Національний технічний університет України
«Київський політехнічний інститут»

О.І Лисенко, І.В. Алєксєєва

ДОСЛІДЖЕННЯ ОПЕРАЦІЙ
КОНСПЕКТ ЛЕКЦІЙ

Київ – 2016

Дослідження операцій. Конспект лекцій / Уклад.: О.І. Лисенко, І.В. Алєксєєва, – К: НТУУ «КПИ», 2016. – 196 с.

Гриф надано Вченою радою Навчально-наукового
Інституту телекомунікаційних систем НТУУ «КПІ»
(протокол № 6 від 30.05.2016)

Навчальне видання

Дослідження операцій
Конспект лекцій

для студентів технічних спеціальностей

Укладачі: *Лисенко Олександр Іванович*, д-р техн. наук, професор
Алєксєєва Ірина Віталіївна, канд. фіз.-мат. наук, доцент

Відповідальний редактор: *В.С. Явіся,* канд. техн. наук, доцент

Рецензенти: *О.В. Барабаш*, д-р техн. наук, професор
(Державний університет телекомунікацій)

В.Б. Кисельов, д-р техн. наук, професор
(Академія муніципального управління)

Зміст

Передмова

Методи дослідження операцій, які включають в себе методи математичного моделювання та оптимізації, складають фундамент прикладної математичної підготовки студентів **освітньо-кваліфікаційного рівня бакалавр за напрямком** 6.050903 «Телекомунікації» (спеціальності 7(8).05090302 «Телекомунікаційні системи та мережі, 7(8).05090303 «Технології та засоби телекомунікацій»).

В конспекті лекцій «Дослідження операцій» викладено основні положення теорії детермінованого лінійного та нелінійного програмування; методів чисельного пошуку екстремумів опуклих функцій на опуклих множинах, методів параметричного та цілочисельного програмування, термінологія та визначення основних понять теорії дослідження операцій в обсязі достатнім для застосування отриманих базових знань, вмінь та навичок для практичного використання математичних моделей детермінованих явищ, які мають місце в телекомунікаціях.

Вивчення наведеного у посібнику матеріалу призводить до формування фундаментальних теоретичних знань з математичного моделювання та оптимізації, які використовуються при дослідженні операцій, а також прикладних практичних навиків із застосування системи комп'ютерної математики MATLAB для побудови комп'ютерних математичних моделей та кількісного розв'язання оптимізаційних задач як на попередніх етапах проектування систем, пристроїв та засобів телекомунікацій, так и у реальному часі.

Посібник містить значну кількість прикладів та завдань для самостійного вивчення матеріалу, може бути використаний: для вивчення курсу дослідження операцій: математичне моделювання та оптимізація в телекомунікаціях; курсового та дипломного проектування.

Посібник дозволяє студентам вивчити:

– термінологію, визначення, основні поняття, символічне позначення основних операцій та їх зміст, що використовуються в детермінованих умовах ;
– постановки задач математичного програмування, критерії та методи технічної та економічної оптимізації складних систем, що найбільш поширені в телекомунікаціях;
– чисельні методи в задачах лінійного та нелінійного програмування та формувати прикладні практичні навички фізичної інтерпретації результатів обчислювальних експериментів та їх візуалізації;
– інформаційні технології розв'язання класичних задач лінійного, нелінійного, дискретного(у тому цислі і цілочисельного), програмування та дослідження операцій;
– прийоми формалізації задач структурно-функціонального аналізу, загальної стратегії їх розв'язання та системної оптимізації складних конструктивних елементів в телекомунікаціях.

Лекція 1. Загальна методологія математичного програмування та дослідження операцій

1.1. Основні поняття та визначення. Ефективність операцій

Означення 1.1. *Операція* – будь-яка дія або система дій, що об'єднані єдиним задумом і спрямовані на досягнення визначеної цілі.

Означення 1.2. Конкретну реалізацію послідовності дій у просторі і часі називають *сценарієм операції*.

Означення 1.3. Під назвою *дослідження операцій в телекомунікаціях* розуміють застосування математичних методів для кількісного обґрунтування рішень стосовно розвитку інформаційно-телекомунікаційних технологій та підвищення ефективності їх використання.

Приклади операцій у телекомунікаційних системах:

- система заходів щодо підвищення надійності телекомунікаційних систем і пристроїв;
- система заходів зниження інформаційних ризиків (пошкодження інформації, знищення інформації, несанкціонований доступ до конфіденційної інформації);
- розміщення базових станцій стільникового зв'язку;
- розміщення замовлень на виготовлення апаратури;
- створення угруповання навігаційних та зв'язних супутників Землі;
- створення системи гарантійного обслуговування телекомунікаційних систем та мереж ;
- оптимізація параметрів та структури безпроводових сенсорних мереж.

Означення 1.4. Характеристики дій (фактори операцій), які змінювати в умовах операцій неможливо вважаються *параметрами операцій*.

Означення 1.5. Характеристики дій, за рахунок зміни яких проявляється позитивний результат називають *змінними* або *керуючими параметрами*.

Таким чином, операціями завжди можливо керувати. Будь-який вибір, в залежності від дослідника керуючих параметрів носить назву розв'язок (рішення, план операції чи стратегія).

Розв'язки можуть бути такими, що задовольняють чи не задовольняють технічним та (або) економічним, та (або) екологічним, та (або) соціальним обмеженням, і відповідно носять назву *допустимі* або *недопустимі*. Зрозуміло, що останні відкидають, а серед перших обирають оптимальні, тобто ті, які у деякому розумінні мають перевагу над іншими.

Основна задача дослідження операцій – попереднє кількісне обґрунтування оптимальних рішень. Під прийняттям рішення в технічних системах розуміють остаточне затвердження усіх дій, що складають операцію на основі запропонованого оптимального розв‘язку та із врахуванням додаткової інформації. Існує поняття «особа, що приймає рішення» (ОПР).

ОПР – це фізична особа або колектив фахівців. Для прийняття рішення ОПР використовує спеціальні методи і методики. Методи дослідження операцій складають теоретичну основу процесу підготовки та прийняття рішення. Поряд із основною задачею (обґрунтування оптимальних розв‘язків), при дослідженні операцій досліджують додаткові задачі:

1) порівняльний аналіз різних варіантів організації операцій;
2) оцінка впливу на результат операцій різноманітних змін параметрів операцій;
3) дослідження операцій на предмет виявлення критичних елементів (критичних параметрів операцій), тобто таких складових операцій, порушення запланованого функціонування яких, призводить до суттєвого погіршення остаточних результатів операцій.

Допоміжні задачі набувають особливого значення, коли дану операцію необхідно розглядати не ізольованою, а як елемент цілісної системи операцій. Дослідження цілісної системи операцій базується на так званому системному підході, який вимагає комплексного врахування взаємної залежності та обумовленості складових елементів системи операцій. Для отримання розумних(раціональних) рішень необхідно вміло використовувати як прийоми агрегування, тобто укрупнення операцій, так і їх декомпозиції, тобто розбиття на окремі складові.

Припустимо, що нам вдалося виділити окрему операцію. Метою операції є досягнення її найбільшої ефективності. Під ефективністю операції розуміємо ступінь її налаштованості на виконання задачі, що перед нею стоїть. Чим краще організована операція, тим вона ефективніша.

Означення 1.6. *Критерієм оцінки* або *показником ефективності* або *цільовою функцією* W називається функція фізичної величини чи сукупності фізичних величин, обчислення або вимірювання якої дозволяє кількісно оцінити результат виконання операції.

Порівнюючи показники ефективності для різного складу дій потрібно обрати найкращий спосіб для організації операції. Показник ефективності може бути задано аналітично, алгоритмічно або його можливо вимірювати. Конкретний вигляд показника ефективності, яким слід користуватися при чисельній оцінці ефективності, залежить від спеціальних властивостей конкретної операції, її цільової спрямованості, а також від задачі дослідження. Ця задача може бути поставлена в різних формах: детермінованій або із урахуванням елементів невизначеності.

1.2. Математичні моделі операцій

1.2.1. Загальні відомості про математичні моделі операцій

Загальних способів побудови математичних моделей операцій не існує. В кожному випадку модель будують виходячи із цільової спрямованості операцій і задачі наукового дослідження з урахуванням необхідної точності рішення, а також точністю із якою відомі вихідні дані. При побудові математичної моделі явище (в даному випадку операція) деяким чином спрощується, схематизується. З безлічі факторів, що впливають на явище виділяють порівняно невелику кількість найважливіших, а потім отримана схема опису явища повинна бути представлена в найбільш відповідних для неї математичних термінах (як кажуть, із використанням адекватного математичного апарату). В результаті встановлюються кількісні зв'язки між умовами операції, її параметрами та результатом операції, який оцінюється показником ефективності або сукупністю показників, якщо їх в даній задачі декілька.

Побудова математичної моделі – це мистецтво, яке вдосконалюється із досвідом. В складних випадках, коли немає єдиного рішення щодо математичної моделі операції, використовується прийом, який отримав назву конкуренція моделей. Тобто, одне і те саме явище досліджується із використанням декількох математичних моделей. Після співставлення результатів теоретичних досліджень з експериментальними даними обирається найкраща за критерієм мінімального відхилення від експериментальних даних модель.

Для складних задач дослідження операцій характерною особливістю є корекція математичної моделі після кожного циклу дослідження. Побудова математичної моделі – найбільш важлива і відповідальна частина дослідження, яка вимагає як глибоких математичних знань, так і знань предметної області, тобто суті явища, яке моделюється.

Математичні моделі, які використовуються в задачах дослідження операцій можливо поділити на два класи:

1) аналітичні;

2) алгоритмічні.

Перші встановлюють формульні (аналітичні) залежності між показником ефективності, змінними і параметрами задачі. Зазвичай ці залежності записують у вигляді рівнянь або нерівностей: алгебраїчних, звичайних диференційних або із частковими похідними. За допомогою аналітичних моделей вдається із задовільною точністю описати лише прості операції, де кількість елементів, що взаємодіють порівняно невелика.

В операціях значного масштабу, де взаємодіють багато елементів, що перебувають під дією випадкових факторів, доцільно застосовувати алгоритмічні математичні моделі, які дозволяють виконувати імітаційне моделювання операцій. Суть цього моделювання полягає в тому, що процес розвитку операції імітується на комп'ютері із усіма випадковостями, які цю операцію супроводжують.

Найкращі результати дослідження операцій можливо отримати при сумісному застосуванні аналітичних та імітаційних моделей: проста аналітична модель дозволяє розібратись у основних закономірностях явища, з'ясувати напрямок його розвитку, а подальшого уточнення кількісних значень показника ефективності операції можливо досягти імітаційним моделюванням.

1.2.2. Детерміновані аналітичні моделі операцій

Розглянемо задачу дослідження операції в загальній постановці, тобто без врахування специфічних властивостей конкретної операції та її цілей. Припустимо, що необхідно виконати деяку операцію, тобто керований захід, на остаточний результат якого можливо вплинути обираючи змінні характеристики цієї операції $x_1, x_2, \ldots, x_n$.

Ефективність операції оцінюється за допомогою чисельного показника W. Оптимальним розв'язком задачі вважаються ті значення $x_1, x_2, \ldots, x_n$, при яких W досягає глобального максимуму. В літературі позначають:

$$M = \sup_{X \in G} W,$$

де $X = \left[x_1, \ldots, x_n\right]^T, G -$ область допустимих розв'язків (ОДР).

В тому випадку коли необхідно досягти глобального мінімуму:

$$m = \inf_{X \in G} W.$$

Цю задачу можливо переформулювати в задачу пошуку глобального максимуму, розглядаючи замість показника ефективності W показник ефективності $-W$.

Якщо математична модель побудована, то це означає:

1) відомі фактори операції $\alpha_1, \alpha_2, \ldots, \alpha_m$, на які дослідники впливати не можуть;

2) визначено фізичний зміст і математична формалізація факторів операції $x_1, x_2, \ldots, x_n$, на які дослідники можуть впливати;

3) визначено спосіб отримання чисельного значення показника ефективності операції W, яке буде залежати від $\alpha_1, \alpha_2, \ldots, \alpha_m$ та $x_1, x_2, \ldots, x_n$;

4) з'ясовано ОДР G, тобто умови обмежень, що діють в даній операції

$$F_j\left(x_1, x_2, \ldots, x_n\right)\left\{\le 0, = 0, \ge 0\right\} \qquad j = \overline{1, q}.$$

Загальна детермінована поставка задачі дослідження операції формулюється наступним чином: при заданих умовах (4) знайти такі розв'язки (2), при яких показник ефективності (3) досягає глобального максимуму.

Специфічним для детермінованої задачі дослідження операції, порівняно із задачами пошуку екстремумів гладкої функції (допускається існування похідних та частинних мішаних похідних високих порядків), є наявність обмежень-нерівностей та недиференційовність, в деяких випадках, показника ефективності за елементами розв'язку $x_1, x_2, \ldots, x_n$.

Загальних математичних методів знаходження екстремумів функцій при наявності довільних обмежень не існує, але коли показник ефективності та обмеження мають специфічні властивості, існують спеціальні методи розв'язання таких задач. Якщо показник ефективності W залежить від $x_1, x_2, \ldots, x_n$ лінійно і обмеження (4) мають вид лінійних алгебраїчних рівнянь або нерівностей, то глобальний максимум W знаходимо із використанням спеціального математичного апарату, що носить назву *лінійне програмування*. Якщо обмеження (4) та показник ефективності W мають інші властивості (наприклад, опуклі, сепарабельні, квадратичні), застосовується математичний апарат *опуклого або сепарабельного, або квадратичного програмування.* Якщо операція, за фізичним змістом операції, розділяється на декілька кроків чи станів, а показник ефективності W дорівнює сумі показників W_k, досягнутих на окремих етапах виконання операції, то для пошуку оптимального рішення можливо застосовувати метод *динамічного програмування*.

Зауваження 1.1. 1. Для розв'язання задачі пошуку екстремуму показника ефективності W, зазвичай, використовують чисельні методи, які попередньо перевіряють (тестують) на розв'язанні задач пошуку екстремумів функцій із спеціальними властивостями, наприклад, функції Розенброка (рис.1.1). Ці функції дозволяють з'ясувати досконалість чисельних методів з точки зору уникнення зациклювань та спроможності обходити так звані «складки» та точки стаціонарності.
2. Показник ефективності може мати на ОДР G декілька локальних екстремумів (рис.1.2).

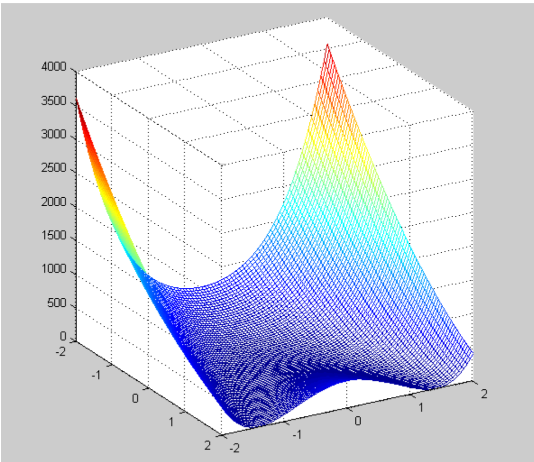


Рис.1.1. Функція Розенброка

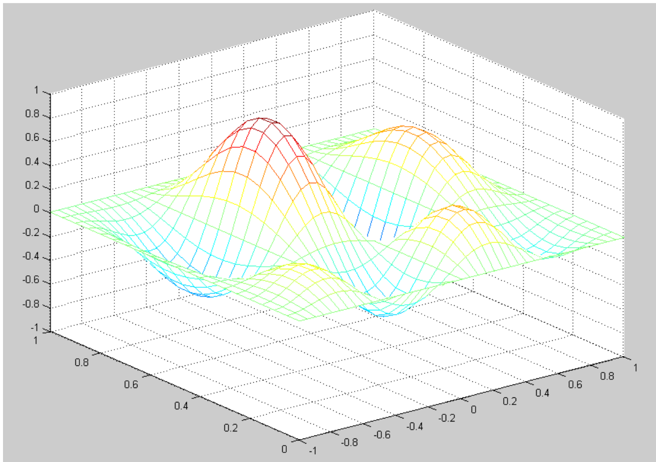


Рис. 1.2. Приклад багатоекстремального показника ефективності

1.2.3. Математичні моделі операцій із врахуванням невизначеності

Типовим для практики є випадок, коли не всі параметри операції, що не може змінювати дослідник, відомі, тобто є такі з них, які включають в себе елемент невизначеності. Наприклад, невизначеність параметрів каналів зв'язку може бути викликана змінами стану атмосфери(зміни погоди або сонячної активності) або дією штучних електромагнітних завад в телекомунікаційних системах спеціального призначення.

В умовах невизначеності до складу математичної моделі операції входять 3-и категорії факторів(параметрів) операції:

1) $\alpha_1,\alpha_2,\ldots,\alpha_m$, які відомі заздалегідь і не можуть бути змінені;

2) $Y_1,Y_2,\ldots,Y_k$, які є невідомими, і впливати на них дослідник не може;

3) $x_1,x_2,\ldots,x_n$, які необхідно обрати для надання операції бажаних властивостей.

Від цих факторів залежить показник ефективності W та (або) функції, що описують ОДР G·

Якби фактори $Y_1,Y_2,\ldots,Y_k$ були відомі, то можливо було б обрати такі фактори $x_1,x_2,\ldots,x_n$, при яких W досягав би глобального максимуму. Але яких саме значень $y_1,y_2,\ldots,y_k$ набудуть $Y_1,Y_2,\ldots,Y_k$ при реалізації операції невідомо і заздалегідь передбачити значення W теж буде неможливо.

Постановка задачі дослідження операції в умовах невизначеності набуває вигляду:

при заданих обмеженнях, фіксованих параметрах $\alpha_1,\alpha_2,\ldots,\alpha_m$, із урахуванням невідомих факторів $Y_1,Y_2,\ldots,Y_k$ знайти такі елементи розв‘язку $x_1,x_2,\ldots,x_n$, які «по можливості» дозволяють показнику ефективності W досягти, як кажуть, «розумного» (раціонального) або вигідного значення.

Методи розв’язання задач дослідження операцій в такій постановці суттєво залежать від:

- природи факторів $Y_1,Y_2,\ldots,Y_k$;
- об’єму відомостей про математичну модель факторів $Y_1,Y_2,\ldots,Y_k$.

Невизначеність $Y_1,Y_2,\ldots,Y_k$ можливо класифікувати наступним чином:

1) $Y_1,Y_2,\ldots,Y_k$ – випадкові величини або випадкові функції із відомими статистичними даними, які дозволяють ідентифікувати їх сумісну функцію розподілу або сумісну щільність ймовірності;

2) $Y_1,Y_2,\ldots,Y_k$ – невідомі фактори, які неможливо вивчати за допомогою статистичних методів, тому що відсутні відповідні статистичні дані, або явища, із якими пов’язані $Y_1, \ldots,Y_k$ взагалі не мають властивості статистичної стійкості;

3) $Y_1,Y_2,\ldots,Y_k$ – невідомі фактори, невизначеність яких пов’язана із діями «супротивника». Така ситуація є характерною для конкурентної боротьби на ринку телекомунікаційних послуг та в задачах захисту інформації.

В першому випадку математична модель задачі модифікується за рахунок зміни вигляду показника ефективності за допомогою одного з двох прийомів. Перший прийом: штучне зведення задачі з невизначеністю до детермінованої задачі дослідження операції шляхом заміни $Y_1,Y_2,\ldots,Y_k$, що входять до складу показника ефективності, на їх відповідні математичні сподівання $M[Y_1],\ldots,M[Y_k]$. Другий прийом: «оптимізація в середньому», коли в якості показника ефективності обирається математичне сподівання:

$$\bar{W} = M[W] = \int \int\limits_{-\infty}^{\infty} \ldots \int W(y_1,\ldots,y_k,x_1,\ldots,x_k) f(y_1,\ldots,y_k) dy_1 \cdot \ldots \cdot dy_k. \quad (1.1)$$

В другому випадку, базуючись на знанні діапазону значень, обчислюються так звані локально-оптимальні рішення: оптимальне рішення для фіксованих значень $y_1,\ldots,y_k$ із відомого діапазону $y_1 \in [y_{min};y_{max}],\ldots,y_k \in [y_{min};y_{max}]$ відповідно, які в подальшому аналізують і знаходять деякі значення $x_1,x_2,\ldots,x_n$, які є компромісним розв‘язком, що не є строго оптимальним ні для однієї з досліджуваних вище умов локальної оптимальності, але є прийнятними для розв’язання задачі в цілому. Таких прийнятних розумних розв‘язків може бути декілька.

В третьому випадку для пошуку розв‘язку використовується так звана теорія ігор – математична теорія конфліктних ситуацій. Вважаючи, що протидіюча нам у конфлікті сторона завжди прагне максимізувати наші збитки, то зрозумілою з нашого боку є стратегія (тобто таке рішення), яка мінімізує максимум нанесених збитків. Кажуть, що має місце так звана мінімаксна стратегія.

1.2.4. Математичні моделі операцій із декількома показниками ефективності

В практиці телекомунікаційних систем достатньо часто зустрічаються випадки, коли ефективність операції оцінюють не по одному, а по декількох показниках ефективності $W_1,W_2,\ldots,W_l$, приміром, надійнісних, вартісних (економічних), технічних. Тобто, математична модель операції в задачі із декількома показниками ефективності, як в детермінованому випадку, так і у випадку з невизначеністю, відрізняється від розглянутих тим, що замість скалярного показника ефективності $W(x_1,\ldots,x_n)$, або $W(Y_1,\ldots,Y_k,x_1,\ldots,x_n)$ використовується векторний показник ефективності:

$$\Omega = [W_1(x_1,\ldots,x_n),\ldots,W_l(x_1,\ldots,x_n)]^T$$

або

$$\Omega = [W_1(Y_1,\ldots,Y_k,x_1,\ldots,x_n),\ldots,W_l(Y_1,\ldots,Y_k,x_1,\ldots,x_n)]^T.$$

Бажаною є постанова задачі дослідження операцій, що вимагає пошуку такого розв‘язку $x_1,\ldots,x_n$, при якому усі складові векторного показника ефективності досягають свого екстремального значення. В загальному випадку такого розв‘язку не існує, але використовуючи методи розв’язання багатокритеріальних задач, або, як ще кажуть, методи оптимізації за векторним критерієм, можливо побудувати процедури вилучення неконкурентноздатних варіантів рішення і тим самим звузити множину розв‘язків до меж, в яких виконується подальший вибір раціонального рішення.

Лекція 2. Структура методики дослідження операцій

2.1. Склад методики дослідження операцій

Складом методики дослідження операцій є наступні заходи:

1) Визначення цілей.
2) Складання плану розробки проекту, операції.
3) Формулювання проблем.
4) Побудова моделі.
5) Розробка обчислювального методу.
6) Розробка технічного завдання на програмування, програмування та відлагодження програми.
7) Збір даних.
8) Перевірка моделі.
9) Реалізація результатів.

Розглянемо їх детальніше.

1 Першочергова ціль будь-якого дослідження операцій полягає в тому, щоб з'ясувати, що очікує отримати керівник операції в результаті її проведення, тобто, які передбачувані результати проведення операції можна очікувати.

Цілі дослідження треба формулювати, виходячи з суті рішення або рішення, на яке орієнтована дана робота. Цілі не треба формулювати ані занадто вузько, ані занадто широко. Неправильне і неточне формулювання цілей може призвести дослідників до розв'язання невірно поставленої задачі.

