = -x_2, \\ \dot{x}_2 = x_1 + 2u; \end{cases} \quad x_1(0) = 2, \quad x_2(0) = 1;$$

$$-1 \leq u \leq 2; \quad t \in [0, 4]; \quad \mathcal{J}(u) = x_2(4) + \int_0^4 u \, dt.$$

5.7.2 Задача синтезу оптимального керування

Розглянемо матеріальну точку масою m , яка рухається по прямій без спротиву. Поставимо задачу переводу такої точки із довільного положення x_0 до початку координат за мінімальний час. Початкову швидкість v_0 точки вважаємо відомою. Точка керується деякою зовнішньою силою F , яку можна змінювати у певному діапазоні $[F_1, F_2]$.

Формалізуємо задачу. Із другого закону Ньютона для класичної механіки випливає, що $\vec{F} = m\vec{a}$, де $\vec{a}$ — прискорення точки. Нехай $x(t)$ — координата точки у момент часу t . Тоді її швидкість у момент t є $\dot{x}(t)$, а прискорення — $\ddot{x}(t)$. Перепишемо рівняння руху у скалярній формі:

$$m\ddot{x}(t) = F(t), \quad F(t) \in [F_1, F_2], \quad t \in [0, T] \quad (5.26)$$

при початкових умовах $x(0) = x_0$, $\dot{x}(0) = v_0$. Тут T — час закінчення руху, який потрібно визначити. При цьому $x(T) = 0$, $\dot{x}(T) = 0$.

Перепишемо рівняння (5.26) у вигляді $\ddot{x}(t) = \frac{F(t)}{m}$ та введемо позначення: $x(t) \equiv x_1(t)$, $\dot{x} \equiv x_2(t)$, $u(t) \equiv \frac{1}{m}F(t)$. Отримаємо систему

$$\begin{cases} \dot{x}_1(t) = x_2(t), \\ \dot{x}_2(t) = u(t) \end{cases} \quad (5.27)$$

з початковими умовами

$$x_1(0) = x_0, \quad x_2(0) = v_0 \quad (5.28)$$

та обмеженнями на функцію $u(t)$

$$u(t) \in \left[\frac{F_1}{m}, \frac{F_2}{m} \right]. \quad (5.29)$$

Виходячи з умови задачі, у кінцевий момент часу T координата та швидкість точки

повинні дорівнювати нулю, тобто

$$x_1(T) = 0, \quad x_2(T) = 0, \quad (5.30)$$

а час руху повинен бути мінімальним:

$$T \rightarrow \min. \quad (5.31)$$

Задача (5.27)–(5.31) є задачею оптимального керування. Потрібно знайти керування $u(t)$, що задовольняє обмеженням (5.29), яке переводить об'єкт, що описується системою (5.27), з початкового положення (5.28) у кінцеве положення (5.30) за мінімальний час (5.31).

Завдання Розв'язати задачу вигляду (5.27)–(5.31) при заданих початкових умовах і обмеженнях на керування. Проілюструвати розв'язок графічно.