2 Другий етап дослідження полягає у складанні плану виконання проекту операції, тобто установленню необхідних термінів завершення певних видів робіт. Це – одна з форм контролю за ходом розробки проекту. Як документ, план розробки проекту операції являє собою календарний графік виконання його етапів. Етапи можуть деталізуватися до рівня окремих завдань. Наприклад, етап розробки обчислювального методу може мати такі завдання: розробку методу розв'язання для кожної підмоделі задачі; опис і документальне оформлення методів розв'язання; перевірку запропонованих методів на вибраних задачах невеликої розмірності; внесення уточнень та змін щодо методів розв'язання на підставі результатів пробних розрахунків тощо. При складанні плану треба також приділяти увагу розподілу робіт між окремими виконавцями.

3 Формулювання проблеми – наступний етап дослідження. Він містить не тільки обговорення з керівником операції цілей дослідження, а й збір даних, які надають можливість уявити суть проблеми, що мало місце у минулому, чого треба очікувати в майбутньому, який характер співвідношень між змінними досліджуваної задачі. На основі цих результатів формулюється загальна схема побудови моделі і визначаються напрям усієї наступної роботи.

Одним із питань, які пов'язані з формулюванням проблеми, є визначення того, чи можна усю проблему представити у вигляді окремих підпроблем, щоб паралельно або послідовно дослідити їх незалежно одна від одної (декомпозиція). Друге питання пов'язане з визначенням ступені деталізації моделі, що розробляється. Останнє залежить від обсягів виділених коштів, календарного плану розробки проекту, цілі дослідження.

Наступна фаза стосується галузі застосування та розмірності моделі, що розробляється, визначенню керованих і некерованих змінних, технологічних параметрів операції, показників ефективності, які нададуть змогу оцінити конкретні розв'язки розглянутої проблеми.

4 Четвертий етап дослідження пов'язаний із побудовою моделі. Вона відображає взаємозв'язок між керованими змінними, некерованими змінними, технологічними параметрами і показниками ефективності. Правильно побудована модель – основна умова успішної розробки проекту операції.

Приступаючи до розробки моделі, насамперед, треба з'ясувати питання про можливість використання тих чи інших показників і співвідношень у рамках моделі. Існує декілька різних типів співвідношень, які формують модель: співвідношення, які виходять із визначень, емпіричні співвідношення, нормативні співвідношення. Крім того, треба зібрати та ретельно проаналізувати великий об'єм даних.

На кінцевому етапі побудови моделі досліднику треба дати точне аналітичне або алгоритмічне формулювання досліджуваної проблеми.

5 Разом з роботою з побудови моделі необхідно вибрати або розробити чисельний метод розв'язання. Для цього треба з'ясувати такі моменти:

- чи треба використовувати імітаційне моделювання або будь-який із методів оптимізації:
- чи повинна модель враховувати випадковий характер деяких змінних або ж достатньо використовувати детермінований підхід;
- чи треба враховувати нелінійність певних співвідношень, чи достатньо обмежитися їх лінійною апроксимацією;
- чи можна використовувати існуючі методи розв'язання, або треба розробити новий метод.

Отже, необхідно з'ясувати, які треба зробити припущення та який метод розробити, щоб застосування моделі було практично виправданим у відношенні до використовуваних обчислювальних процедур. Цей етап також містить перевірку запропонованого чисельного методу із використанням модельних(тобто таких, точний розвязок яких відомий) тестових задач, перевіряється можливість використання і коректність розроблених методів розв'язання.

6 Створення програм для ЕОМ у багатьох випадках є складовою частиною дослідження. Розробка технічного завдання на програмування повинна виконуватися ретельно, що дасть змогу забезпечити більш якісне документальне оформлення програм, результатом чого дослідження стає значною мірою орієнтованим на користувача, на задоволення його потреб. Треба звернути увагу на

одну з робіт, які проводяться на цьому етапі, складання вхідних форм та вихідних (документів), їх обговорення та узгодження з керівництвом та виконуючим персоналом.

Вхідні форми – бази даних, які надають можливість забезпечити користувача інформацією, що швидко підготовлюється і легко оновлюється. Наглядні вихідні форми повинні дати користувачу зрозумілу, добре підібрану й зручно розташовану інформацію.

Що стосується саме програмування і відлагодження, то в багатьох випадках проект з дослідження операцій не потребує розробки машинних програм, а за потреби завжди можна використати існуючі програми.

7 На цьому етапі здійснюється збір та аналіз даних, які є необхідними для перевірки правильності моделі та практичного застосування результатів дослідження операцій, тоді як на попередніх етапах збір даних переслідував цілі, що були пов'язані, перш за все, з формулюванням проблеми та побудовою моделі. Тому проблема відсутності даних не є перешкодою до виконання продуктивних операційних досліджень, оскільки математична модель є засобом, який дозволяє обійти труднощі отримання відповідних оцінок шляхом зведення їх до більш простих вимірювань.

Використання моделей, які розробляються при дослідженні операцій, допомагає в процесі прийняття рішень. Розв'язок проблеми зводиться до найпростіших вимірювань, встановленню вихідних змінних і показників ефективності, які є функціями цих змінних. У цьому випадку може бути потрібно більше даних, але отримати їх значно простіше, а вимоги до їх точності будуть менш жорсткими.

8 Етап перевірки моделі містить дві фази: визначення способів перевірки і здійснення самої перевірки. На першій – вибираються аналітичні й експериментальні методи перевірки несуперечності, чутливості, реалістичності та роботоздатності моделі. Для здійснення перевірки моделі будуть необхідні дані, які отримані на попередньому стані. Результати цієї роботи можуть призвести до необхідності перебудови моделі та, відповідно, до складання нових програм.

9 Отримані результати дослідження операцій треба представити рядом робочих процедур, які можна легко зрозуміти і застосувати діючій стороні. Це найважливіший етап, яким завершується операційне дослідження, його можна розглядати як самостійну задачу.

2.2. Типові класи задач дослідження операцій

За своєю змістовною постановкою множину задач дослідження операцій можна розбити на ряд класів, до яких зводиться більшість з них.

Основні класи задач дослідження операцій:

- управління запасами;
- розподілу ресурсів;
- ремонту та заміни обладнання;
- масового обслуговування;

- упорядкування та координації;
- вибору маршруту;
- пошуку;
- змагальні;
- комбіновані.

У дійсності задачі відповідних класів «виникають» одна з одної відповідно до того, як поширюється уявлення про досліджувану операцію.

Надамо стислу характеристику перелічених класів задач.

Задачі управління запасами. Можна виділити, принаймні, чотири основні причини, що призводять до необхідності створення запасів:

- необхідність гарантувати безперебійність виробничого процесу;
- періодичність виробництва окремих матеріальних ресурсів у постачальників;
- особливості транспортування від постачальників до споживачів (невідповідність вантажопідйомності транспортних засобів і розмірів споживання);
- незбіжність ритму виробництва та постачання виробничих ресурсів із ритмом їх споживання.

Задача управління запасами в загальному випадку формулюється так. Існують певні запаси, витрати на зберігання яких є функцією їх величини. Відомі також витрати на доставку ресурсів. Треба визначити оптимальний розмір поставки, частоту та терміни надходження ресурсів, щоб сумарні витрати були мінімальними. Критерієм оптимальності є сума витрат на зберігання і поставку ресурсів.

Задачі управління запасами можна класифікувати:

- за кількістю періодів управління (поповнення запасів) – на одноперіодні та багатоперіодні;
- за характером поповнення запасів – із неперервною системою поповнення запасів (миттєве) і періодичне (із затримкою);
- за урахуванням попиту – на детерміновані і ймовірнісні (статистичні);
- за кількістю типів ресурсів – на однопродуктові і багатопродуктові;
- за видом цільової функції – на задачі з пропорційними та непропорційними втратами.

Задачі розподілу ресурсів пов'язані з розподілом обмежених ресурсів по роботах, які треба виконати. При цьому можуть бути задані як роботи, так і ресурси, або тільки роботи. Перекидання, передавання ресурсів з однієї роботи на іншу, якщо не всі роботи можна виконати найефективніше через відсутність ресурсів, призводить до зменшення спільної ефективності усіх робіт, взятих разом. Тому задачі розподілу ресурсів полягають у відшуканні такого розподілу ресурсів, при якому максимізується спільний прибуток або результат чи мінімізуються витрати. Задачі розподілу ресурсів дуже різноманітні за змістом і багато з них має спеціальну назву: транспортна задача; задача про призначення; задача про суміші; задача вибору оптимальних технологій тощо.

Задачі ремонту та заміни обладнання виникають у тих випадках, коли технічні характеристики працюючого обладнання погіршуються за рахунок старіння, спрацювання та інших причин. Це призводить до необхідності заміни обладнання з метою зменшення сумарних витрат на експлуатацію або попередження нового виходу з ладу. В деяких випадках виникає потреба раціональної організації профілактичного обслуговування, тобто попереджувально-відновлювального ремонту.

Задачі цього класу розподіляють таким чином:

1. За характером заміни обладнання:
   - задача заміни обладнання довгострокового використання;
   - задача заміни обладнання з метою попередження відмов;
   - задачі вибору оптимального плану попереджувального ремонту та профілактичного обслуговування.
2. За характером урахування витрат на обладнання – на дискретні та неперервні.
3. За виходом із ладу обладнання – на детерміновані та випадкові.
4. За стратегією заміни обладнання – на планові та змішані.
5. За часом урахування витрат на обладнання – з приведенням та без приведення витрат більш пізніх років до розрахункового.

Задачі масового обслуговування умовно поділяють на задачі аналізу та задачі синтезу – оптимізації систем масового обслуговування. Задачі аналізу припускають оцінку ефективності функціонування систем при незмінних, наперед відомих вихідних характеристиках системи масового обслуговування: структурі системи; дисципліни обслуговування; потоках вимог і законах розподілу часу їх обслуговування. Задачі синтезу спрямовані на пошук оптимальних параметрів системи масового обслуговування (вибір числа каналів, їх послідовності, включення до роботи, пропускної здатності) і характеристик функціонування (формування вхідного потоку вимог, вибір дисципліни обслуговування тощо).

Задачі упорядкування та координації. Задачі упорядкування пов'язані з визначенням оптимальної послідовності обробки виробів, масивів інформації тощо. Задачі координації відносяться до комплексів операцій, та складаються з певної сукупності окремих операцій, які повинні виконуватися за часом у заданій послідовності. Це – задачі сіткового планування й управління. В цьому класі задач розглядаються співвідношення між строками закінчення комплексу операцій та моментами початку усіх операцій комплексу.

Задачі вибору маршруту зустрічаються при дослідженні різноманітних процесів на транспорті, в системах зв'язку. Типова задача полягає у відшуканні найкращого маршруту, який пов'язує декілька пунктів. На допустимі маршрути може накладатися ряд обмежень, коли забороняється повернення до пройденого пункту або у кожному пункті можна побути тільки одного разу.

Серед цих задач найбільш відомими є: задача вибору найкоротшого шляху між довільними пунктами; задача комівояжера; задача про максимальний потік.

Задачі пошуку складаються у відшуканні найкращого засобу отримання інформації, яка однозначно визначала б розв'язок. Критерієм у такій задачі є мінімум витрат двох видів: вартості отримання інформації й ціни помилки. У першому випадку йдеться про вартість вибірки, тобто вартості вибору спостережень, у другому – про помилки двох видів: помилки вибірки (виявлення того, що насправді є відсутнім) і помилки спостереження (пропускання того, що насправді має місце). Якщо на проведення пошуку виділені фіксовані ресурси, то чим більше буде розмір вибірки, тим менше обсяг ресурсів на кожне спостереження. Таким чином, при бажанні зменшити помилки вибірки, як правило, зростає помилка спостереження і навпаки.

В обмеженій задачі пошуку обсяг ресурсів, які виділені на пошук, є заданим, і задача полягає у розробці плану пошуку, який мінімізує ціну помилки. У загальній задачі кількість ресурсів можна змінювати таким чином, що її метою є мінімізація сумарних витрат ресурсів і ціни помилки.

Змагальні задачі – клас задач дослідження операцій, що виникають під час прийняття рішень в умовах конфліктів, незбігу інтересів осіб. Особливе місце в дослідженні проблем конфлікту займає вибір і порівняльний аналіз можливих (допустимих) способів поведінки сторін, що дає основу для прийняття кожною стороною розумних рішень відносно своїх дій. Особи, які приймають рішення, повинні врахувати не тільки свої цілі, але й цілі, які переслідують інші учасники конфлікту. Відповідну інформацію вдається отримати не завжди, що складає додаткові труднощі як для дослідників, так і осіб, які приймають остаточне рішення.

Комбіновані задачі містять декілька типових задач одночасно.

Лекція 3. Предмет математичного програмування

3.1. Загальна задача математичного програмування

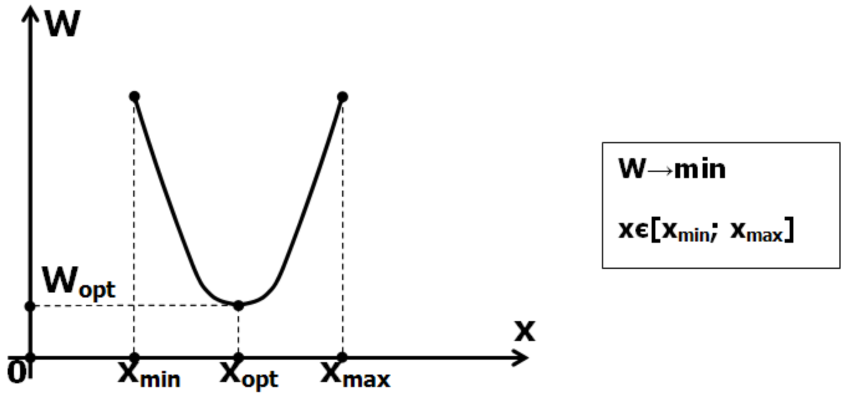
Послідовність розв’язання задачі пошуку найкращих дій за заданим показником ефективності W (задача дослідження операцій) складається з наступних етапів:

1) якісна (неформальна, змістова, вербальна) постановка задачі дослідження операцій;
2) побудова математичної моделі;
3) математична постановка задачі;
4) розробка методу розв’язання задачі (метод оптимізації);
5) розробка методики та алгоритму реалізації запропонованого методу;
6) розробка комп’ютерної програми;
7) розрахунок (чисельний експеримент), імітаційне моделювання;
8) інтерпретація отриманого результату.

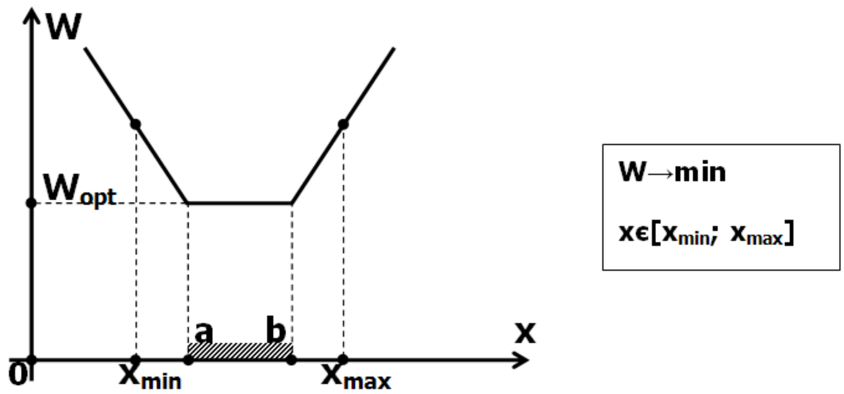
Загальною *задачею математичного програмування* є знаходження глобального екстремуму показника ефективності W на області допустимих значень G. Тобто математичне програмування дозволяє виконати 4) та 5) етапи задачі дослідження операцій.

Окрім терміну математичне програмування, ще використовують термін математична модель оптимізації, або математична модель розв’язання задач на екстремум. Набір змінних $X = [x_1, x_2, \ldots, x_n]^T$, який задовольняє обмеженням G, називають *планом задачі математичного програмування*.

Система обмежень повинна бути сумісною, інакше множина планів буде порожньою множиною. Множина планів може бути як обмеженою, так і не обмеженою. План, що надає показнику ефективності оптимальне значення називається оптимальним, оптимальний розв’язок не завжди єдиний (рис. 3.1).



а) точка оптимуму єдина



б) $x_{opt} \in [a;b]$ – континуум

Рис. 3.1. Графічна ілюстрація можливостей існування єдиного та безлічі оптимальних планів

3.2. Класифікація задач математичного програмування

Задачі математичного програмування поділяються на класичні та некласичні.

Ознаками віднесення задач математичного програмування до класичних є вимоги:

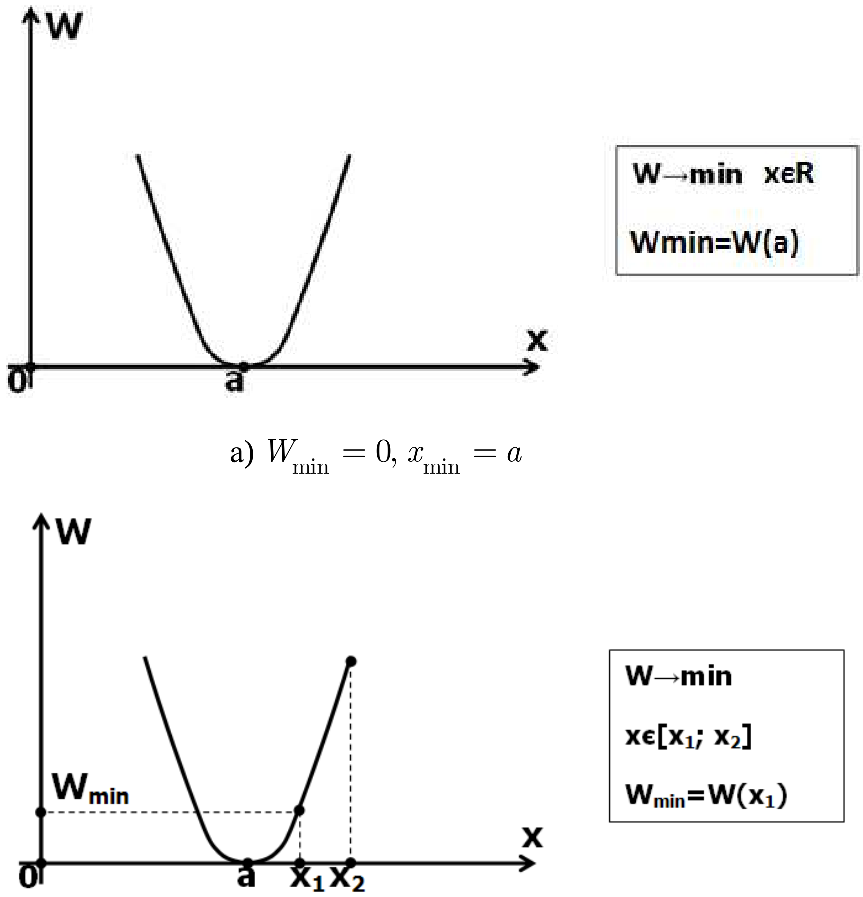
- неперервності показника ефективності та обмежень по відношенню до керуючих змінних;
- існування за керуючими змінними перших та других частинних та мішаних частинних похідних від показника ефективності.

Простою ознакою некласичності даних задач математичного програмування є вимога стосовно дискретності керуючих змінних.

Класичні задачі можливо поділити на 2 підкласи:

1) задачі пошуку безумовного екстремуму;

2) задачі пошуку умовного екстремуму (рис.3.2).



а) $W_{\min} = 0,\ x_{\min} = a$

б) $W_{\min} = W\left(x_1\right),\ x_{\min} = x_1$

Рис. 3.2. Графічна ілюстрація класифікації класичних задач математичного програмування: а) безумовний екстремум; б) умовний екстремум.

Некласичні задачі математичного програмування поділяють на спеціальні та неспеціальні. До основних типів спеціальних задач математичного програмування відносяться:

1) Задача лінійного програмування (ЗЛП).

Загальна постановка ЗЛП має вигляд:

$$W = \sum_{j=1}^{n} c_j x_j \to \max_{X \in G}, \tag{3.1}$$

де $X = [x_1, x_2, \ldots, x_n]^T$, G – область допустимих розв‘язків, яка задається обмеженнями

$$\sum_{j=1}^{n} a_{ij} x_j \le b_i, \quad i = \overline{1, m},\ x_j \in \mathbb{R}, \tag{3.2}$$

$$x_j \ge 0, \quad j = \overline{1, n}. \tag{3.3}$$

В записаних виразах c_j, a_{ij}, b_i $\left(i = \overline{1, m},\ j = \overline{1, n}\right)$ – параметри, x_j $\left(j = \overline{1, n}\right)$ – керуючі змінні.

2) Задача квадратичного програмування (ЗКП).

Загальна постановка ЗКП відрізняється від задачі лінійного програмування формою запису показника ефективності:

$$W = \sum_{j=1}^{n} c_j x_j + \frac{1}{2}\sum_{k=1}^{n}\sum_{j=1}^{n} d_{k_j} x_k x_j \to \max_{X \in G}, \qquad (3.4)$$

де d_{k_j} $\left(k = \overline{1,n},\ j = \overline{1,n}\right)$ параметри.

3) Задача сепарабельного програмування (ЗСП).

Загальна постановка ЗСП відрізняється від задач лінійного програмування формою запису показника ефективності:

$$W = \sum_{j=1}^{n} f_j\left(x_j\right) \to \max_{X \in G}$$ адитивний показник ефективності

або

$$W = \prod_{j=1}^{n} f_j\left(x_j\right) \to \max_{X \in G}$$ мультиплікативний показник ефективності

де $f_j\left(x_j\right)$ $\left(j = \overline{1,n}\right)$ – відомі функції.

4) Задача геометричного програмування (ЗГП).

Загальна постановка ЗГП набуває вигляду:

$$W = \sum_{k=1}^{L} c_k \left(\prod_{j=1}^{n} x_j^{\alpha_{k_j}}\right) \to \max_{X \in G}, \qquad (3.5)$$

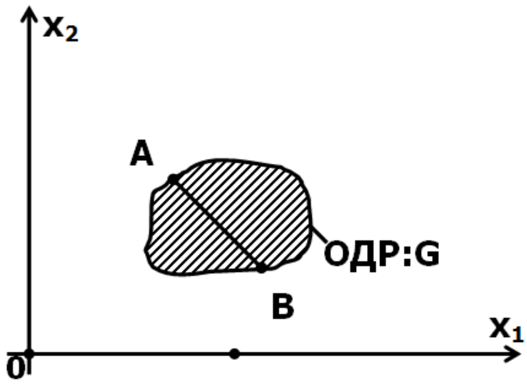
де G – область допустимих розв‘язків, що задається нерівностями:

$$\sum_{k=1}^{L} c_{ik} \left(\prod_{j=1}^{n} x_j^{\alpha_{ik_j}}\right) \le b_i,\ i = \overline{1,m},\ x_j \ge 0,\ j = \overline{1,n}, \qquad (3.6)$$

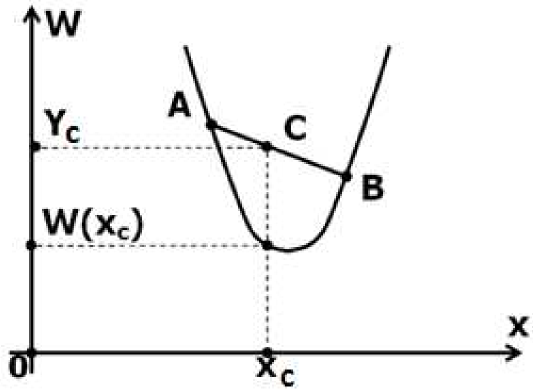
де $c_k, c_{ik}, \alpha_{k_j}, \alpha_{ik_j}, b_i$ $\left(k = \overline{1,L},\ i = \overline{1,m},\ j = \overline{1,n}\right)$ – параметри.

5) Задача опуклого програмування (ЗОП).

Загальна постановка ЗОП полягає в тому, що показник ефективності і обмеження є опуклими функціями.



а) точки А, В належать границі області допустимих розв‘язків (ОДР). Якщо всі точки відрізка АВ належать ОДР при будь якому розташуванні А, В на границі ОДР, то ОДР – опукла.



б) W – показник ефективності, опуклий (вигнутий) донизу. Показник ефективності вважається опуклим донизу, якщо при будь-якому розташуванні точок А та В ордината точки С буде більша за значення показника ефективності в цій точці: $Y_c > W(x_c)$.

Рис. 3.3. Графічна ілюстрація щодо пояснення змісту терміну «опукла область допустимих рішень» та «опуклий показник ефективності»

6) **Задача дискретного програмування (ЗДП),** яка виникає у випадку, якщо будь-яку із попередніх задач доповнити умовою, що хоча б одна із компонент $X = [x_1, x_2, \ldots, x_n]^T$ може приймати окремі значення, наприклад, цілочисельні.

7) **Задача стохастичного програмування (ЗСтП),** яка виникає в будь якому із попередніх випадків, коли всі або деякі, перелічені в пунктах 1, 2, 3, 4, параметри є випадковими величинами.

Всі інші види задач математичного програмування відносяться до неспеціальних, некласичних задач математичного програмування.

3.3. Теорема про достатні умови глобального максимуму

Означення 3.1. В загальній задачі математичного програмування вектор змінних $\hat{X}$ є *точкою глобального максимуму*, якщо він належить області існування планів задачі G і показник ефективності набуває на ньому значення не менше, ніж в будь якій допустимій точці:

$$\hat{X} \in G \text{ та } W(\hat{X}) \geq W(X)\ \forall X \in G.$$

Означення 3.2. Якщо $W(\hat{X}) > W(X)\ \forall X \in G$, то *глобальний максимум* називається *строгим* або *сильним*.

Означення 3.3. Вектор змінних є *точкою локального максимуму*, якщо він належить допустимій множині та на ньому досягається значення показника ефективності більше або рівне значенню показника ефективності в деякому малому околі цього вектора

$$\hat{X} \in G\ ,\ W(\hat{X}) \geq W(X) \quad \forall X \in G \cap M_\varepsilon(\hat{X}),$$

де $M_\varepsilon\left(\hat{X}\right)$ – ε-окіл вектора $\hat{X}$, який в даному випадку є множиною точок, що задовольняють умові

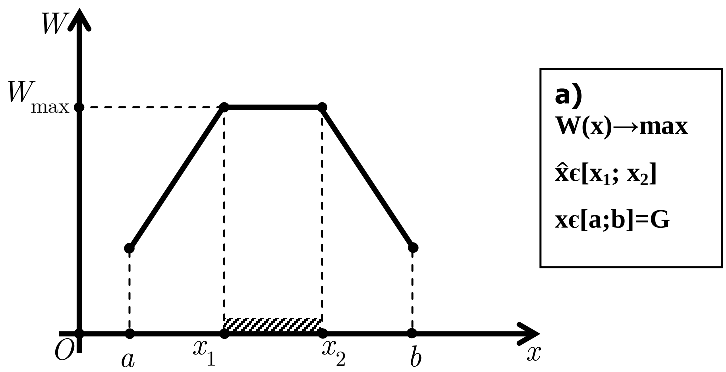
$$\left|X-\hat{X}\right|=\sqrt{\sum_{j=1}^{n}\left(x_j-\hat{x}_j\right)^2}\le\varepsilon .$$

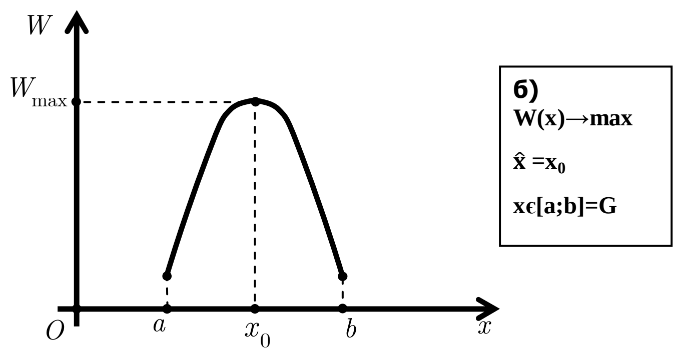
Означення 3.4. Якщо $W\left(\hat{X}\right)>W\left(X\right)\ \forall X\in G\cap M_\varepsilon\left(\hat{X}\right)$, то *локальний максимум* називається *строгим* або *сильним*.

Теорема 3.1. Припустимо, що допустима множина G обмежена і замкнена(границя входить до складу ОДР G), ще й опукла, а неперервний показник ефективності є опуклим догори на G. Тоді локальний максимум є глобальним. Якщо $W\left(X\right)$ строго вигнута догори функція, то розв'язок задачі пошуку максимуму єдиний (існує єдиний глобальний максимум).

Достатні умови існування глобального максимуму формулює теорема Веєрштраса.

Теорема 3.2. Припустимо, що допустима множина G – обмежена, замкнена і опукла, тоді неперервний показник ефективності $W\left(X\right)$ досягає глобальний максимум у внутрішній або граничній точці множини G.





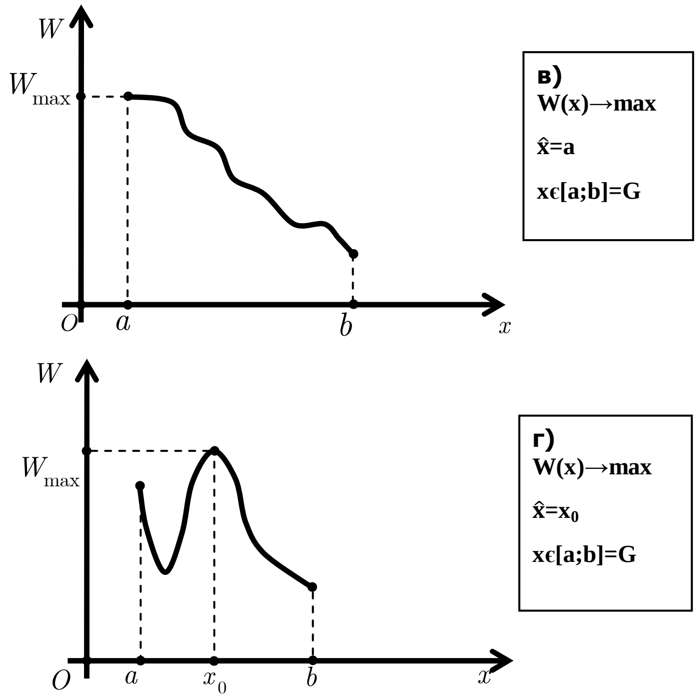


Рис. 3.4. Графічна ілюстрація теоремі про глобальний максимум вигнутої вгору функції (а,б) та теореми Веєрштраса (в,г).

Головною специфічною рисою некласичних задач математичного програмування є те, що вони мають яскраво означену прикладу спрямованість. При математичній постановці задач, які спрямовані на розв‘язання конкретних практичних завдань, зазвичай виникає велика кількість специфічних обмежень, які класична теорія не враховує.

Лекція 4. Типові задачі математичного програмування

4.1. Приклади типових задач математичного програмування

Кожна з наведених нижче задач має свою назву, що відображає її первинне походження, яке в багатьох випадках не має ніякого відношення у прямій постановці до телекомунікаційних систем та засобів телекомунікацій, але прикладна привабливість цих задач для телекомунікацій полягає в тому, що більшість телекомунікаційних задач, які розв’язуються на рівні бізнесуправління, може бути зведена до цих класичних прикладів. Типові приклади виконують як функцію прототипа-аналога так і накопичувача знань щодо об’єктно-орієнтованого способу розв’язання конкретних прикладних телекомунікаційних задач. Зводячи технічні задачі до прикладів, що будуть наведені, ми отримаємо вже відпрацьований ефективний математичний апарат та алгоритмічний інструмент розв’язання своїх специфічних телекомунікаційних задач, тобто ці приклади виконують функцію підказки, що вказують прийоми формалізації специфічних актуальних телекомунікаційних задач.

4.1.1. Задача про розкрій

На фірмі, що спеціалізується на виробництві корпусів для системних блоків, листи металу можуть розкроюватися декількома способами. Якщо лист розкроїти за j-м способом $\left(j=\overline{1,n}\right)$, то отримаємо a_{ij} корпусів i-го виду$\left(i=\overline{1,m}\right)$, при цьому величина відходів з одного листа дорівнює $c_j\left(\text{м}^2\right)$.

Необхідно знайти, яку кількість листів металу необхідно розкроювати за кожним із способів для того, щоб отримати корпусів i-го виду не менше ніж b_i із мінімальною кількістю відходів.

Якщо через x_j позначити кількість листів металу, розкроєних j-м способом, то математична модель та постановка задачі набуває вигляду:

$$W=\sum_{j=1}^{n} c_j x_j \rightarrow \min_{X\in G}, \tag{4.1}$$

де ОДР G являє собою умови стосовно виконання обмежень на виробництво корпусів j-го типу:

$$\sum_{j=1}^{n} a_{ij} x_j \geq b_i \ \left(i=\overline{1,m}\right) \tag{4.2}$$

та на x_j $\left(j=\overline{1,n}\right)$, які випливають із фізичного змісту цих змінних:

$$x_j \geq 0,\ x_j \in \mathbb{N}_0,\ X=\left[x_1,\ldots,x_n\right]^T\ \left(j=\overline{1,n}\right). \tag{4.3}$$

4.1.2. Задача виробничого планування або задача оптимального використання ресурсів виробництва

Для виробництва продукції j-го виду $\left(j=\overline{1,n}\right)$ фірма має обмежені ресурси (виробничі приміщення, спеціалізовані прилади, оргтехніка, витратні матеріали, кількість фахівців, фінансові ресурси тощо) b_i $\left(i=\overline{1,m}\right)$. Витрати ресурсів i-го виду на виготовлення одиниці продукції j-го виду дорівнюють a_{ij}.

Необхідно знайти скільки та якої продукції виробляти, щоб отримати максимальний прибуток. Вважаємо, що збут продукції кожного виду відбувається повністю.

Якщо позначити через x_j, P_j відповідно об'єм продукції j-го виду та прибуток від його реалізації, то математична модель постановки задачі набуває вигляду

$$W=\sum_{j=1}^{n} P_j x_j \rightarrow \max_{X\in G} \tag{4.4}$$

за умови виконання обмежень G по ресурсах:

$$\sum_{j=1}^{n} a_{ij} x_j \leq b_i \ \left(i=\overline{1,m}\right) \tag{4.5}$$

із урахуванням фізичного змісту об’ємів виробництва

$$x_j \geq 0,\ x_j \in \mathbb{N}_0,\ X = \left[x_1,\ldots,x_n\right]^T\ \left(j = \overline{1,n}\right). \tag{4.6}$$

4.1.3. Задачі про суміші або задача про дієту

Сучасний фітотерапевтичний центр виконує лікувальні роботи на основі препаратів природного походження. Кожен вид рослини характеризується відповідним вмістом лікувальних елементів. Відомо вміст i-го елемента в одиниці j-го рослинного препарату a_{ij} $\left(i = \overline{1,m},\ j = \overline{1,n}\right)$, а також вартість c_j одиниці препарату j-го виду і добова мінімальна лікувальна доза i-го препарату b_i.

Необхідно скласти лікувально-профілактичний раціон мінімальної вартості. Якщо позначити через x_j кількість препарату j-го виду, що включено до складу раціону, то математична модель та постановка задачі набувають вигляду:

$$W = \sum_{j=1}^{n} c_j x_j \to \min_{X \in G} \tag{4.7}$$

за умови виконання обмежень G на добову дозу

$$\sum_{j=1}^{n} a_{ij} x_j \geq b_i\ \left(i = \overline{1,m}\right) \tag{4.8}$$

та фізичний зміст кількості препарату

$$x_j \geq 0,\ x_j \in \mathbb{R}\ \left(j = \overline{1,n}\right). \tag{4.9}$$

4.1.4. Задача про призначення

Фірма, що займається гарантійним обслуговуванням та ремонтом комп’ютерної техніки має у своєму розпорядженні n ремонтних бригад. Відома продуктивність c_{ij} кожної i-ої бригади $\left(i = \overline{1,m}\right)$ при виконанні j-ої роботи.

Потрібно так розподілити бригади за роботами, щоб досягти максимальної сумарної виробничої потужності. Зрозуміло, що i-та бригада може виконувати в деякому інтервалі часу лише одну задану роботу.

Таблиця 4.1.

Пояснення щодо можливого робочого завантаження бригад

Номер бригади	Номер роботи							
	1	2	$\cancel{3}$	4	$\cancel{5}$	6	$\cancel{7}$	$\cancel{8}$
1	0	1	0	0	0	0	0	0
2	0	0	0	0	0	1	0	0
3	1	0	0	0	0	0	0	0
4	0	0	0	1	0	0	0	0

Якщо нас цікавить тільки заданий інтервал часу, то немає сенсу розглядати ті роботи, що протягом всього інтервалу часу не виконуються. Або розглядати ті бригади, на які в даному інтервалі часу роботи не вистачило (табл. 4.1), тобто при побудові математичної моделі задачі вважаємо, що кількість робіт, які необхідно виконати дорівнює кількості бригад (всі роботи обов’язково виконуються, всі бригади працюють) (табл. 4.2).

Математичну модель описаного явища можливо представити матрицею (див. табл. 4.2).

Відповідно до змісту задачі про призначення, в окремому стовпці, наприклад третьому, чи в окремому рядку, наприклад другому, може бути лише одна одиниця, яку інтерпретуємо як призначення другої бригади для виконання роботи 3-го типу.

Таблиця 4.2.

Побудова матриці математичної моделі

Номер бригади	Номер роботи j					
	1	2	…	$n-2$	$n-1$	n
1	1	0	…	0	0	0
2	0	1	…	0	0	0
…	…	…	…	…	…	…
$n-2$	0	0	0	0	0	1
$n-1$	0	0	0	0	1	0
n	0	0	0	1	0	0

Остаточно математична модель та постановка задачі набувають вигляду:

$$W = \sum_{i=1}^{n}\sum_{j=1}^{n} c_{ij} x_{ij} \to \max_{X \in G} \quad (4.10)$$

за умови виконання обмежень G:

$$x_{ij} \in \{0,1\},\ \sum_{i=1}^{n} x_{ij} = 1,\ \sum_{j=1}^{n} x_{ij} = 1, \quad (4.11)$$

де $X = \left[x_{11}, x_{12}, \ldots, x_{ij}, \ldots, x_{nn}\right]^T$ – кількість компонент вектора дорівнює n^2.

4.1.5. Задача оптимального розміщення

Крупною телекомунікаційною компанією укладені договори на поставку обладнання замовникам в заданому асортименті, об’ємі та строках. Для виконання договірних зобов’язань керівництво компанії опрацьовує заходи щодо розширення виробництва на існуючих підприємствах за рахунок їх реконструкцій, а також завдяки будівництву нових виробничих приміщень.

Необхідно визначити об’єми виробництва продукції. На існуючих підприємствах і тих, що реконструюються та будуються, а також обсяги поставок продукції від підприємств-постачальників до споживачів і при цьому сумарні витрати на виробництво і доставку продукції повинні бути мінімальними. Побудову математичної моделі почнемо із введення та пояснення фізичного змісту наступних позначень:

i – вид продукції, яка виробляється $\left(i=\overline{1,m}\right)$;

j – номер підприємства, яке виробляє продукцію $\left(j=\overline{1,n}\right)$;

k – номер споживача продукції $\left(k=\overline{1,l}\right)$;

b_{ij} – об'єм продукції i-го виду, який виробляє j-те підприємство протягом заданого часу (виробнича потужність підприємства);

c_{ij} – вартість виробництва продукції i-го виду, виробленої j-им підприємством;

p_{ik} – об'єм поставки продукції i-го виду k-му споживачу згідно договорів;

x_{ij} – об'єм виробництва продукції i-го виду на j-му підприємстві;

x_{ijk} – об'єм поставки продукції i-го виду, виробленого j-им підприємством k-му споживачу;

W – сумарні виробничі та транспортні витрати, тобто показник ефективності.

Математична модель і постановка задачі набувають вигляду:

$$W=\sum_{i=1}^{m}\sum_{j=1}^{n}c_{ij}x_{ij}+\sum_{i=1}^{m}\sum_{j=1}^{n}\sum_{k=1}^{l}c_{ijk}x_{ijk}\to\min_{X\in G} \tag{4.12}$$

за умови виконання обмежень G, які накладаються на виробничі потужності кожного підприємства:

$$\sum_{k=1}^{l}x_{ijk}\le b_{ij}\ \left(i=\overline{1,m},\ j=\overline{1,n}\right), \tag{4.13}$$

на баланс виробництва і споживання продукції:

$$x_{ij}=\sum_{k=1}^{l}x_{ijk}\quad\left(i=\overline{1,m},\ j=\overline{1,n}\right) \tag{4.14}$$

на задоволення запиту споживачів:

$$\sum_{j=1}^{n}x_{ijk}=p_{ik}\quad\left(i=\overline{1,m},\ k=\overline{1,l}\right) \tag{4.15}$$

та невід'ємність об'єму поставок і виробництва продукції:

$$x_{ij}\ge 0,\ x_{ijk}\ge 0\quad\left(i=\overline{1,m},\ j=\overline{1,n},\ k=\overline{1,l}\right), \tag{4.16}$$

$$X = \left[\overbrace{x_{11}, \ldots, x_{ij}, \ldots, x_{mn}}^{m \times n}, \overbrace{x_{111}, \ldots, x_{mnl}}^{m \times n \times l} \right]^T .$$

4.1.6. Задача про розподіл ресурсів

В телекомунікаційній компанії є технічні і людські ресурси $R_1, R_2, \ldots, R_m$ (наприклад, кількість каналів та серверів на їх обслуговування, кількість бригад монтажу та ремонту обладнання, кількість бригад технологічної підтримки абонентів і т.д.) для надання телекомунікаційних послуг у кількості відповідно $b_1, \ldots, b_m$ одиниць. За допомогою цих ресурсів можливо надати послуги $T_1, \ldots, T_n$ (голосовий зв'язок, визначення координат мобільної станції, можливість прийому зображень у реальному часі, можливість забезпечити безпроводовий інтернет і т.д.). Для забезпечення послуги T_j необхідно a_{ij} одиниць ресурсу R_i $\left(i = \overline{1,m};\ j = \overline{1,n}\right)$. Кожна одиниця ресурсу R_i коштує d_i $\left(i = \overline{1,m}\right)$. Кожну послугу T_j можливо реалізувати за ціною c_j $\left(j = \overline{1,n}\right)$. На кожний вид послуги є свій запит: відомо, що ринок телекомунікаційних послуг може скористатись не більше, ніж K_j одиницями послуги T_j $\left(j = \overline{1,n}\right)$. Задача полягає в тому, щоб знайти які послуги і в якій кількості реалізувати для отримання максимального прибутку.

Побудуємо математичну модель задачі. Позначимо:

$x_1, x_2, \ldots, x_n$ – кількісний вираз послуги відповідно $T_1, T_2, \ldots, T_n$, які заплановано «виробити».

Умови запиту та фізичний зміст накладають на x_j $\left(j = \overline{1,n}\right)$ обмеження $0 \le x_j \le K_j$ $\left(j = \overline{1,n}\right)$.

Обмеження виду

$$\sum_{j=1}^{n} a_{ij} x_j \le b_i \ \left(i = \overline{1,m}\right)$$

формалізують вимогу не перевищення наявного запасу відповідного ресурсу.

Обчислимо прибуток L в залежності від елементів рішення $x_1, \ldots, x_n$. Собівартість одиниці послуги T_j дорівнює

$$S_j = a_{1j} d_1 + a_{2j} d_2 + \ldots + a_{nj} d_n = \sum_{i=1}^{m} a_{ij} d_i \ \left(j = \overline{1,n}\right)$$

Чистий прибуток q_j, який буде отримано від реалізації одиниці послуги T_j дорівнює різниці між її ціною для продажу c_j та собівартістю S_j:

$$q_j = c_j - S_j \ \left(j = \overline{1,n}\right).$$

Загальний чистий прибуток від реалізації усіх послуг дорівнює значенню виразу:

$$W = \sum_{j=1}^{n} q_j x_j \ .$$

Математична модель і постановка задачі набувають вигляду:

$$W = \sum_{j=1}^{n} q_j x_j \to \max_{X \in G} \tag{4.17}$$

за умови виконання обмежень G:

$$\begin{aligned} & 0 \le x_j \le K_j \ \left(j = \overline{1,n}\right), \\ & \sum_{j=1}^{n} a_{ij} x_j \le b_i \ \left(i = \overline{1,m}\right), \end{aligned} \tag{4.18}$$

$$X = \left[x_1, \ldots, x_n\right]^T .$$

Цей запис означає: знайти такі невід'ємні та обмежені зверху значення $x_j \ \left(j = \overline{1,n}\right)$, при яких задовольняються ресурсні обмеження та максимізується показник ефективності.

4.1.7. Задача «про перевезення»

Описова постановка задачі полягає в тому, що фірма, яка працює за технологією інтернет-магазину має на всій території України m складських приміщень $c_1, \ldots, c_m$ та n пунктів споживання продукції $\Pi_1, \ldots, \Pi_n$.

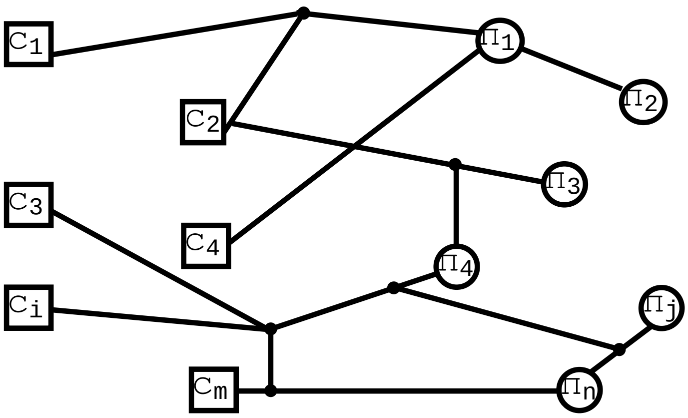


Рис. 4.1. Схема розташування складів $c_i \ \left(i = \overline{1,m}\right)$ та споживачів $\Pi_j \ \left(j = \overline{1,n}\right)$ та мережі транспортних сполучень

Суть задачі полягає у складанні плану перевезень із складів $c_i\left(i=\overline{1,m}\right)$ в пункти споживання Π_j $\left(j=\overline{1,n}\right)$ деякого телекомунікаційного обладнання. На складах c_i $\left(i=\overline{1,m}\right)$ є запаси цього обладнання у кількості a_i одиниць. Користувачі послуг Π_j інтернет-магазину надали заявок відповідно b_j одиниць обладнання. Заявки можливо виконати, якщо сума усіх заявок не перевищує суму усіх запасів $\sum_{j=1}^{n} b_j \leq \sum_{i=1}^{m} a_i$.