Варіант 1 $x_1(0) = -1, \quad x_2(0) = 0; \quad -1 \leq u \leq 1$

Варіант 2 $x_1(0) = 3, \quad x_2(0) = 2; \quad -1 \leq u \leq 2$

Варіант 3 $x_1(0) = -1, \quad x_2(0) = 1; \quad -3 \leq u \leq 1$

Варіант 4 $x_1(0) = 1, \quad x_2(0) = 0; \quad -2 \leq u \leq 1$

Варіант 5 $x_1(0) = 1, \quad x_2(0) = 0; \quad |u| \leq \frac{1}{2}$

Варіант 6 $x_1(0) = 1, \quad x_2(0) = 0; \quad -2 \leq u \leq 1$

Варіант 7 $x_1(0) = 1, \quad x_2(0) = 0; \quad -1 \leq u \leq 3$

Варіант 8 $x_1(0) = 1, \quad x_2(0) = 1; \quad -3 \leq u \leq 1$

Варіант 9 $x_1(0) = 1, \quad x_2(0) = 0; \quad |u| \leq 2$

Варіант 10 $x_1(0) = -2, \quad x_2(0) = 0; \quad -1 \leq u \leq 1$

Варіант 11 $x_1(0) = 1, \quad x_2(0) = 3; \quad |u| \leq 2$

Варіант 12 $x_1(0) = -1, \quad x_2(0) = 0; \quad |u| \leq 2$

Варіант 13 $x_1(0) = -1, \quad x_2(0) = 1; \quad |u| \leq 1$

Варіант 14 $x_1(0) = -2, \quad x_2(0) = 0; \quad |u| \leq 2$

Варіант 15 $x_1(0) = -1, \quad x_2(0) = 0; \quad -3 \leq u \leq 2$

Варіант 16 $x_1(0) = 4, \quad x_2(0) = 2; \quad -2 \leq u \leq 1$

Варіант 17 $x_1(0) = 2, \quad x_2(0) = 0; \quad |u| \leq 2$

Варіант 18 $x_1(0) = -2, \quad x_2(0) = 2; \quad -1 \leq u \leq 3$

Варіант 19 $x_1(0) = 1, \quad x_2(0) = 0; \quad |u| \leq \frac{3}{2}$

Варіант 20 $x_1(0) = 2, \quad x_2(0) = 0; \quad |u| \leq 2$

Варіант 21 $x_1(0) = -1, \quad x_2(0) = 2; \quad |u| \leq 1$

Варіант 22 $x_1(0) = -1, \quad x_2(0) = 0; \quad -2 \leq u \leq 1$

Варіант 23 $x_1(0) = 5, \quad x_2(0) = 1; \quad |u| \leq 1$

Варіант 24 $x_1(0) = 2, \quad x_2(0) = 2; \quad -2 \leq u \leq 1$

Варіант 25 $x_1(0) = -2, \quad x_2(0) = 1; \quad -1 \leq u \leq 2$

Рекомендована література

- [1] Бейко И. В., Бублик Б. Н., Зинько П. Н. Методы и алгоритмы. Решение задач оптимизации. — К.: Вища школа, 1983. — 511 с.
- [2] Васильев Ф. П. Численные решения экстремальных задач. — М.: Наука, 1983. — 549 с.
- [3] Габасов Р., Кириллова Ф. М. Методы оптимизации. — Минск: БГУ, 1981. — 350 с.
- [4] Гельфанд И. М., Фомин С. В. Вариационное исчисление. — М.: Физ-мат литература, 1961. — 228 с.
- [5] Ф. Гилл, У. Мюррей, М. Райт. Практическая оптимизация. — М.: Мир, 1985. — 509 с.
- [6] Карманов В. Г. Математическое программирование. — М.: Наука, 1986. — 285 с.
- [7] Кічмаренко О. Д., Осадчий О. К., Яровий А. Т. Методичні вказівки і учбові завдання на тему «Транспортна задача і задача про призначення» для самостійної роботи з дисципліни «Методи оптимізації і дослідження операцій» студентам фаху «Прикладна математика». — Одеса: Астропринт, 2006. — 33 с.
- [8] Моклячук М. П. Варіаційне числення. Екстремальні задачі. — К.: Либідь, 1994. — 327 с.
- [9] Полак Э. Численные методы оптимизации. Единый подход. — М.: Мир, 1974. — 376 с.
- [10] Поляк Б. Т. Введение в оптимизацию. — М.: Наука, 1983. — 384 с.
- [11] Пшеничный Б. Н., Данилин Ю. М. Численные методы в экстремальных задачах. — М.: Физ-мат литература, 1975. — 319 с.
- [12] Сухарев А. Г., Тимохов А. В., Федоров В. В. Курс методов оптимизации. — М.: Нау-