Склади пов'язані із пунктами споживання Π_j $\left(j=\overline{1,n}\right)$ мережею доріг із визначеними тарифами перевезення. Вартість перевезення одиниці товару зі складу c_i у пункт споживання Π_j дорівнює v_i $\left(i=\overline{1,m}; j=\overline{1,n}\right)$. Потрібно скласти план перевезень, тобто вказати з якого складу в які пункти і яку кількість обладнання необхідно спрямувати таким чином, щоб заявки були виконані, а загальні витрати на всі перевезення були мінімальні.

Побудуємо математичну модель задачі. Позначимо x_{ij} – кількість одиниць обладнання, яка спрямовується зі складу c_i у пункт Π_j.

Розв'язок (план перевезень) складається із $n \times m$ чисел, які представимо у вигляді прямокутної матриці:

$$P=\begin{bmatrix} x_{11} & x_{12} & \dots & x_{1n} \\ x_{21} & x_{22} & \dots & x_{2n} \\ \dots & \dots & \dots & \dots \\ x_{m1} & x_{m2} & \dots & x_{mn} \end{bmatrix} = \left\{x_{ij}\right\}\Big|_{\substack{i=\overline{1,m} \\ j=\overline{1,n}}}.$$

Потрібно обрати такі невід'ємні значення змінних x_{ij}, що є елементами матриці перевезень P, щоб були виконані наступні умови:

1. Ємність складу перевищувати не можна (це означає, що загальна кількість обладнання, яке було взяте із кожного складу не повинна перевищувати його запасів на цьому складі):
$$\sum_{j=1}^{n} x_{ij} \leq a_i \ \left(i=\overline{1,m}\right).$$
2. Заявки, сформовані споживачами, повинні бути виконаними:
$$\sum_{i=1}^{m} x_{ij} = b_j \ \left(j=\overline{1,n}\right).$$
3. Загальна вартість перевезень повинна обчислюватись за формулою:

$$W = \sum_{i=1}^{m}\sum_{j=1}^{n} v_{ij} x_{ij}.$$

Постановка задачі формулюється наступним чином: знайти такий план перевезень P, щоб вартість W була мінімальною і всі обмеження були виконаними, за умови, коли сума усіх заявок дорівнює сумі усіх запасів

$$\sum_{j=1}^{n} b_j = \sum_{i=1}^{m} a_i,$$

тобто з кожного складу буде вивезено усе, що на ньому є, і нерівність пункту 1 перетворюється у рівність:

$$\sum_{j=1}^{n} x_{ij} = a_i \ \left(i = \overline{1,m}\right).$$

Постановка задачі математичного програмування в цьому випадку набуває вигляду:

$$W = \sum_{i=1}^{m}\sum_{j=1}^{n} v_{ij} x_{ij} \to \min_{X \in G} \tag{4.19}$$

за умови виконання обмежень G:

$$\begin{gathered} x_{ij} \ge 0 \ \left(i = \overline{1,m};\ j = \overline{1,n}\right), \\ \sum_{j=1}^{n} x_{ij} = a_i \ \left(i = \overline{1,m}\right), \\ \sum_{i=1}^{m} x_{ij} = b_j \ \left(j = \overline{1,n}\right), \\ \sum_{j=1}^{n} b_j = \sum_{i=1}^{m} a_i\ . \end{gathered} \tag{4.20}$$

Ця задача ще має назву «транспортна задача із правильним балансом».

4.1.8. Задача «про виробництво складного обладнання»

Протягом часу T планується виробництво складного телекомунікаційного обладнання, кожен компонент якого складається із n елементів $E_1, E_2, \ldots, E_n$. Замовлення на виготовлення цих елементів можливо розмістити на m різних підприємствах $\Pi_1, \Pi_2, \ldots, \Pi_m$. За інтервал часу T на підприємстві Π_i можливо виробити a_{ij} елементів типу E_j $\left(i = \overline{1,m};\ j = \overline{1,n}\right)$.

Зважаючи на те, що нас цікавить лише певний комплект елементів, який дозволяє виконати повну зборку відповідного пристрою, сформулюємо наступну мету операції: необхідно розподілити замовлення по підприємствах таким чином, щоб кількість повних комплектів, виготовлених за час T була максимальною. Розподілити замовлення означає визначити для кожного підприємства $П_i$ $\left(i=\overline{1,m}\right)$ частку часу x_{ij} від загального часу T, яку це підприємство буде працювати виробляючий елемент E_j $(j=\overline{1,n})$.

Враховуючи, що

$$x_{ij}=\frac{\tau_{ij}}{T},$$

де τ_{ij}– час, який повинно витратити підприємство $П_i$ на виробництво елемента E_j, а також той факт, що

$$\sum_{\mathrm{j=1}}^{\mathrm{n}}\tau_{\mathrm{ij}}\leq \mathrm{T}$$

отримаємо обмеження на змінну x_{ij}:

$$\sum_{j=1}^{n}x_{ij}\leq 1\ \left(i=\overline{1,m}\right).$$

Підкреслимо, що при виконанні загальної задачі враховуються лише повні комплекти, що складаються із усіх елементів $E_1,\ldots,E_n$. Тому визначимо кількість повних комплектів, які зможуть виробляти усі підприємства разом за час T. Зрозуміло, що загальна кількість елементів типу E_j обчислюється за виразом:

$$N_j=\sum_{i=1}^{m}a_{ij}x_{ij}\ \left(j=\overline{1,n}\right).$$

Показник ефективності – це кількість повних комплектів, що дорівнює кількості елементів, яких вироблено найменше:

$$W=\min_{j\in\left\{1,\ldots,n\right\}}\sum_{i=1}^{m}a_{ij}x_{ij}.$$

Математична постановка задачі набуває вигляду:

$$W=\min_{j\in\left\{1,\ldots,n\right\}}\sum_{i=1}^{m}a_{ij}x_{ij}\to\max_{X\in G}. \tag{4.21}$$

при виконанні обмежень:

$$\begin{gathered}\sum_{j=1}^{n}x_{ij}\leq 1\ \left(i=\overline{1,m}\right),\\ x_{ij}\geq 0\ \left(i=\overline{1,m};\ j=\overline{1,n}\right),\end{gathered} \tag{4.22}$$

де $X=\left[x_{11},\ldots,x_{ij},\ldots,x_{mn}\right]^T$.

Із наведених прикладів видно, що більшість задач сформульовані із використанням лише лінійних функцій від елементів рішення як у показнику ефективності, так і в обмеженнях. Нагадаємо, що за цими елементами необхідно було знайти максимум або мінімум показника ефективності. Такі задачі в математичному програмуванні виділені в окремий клас задач, що отримав назву **задачі лінійного програмування**. Перейдемо до їх детального вивчення.

Лекція 5. Задачі лінійного програмування

5.1. Поняття про лінійне програмування

Найбільш розвинутим розділом математичного програмування є лінійне програмування.

Лінійне програмування вивчає важливу для практики задачу відшукання максимуму (мінімуму) лінійної функції при наявності обмежень у вигляді лінійних нерівностей або рівнянь.

Оскільки прикладна сторона математичного програмування – це оптимізація рішень взагалі і в техніці зокрема, то лінійне програмування ще називають лінійною оптимізацією. Алгоритми і програми створені на базі методів лінійного програмування і представлені у персональному комп’ютері, у зручній для їх використання формі, утворюють так звані системи комп’ютерної математики. Ці системи є невід’ємними елементами сучасних інформаційних технологій пошуку оптимального розв‘язку. Перейдемо до розгляду ключового елементу цієї технології, а саме методів розв’язання задач лінійного програмування.

5.2. Форми запису задачі лінійного програмування

Загальна задача лінійного програмування полягає в пошуку максимуму (мінімуму) лінійної функції.

$$W = c_1x_1 + c_2x_2 + \ldots + c_nx_n = \sum_{j=1}^{n} c_jx_j \to \max_{X\in G} (\min) \qquad (5.1)$$

за умови використання обмежень:

$$a_{i1}x_1 + \ldots + a_{in}x_n = \sum_{j=1}^{n} a_jx_j \le b_i \quad \left(i = \overline{1,k}\right), \qquad (5.2)$$

$$a_{i1}x_1 + \ldots + a_{in}x_n = \sum_{j=1}^{n} a_jx_j = b_i \quad \left(i = \overline{k+1,m}\right), \qquad (5.3)$$

$$x_j \ge 0 \quad \left(i = \overline{1,n}\right), \qquad (5.4)$$

які утворюють область допустимих розв‘язків G, якій повинен задовольняти вектор керуючих змінних $X = \left[x_1,\ldots,x_n\right]^T$.

Нагадаємо, що функція W називається показником ефективності або цільовою функцією, або критерієм оптимальності, або лінійною формою. Сукупність значень невідомих керуючих змінних $X=\left[x_1,\ldots,x_n\right]^T$, що задовольняють умовам (5.2), (5.3), (5.4) задачі (5.1) – (5.4) має назву розв‘язку. Розв‘язок $X=\left[x_1,\ldots,x_n\right]^T$ називається оптимальним, якщо він максимізує (мінімізує) значення показника ефективності.

Якщо в задачі (5.1) – (5.4) виключити обмеження (5.3), то задача лінійного програмування буде називатися *симетричною задачею лінійного програмування*.

Якщо в задачі (5.1) – (5.4) виключити обмеження (5.2), то отримана задача лінійного програмування буде називатися *канонічною* або *основною задачею лінійного програмування*.

Виконуючи формальні перетворення,можливо перейти від симетричної задачі лінійного програмування, до основної і навпаки. Перехід від нерівності до рівності виконується завдяки використання допоміжних додатних змінних:

$$a_{i1}x_1+\ldots+a_{in}x_n\le b_i \Leftrightarrow \begin{cases} a_{i1}x_1+\ldots+a_{in}x_n+y_i=b_i, \left(i=\overline{1,k}\right) \\ y_i\ge 0, \end{cases}$$

або

$$a_{i1}x_1+\ldots+a_{in}x_n\ge b_i \Leftrightarrow \begin{cases} a_{i1}x_1+\ldots+a_{in}x_n-y_i=b_i, \left(i=\overline{1,k}\right) \\ y_i\ge 0. \end{cases}$$

Перехід від рівняння до нерівності виконується шляхом заміни одного рівняння двома нерівностями.

$$a_{i1}x_1+\ldots+a_{in}x_n=b_i \Leftrightarrow \begin{cases} a_{i1}x_1+\ldots+a_{in}x_n\ge b_i, \\ a_{i1}x_1+\ldots+a_{in}x_n\le b_i \left(i=\overline{k+1,m}\right). \end{cases}$$

Підкреслимо, що у випадку, коли за фізичним змістом задачі (5.1) – (5.4) деяка змінна x_s може мати будь який знак, то необхідно ввести в дослідження дві допоміжні невід’ємні змінні і цю змінну замінити їх різницею:

$$x_s\in\mathbb{R} \Leftrightarrow x_s=u_s-v_s, \left(u_s\ge 0\ , v_s\ge 0\right).$$

5.3. Геометрична інтерпретація задач лінійного програмування

Поставимо задачу лінійного програмування у вигляді

$$W=c_1x_1+c_2x_2+\ldots+c_nx_n \to \max_{X\in G}(\min) \tag{5.5}$$

за умови виконання обмежень G:

$$\begin{cases} a_{11}x_1 + \ldots + a_{1n}x_n \le b_1, \\ \ldots \\ a_{m1}x_1 + \ldots + a_{mn}x_n \le b_m,\ x_j \ge 0\ \left(j = \overline{1,n}\right) \end{cases} \tag{5.6}$$

Кожна із нерівностей системи (5.6) в Евклідовому просторі є півпростором із граничними гіперплощинами

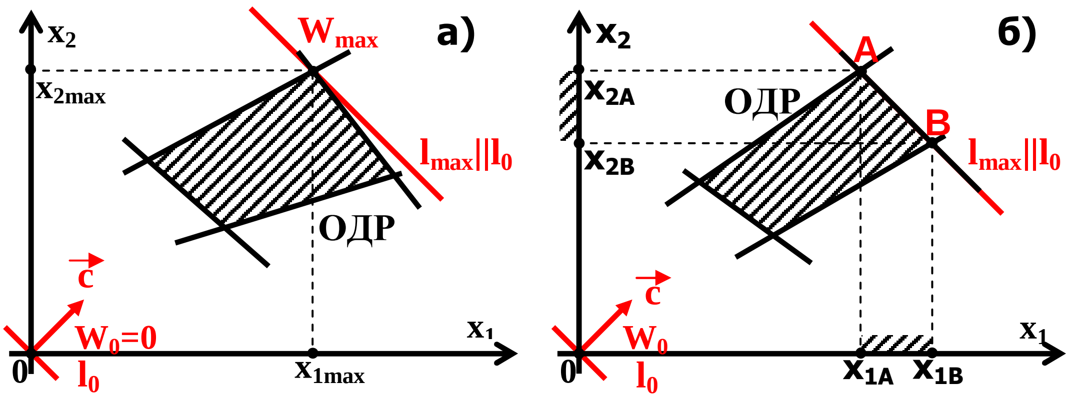
$$a_{i1}x_1 + \ldots + a_{in}x_n = b_i\ \left(i = \overline{1,m}\right),$$
$$x_j = 0\ \left(j = \overline{1,n}\right).$$

Припустимо, що система нерівностей (5.6) сумісна. Результат її розв'язання утворює область допустимих розв'язків, яка, як відомо із курсу лінійної алгебри, буде обмежена опуклим багатогранником, грані якого можуть співпадати із частинами гіперплощин, а вершини утворюються як точки перетину цих гіперплощин. Будь-яка внутрішня і гранична точка області допустимих розв'язків є допустимим розв'язком задачі лінійного програмування. Якщо прирівняти показник ефективності до нуля, то отримаємо рівняння гіперплощини в n-вимірному Евклідовому просторі

$$W = c_1x_1 + c_2x_2 + \ldots + c_nx_n = 0,$$

яка проходить через початок координат і є перпендикулярною до вектора-градієнта $\vec{C} = \left[c_1, \ldots, c_n\right]^T$, складеного із коефіцієнтів показника ефективності. Напрям вектора $\vec{C}$ вказує напрям зростання функції W. Тому, для пошуку максимального значення функції W необхідно пересувати гіперплощину в напрямку вектора $\vec{C}$ так, щоб область допустимих розв'язків і ця пересунута гіперплощина мали хоча б одну спільну точку.

На прикладі двовимірного Евклідового простору проілюструємо малюнками типові випадки, що трапляються при розв'язанні задач лінійного програмування (рис.5.1).



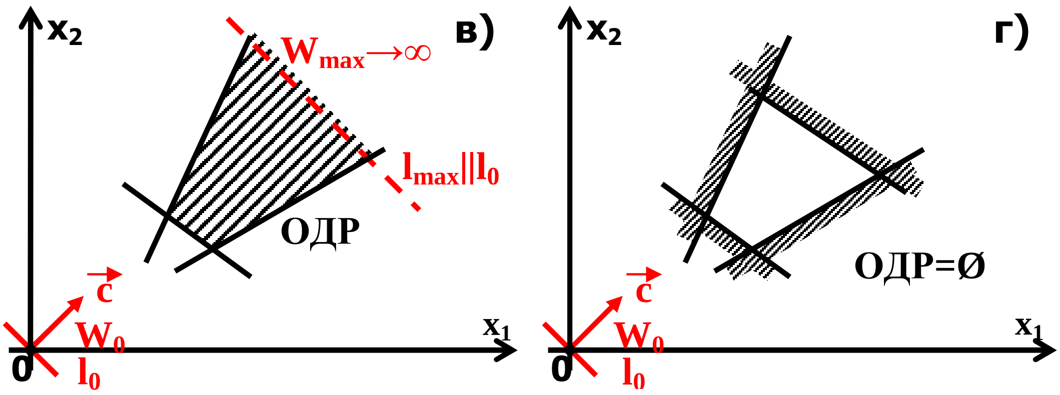


Рис. 5.1. Графічний образ розташування основної площини $W = c_1x_1 + c_2x_2 + \ldots + c_nx_n$ та області допустимих розв'язків: а) задача лінійного програмування має єдиний розв'язок; б) задача лінійного програмування має безліч оптимальних розв'язків; в) задача лінійного програмування має абстрактний розв'язок (цей розв'язок не має фізичного змісту тому, що не враховано усі ресурсні обмеження, які не дозволяють досягти нескінченно великого значення показника ефективності); г) система обмежень несумісна $G = \varnothing$: задача лінійного програмування розв'язків не має.

Вершини області допустимих розв'язків називають *опорними точками*, а розв'язки, що існують в цих вершинах – *опорними розв'язками*.

5.4. Графічний метод розв'язання задач лінійного програмування

Приклад 5.1. Розглянемо приклад виробничого планування, при $n = 2$. Припустимо, що телекомунікаційна фірма виготовляє вироби двох видів:

A – супутник зв'язку,
B – навігаційний супутник.
Для їх виробництва використовуються ресурси:
C – базові модулі сонячних батарей;
D – функціонально необхідні елементи життєзабезпечення супутника;
E – модулі системи орієнтації, стабілізації та навігації;
в об'ємах відносно: $C = 600$, $D = 800$, $E = 240$.
Норми витрат ресурсів наведено в таблиці 3.1.
Прибуток від реалізації виробу A складає 40 у.о., B – 50 у.о.
Знайти об'єм виробництва, що забезпечує максимальний прибуток.

Таблиця 5.1

Норми витрат ресурсів

	A	B
C	24	8
D	8	8
E	3	8

Побудова математичної моделі.

Позначимо $x_{1,2}$ – об'єми виробництва виробу A та B. Тоді прибуток фірми від реалізації x_1 виробу A та x_2 виробу B складатиме $W = 40x_1 + 50x_2$.

Обмеження G (область допустимих розв'язків), пов'язані із ресурсами та фізичним змістом змінних, мають вигляд:

$$\begin{cases} 24x_1 + 8x_2 \le 600; \\ 8x_1 + 8x_2 \le 480; \\ 3x_1 + 8x_2 \le 240; \\ x_1 \in N_0,\ x_2 \in N_0, \end{cases}$$

де $N_0 = \{0, 1, 2, \ldots\}$. Знехтуємо умовою цілочисельності розв'язків, а залишимо умову їх невід'ємності, тобто розв'яжемо так звану послаблену задачу.

Математична постановка задачі: необхідно максимізувати показник ефективності

$$W = 40x_1 + 50x_2 \to \max_{x_{1,2} \in G}$$

за умови виконання записаних обмежень G. Як бачимо, задача лінійного програмування сформульована як симетрична задача лінійного програмування.

❍ Знайдемо область допустимих розв'язків цієї задачі. Для цього розв'яжемо нерівності відносно змінної x_2:

$$\begin{cases} x_2 \le -3x_1 + 75; \\ x_2 \le -x_1 + 60; \\ x_2 \le -\,{}^{3}\!/_{8}\, x_1 + 30; \\ x_{1,2} \ge 0. \end{cases}$$

Границі:

$$\begin{aligned} x_2 &= -3x_1 + 75, \\ x_2 &= -x_1 + 60, \\ x_2 &= -\frac{3}{8}x_1 + 30, \\ x_1 &= 0,\ x_2 = 0. \end{aligned}$$

Виконаємо побудову області допустимих значень (рис.5.2) та знайдемо положення основної прямої

$$W = 40x_1 + 50x_2 = 0 \Rightarrow l_0 : x_2 = -\frac{4}{5}x_1.$$

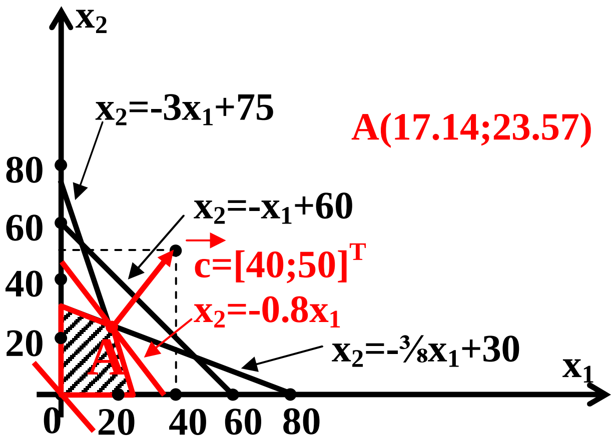


Рис. 5.2. Графічна ілюстрація знаходження області допустимих значень

Знайдемо W_{max}:

$$x_2 = -\frac{4}{5}x_1 + \left(\frac{W}{50}\right) = -\frac{4}{5}x_1 + W^*,$$

де W^* – відрізок, який пряма відтинає на осі ординат. Чим більше значення W^*, тим більше W тому, що $W^* = \dfrac{W}{50}$.

Знайдемо $\left(x_{1_{max}}, x_{2_{max}}\right)$ як точку перетину границь:

$$\begin{cases} x_2 = -\dfrac{3}{8}x_1 + 30, \\ x_2 = -3x_1 + 75, \end{cases} \Leftrightarrow \begin{cases} x_1 = 17\dfrac{1}{7} \approx 17,14, \\ x_2 = 23\dfrac{4}{7} \approx 23,57. \end{cases}$$

$$W_{\max} = 40 \cdot 17,14 + 50 \cdot 23,57 = 1864,1.$$

Обчислимо оптимальний розв‘язок із урахуванням умови цілочисельності та невід‘ємності змінних x_1 та x_2 (рис.5.3).

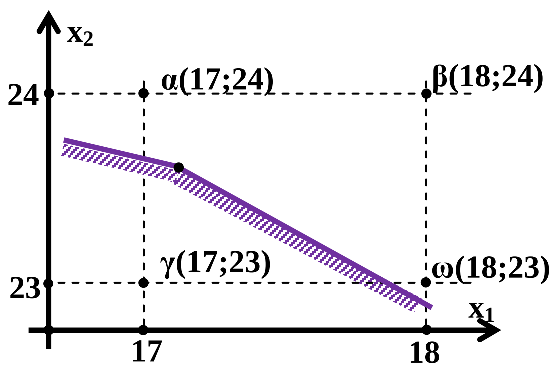


Рис. 5.3. Розташування оптимального розв‘язку поблизу цілочисельних значень

З рисунку 5.3 (із результатів обчислення можливості попадання $\alpha, \beta, \gamma, \omega$ в область допустимих значень) отримаємо наступний результат: лише γ (17,23) належить області допустимих значень, тобто

$$\mathrm{W}_{\max} = 40 \cdot 17 + 50 \cdot 23 = 1830$$

і досягається це значення критерію при випуску продукції: 17 одиниць типу A і 23 одиниць типу B.●

Приклад 5.2. Знайти оптимальний розв‘язок основної задачі лінійного програмування:

$$W = -5x_1 - 2x_2 - 18 \to \min_{X \in G},$$

де ОДР G задана наступними обмеженнями:

$$\begin{cases} x_1 - x_2 + x_3 = 4, \\ 2x_1 - x_2 - x_3 - x_4 = -5, \\ x_1 + x_2 - x_5 = -4, \\ x_2 + x_6 = 5, \\ 2x_1 - 2x_2 - x_6 + 2x_7 = 7, \end{cases}$$

при умові, що $x_i \geq 0 \left(i = \overline{1,7}\right)$, $X = \left[x_1, \ldots, x_7\right]^T \in \mathbb{R}^7$.

❍ Кількість керуючих змінних $n = 7$, а кількість обмежень-рівностей $m = 5$, тому $n - m = 7 - 5 = 2$. Отже, використаємо графічно-аналітичний метод розв‘язання задачі лінійного програмування:

– побудуємо ОДР G;

– графічним способом знайдемо точку, що відповідає мінімальному значенню критерію W (оптимальний розв‘язок);

– обчислимо оптимальний розв‘язок та мінімальне значення критерію.

1) Дослідження системи рівнянь на сумісність.

Із курсу лінійної алгебри відомо: для сумісності системи лінійних рівнянь необхідно і достатньо, щоб ранг матриці системи r_A дорівнював рангу розширеної матриці r_{AB}. Запишемо систему обмежень у матричному вигляді:

$$A \cdot X = B.$$

де $A = \begin{bmatrix} 1 & -1 & 1 & 0 & 0 & 0 & 0 \\ 2 & -1 & -1 & -1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 2 & -2 & 0 & 0 & 0 & -1 & 2 \end{bmatrix}$, $B = \begin{bmatrix} 4 \\ -5 \\ -4 \\ 5 \\ 7 \end{bmatrix}$, $\left[AB\right] = \begin{bmatrix} & 4 \\ & -5 \\ A & -4 \\ & 5 \\ & 7 \end{bmatrix}$.

A – матриця системи, B – матриця вільних коефіцієнтів, $\left[AB\right]$ – розширена матриця системи.

Ранг матриці це найбільший порядок відмінного від нуля мінора:

$\Delta_1 = 1 \neq 0$,

$$\Delta_2 = \begin{vmatrix} 1 & -1 \\ 2 & -1 \end{vmatrix} \neq 0,$$

$$\Delta_3 = \begin{vmatrix} 1 & -1 & 1 \\ 2 & -1 & -1 \\ 1 & 1 & 0 \end{vmatrix} = 1\cdot(-1)\cdot 0 + (-1)\cdot(-1)\cdot 1 + 2\cdot 1\cdot 1 -$$

$$-1\cdot(-1)\cdot 1 - 1\cdot 1\cdot(-1) - 2\cdot(-1)\cdot 0 = 5 \neq 0,$$

$$\Delta_4 = \begin{vmatrix} 1 & -1 & 1 & 0 \\ 2 & -1 & -1 & -1 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{vmatrix} = 1\cdot\begin{vmatrix} 1 & 1 & 0 \\ 2 & -1 & -1 \\ 1 & 0 & 0 \end{vmatrix} = -1\cdot(-1)\begin{vmatrix} 1 & 1 \\ 1 & 0 \end{vmatrix} = -1 \neq 0,$$

$$\Delta_5 = \begin{vmatrix} 1 & -1 & 1 & 0 & 0 \\ 2 & -1 & -1 & -1 & 0 \\ 1 & 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & 0 & 0 \\ 2 & -2 & 0 & 0 & 0 \end{vmatrix} = (-1)\cdot\begin{vmatrix} 1 & -1 & 1 & 0 \\ 2 & -1 & -1 & -1 \\ 0 & 1 & 0 & 0 \\ 2 & -2 & 0 & 0 \end{vmatrix} =$$

$$= -(-1)^{2+4}\cdot(-1)\cdot\begin{vmatrix} 1 & -1 & 1 \\ 0 & 1 & 0 \\ 2 & -2 & 0 \end{vmatrix} = \begin{vmatrix} 0 & 1 \\ 2 & -2 \end{vmatrix} = -2 \neq 0.$$

Таким чином маємо, що

$$r_A = m = 5.$$

Визначників більшого порядку немає, тому що матриця A є прямокутною матрицею розміру 5×7. Зрозуміло, що розширена матриця системи матиме розмір 5×8, а це означає що найбільший порядок визначника розширеної матриці співпадає із найбільшим порядком визначника системи і дорівнює 5.

Висновок: $r_A = r_{AB} = 5$, система обмежень у формі рівностей сумісна.

2) Вибір вільних та базисних змінних.

Враховуючи той факт, що $n = 7$, $m = 5$, тобто $n - m = 2$, можливо обрати дві змінні, наприклад, x_1 та x_2 в якості вільних і 5 змінних x_3, x_4, x_5, x_6, x_7 в якості базисних. Тоді із системи обмежень знаходимо:

$$\begin{cases} x_3 = -x_1 + x_2 + 4 \geq 0, \\ x_4 = 3x_1 - 2x_2 + 1 \geq 0, \\ x_5 = x_1 + x_2 + 4 \geq 0, \\ x_6 = -x_2 + 5 \geq 0, \\ x_7 = \frac{1}{2}\left(-2x_1 + 2x_2 + x_6 + 7\right) = -x_1 + \frac{1}{2}x_2 + 6 \geq 0. \end{cases}$$

Враховуючи невід'ємність змінних будуємо область допустимих розв‘язків, аналітичний вираз для якої в залежності від змінних x_1 та x_2 має вигляд:

$$\begin{cases} x_2 \geq x_1 - 4, \\ x_2 \leq \frac{3}{2}x_1 + \frac{1}{2}, \\ x_2 \geq -x_1 - 4, \\ x_2 \leq 5\ , \\ x_2 \geq 2x_1 - 12,\ x_{1,2} \geq 0. \end{cases}$$

Знайдемо положення основної гіперплощини:

$$W = -5x_1 - 2x_2 - 18 \Leftrightarrow x_2 = -\frac{5}{2}x_1 - \underbrace{\frac{W+18}{2}}_{b};$$

$$\begin{cases} b = 0\ , \\ x_2 = -\frac{5}{2}x_1, \end{cases}$$

$\vec{c} = \left[-5, -2\right]$ – вектор градієнту.

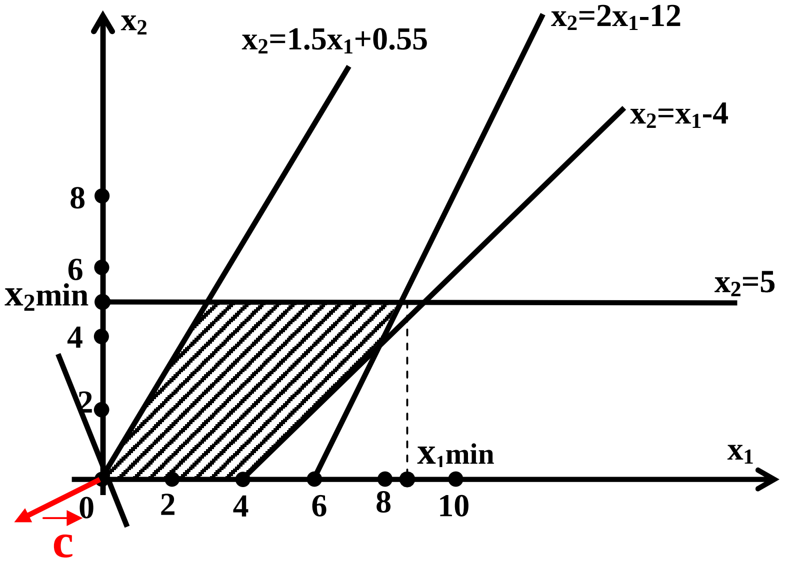


Рис. 5.4. Графічна ілюстрація побудови області допустимих розв‘язків

Оптимальне (мінімальне) значення показника ефективності дорівнює:

$$W_{\min} = -70,5$$

і досягається це значення у точці $\left(x_{1min}; x_{2min}\right) = \left\{\left(8,5; 5\right)\right\}$.●

Висновок: наведена у п.5.3 геометрична інтерпретація задачі лінійного програмування та приклади дозволяють стверджувати, що:

1) ***Принцип 1.***

Оптимальний розв‘язок, якщо він існує, розташований в одній із «опорних точок». За аналітичною ознакою опорна точка – це така точка, в якій хоча б $k = n - m$ (n - кількість невідомих, m - кількість сумісних обмежень рівностей) змінних перетворюються в 0. Геометрично опорна точка визначається як кутова точка ОДР G. Проілюструємо принцип 1 графічними образами (рис. 5.5) для випадку $k = n - m = 2$. Зрозуміло, що у загальному випадку, коли $k = 2$ всі змінні $x_3, x_4, ..., x_n$ можна виразити через x_1 та x_2, що дозволяє критерій

$$W = c_1x_1 + c_2x_2 + c_3x_3 + ... + c_nx_n$$

записати у вигляді:

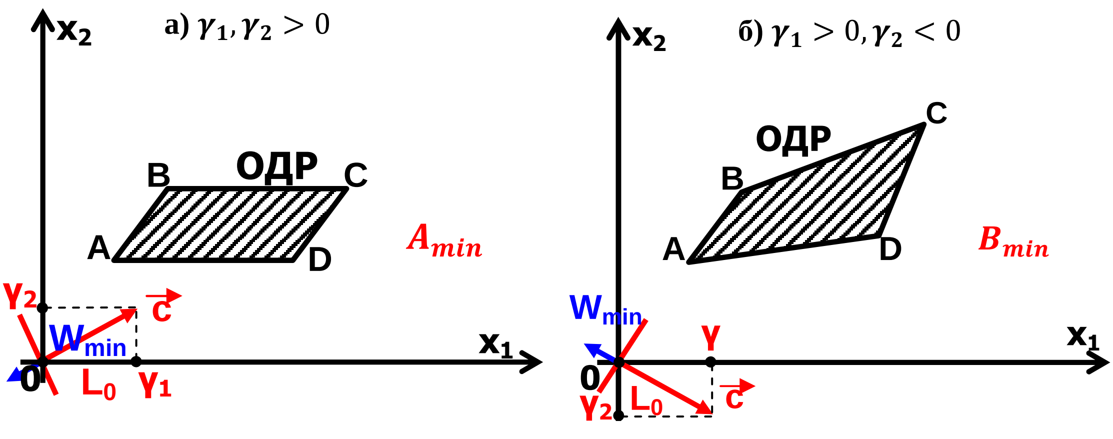
$$W = \gamma_1x_1 + \gamma_2x_2 + \gamma_0.$$

Припустимо, що необхідно знайти оптимальне значення критерію W. Якщо $\gamma_0 = 0$, то основна пряма L_0:

$$\gamma_1x_1 + \gamma_2x_2 = 0$$

проходить через початок координат.

Розглянемо за допомогою рисунків 5.5 а), б), в), г) напрямки переміщення цієї прямої в площині двох змінних $x_{1,}$ x_2, які дозволяють знайти найменше значення показника ефективності.



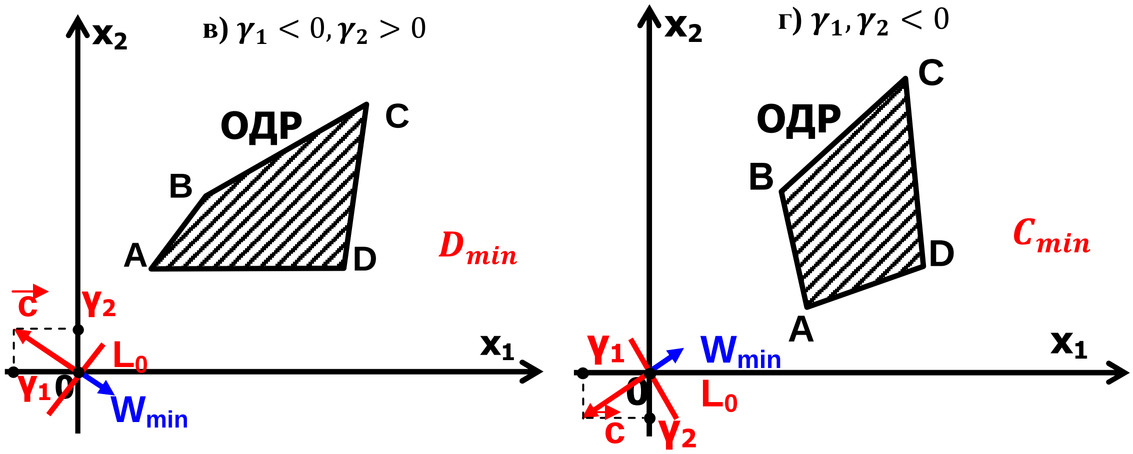


Рис. 5.5. Графічна ілюстрація положення основної прямої та напрямку її пересування при пошуку найменшого значення в точці, яка задовольняє область допустимих розв'язків

2) ***Принцип 2.***

Для пошуку оптимального розв'язку необхідно переходити від однієї опорної точки до іншої, пересуваючись в напрямку зменшення або збільшення показника ефективності в залежності від вимог задачі лінійного програмування.

3) ***Принцип 3.***

Застосування геометричного методу розв'язання задачі лінійного програмування ускладнюється, коли $n-m=3$. Геометричний метод зовсім втрачає придатність до розв'язання практичних задач при $n-m>3$. Тому, у довільному випадку застосовуються чисельні методи пошуку екстремуму показника ефективності, які ґрунтуються на записаних вище принципах (1) та (2).

Універсальним методом розв'язання задач лінійного програмування є **симплекс-метод**.

Лекція 6. Симплекс метод розв'язання задачі лінійного програмування

6.1. Поняття про симплекс метод

У назві основного методу розв'язання основної задачі лінійного програмування використовується термін "симплекс", що означає n-вимірний тетраедр або n-вимірний трикутник.

Симплекс-метод знаходження локального мінімуму будь-якої функції декількох змінних лінійної або нелінійної запропоновано Нелдером і Мідом. Цей метод хоча і містить у своїй назві слово «симплекс» не має нічого спільного із симплекс методом розв'язання задачі лінійного програмування. Суть методу Нелдера – Міда полягає у спеціальній процедурі обчислення координат вершин цього n-вимірного трикутника для наступної ітерації (наближення) в залежності від результату порівняння значень показника ефективності у вершинах n-вимірного трикутника, координати яких можуть бути обчислені у попередній ітерації. "Найгірша" вершина, в якій показник ефективності приймає найбільше значення, якщо відбувається пошук мінімального значення показника ефективності, відкидається і замінюється новою.

Координати нової вершини отримують, наприклад, наступним прийомом: «відображенням» старої вершини відносно прямої, що проходить через дві інші вершини. Окрім «відображення» для пошуку координат нової вершини використовуються так звані процедури "продовження", "стискання" або "скорочення". В результаті застосування означених прийомів та процедур значення показника ефективності у вершинах трикутників на кожній ітерації зменшується і при цьому зменшується «розмір» самого n-вимірного трикутника, стискаючись поступово до точки мінімального значення показника ефективності (рис. 6.1).

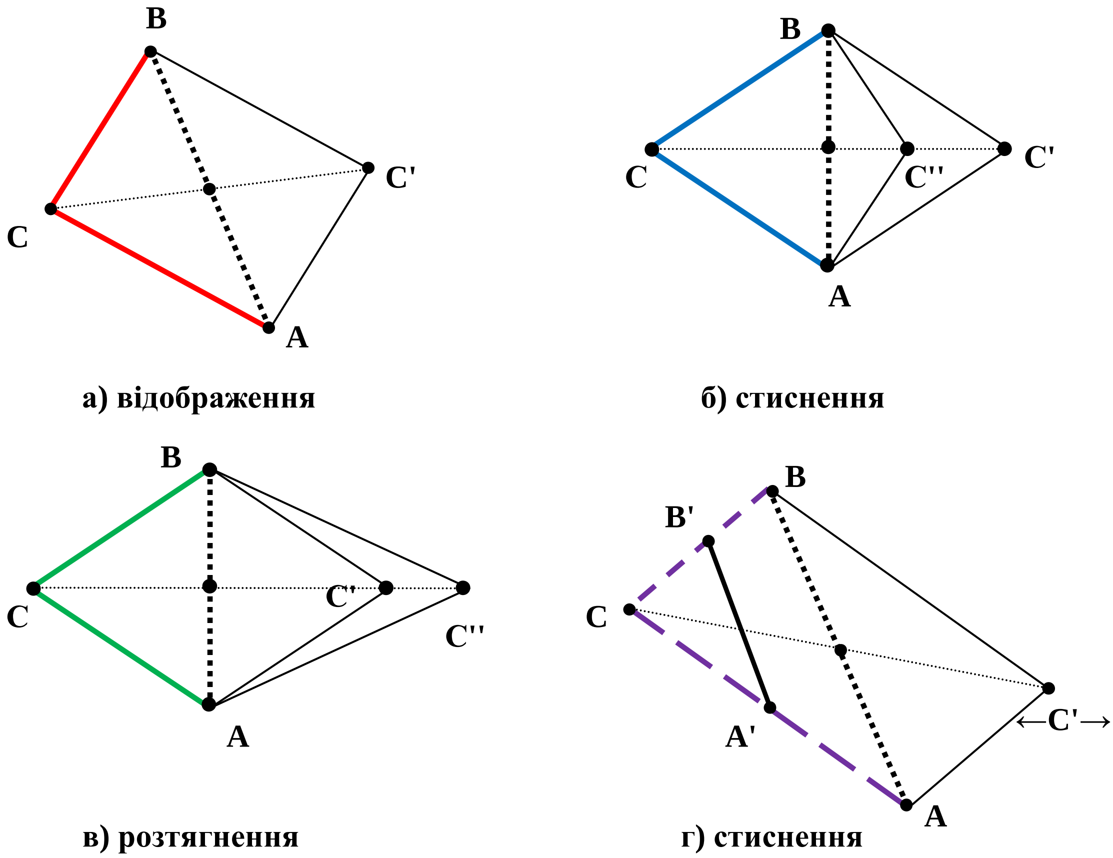


Рис. 6.1. Графічна ілюстрація прийомів симплекс методу Нелдера-Міда

Перейдемо до розгляду симплекс методу розв‘язання задач лінійного програмування.

6.2. Приведення стандартної форми обмежень нерівностей до обмежень рівностей (рівнянь обмежень) основної задачі лінійного програмування

Стандартною формою обмежень нерівностей вважається система обмежень вигляду:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n \le b_1, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n \le b_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n \le b_m, \end{cases} \tag{6.1}$$

де $x_j \ge 0 \left(j = \overline{1,n}\right)$. Усі нерівності системи лінійно незалежні.

Цю систему можливо перетворити в обмеження рівності за допомогою додаткових невід'ємних змінних, а саме:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n \le b_1, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n \le b_m, \end{cases} \Leftrightarrow$$

$$\Leftrightarrow \begin{cases} y_1 = b_1 - a_{11}x_1 - a_{12}x_2 - \ldots - a_{1n}x_n, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ y_m = b_m - a_{m1}x_1 - a_{m2}x_2 - \ldots - a_{mn}x_n, \end{cases}$$

де $x_j \ge 0 \ \left(j = \overline{1,n}\right)$; $y_i \ge 0 \ \left(i = \overline{1,m}\right)$.

Якщо до записаних обмежень рівностей додати показник ефективності

$$W = c_1x_1 + c_2x_2 + \ldots + c_nx_n \to \min_{\substack{x_j \ge 0\left(j=\overline{1,n}\right) \\ y_i \ge 0\left(i=\overline{1,m}\right)}}, \tag{6.2}$$

який необхідно мінімізувати за невід'ємними змінними $x_j \ge 0 \ \left(j = \overline{1,n}\right)$ та $y_i \ge 0 \ \left(i = \overline{1,m}\right)$, то отримаємо основну задачу лінійного програмування.

Отже, в результаті переходу до основної задачі лінійного програмування маємо:

1) Рівняння обмежень задані у формі, де базисні (залежні) змінні $y_i \ge 0 \ \left(i = \overline{1,m}\right)$ виражені через незалежні (вільні) $x_j \ge 0 \ \left(j = \overline{1,n}\right)$ змінні.

2) Загальна кількість змінних дорівнює «$n + m$», де n – кількість початкових, m – додаткових змінних.

3) Показник ефективності явно залежить від початкових змінних. Вважаємо, що усі коефіцієнти при додаткових змінних показника ефективності дорівнюють 0.

Приклад 6.1. Дана задача лінійного програмування:

$$W = x_1 - 2x_2 - 3x_3 \to \min_{x_1, x_2, x_3 \ge 0}$$

за умови виконання обмежуючих нерівностей

$$\begin{cases} 2x_1 - x_2 + x_3 \le 5, \\ x_3 - 2x_2 \le -1, \\ x_5 - 2x_4 + x_1 \le -1, \\ x_5 - x_1 \le 0, \\ x_3 + 2x_1 + x_5 \ge 0. \end{cases}$$

Потрібно привести цю задачу до вигляду основної задачі лінійного програмування.

❍ 1. Приведемо обмеження нерівності до стандартної форми

$$\begin{cases} 2x_1 - x_2 + x_3 \le 5, \\ x_3 - 2x_2 \le -1, \\ x_5 - 2x_4 + x_1 \le -1, \\ x_5 - x_1 \le 0, \\ -x_3 - 2x_1 - x_5 \le 0. \end{cases}$$

2. Використовуємо для отримання обмежень рівностей додаткові змінні

$$\begin{cases} y_1 = 5 - \left(2x_1 - x_2 + x_3\right), \\ y_2 = -1 - \left(x_3 - 2x_2\right), \\ y_3 = -1 - \left(x_5 - 2x_4 + x_1\right), \\ y_4 = 0 - (x_5 - x_1), \\ y_5 = 0 - (-x_3 - 2x_1 - x_5); \end{cases} \Leftrightarrow \begin{cases} y_1 = 5 - 2x_1 + x_2 - x_3, \\ y_2 = -1 - x_3 + 2x_2, \\ y_3 = -1 - x_5 + 2x_4 - x_1, \\ y_4 = -x_5 + x_1, \\ y_5 = x_3 + 2x_1 + x_5. \end{cases}$$

3. Постановка задачі набуває вигляду

$$W = x_1 - 2x_2 - x_3 \to \min_{\substack{x_{1,2,3} \ge 0 \\ y_{1,2,3} \ge 0}}$$

за умови виконання обмежень (2). ●

6.3. Основні прийоми та способи симплекс методу розв'язання задач лінійного програмування

Для розв'язання основної задачі лінійного програмування використовуємо наведені в розділі 5.4 принципи побудови оптимального розв‘язку.

Прийоми та способи симплекс-методу розв'язання задачі лінійного програмування викладені у припущенні, що в задачі лінійного програмування використовується n змінних та m незалежних лінійних рівнянь-обмежень.

Прийом 1. Вибір вільних змінних.

Оберемо будь-які k змінних $k = n - m$ ($r = m$ - ранг системи обмежень рівностей) в якості вільних і представимо через них m базисних змінних, що залишилися. Припустимо, що у якості вільних обрані перші k змінних, тобто $x_1, x_2, \ldots, x_k$, а інші m представимо через них:

$$\begin{cases} x_{k+1} = \alpha_{k+1,1} x_1 + \ldots + \alpha_{k+1,k} x_k + \beta_{k+1}, \\ \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \quad x_n = \alpha_{n,1} x_1 + \ldots + \alpha_{n,k} x_k + \beta_n, \end{cases}$$

Тобто маємо

$$\underbrace{1, 2, \ldots, k,}_{k\text{ вільних змінних}} \quad \underbrace{k+1, \ldots, n}_{m\text{ базисних змінних}}.$$

Якщо покласти, що всі вільні змінні дорівнюють $x_1 = x_2 = \ldots = x_k = 0$, то отримаємо координати вершини симплексу:

$$x_{k+1} = \beta_{k+1}, \ldots, x_n = \beta_n .$$

Цей розв'язок може бути допутимим, якщо всі вільні коефіцієнти – невід'ємні, або недопустимим, якщо серед коефіцієнтів є хоча б одне від'ємне число.