- ка, Физ-мат литература, 1986. — 352 с.
- [13] Химмельблау Д. Прикладное нелинейное программирование. — М.: Мир. 1975. — 534 с.
- [14] Яровой А. Т. Минимизация функций с ограничениями. Лабораторная работа № 7 по дисциплине «Методы оптимизации». — Одесса: ОГУ, 1989. — 70 с.
- [15] Яровий А. Т. Методичні вказівки для самостійної роботи з дисципліни «Методи оптимізації» студентам фаху «Прикладна математика». — Одеса: Астропринт, 2002. — 33 с.
- [16] Яровий А. Т., Бенумерова Т. В., Осадчий О. К. Методичні вказівки для самостійної роботи з дисципліни «Дослідження операцій». Цілочисельне програмування. — Одеса: Астропринт, 2006. — 36 с.
- [17] Яровий А. Т., Бенумерова Т. В. Методичні вказівки для самостійної роботи з дисципліни «Методи оптимізації і дослідження операцій» студентам фаху «Прикладна математика». Тема: «Симплекс-метод і аналіз моделей на чутливість». — Одеса: 2007. — 17 с.
- [18] Яровий А. Т., Бенумерова Т. В. Методичні вказівки для самостійної роботи з дисципліни «Методи оптимізації і дослідження операцій» студентам фаху «Прикладна математика». Параметричне лінійне програмування. — Одеса: 2008. — 26 с.
- [19] Яровий А. Т., Бенумерова Т. В. Методичні вказівки для самостійної роботи з дисципліни «Методи оптимізації». Тема: «Модифікований симплекс-метод». Одеса: Демидов, 2009. — 16 с.
- [20] Яровий А.Т., Бенумерова Т.В. Методичні вказівки по виконанню лабораторної роботи «Розв'язання задач лінійного програмування» з дисципліни «Методи оптимізації». — Одеса: 2010. — 48 с.
- [21] Яровий А. Т., Бенумерова Т. В. Методичні вказівки по виконанню лабораторної роботи «Аналіз моделей на чутливість» з дисципліни «Методи оптимізації». — Одеса: ЧП Демидов, 2011. — 23 с.
- [22] Яровий А. Т. Методичні вказівки для самостійної роботи з дисципліни «Методи оптимізації» студентам фаху «Математика» на тему: «Двоїстість, двоїстий симплекс-метод». — Одеса: Демидов, 2012. — 32 с.

ДЛЯ НОТАТОК

Навчальне видання

ЯРОВИЙ Анатолій Трохимович

СТРАХОВ Євген Михайлович

МЕТОДИ ОПТИМІЗАЦІЇ ТА ВАРІАЦІЙНЕ ЧИСЛЕННЯ

Підписано до друку 11.05.2017 р.

Формат 60×84/16. Папір офсетний.

Ум. друк. арк. 9,07.

Наклад 100 прим.

ФОП «Сухачов»

65023, м. Одеса, вул. Садова, 14

Свідоцтво про внесення до Державного реєстру

суб'єктів видавничої справи ДК № 4362 від 13.07.2012 р.

Тел.: (098) 457-75-00, (048) 706-78-70, (067) 798-05-70, (094) 931-98-70

E-mail: osvita.odessa@gmail.com, www.svitknigi.com

Видавництво «Освіта України»

04136, м. Київ, вул. Маршала Гречка, 13, оф. 218

Свідоцтво про внесення до Державного реєстру

суб'єктів видавничої справи ДК № 1957 від 23.04.2009 р.

Тел./факс (044) 507-04-52, (044) 228-81-29, (063) 237-59-92

E-mail: osvita2005@mail.ru, www.rambook.ru

Видавництво «Освіта України» запрошує авторів до співпраці

з випуску видань, що стосуються питань управління,

модернізації, інноваційних процесів, технологій,

методичних і методологічних аспектів освіти та навчального процесу

у вищих навчальних закладах.

Надаємо всі види видавничих та поліграфічних послуг.