Припустимо, що розв'язок є допустимий, тобто знайдено опорний розв'язок. При цьому виникає питання, чи є розв'язок оптимальним? Для перевірки опорного розв'язку на оптимальність, представимо показник ефективності як функцію, яка залежить від вільних змінних:

$$W = c_1 x_1 + \ldots + c_k x_k + c_{k+1} x_{k+1} + \ldots + c_n x_n \Leftrightarrow$$

$$W = \gamma_0 + \gamma_1 x_1 + \ldots + \gamma_k x_k \to \min_{x_j \geq 0,\ \left(j=\overline{1,k}\right)} .$$

Враховуючи, що в опорному розв'язку $x_j = 0 \ \left(j = \overline{1,k}\right)$, тоді $W = \gamma_0$.

Проаналізуємо, чи можливо зменшити показник ефективності, збільшивши які-небудь змінні $x_1, \ldots, x_k$ (зменшувати їх неможливо, тому що всі ці змінні, при отриманні оптимального розв'язку дорівнюють 0, а від'ємні значення недопустимі).

Якщо всі коефіцієнти $\gamma_1, \ldots, \gamma_k$ додатні, то збільшуючи будь-які змінні $x_1, \ldots, x_k$ порівняно із 0 неможливо зменшити показник ефективності. Тому знайдений опорний розв'язок є оптимальним.

Прийом 2. Покращення опорного розв'язку.

Якщо серед коефіцієнтів $\gamma_1, \ldots, \gamma_k$ є від'ємні, то збільшуючи ті змінні, при яких коефіцієнти від'ємні, досягаємо покращення розв'язку, тобто зменшення показника ефективності.

Припустимо, що є єдиний від'ємний коефіцієнт γ_1. Необхідно виконати дві дії:

1) Збільшити x_1, але так, щоб жодна із базисних змінних $x_{k+1}, \ldots, x_n$ не стала від'ємною, якщо деякий із коефіцієнтів $\alpha_{k+11}, \ldots, \alpha_{n1}$ при x_1 від'ємний. Збільшувати x_1 можливо без обмежень, якщо всі коефіцієнти при x_1 у виразах для обчислення базисних змінних додатні. Але в цьому випадку показник ефективності W прямує до $-\infty$ при $x_1 \to +\infty$, тобто оптимального розв'язку, який має фізичний зміст, не існує. Існує абстрактний розв'язок. Розв'язання задачі припиняється, необхідно переформулювати постановку задачі.

2) Виключити x_1 із списку вільних змінних і вставити у список базисних, а із списку базисних виключити ту змінну, припустимо x_L, яка першою досягне значення 0 при збільшенні x_1. Випишемо рівняння для x_L:

$$x_L = \alpha_{L1} \cdot x_1 + \ldots + \alpha_{Lk} \cdot x_k + \beta_L$$

в якому покладемо, що $x_2 = 0, \ldots, x_k = 0, \; x_L = 0$.

Тоді

$$x_1 = -\frac{\beta_L}{\alpha_{L1}} \geq 0$$

тому, що $\alpha_{L1} < 0, \beta_L \geq 0$.

Отримане значення x_1 – це те значення, при якому $x_L = 0$. Взагалі першою досягне нуля та змінна із складу $x_{k+1}, \ldots, x_n$ для якої $-\frac{\beta_L}{\alpha_{L1}}$ або $\left|\frac{\beta_L}{\alpha_{L1}}\right|$ буде найменшим ($L \in \left[k+1, n\right], \; L \in \mathbb{N}$). Припустимо, що це x_r, $r \in \left[k+1, n\right]$, $r \in \mathbb{N}$.

Прийом 3. Зміна складу вільних змінних.

Обираємо новий склад вільних змінних $x_2, x_3, \ldots, x_k, x_r$ та базисних $x_1, x_{k+1}, \ldots, x_{r-1}, x_{r+1}, \ldots, x_n$. Обчислимо нові базисні змінні через нові вільні та перевіряємо умову невід'ємності базисних змінних при нульових вільних змінних, тобто з'ясуємо чи є розв'язок опорним. Перевіряємо отримані опорні розв'язки на оптимальність, використовуючи прийоми 1 та 2.

Висновок: практична реалізація симплекс методу потребує розробки двох алгоритмів:
1. Визначення опорного розв‘язку основної задачі лінійного програмування.
2. Визначення оптимального розв‘язку основної задачі лінійного програмування.

Приклад 6.2. Пошук оптимального розв‘язку шляхом поступового покращення результату із використанням трьох, описаних вище прийомів.

Постановка задачі:

$$W = 5x_1 - 2x_3 \rightarrow \min_{X \in G} \tag{6.3}$$

при умові виконання обмежень G:

$$\begin{cases} -5x_1 - x_2 + 2x_3 \leq 2, \\ -x_1 + x_3 + x_4 \leq 5, \\ 3x_1 + 5x_4 \leq 7, \\ x_i \geq 0, \left(i = 1, 2, 3, 4\right), \end{cases} \tag{6.4}$$

$$X = [x_1, x_2, x_3, x_4]^T .$$

Методика розв‘язання задачі:

1) виконати перехід до стандартного виду основної задачі лінійного програмування;

2) перевірити систему обмежень-рівностей на сумісність;

3) виконати обчислення оптимального розв‘язку із використанням прийомів симплекс методу.

❍ 1. Перехід до стандартного вигляду задачі лінійного програмування:

$$(6.4) \Leftrightarrow \begin{cases} -5x_1 - x_2 + 2x_3 + y_1 = 2; \\ -x_1 + x_3 + x_4 + y_2 = 5; \\ 3x_1 + 5x_4 + y_3 = 7; \\ x_{1,2,3,4} \geq 0,\ y_{1,2,3} \geq 0, \end{cases} \tag{6.5}$$

де (6.5) задає ОДР G.

2. Перевірка обмежень-рівностей на сумісність:

$$(6.5) \Leftrightarrow \begin{cases} -5 \cdot x_1 - 1 \cdot x_2 + 2 \cdot x_3 + 0 \cdot x_4 + y_1 = 2, \\ -1 \cdot x_1 + 0 \cdot x_2 + 1 \cdot x_3 + 1 \cdot x_4 + y_2 = 5, \\ -3 \cdot x_1 + 0 \cdot x_2 + 0 \cdot x_3 + 5 \cdot x_4 + y_3 = 7, \end{cases} \tag{6.6}$$

де

$$A = \begin{bmatrix} -5 & -1 & 2 & 0 & 1 \\ -1 & 0 & 1 & 1 & 1 \\ -3 & 0 & 0 & 5 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 2 \\ 5 \\ 7 \end{bmatrix}.$$

Ранг матриці $A : r_A = 3$, що дорівнює m – кількості обмежень-рівностей. Отже, система обмежень сумісна та лінійно залежних обмежень немає.

Підкреслимо, що зазвичай сумісності та незалежності обмежень можливо досягти ще при побудові математичної моделі задачі, якщо враховувати фізичний зміст цих обмежень.

3. Обчислення оптимального розв‘язку задачі лінійного програмування симплекс методом почнемо з вибору базисних та вільних змінних:

$$\begin{cases} y_1 = 5x_1 + x_2 - 2x_3 + 2, \\ y_2 = x_1 - x_3 - x_4 + 5, \\ y_3 = 3x_1 - 5x_4 + 7, \\ x_{1,2,3,4} \geq 0\ , \\ y_{1,2,3} \geq 0, \end{cases} \tag{6.7}$$

де система (6.7) задає ОДР G, на якій оптимізується критерій

$$W = 5x_1 - 2x_3 \to \min_{\left[X^T\ Y^T\right]^T \in G},$$

$X = \left[x_1, x_2, x_3, x_4\right]^T, Y = \left[y_1, y_2, y_3\right]^T$.

Загальна кількість змінних в переформульованій задачі $n = 7$, $m = 3$ – кількість сумісних, лінійно незалежних обмежень-рівностей. Кількість вільних змінних $k = 7 - 3 = 4$.

Прийом 1.

Обираємо в якості вільних змінних $x_{1,2,3,4}$ і припускаємо, що вони дорівнюють нулю. В результаті отримаємо розв'язок:

$$X_1 = \left[x_1 = 0; x_2 = 0; x_3 = 0; x_4 = 0; y_1 = 2; y_2 = 5; y_3 = 7\right],$$

який можливо вважати опорним.

При цьому $W\left(X_1\right) = 0$.

Прийом 2.

2.1) Коефіцієнт при x_3 виразу для обчислення показника ефективності від'ємний, тому за рахунок збільшення x_3 можливо зменшити W. Обираємо в якості нової вільної змінної ту змінну (серед $y_{1,2,3}$), для якої модуль відношення вільних коефіцієнтів до коефіцієнта при x_3 найменший:

$$\left|\frac{2}{-2}\right| = 1 < \left|\frac{5}{-1}\right|.$$

Обираємо y_1. Таким чином, вільними змінними є x_1, x_2, x_4, y_1. Обчислимо нові базисні змінні x_3, y_2, y_3:

$$\begin{cases} x_3 = -\dfrac{1}{2}y_1 + \dfrac{5}{2}x_1 + \dfrac{1}{2}x_2 + 1, \\ y_2 = \dfrac{1}{2}y_1 - \dfrac{3}{2x_1} - \dfrac{1}{2}x_2 - x_4 + 4, \\ y_3 = 3x_1 - 5x_4 + 7. \end{cases}$$

Перевіряємо знайдений розв'зок, чи є він опорним.

$$X_2 = \left[y_1 = 0; x_{1,2} = 0; x_3 = 1; y_2 = 4; y_3 = 7\right]$$

X_2 – опорний розв'язок. Спостерігається зменшення значення критерію:

$$W\left(X_2\right) = y_1 - x_2 - 2 = -2.$$

2.2) Коефіцієнт при x_2 від'ємний. Можливо зменшити значення показника ефективності завдяки збільшенню x_2. Аналогічно викладеному у 2.1) визначаємо нові вільні та базисні змінні:

$$\begin{cases} x_3 = -y_2 + x_1 - x_4 + 5, \\ x_2 = y_1 - 2y_2 - 3x_1 - 2x_1 - 2x_4 + 8, \\ y_3 = 3x_1 - 5x_4 + 7, \end{cases}$$

$$X_3 = \left[\underbrace{y_1 = 0; x_1 = 0; x_4 = 0; y_2 = 0}_{\text{вільні змінні}}; \underbrace{x_2 = 8; x_3 = 5; y_3 = 7}_{\text{базисні змінні}} \right],$$

$$W\left(X_3\right) = 2y_2 + 3x_1 + 2x_4 - 10 = -10.$$

Всі коефіцієнти при вільних змінних у виразі для обчислення показника ефективності невід'ємні, тому досягати зменшення показника ефективності за рахунок додаткових вільних змінних неможливо. Досягнуто глобального мінімуму

$$W = -10$$

в точці $x_1 = 0$, $x_4 = 0$, $y_1 = 0$, $y_2 = 0$, $x_2 = 8$, $x_3 = 5$, $y_3 = 7$.●

Зауваження 6.1. В наведених прикладах при застосуванні прийомів симплекс методу кожного разу виникали розв'язки, які були невід'ємними, тобто опорними. В більшості практичних випадків виникають ситуації, коли при обраних вільних змінних з нульовими значеннями виникають від'ємні значення базисних змінних, тобто не виконується умова невід'ємності змінних. Тому у наступних лекціях розглянемо спеціальний алгоритм обчислення опорного розв'язку.

Лекція 7. Алгоритм пошуку опорних розв'язків основної задачі лінійного програмування

7.1. Табличний алгоритм заміни базисних змінних (стандартні таблиці)

Як було показано раніше, прийоми реалізації симплекс-методу розв'язання задачі лінійного програмування базуються на багатократних перераховуваннях системи рівнянь обмежень основної задачі лінійного програмування відносно нових вільних та базисних змінних. З метою прискорення цих розрахунків розглянемо спеціальний табличний алгоритм. Припустимо, що необхідно виконати мінімізацію показника ефективності, представленого лінійною функцією

$$W = c_0 + c_1x_1 + c_2x_2 + c_3x_3 \to \min_{x_{1,2,3} \geq 0}$$

за умови виконання обмежень

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 \leq b_1, \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 \leq b_2, \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 \leq b_3. \end{cases}$$

Представимо цю задачу у вигляді основної задачі лінійного програмування у формі зручній для застосування стандартизованого табличного алгоритму:

$$W = c_0 - \left(\left(-c_1\right)x_1 + \left(-c_2\right)x_2 + \left(-c_3\right)x_3\right) = c_0 - \left(\gamma_1 x_1 + \gamma_2 x_2 + \gamma_3 x_3\right) \to \min_{x_{1,2,3} \geq 0}$$

$$\begin{cases} y_1 = b_1 - \left(a_{11}x_1 + a_{12}x_2 + a_{13}x_3\right), \\ y_2 = b_2 - \left(a_{21}x_1 + a_{22}x_2 + a_{23}x_3\right), \\ y_3 = b_3 - \left(a_{31}x_1 + a_{32}x_2 + a_{33}x_3\right), \end{cases}$$

де $\gamma_1 = -c_1$, $\gamma_2 = -c_2$, $\gamma_3 = -c_3$.

Складемо таблиці (табл. 7.1 та 7.2), в яких позначено: ПЕ - показник ефективності; БЗ – базисні змінні; ВЕ – вільний елемент.

Таблиця 7.1

Вихідне представлення інформації

ПЕ, БЗ	ВЕ	Коефіцієнти при вільних змінних		
		x_1	x_2	x_3
W	c_0	γ_1	γ_2	γ_3
y_1	b_1	a_{11}	a_{12}	a_{13}
y_2	b_2	a_{21}	a_{22}	a_{23}
y_3	b_3	a_{31}	a_{32}	a_{33}

Припустимо, що необхідно виконати заміну базисної змінної y_3 на x_2 і, навпаки, вільної змінної x_2 на y_3 . В аналітичному вигляді така заміна змінних виконується із використанням рівняння:

$$x_2 = \frac{b_3}{a_{32}} - \left(\frac{a_{31}}{a_{32}} \cdot x_1 + \frac{1}{a_{32}} \cdot y_3 + \frac{a_{33}}{a_{32}} \cdot x_3\right),$$

в якому елемент a_{32} називають ***розв'язувальним елементом*** і, відповідно, у стандартизованій таблиці рядок та стовпець називають ***розв'язувальними рядком*** та ***стовпцем*** (див. табл. 7.2). Якщо вираз для x_2 підставити у формули для обчислення W , y_1, y_2, то отримаємо таблицю перерахунку коефіцієнтів рівнянь обмежень для нових базисних змінних y_1, y_2, x_2 та нових вільних змінних x_1, y_3, x_3 (див. табл. 7.2).

Таблиця 7.2

Формули перерахунку коефіцієнта після заміни змінних

ПЕ БЗ	ВЕ	Коефіцієнти при вільних змінних		
		x_1	y_3	x_3
W	$c_0 + b_3\left(-\frac{\gamma_2}{a_{32}}\right)$	$\gamma_1 + a_{31}\left(-\frac{\gamma_2}{a_{32}}\right)$	$-\frac{\gamma_2}{a_{32}}$	$\gamma_3 + a_{33}\left(-\frac{\gamma_2}{a_{32}}\right)$

y_1	$b_1 + b_3\left(-\frac{a_{12}}{a_{32}}\right)$	$a_{11} + a_{31}\left(-\frac{a_{12}}{a_{32}}\right)$	$-\frac{a_{12}}{a_{32}}$	$a_{13} + a_{33}\left(-\frac{a_{12}}{a_{32}}\right)$
y_2	$b_2 + b_3\left(-\frac{a_{22}}{a_{32}}\right)$	$a_{21} + a_{31}\left(-\frac{a_{22}}{a_{32}}\right)$	$-\frac{a_{22}}{a_{32}}$	$a_{23} + a_{33}\left(-\frac{a_{22}}{a_{32}}\right)$
x_2	$b_3\left(\frac{1}{a_{32}}\right)$	$a_{31}\left(\frac{1}{a_{32}}\right)$	$\frac{1}{a_{32}}$	$a_{33}\left(\frac{1}{a_{32}}\right)$

В загальному випадку алгоритм перетворення x_j в y_i і навпаки із використанням стандартної таблиці складається з наступних дій:

1. Виділення в стандартній таблиці розв'язувального елемента a_{ij}. Обчислення зворотної величини $\lambda = \frac{1}{a_{ij}}$. Для наочності обчислень пропонується записати λ у правому нижньому куті фрагменту таблиці, де розташований розв'язувальний елемент.
2. Множення усіх елементів розв'язувального рядка (крім a_{ij}) на λ. Результат записується у відповідному фрагменті таблиці справа знизу.
3. Множення усіх елементів розв'язувального стовпця (окрім a_{ij}) на $-\lambda$. Результат записується у відповідному фрагменті таблиці справа знизу.
4. Виділення у розв'язувальному рядку усіх попередніх, а у розв'язувальному стовпці – усіх нових елементів за виключенням самого розв'язувального елемента a_{ij}.
5. Обчислення добутку нового елемента розв'язувального стовпця на попередній елемент розв'язувального рядка. Для всіх попередніх елементів таблиці окрім тих, що розташовані у розв'язувальних рядку та стовпці. Запис добутку у відповідний фрагмент таблиці.
6. Формулювання остаточного результату:

   1) заміна x_j на y_i і навпаки;

   2) залишити у розв'язувальних стовпці та рядку лише нижні елементи фрагментів таблиці;

   3) заміна інших елементів таблиці сумою попередніх чисел та результатів розрахунку згідно пункту 5.

Приклад 7.1. Необхідно виконати заміну змінних у наступній задачі лінійного програмування:

$$W = -x_1 - x_3 + 2 \to \min_{x_{1,2,3} \geq 0}$$

при виконанні умов:

$$\begin{cases} x_1 - x_2 + 2x_3 - 5 \geq 0, \\ 2x_1 - x_2 + 1 \geq 0, \\ 2x_2 - x_3 - 1 \geq 0. \end{cases}$$

❍ Переформулюємо критерій та обмеження нерівності у вигляді:

$$W = 2 - (x_1 + x_3) \to \min_{\substack{x_{1,2,3} \geq 0 \\ y_{1,2,3} \geq 0}}$$

$$\begin{cases} y_1 = -5 - \left(-x_1 + x_2 - 2x_3\right), \\ y_2 = 1 - \left(-2x_1 + x_2\right), \\ y_3 = -1 - \left(-2x_2 + x_3\right). \end{cases}$$

Представимо вихідні дані у вигляді таблиці 7.3 та виконаємо заміну змінних, наприклад y_2 на x_1 у рядку і x_1 на y_2 у стовпчику, із використанням наведеного вище алгоритму перетворення x_j в y_i і навпаки.

Таблиця 7.3

Вихідні дані

ПЕ БЗ	ВЕ	Коефіцієнти при вільних змінних		
		$x_1 \leftrightarrow y_2$	x_2	x_3
W	$2 \quad {}^{1}/_{2}$	$1 \quad {}^{1}/_{2}$	$0 \quad {}^{1}/_{2}$	$1 \quad 0$
y_1	$-5 \quad -{}^{1}/_{2}$	$-1 \quad -{}^{1}/_{2}$	$1 \quad -{}^{1}/_{2}$	$-2 \quad 0$
$y_2 \leftrightarrow x_1$	$1 \quad -{}^{1}/_{2}$	$-2 \quad -{}^{1}/_{2}$	$1 \quad -{}^{1}/_{2}$	$0 \quad 0$
y_3	$-1 \quad 0$	$0 \quad 0$	$-2 \quad 0$	$1 \quad 0$

Результат заміни змінних представлено у табл.7.4.

Таблиця 7.4

Результат заміни змінних

ПЕ БЗ	ВЕ	Коефіцієнти при вільних змінних		
		y_2	x_2	x_3
W	5/2	1/2	1/2	1
y_1	-11/2	-1/2	1/2	-2
x_1	-1/2	-1/2	-1/2	0
y_3	-1	0	-2	1

●

7.2. Алгоритм пошуку опорного розв‘язку за допомогою стандартних таблиць

Основна задача лінійного програмування у стандартній формі має вигляд:

$$W = c_0 - \left(\gamma_1 x_1 + \gamma_2 x_2 + \ldots \gamma_n x_n\right) \to \min_{\substack{x_i \ge 0 \left(i=\overline{1,n}\right) \\ y_j \ge 0 \left(j=\overline{1,m}\right)}}$$

за умови виконання обмежень рівностей:

$$y_j = b_j - \left(a_{j1} + a_{j2} + \ldots + a_{jn} x_n\right),\ j = \overline{1,m}\,.$$

В кожній вершині області допустимих розв‘язків розташовані опорні розв‘язки, для яких характерним є те, що хоча б n змінних дорівнюють 0, а інші змінні невід’ємні.

Покладемо $x_i = 0\ \left(i = \overline{1,n}\right)$, $y_j = b_j\ \left(j = \overline{1,m}\right)$. Тоді ознакою того, що записаний вище розв‘язок

$$\begin{cases} x_i = 0\ \left(i = \overline{1,n}\right), \\ y_j = b_j\ \left(j = \overline{1,m}\right) \end{cases}$$

є опорним, слід вважати невід’ємність вільних елементів $b_j \ge 0\ \left(j = \overline{1,m}\right)$.

Якщо ця умова не виконується, то необхідно перейти від розв‘язку, що не задовольняє область допустимих розв‘язків до опорного. Алгоритм пошуку опорного розв‘язку повинен призводити на кожній ітерації або до зменшення кількості від’ємних вільних коефіцієнтів обмежень рівностей, які задають рівняння для пошуку базисних змінних, або зменшувати модуль цих від’ємних коефіцієнтів, і тим самим, поступово, наближати базисні змінні до границі області допустимих розв‘язків.

Розглянемо алгоритм пошуку опорного розв‘язку, який використовує стандартні таблиці.

Прийом 1. Дослідити область допустимих розв‘язків на існування.

Знайти в стандартній таблиці від’ємний вільний елемент для будь-якої базисної змінної. Якщо в знайденому рядку немає від’ємних коефіцієнті при вільних змінних, то це ознака відсутності розв‘язку тому, що

$$y_j = b_j - \left(a_{j1} x_1 + a_{j2} x_2 + \ldots + a_{jn} x_n\right) < 0$$

при $b_j < 0$ та $x_i \ge 0$, а змінна y_j повинна бути невід’ємною. Виконати умову невід‘ємності неможливо.

Прийом 2. Пошук розв’язувального елементу.

Припустимо, що у рядку із від’ємним вільним коефіцієнтом є від’ємний коефіцієнт при деякій вільній змінній. Розглядаючи стовпець з від’ємним коефіцієнтом при деякій вільній змінній, знаходимо в ньому елемент однакового з вільним коефіцієнтом знаку, але такий, щоб відношення вільного коефіцієнта до нього було мінімальним. Цей елемент і є розв’язувальним.

Прийом 3. Виконуємо заміну змінних із застосуванням стандартних таблиць.

Розглянемо на прикладі застосування алгоритму пошуку опорного рішення із використанням стандартних таблиць.

Приклад 7.2. Знайти, якщо він існує, опорний розв‘язок основної задачі лінійного програмування із наступними рівняннями обмеженнями:

$$\begin{cases} y_1 = 1 - \left(-x_1 - 2x_2 + x_3\right), \\ y_2 = -5 - (-2x_1 + x_2 - x_3), \\ y_3 = 2 - \left(x_1 + x_2\right), \\ y_4 = 1 - \left(-x_2 + x_3\right). \end{cases}$$

❍Запишемо вихідні дані у вигляді симплекс таблиці.

Таблиця 7.5

Вихідні дані

БЗ	ВЕ	Коефіцієнти при вільних змінних		
		$x_1 \leftrightarrow y_3$	x_2	x_3
y_1	1 / 2	-1 / 1	-2 / 1	1 / 0
y_2	-5 / 4	-2 / 2	1 / 2	-1 / 0
y_3	2 / 1	1 / 1	1 / 1	0 / 0
y_4	1 / 0	0 / 0	-1 / 0	1 / 0

Обираємо другий рядок та перший стовпець, тому що у другому рядку є від’ємний коефіцієнт -2 та від’ємний вільний коефіцієнт -5. Зупинимось на стовпчику із коефіцієнтом -2 та обчислимо відношення

$$\frac{-5}{-2} = 2{,}5 \geq \frac{2}{1} = 2,$$

тобто в якості розв’язувального елементу вибираємо $a_{31} = 1$.

Заповнимо таблицю 7.6, яка буде таблицею першої ітерації до опорного розв’язку.

Повторно застосовуємо алгоритм пошуку опорного розв’язку із використанням стандартних таблиць (див. табл. 7.6).

Таблиця 7.6

Перша ітерація

БЗ	ВЕ	Коефіцієнти при вільних змінних		
		y_3	x_2	$x_3 \leftrightarrow y_2$
y_1	3 -1	1 2	-1 3	1 1
$y_2 \leftrightarrow x_3$	-1 1	2 -2	3 -3	-1 -1
x_1	2 0	1 0	1 0	0 0
y_4	1 -1	0 2	-1 3	1 1

Заповнимо табл.7.7, яка буде таблицею другої ітерації до опорного розв'язку

Таблиця 7.7

Друга ітерація

БЗ	ВЕ	Коефіцієнти при вільних змінних		
		y_3	x_2	y_2
y_1	2	3	2	1
x_3	1	-2	-3	-1
x_1	2	1	1	0
y_4	0	2	2	1

В табл. 7.7 всі вільні елементи невід'ємні. Опорний розв'язок знайдено.●

Приклад 5.3. Знайти, якщо він існує, опорний розв'язок задачі лінійного програмування із наступними рівняннями обмеженнями:

$$\begin{cases} y_1 = -4 - \left(-x_1 + 2x_2\right), \\ y_2 = -3 - \left(x_1 - x_2 + x_3\right), \\ y_3 = -10 - \left(2x_1 - x_2 + x_3\right), \\ y_4 = -2 - \left(-x_1 + x_2\right). \end{cases}$$

❍ Обираємо, наприклад, перший рядок та перший стовпець із від'ємним коефіцієнтом (табл. 7.8).

Як на початку розв'язання задачі лінійного програмування, так і в процесі отримання кожної наступної ітерації контролюємо виникнення умови, коли вільний елемент будь-якої базисної змінної від'ємний, а серед коефіцієнтів при вільних змінних немає від'ємного.

Таблиця 7.8

Вихідні дані

БЗ	ВЕ	Коефіцієнти при вільних змінних		
		$x_1 \leftrightarrow y_4$	x_2	x_3
y_1	-4 2	-1 -1	2 -1	0 0
y_2	-3 -2	1 1	-1 1	1 0
y_3	-10 -4	2 2	-1 2	1 0
y_4	-2 2	-1 -1	1 -1	0 0

Таблиця 7.9

Перша ітерація

БЗ	ВЕ	Коефіцієнти при вільних змінних		
		y_4	x_2	x_3
y_1	-2	-1	1	0
y_2	-5	1	0	1
y_3	-14	2	1	1
x_1	2	-1	-1	0

Так в рядках 2 та 3 маємо:

$$y_2 = -5 - \left(y_4 + x_3\right) < 0 \ ,$$

$$y_3 = -14 - \left(2y_4 + x_2 + x_3\right) < 0 \ \ \forall x_2, x_3 \geq 0, y_4 \geq 0.$$

Тобто розв‘язку, який належить ОДР за ознакою невід’ємності управляючих змінних, не існує. Отже, область допустимих розв‘язків – порожня множина. Як бачимо, з’ясування питання про існування допустимих розв‘язків відбувається в процесі пошуку опорного розв‘язку.●

Лекція 8. Алгоритм пошуку оптимального розв‘язку основної задачі лінійного програмування

8.1. Алгоритм пошуку оптимального розв‘язку за допомогою стандартних таблиць

Раніше було розглянуто основні складові (прийоми симплекс методу) пошуку оптимального опорного рішення задачі лінійного програмування, тобто задачі знаходження мінімуму показника ефективності:

$$W = c_0 - \left(\gamma_1 x_1 + \ldots + \gamma_n x_n\right) \to \min_{\substack{x_i \ge 0 \left(i=\overline{1,n}\right) \\ y_j \ge 0 \left(j=\overline{1,m}\right)}}$$

при виконанні рівнянь-обмежень

$$y_i = b_j - \left(a_{j1} x_1 + \ldots + a_{jn} x_n\right),\ j = \overline{1,m}.$$

Покажемо на прикладі, як у зручній формі табличного алгоритму заміни змінних виконувати оптимізацію, тобто розв'язувати основу задачу лінійного програмування в цілому.

Приклад 8.1. Розв'язати основну задачу лінійного програмування

$$W = 0 - \left(-x_1 + 2x_2 + x_3\right) \to \min_{\substack{x_{1,2,3} \ge 0 \\ y_{1,2,3,4} \ge 0}}$$

при виконанні рівнянь-обмежень

$$\begin{cases} y_1 = 2 - (x_1 + x_2 - 2x_3), \\ y_2 = 1 - \left(x_1 - x_2 + x_3\right), \\ y_3 = 5 - \left(x_2 + x_3\right), \\ y_4 = 2 - (2x_1 - x_2). \end{cases}$$

❍При $x_1 = x_2 = x_3 = 0$, $y_1 = 2, y_2 = 1, y_3 = 5, y_4 = 2$, тобто опорний розв‘язок існує, але він не є оптимальним, тому що коефіцієнти при x_2 та x_3, у виразі для W – додатні. Це означає, що за рахунок збільшення x_2 або x_3 можливо зменшити значення показника ефективності. Перейдемо до застосування алгоритму стандартних таблиць (табл. 8.1).

В першому рядку табл.8.1, що відображає вихідні данні, показник ефективності має додатні коефіцієнти при вільних змінних x_2 та x_3. Тому можливо обрати будь який із стовбців в якості розв'язувального. Нехай це буду стовбець з x_3, тобто x_3 виключимо із складу вільних змінних. Розв'язувальний елемент в обраному стовбці повинен буди додатний. Таких коефіцієнтів два: один у рядку з y_2 та інший у рядку з y_3. Оберемо як розв'язувальний елемент той, для якого відношення вільного елементу до елементу обраного в якості розв'язувального найменше.

Таблиця 8.1

Вихідні дані

ПЕ БЗ	ВЕ	Коефіцієнти при вільних змінних		
		x_1	x_2	$x_3 \leftrightarrow y_2$
W	0 -1	-1 -1	2 1	1 -1
y_1	2 2	1 2	1 -2	-2 2
$y_2 \leftrightarrow x_3$	1 1	1 1	-1 -1	1 1
y_3	5 -1	0 -1	1 1	1 -1
y_4	2 0	2 0	-1 0	0 0

Таблиця 8.2

Перша ітерація

ПЕ БЗ	ВЕ	Коефіцієнти при вільних змінних		
		x_1	$x_2 \leftrightarrow y_3$	y_2
W	-1 -6	-2 3/2	3 -3/2	-1 3/2
y_1	4 2	3 -1/2	-1 1/2	2 -1/2
x_3	1 2	1 -1/2	-1 1/2	1 -1/2
$y_3 \leftrightarrow x_2$	4 2	-1 -1/2	2 1/2	-1 -1/2
y_4	2 2	2 -1/2	-1 1/2	0 -1/2

В таблиці, отриманій після першої ітерації (табл. 8.2), у рядку «Показник ефективності» тільки при x_2 коефіцієнт додатний, тому x_2 необхідно вивести зі складу вільних змінних. В якості розв’язувального елементу вибираємо $a_{32} = 2$. В подальшому, для отримання другої та третьої ітерацій, застосовуємо табличний алгоритм заміни базисних змінних (табл. 8.3 та 8.4).

Таблиця 8.3

Друга ітерація

ПЕ БЗ	ВЕ	Коефіцієнти при вільних змінних		
		x_1	y_3	$y_2 \leftrightarrow y_1$
W	-7 -2	-1/2 -5/6	-3/2 -1/6	1/2 -1/3
$y_1 \leftrightarrow y_2$	6 4	5/2 5/3	1/2 1/3	3/2 2/3
x_3	3 -2	1/2 -5/6	1/2 -1/6	1/2 -1/3
x_2	2 2	-1/2 5/6	1/2 1/6	-1/2 1/3
y_4	4 2	3/2 5/6	1/2 1/6	-1/2 1/3

Таблиця 8.4.

Третя ітерація

ПЕ БЗ	ВЕ	Коефіцієнти при вільних змінних		
		x_1	y_3	y_1
W	-9	-4/3	-5/3	-1/3
y_2	4	5/3	1/3	2/3
x_3	1	-1/3	1/3	-1/3
x_2	4	1/3	2/3	1/3
y_4	6	7/3	2/3	1/3

Як бачимо опорний розв’язок

$$x_1 = 0,\ x_2 = 4,\ x_3 = 1,\ y_1 = 0,\ y_2 = 4,\ y_3 = 0,\ y_4 = 6$$

є оптимальним, тому що всі коефіцієнти при вільних змінних у виразі для обчислення показника ефективності від’ємні. Остаточно отримаємо оптимальне(мінімальне) значення показника ефективності

$$\mathrm{W}_{\min} = -9. \bullet$$

Зауваження 8.1. Якщо у кожному стовпці із додатнім коефіцієнтом у рядку показника ефективності не існує додатного елемента, то це означає, що не дивлячись на існування області допустимих розв‘язків обмеженого розв’язку задача не має: при необмеженому збільшенні вільної змінної показник ефективності $W \to -\infty$.

Алгоритм пошуку оптимального розв'язку основної задачі лінійного програмування із використанням симплекс методу:

1. Якщо всі вільні елементи (виключаючи рядок W) у симплекс таблиці (стандартній таблиці) невід'ємні і при цьому у рядку W (без урахування вільного елемента) немає жодного додатного коефіцієнта, то оптимальний розв'язок досягнуто.
2. Якщо в рядку W серед коефіцієнтів вільних змінних є додатній, а у стовпці, що йому відповідає, немає жодного додатного коефіцієнту (елементу) , то лінійна функція W не обмежена знизу і оптимального розв'язку не існує.
3. Якщо у стовпці є додатні коефіцієнти, то вільну змінну, яка відповідає цьому стовпцю необхідно поміняти місцями із тією базисною змінною, для якої відношення вільного елементу до обраного додатного коефіцієнту є найменшим.

Зауваження 8.2. При розв'язанні задачі лінійного програмування може виникнути ситуація коли при заміні змінних не відбувається зміни значення показника ефективності (вироджена задача лінійного програмування). Для припинення процедури пошуку W_{min} можливо використати ознаку того, що показник ефективності залишається незмінним після спеціально встановленої кількості циклів, або ознаку того, що показник ефективності вже неможливо зменшити, тому що немає додатних коефіцієнтів у рядку W симплекс таблиці.

Зауваження 8.3. При розв'язанні задачі лінійного програмування може виникнути зациклювання, суть якого полягає в тому, що після декількох замін змінних відбувається повернення до початкових значень опорного розв'язку та показника ефективності. Для подолання зациклювання необхідно після першого повернення до початкових умов оптимізації замінити розв'язувальний елемент.

Приклад 8.2. Розв'язати основну задачу лінійного програмування

$$W = 2x_1 - x_2 \to \min_{\substack{x_{1,2,3,4}\ge 0\\ y_{1,2,3}\ge 0}}$$

при виконанні рівнянь-обмежень

$$\begin{cases} y_1 = x_1 - x_2, \\ y_2 = -x_2 + x_3 + 2, \\ y_3 = x_3 + x_4 + 1. \end{cases}$$

❍ 1) Перейдемо до стандартної форми задачі лінійного програмування:

$$W = 0 - \left(-2x_1 + x_2\right) \to \min_{\substack{x_{1,2,3,4} \geq 0 \\ y_{1,2,3} \geq 0}}$$

$$\begin{cases} y_1 = 0 - \left(-x_1 + x_2\right), \\ y_2 = 2 - \left(x_2 - x_3\right), \\ y_3 = 1 - \left(-x_3 - x_4\right). \end{cases}$$

2) Із стандартної форми задачі лінійного програмування (див. п.1) випливає, що вільні коефіцієнти невід'ємні. Тобто при $x_1 = x_2 = x_3 = 0$ має місце опорний розв'язок. Таким чином, опорний розв'язок шукати не потрібно. Переходимо до алгоритму покращення значень показника ефективності на опорних розв'язках, тобто до алгоритму пошуку оптимального розв'язку за допомогою стандартних таблиць (табл. 8.5 та 8.6).

Таблиця 8.5

Вихідні дані

ПЕ БЗ	ВЕ	Коефіцієнти при вільних змінних			
		x_1	$x_2 \leftrightarrow y_1$	x_3	x_4
W	0 0	-2 1	1 -1	0 0	0 0
$y_1 \leftrightarrow x_2$	0 0	-1 -1	1 1	0 0	0 0
y_2	2 0	0 1	1 -1	-1 0	0 0
y_3	1 0	0 0	0 0	-1 0	-1 0

Таблиця 8.6

Перша ітерація

ПЕ БЗ	ВЕ	Коефіцієнти при вільних змінних			
		x_1	y_1	x_3	x_4
W	0	-1	-1	0	0
x_2	0	-1	1	0	0
y_2	2	1	-1	-1	0
y_3	1	0	0	-1	-1

Як бачимо:

1. Показник ефективності не змінився і дорівнює $W = 0$;

2. У рядку із W у таблиці вже не існує додатних коефіцієнтів.

Тому робимо висновок, що оптимальний розв‘язок досягнуто в точці

$$x_1 = x_2 = x_3 = x_4 = y_1 = 0, x_2 = 2, y_3 = 1.\bullet$$

8.2. Розв'язання задачі лінійного програмування із використанням системи комп'ютерної математики Mathlab

Постановка задачі:

$$W = f^T \cdot X \to \min_{X \in G} ,$$

де ОДР G задана системою нерівностей

$$\begin{cases} A \cdot X \le b, \\ A_{eq} \cdot X = b_{eq}, \\ lb \le X \le Ub. \end{cases}$$

у якій прийняті наступні позначення:

X – вектор змінних, за якими виконується оптимізація;

f – матриця-стовпець коефіцієнтів показника ефективності;

A – прямокутна матриця;

b – матриця-стовпець вільних елементів (коефіцієнтів) у обмеженнях нерівностях;

A_{eq} – прямокутна матриця;

b_{eq} – матриця-стовпець вільних елементів (коефіцієнтів) у обмеженнях рівностях;

lb, Ub – матриці-стовпці покомпонентних обмежень вектора X.

Методика застосування функції $linprog()$:

% виконати математичну постановку прикладної задачі;

% задати матриці $f, A, b, A_{eq}, b_{eq}, lb, Ub$;

% звернення до функції розв'язок $linprog()$ та отримання результату

$$\left[X, W\right] = linprog\left(f, A, b, A_{eq}, b_{eq}, lb, Ub\right)$$

% якщо якесь із обмежень не задається, то замість цих аргументів при використанні функції $linprog()$ задають порожні масиви у вигляді [].

Лекція 9. Транспортна задача лінійного програмування

9.1. Постановка транспортної задачі

Симплекс-метод – універсальний метод розв'язку задачі лінійного програмування, але існує клас задач лінійного програмування, які дозволяють отримати розв'язок простішими методами. Найбільш відомою є *транспортна задача*. Класична транспортна задача лінійного програмування формулюється наступним чином: існує m пунктів відправлення $A_1, A_2, \ldots, A_m$, в яких зосереджено запаси деякого товару у кількості відповідно $a_1, a_2, \ldots, a_m$ одиниць. Крім того є n пунктів призначення $B_1, B_2, \ldots, B_n$, що надали заявки на отримання товару відповідно у кількості $b_1, b_2, \ldots, b_n$ одиниць. Вважаємо, що сума усіх заявок дорівнює сумі усіх товарів:

$$\sum_{i=1}^{m} a_i = \sum_{j=1}^{n} b_j. \tag{9.1}$$

Відомою є вартість c_{ij} – перевезення одиниці товару від кожного пункту відправлення A_i до кожного пункту призначення B_j. Таблиця вартості перевезення відома:

$$C = \begin{bmatrix} c_{11} & \cdots & c_{1n} \\ \vdots & \ddots & \vdots \\ c_{m1} & \cdots & c_{mn} \end{bmatrix}. \tag{9.2}$$

Необхідно скласти такий план перевезень, при якому всі заявки були б виконані і при цьому загальна вартість всіх перевезень була б мінімальною. При такій постановці задачі показником ефективності плану перевезень є вартість. Тому поставлену задачу називають транспортною задачею за критерієм вартості.

Дамо математичну постановку цієї задачі. Позначимо x_{ij} $\left(x_{ij} \geq 0\right)$ – кількість товару, який відправляється із i-го пункту відправлення A_i до j-го пункту призначення B_j $\left(i = \overline{1, m},\ j = \overline{1, n}\right)$. На змінні x_{ij} $\left(i = \overline{1, m},\ j = \overline{1, n}\right)$ накладаються наступні умови (обмеження):

1) Сумарна кількість товару, що відправляється із кожного пункту відправлення в усі пункти призначення повинна дорівнювати запасу в даному пункті.

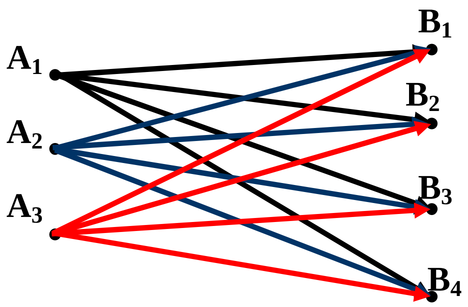


Рис.9.1. Графічний образ пояснення стосовно виникнення обмежень

Ця умова дає m рівнянь-обмежень:

$$\begin{cases} \sum_{j=1}^{n} x_{1j} = a_1, \\ \dots\dots\dots\dots\dots \\ \sum_{j=1}^{n} x_{mj} = a_m. \end{cases} \tag{9.3}$$

2) Сумарна кількість товару, що поступає в кожен пункт призначення із усіх пунктів відправлення повинна дорівнювати замовленню, що надійшло від даного пункту. Ця умова дає n рівнянь-обмежень.

$$\begin{cases} \sum_{i=1}^{m} x_{i1} = b_1, \\ \dots\dots\dots\dots\dots \\ \sum_{i=1}^{m} x_{in} = b_n. \end{cases} \tag{9.4}$$

3) Сумарна вартість всіх перевезень повинна бути мінімальною

$$W = c_{11}x_{11} + \ldots + c_{mn}x_{mn} = \sum_{i=1}^{m}\sum_{j=1}^{n} c_{ij}x_{ij} \quad \to \min_{x_{ij}\geq 0\ \left(i=\overline{1,m},\, j=\overline{1,n}\right)} \tag{9.5}$$

Показник ефективності (9.5), рівняння-обмеження (9.3), (9.4) є лінійними функціями відносно змінних x_{ij} $\left(i=\overline{1,m},\ j=\overline{1,n}\right)$, тобто задача про транспортні перевезення за вартісним критерієм поставлена як основна задача лінійного програмування. Як і будь яку задачу лінійного програмування, транспортну задачу лінійного програмування можливо розв’язати із використанням симплекс-методу розв’язання задачі лінійного програмування, але розмірність цієї задачі, яка дорівнює кількості змінних параметрів, які потрібно знайти, щоб задовольнити (9.5), дорівнює $m \times n$. Це означає, що при $m = 10$ та $n = 10$ (всього 10 складів та 10 пунктів прийому товару) отримаємо 100 змінних. Для досить простої транспортної задачі лінійного програмування різко зростає кількість невідомих і традиційний, розглянутий вище, симплекс-метод потребує значних обчислювальних потужностей.

Враховуючи той факт, що всі коефіцієнти при змінних в обмеженнях (9.3) та (9.4) дорівнюють 1 або 0 можливо спростити процедуру обчислення оптимального розв‘язку. Перед початком викладення методу розв’язання транспортної задачі з’ясуємо особливості системи рівнянь обмежень (9.3) та (9.4).

Якщо скласти ліві і праві частини цих обмежень, то отримаємо рівняння: сума запасів дорівнює сумі замовлень. Тобто всього лінійно незалежних обмежень в (9.3) та (9.4) буде $(m+n-1)$. Це означає, що ранг системи рівнянь-обмежень дорівнює $r = m+n-1$. Тому можливо розв’язати ці рівняння відносно $(m+n-1)$ базисної змінної, вважаючи інші

$$k = mn - (m+n-1) = m(n-1) - (n-1) = (m-1)(n-1)$$

вільними змінними.

Раніше було показано, що в задачі лінійного програмування оптимальний розв‘язок досягається в одній із вершин області допустимих розв’язків і при цьому хоча б k змінних перетворюються в 0. Це означає, що у випадку транспортної задачі $(m-1)(n-1)$ змінних теж повинні дорівнювати 0 у вершині ОДР (опорній точці). При розв’язанні транспортної задачі зазвичай використовують наступну термінологію:

- перевезення x_{ij} – кількість одиниць вантажу, що перевозять із пункту відправлення A_i до пункту прийому B_j;
- план перевезень $\left(x_{ij}\right)$ – будь-яка сукупність значень x_{ij};
- допустимий план $\left(x_{ij}\right)_{\text{д}}$ – план, який задовольняє балансовим умовам: усі заявки задоволені, усі запаси вичерпані, при цьому $x_{ij} \gg 0$;
- опорний план $\left(x_{ij}\right)_{\text{оп}}$ – допустимий план, в якому відмінні від 0 не більше, ніж $r = m + n - 1$ базисних перевезень x_{ij}, а всі інші дорівнюють 0;
- оптимальний план $\left(x_{ij}\right)_{\text{опт}}$ – опорний план, при якому досягається найменше значення показника ефективності (вартості перевезень).

Перейдемо до викладення методів розв’язання транспортної задачі. Ці методи не потребують використання симплекс-таблиць, але використовують так звану транспортну таблицю (табл. 9.1). До неї записують:

– пункти відправлення і пункти призначення;
– запаси, розташовані у пунктах відправлення;
– заявки, розташовані у пунктах призначення;
– вартість перевезень з кожного пункту відправлення в кожний пункт призначення одиниці вантажу.

Зразок транспортної таблиці — Таблиця 9.1

ПВ	Пункти призначення					Запаси
	B_1	B_2	B_3	…	B_n	a_i
A_1	c_{11}	c_{12}	c_{13}	…	c_{1n}	a_1
A_2	c_{21}	c_{22}	c_{23}	…	c_{2n}	a_2
…	…	…	…	…	…	…
A_m	c_{m1}	c_{m2}	c_{m3}	…	c_{mn}	a_m
b_j	b_1	b_2	b_3	…	b_n	$\sum b = \sum a$

Вартість перевезень c_{ij} розташуємо у правому верхньому куті кожного віконця, а всередині цього віконця будемо записувати перевезення x_{ij}. Пам'ятаючи визначення опорного плану, можливо стверджувати, що у $r = m + n - 1$ віконці будуть записані числа, відмінні від 0. Ці віконця будемо називати базисними, а інші (пусті або нульові) – вільними. Таким чином, розв'язок транспортної задачі можливо знайти, якщо заповнити транспортну таблицю і при цьому виконати наступні умови:

- сума перевезень у кожному рядку транспортної таблиці дорівнює запасу даного пункту відправлення;
- сума перевезень у кожному стовпці транспортної таблиці дорівнює заявці пунктів прийому;
- загальна вартість перевезень повинна бути мінімальною.

Всі наступні дії, пов'язані із пошуком оптимального плану будуть використовувати зручну (наочну) форму представлення інформації про транспортну задачу у вигляді транспортної таблиці.

9.2. Табличний метод пошуку опорного плану (метод "північно-західного кута")

Пошук розв‘язку транспортної задачі, як і будь-якої задачі лінійного програмування, починається із знаходження *опорного розв‘язку* або *опорного плану*. На відміну від загального випадку основної задачі лінійного програмування із довільними обмеженнями і показником ефективності, розв'язок транспортної задачі завжди існує. Дійсно, із фізичних міркувань зрозуміло, що хоч якесь допустиме рішення (допустимий план) повинно існувати. Пояснимо на конкретному прикладі основні прийоми методу пошуку опорного розв‘язку, який отримав назву *метод "північно-західного кута"* (табл.9.2).

Таблиця 9.2

Вихідні дані та приклад побудови опорного розв‘язку

ПВ	Пункти призначення					Запаси
	B_1	B_2	B_3	B_4	B_5	a_i
A_1	10 18	8 27	5 3	6	9	48
A_2	6	7	8 30	6	5	30
A_3	8	7	10 9	8 12	7 6	27
A_4	7	5	4	6	8 20	20
b_j	18	27	42	12	26	125

Припустимо, що умова транспортної задачі задана за допомогою транспортної таблиці. Необхідно знайти опорний розв‘язок (побудувати опорний план). Розв'язання задачі починаємо із лівої верхньої позиції верхнього віконця із координатами $(1;1)$, що відповідає образному сприйняттю цієї позиції як «північно-західного» напрямку на географічній карті північної півкулі. За рахунок пункту відправлення A_1 задовольняємо вимогу пункту призначення B_1 і записуємо перевезення у віконце із координатами $(1;1)$. Потім задовольняємо запит пункту призначення B_2 і далі записуємо залишок у віконце із координатами $(1;3)$. Недовиконання заявки пункту призначення B_3 за рахунок пункту відправлення A_1 доповнюємо поставками із A_2 та A_3 і т.д. Віконця таблиці, в яких стоять ненульові перевезення є базовими. Їх кількість повинна дорівнювати

$$r = m + n - 1 = 3 + 5 - 1 = 8.$$

Вільні (порожні) означають 0 перевезень, тобто це вільні змінні, яких повинно бути щонайменше

$$k = (m-1)(n-1) = (4-1)(5-1) = 12.$$

Таким чином, можна стверджувати, що опорний розв'язок знайдено. Обчислимо його вартість:

$$W = 18 \cdot 10 + 27 \cdot 8 + 3 \cdot 5 + 30 \cdot 8 + 9 \cdot 10 + 12 \cdot 8 + 6 \cdot 7 + 20 \cdot 8 = 1039.$$

Зрозуміло, що знайдене значення W не є оптимальним, хоча б тому, що не використовувався алгоритм покращення опорного розв'язку. Перейдемо до алгоритму покращення опорного плану. Виконаємо так звану циклічну перестановку, наприклад, 18 одиниць вантажу із віконця з координатами $(1;1)$ перемістимо у віконце із координатами $(1;3)$ і далі через віконце $(2;3)$ до $(2;1)$, зберігаючи балансові співвідношення (табл. 9.3).

Таблиця 9.3

Покращення опорного плану

ПВ	Пункти призначення					Запаси
	B_1	B_2	B_3	B_4	B_5	a_i
A_1	10	8 27	5 21	6	9	48
A_2	6 18	7	8 12	6	5	30
A_3	8	7	10 9	8 12	7 6	27
A_4	7	5	4	6	8 20	20
b_j	18	27	42	12	26	125

Показник ефективності для транспортної таблиці набуває значення:

$$W = 18 \cdot 6 + 27 \cdot 8 + 21 \cdot 5 + 12 \cdot 8 + 9 \cdot 10 + 12 \cdot 8 + 6 \cdot 7 + 20 \cdot 8 = 913,$$

що на 126 одиниць менше показника ефективності, розрахованого для попереднього опорного плану. Наведений в табл. 9.3, як приклад, спосіб циклічної перестановки лежить в основі алгоритму оптимізації плану перевезень. Перед тим, як перейти до алгоритму оптимізації, розглянемо так званий "вироджений" план транспортної задачі, коли нулю дорівнюють не лише вільні змінні, але й частина базисних (табл. 9.4).

Кількість базисних змінних згідно заданої транспортної таблиці повинна дорівнювати $r = 4 + 5 - 1 = 8$, але в отриманому за допомогою метода "північно-західного" кута плані перевезень ненульовими є 6 змінних, тобто виродженою є задача, в якій деякі з базисних змінних дорівнюють 0. Для побудови стійкого алгоритму обчислення оптимального розв‘язку бажано на кожній ітерації мати всі базисні змінні ненульовими.

Таблиця 9.4

Приклад "виродженого" плану

ПВ	Пункти призначення					Запаси
	B_1	B_2	B_3	B_4	B_5	a_i
A_1	10	10				20
A_2			20	10		30
A_3				25		25
A_4					20	20
b_j	10	10	20	35	20	95

З метою уникнення ситуації виникнення "виродженого" плану достатньо штучно на незначну величину ε змінити запаси і заявки (табл. 9.5), а потім, після знаходження оптимального плану, покласти $\varepsilon = 0$.

Таблиця 9.5

Приклад уникнення "виродженого" плану

ПВ	Пункти призначення					Запаси
	B_1	B_2	B_3	B_4	B_5	a_i
A_1	10	10	ε			$20 + \varepsilon$
A_2			$20 - \varepsilon$	$10 + \varepsilon$		30
A_3				$25 - \varepsilon$	ε	25
A_4					$20 - \varepsilon$	$20 - \varepsilon$
b_j	10	10	20	35	20	95

В отриманій таблиці маємо розрахункову кількість ненульових базисних змінних. Вона дорівнює 8. В двох наведених вище таблицях не вказана вартість перевезень одиниці товару c_{ij}, тому що нас цікавив лише приклад побудови опорного плану та прийоми уникнення появи "виродженого" плану.

9.3. Метод пошуку оптимального плану шляхом циклічного перерахунку (метод циклічного перерахунку або розподільчий метод)

Циклом у транспортній таблиці будемо називати декілька віконець, з’єднаних між собою уявною неперервною ламаною лінією, яка у кожному спеціально обраному віконці розвертається на 90° таким чином, щоб утворити із початковим віконцем замкнену лінію. Так, наприклад, в таблиці 9.6 показано 2 цикли:

–Ц1 – з 4 вершинами

–Ц2 – з 6 вершинами

Таблиця 9.6

Пояснення щодо методу циклічного перерахунку

ПВ	Пункти призначення					Запаси
	B_1	B_2	B_3	B_4	B_5	a_i
A_1	c_{11} -	c_{12}	c_{13} +	+ c_{14}	c_{15} -	a_1
A_2	c_{21} +	c_{22}	c_{23} -	c_{24}	c_{25}	a_2
A_3	c_{31}	c_{32} +	c_{33}	c_{34} –	c_{35}	a_3
A_4	c_{41}	c_{42} -	c_{43}	c_{44}	c_{45} +	a_4
b_j	b_1	b_2	b_3	b_4	b_5	$\sum_{i=1}^{m} a_i = \sum_{j=1}^{n} b_j$

Кожен цикл має парну кількість вершин та сторін, тобто ланцюгів, що зв’язують ці вершини. Вершини позначимо «$+$» або «$-$» в залежності від збільшення, або зменшення перевезень. Цикл з позначеними вершинами будемо *називати перенесеннями*. Перенести кількість одиниць вантажу по означеному циклу, означає збільшити перевезення, що стоять у додатних вершинах циклу, на деяку кількість одиниць вантажу, а перевезення, що стоять у від’ємних вершинах – зменшити на ту саму кількість. Зрозуміло, що при будь-якому циклічному перенесенні, що зберігає невід’ємність перевезень виконується закон збереження матерії, тобто зберігаються балансові умови, але вартість плану змінюється. Вартістю циклу називається зміна вартості перевезень при переміщенні одиниці вантажу по означеному циклу. Кількісна вартість циклу дорівнює алгебраїчній сумі вартостей, що розташовані у вершинах циклу, із врахуванням знаку вершин. Наприклад, вартість Ц1 та Ц2 можливо обчислити:

$$V_{\text{ц}1} = c_{21} - c_{23} + c_{13} - c_{11},$$

$$V_{\text{ц}2} = c_{32} - c_{42} + c_{45} - c_{15} + c_{14} - c_{34}.$$

Якщо вартість циклу від’ємна, то переміщення по цьому циклу ΔX одиниці вантажу призводить до зменшення загальної вартості перевезень на $\Delta x \cdot V_{\text{ц}}$ одиниць.

Процедура зменшення загальної вартості припиняється тоді, коли циклів із від'ємною вартістю вже не буде. Тобто покращення плану перевезень бути не може. Оптимальний план досягнуто.

При покращенні плану перевезень за рахунок циклічних перенесень, як правило, користуються прийомами симплекс-методу: на кожній ітерації, присвяченій зміні циклу замінюють одну базисну змінну вільною і навпаки; загальна кількість базисних змінних залишається незмінною. Доведено, що для будь-якого вільного віконця транспортної таблиці завжди існує цикл і при тому єдиний, одна з вершин якого лежить у цьому вільному віконці, а всі інші у базисних віконцях.

Приклад 9.1. Розглянемо комплексне застосування методу північно-західного кута та цикличного перерахунку для пошуку оптимального плану транспортної задачі, яка задана за допомогою транспортної таблиці 9.7.

Таблиця 9.7

Вихідні дані прикладу 9.1

ПВ	Пункти призначення				Запаси
	B_1	B_2	B_3	B_4	a_i
A_1	- 10 22 Ц1	+ 7 9	6	8	31
A_2	+ 5	- 6 25	5 23	4	48
A_3	8	7	6 18	7 20	38
b_j	22	34	41	20	117

❍**1.** Складемо опорний план методом північно-західного кута. Кількість базисних змінних повинна дорівнювати

$$r = m + n - 1 = 6,$$

що відповідає плану, наведеному у таблиці 9.1. Задача «не вироджена».
Обчислимо значення показника ефективності

$$W = 22 \cdot 10 + 9 \cdot 7 + 25 \cdot 6 + 23 \cdot 5 + 18 \cdot 6 + 20 \cdot 7 = 796.$$

2. Покращимо план (табл. 9.8), замінивши вільну клітинку з координатами $(2;1)$. Вартість циклу

$$V_{ц1} = 5 + 7 - 10 - 6 = -4.$$

Покращимо план за рахунок переміщення 22 одиниць вантажу, інакше у $(1;1)$ буде від'ємне число, чого неможливо допустити, бо від'ємних перевезень не може бути.

Покращений план має показник ефективності

$$W_1 = 796 - 22 \cdot 4 = 708.$$

Таблиця 9.8

Покращений план за циклом Ц1

ПВ	Пункти призначення				Запаси
	B_1	B_2	B_3	B_4	a_i
A_1	10	7 31	6	8	31
A_2	5 22	6 3	- 5 23 Ц2	+ 4	48
A_3	8	7	+ 6 18	- 7 20	38
b_j	22	34	41	20	117

Сформулюємо правило підбору циклу, що є чотирикутником:

1). Лише в одній вершині розташоване вільне віконце;

2). У вільному віконці завжди «+»;

3). Суми вартості по діагоналі, що включає вільне віконце менше за суму вартості іншої діагоналі.

3. Друга ітерація. Обираємо цикл Ц2 (див. табл. 9.8) з вільним віконцем із координатами (2; 4):

$$V_{ц2} = 4 + 6 - 5 - 7 = -2,$$

$$\Delta x \cdot V_{ц2} = 20 \cdot (-2) = -40,$$

$$W_2 = 708 - 40 = 668.$$

Результат перенесень за Ц2 представлено у табл. 9.9.

Таблиця 9.9

Покращений план за циклом Ц2

ПВ	Пункти призначення				Запаси
	B_1	B_2	B_3	B_4	a_i
A_1	10	7 31	6	8	31
A_2	5 22	6 3	5 3	4 20	48
A_3	8	7	6 38	7	38
b_j	22	34	41	20	117

План, наведений в табл. 9.9 є оптимальним, тому що всі цикли побудовані для вільних клітинок, мають додатну вартість.●

Отже, розподільчий (циклічний) метод полягає в безпосередньому пошуку вільних клітинок транспортної таблиці із від'ємною вартістю циклу і в перенесенні перевезень за цим циклом.

Приклад 9.2. Знайти оптимальний план перевезень для транспортної задачі заданої транспортною таблицею 9.10.

Таблиця 9.10

Вихідні дані для прикладу 9.2

ПВ	Пункти призначення			Запаси
	B_1	B_2	B_3	a_i
A_1	- 10 20 Ц1	7 20	+ 6 ε	40 + ε
A_2	+ 5	6	- 5 23	23
A_3	8	7	6 20 − ε	20 − ε
b_j	20	20	43	83

❍1. Будуємо опорний план методом північно-західного кута та перевіряємо його на невиродженість:

$$r = n + m - 1 = 3 + 3 - 1 = 5 > 4 .$$

Опорний план вироджений.

2. Для уникнення виродженості використовуємо метод збурень запасів та заявок на величину $\pm\varepsilon$, відповідно у першому та третьому рядках.

3. Розподільчим методом покращуємо розв‘язок, використовуючи цикл Ц1 (табл. 9.11):

$$V_{\text{ц}1} = 5 + 6 - 10 - 5 = -4,$$

$$\Delta x \cdot V_{\text{ц}1} = 20 \cdot \left(-4\right) = -80 .$$

Таблиця 9.11

Покращення плану із використанням Ц1

ПВ	Пункти призначення			Запаси
	B_1	B_2	B_3	a_i
A_1	10	7 20	6 20 + ε	40 + ε
A_2	5 20	6	5 3	23
A_3	8	7	6 20 − ε	20 − ε
b_j	20	20	43	83

Отримана транспортна таблиця (див. табл. 9.11) відображає оптимальний план перевезень. Циклів з від’ємною вартістю по відношенню до вільних клітинок немає. Показник ефективності дорівнює

$$W = \left(20 \cdot 5 + 20 \cdot 7 + \left(20 + \varepsilon\right) \cdot 6 + \left(20 - \varepsilon\right) \cdot 6 + 3 \cdot 5\right)\Big|_{\varepsilon=0} = 495 . \bullet$$

Лекція 10. Метод потенціалів розв'язання транспортної задачі

10.1. Обґрунтування методу потенціалів

Розглянутий у лекції 9 розподільчий метод розв'язання транспортної задачі має суттєвий недолік: необхідно знаходити цикли для всіх вільних віконець і обчислювати їх вартість. Для зменшення кількості обчислень за рахунок обчислення вартості лише тих циклів, де ця вартість від'ємна запропоновано *метод потенціалів*.

Ідея методу полягає в наступному. Припустимо, що задана транспортна задача у вигляді

$$W = \sum_{i=1}^{m}\sum_{j=1}^{n} c_{ij}x_{ij} \to \min_{x_{ij} \geq 0\ (i=\overline{1,m};\, j=\overline{1,n})} \tag{10.1}$$

за умови

$$\begin{cases} \sum_{j=1}^{n} x_{ij} = a_i, \\ \sum_{i=1}^{m} x_{ij} = b_j, \end{cases} \tag{10.2}$$

при цьому

$$\sum_{j=1}^{n} b_j = \sum_{i=1}^{m} a_i. \tag{10.3}$$

Уявимо, що кожен з пунктів відправлення A_i та призначення (прийому) B_j сплачують певні кошти у розмірі відповідно α_i та β_j. Позначимо суму коштів $\alpha_i + \beta_j = \widetilde{c_{ij}}$.

Підкреслимо, що α_i, β_j, $\widetilde{c_{ij}}$ – уявні абстрактні вартості і платежі, тобто вони можуть бути як додатними, так і від'ємними або нульовими. Для скорочення запису позначимо всю сукупність платежів (α_i, β_j) . Тоді матриця псевдовартостей матиме вигляд:

$$\begin{bmatrix} \widetilde{c_{11}} = \alpha_1 + \beta_1 & \widetilde{c_{12}} = \alpha_1 + \beta_2 & \ldots & \widetilde{c_{1n}} = \alpha_1 + \beta_n \\ \widetilde{c_{21}} = \alpha_2 + \beta_1 & \widetilde{c_{22}} = \alpha_2 + \beta_2 & \ldots & \widetilde{c_{2n}} = \alpha_2 + \beta_n \\ \ldots & \ldots & \ldots & \ldots \\ \widetilde{c_{m1}} = \alpha_m + \beta & \widetilde{c_{m2}} = \alpha_m + \beta_2 & \ldots & \widetilde{c_{mn}} = \alpha_m + \beta_n \end{bmatrix}.$$

Доведемо теорему про платежі.

Теорема 10.1. Для заданої сукупності платежів (α_i, β_j), загальна псевдовартість перевезень не залежить від плану перевезень, а залежить лише від сукупності платежів

$$\widetilde{W} = \sum_{i=1}^{m}\sum_{j=1}^{n}\widetilde{c}_{ij}x_{ij}.$$

►Розглянемо вираз для обчислення псевдо вартості:

$$\widetilde{W} = \sum_{i=1}^{m}\sum_{j=1}^{n}\widetilde{c}_{ij}x_{ij} = \sum_{i=1}^{m}\sum_{j=1}^{n}(\alpha_i + \beta_j)x_{ij} = \sum_{i=1}^{m}\sum_{j=1}^{n}(\alpha_i)x_{ij} + \sum_{i=1}^{m}\sum_{j=1}^{n}(\beta_j)x_{ij} =$$

$$= \sum_{i=1}^{m}\alpha_i\sum_{j=1}^{n}x_{ij} + \sum_{j=1}^{n}\beta_j\sum_{i=1}^{m}x_{ij} = \sum_{i=1}^{m}\alpha_i a_i + \sum_{j=1}^{n}\beta_j b_j.$$

Як бачимо загальна псевдовартість перевезень не залежить від допустимого плану перевезень, тобто такого плану, що задовольняє системі балансових обмежень і умов $x_{ij} \geq 0$, а залежить лише від запасів a_i, заявок b_j та платежів (α_i, β_j), що і треба було довести◄

З'ясуємо зв'язок між псевдовартістю $\widetilde{c}_{ij}$ перевезення та її істинним значенням . Припустимо, що план перевезень (x_{ij}) невироджений: кількість базисних віконець з додатними перевезеннями у таблиці перевезень дорівнює $m + n - 1$. Покладемо, що в усіх базисних віконцях псевдовартості дорівнюють істинним вартостям

$$\widetilde{c}_{ij} = \alpha_i + \beta_j = c_{ij} \quad \forall x_{ij} > 0.$$

У вільних віконцях $x_{ij} = 0$ і при цьому співвідношення між $\widetilde{c}_{ij}$ та c_{ij} будь-яке. Це співвідношення є індикатором (показником) того, чи є план перевезень, якому відповідає дана таблиця оптимальним, чи його можливо покращити.

Доведемо теорему.

Теорема 10.2. Якщо для всіх базисних віконець допустимого плану $\widetilde{c}_{ij} = c_{ij}$, а для всіх вільних віконець допустимого плану $\widetilde{c}_{ij} \leq c_{ij}$, то план є оптимальним.

Нагадаємо, що для базисних віконець перевезення повинні бути додатними, а для вільних віконець $x_{ij} = 0$.

► Визначимо вартість плану $\left(x_{ij}\right)$ заданого умовами даної теореми та з урахуванням "теореми про платежі":

$$W = \sum_{i=1}^{m}\sum_{j=1}^{n}c_{ij}x_{ij} = \sum_{i=1}^{m}\sum_{j=1}^{n}\widetilde{c}_{ij}x_{ij} = \sum_{i=1}^{m}\alpha_i a_i + \sum_{j=1}^{n}\beta_j b_j = const.$$

Виконаємо заміну плану $\left(x_{ij}\right) \rightarrow (x_{ij}^{*})$, в якому змінимо частково або повністю склад вільних або базисних змінних, але співвідношення між елементами матриці істинних вартостей та псевдовартостей залишилось незмінним, тобто таким, що відповідає умовам теореми. Це означає, що для базисних змінних (x_{ij}^{*}), які теж були базисними у плані (x_{ij}) виконується умова $\widetilde{c_{ij}} = c_{ij}$, а для інших базисних змінних з плану виконується умова $\widetilde{c_{ij}} \geq c_{ij}$.

В результаті отримаємо вираз для обчислення загальної вартості перевезень при новому плані (x_{ij}^{*}):

$$W^{*} = \sum_{i=1}^{m}\sum_{j=1}^{n} c_{ij}x_{ij}^{*} \geq \sum_{i=1}^{m}\sum_{j=1}^{n} \widetilde{c_{ij}}x_{ij} = \sum_{i=1}^{m}\alpha_i a_i + \sum_{j=1}^{n}\beta_j b_j = const,$$

що і треба було довести. ◄

Зауваження 10.1. Ніякими змінами плану перевезень їх підсумкова вартість не може бути зменшена.

Прокоментуємо на прикладі зміст доведеної теореми. Припустимо, що $m = n = 3$. Тоді кількість базисних віконець повинні дорівнювати $m + n - 1 = 5$. Загальна вартість W при виконанні умов теореми обчислюється за формулою:

$$\widetilde{W} = \begin{bmatrix} \widetilde{c_{11}} & \widetilde{c_{12}} & \widetilde{c_{13}} \\ \widetilde{c_{21}} & \widetilde{c_{22}} & \widetilde{c_{23}} \\ \widetilde{c_{31}} & \widetilde{c_{32}} & \widetilde{c_{33}} \end{bmatrix} \cdot \begin{bmatrix} x_{11} & x_{12} & 0 \\ 0 & x_{22} & x_{23} \\ x_{31} & 0 & 0 \end{bmatrix}^{T} \cdot \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} =$$

$$= \begin{bmatrix} c_{11} & c_{12} & \widetilde{c_{13}} \\ \widetilde{c_{21}} & c_{22} & c_{23} \\ c_{31} & \widetilde{c_{32}} & \widetilde{c_{33}} \end{bmatrix} \cdot \begin{bmatrix} x_{11} & 0 & x_{31} \\ x_{12} & x_{22} & 0 \\ 0 & x_{23} & 0 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = c_{11}x_{11} + c_{12}x_{12} + c_{22}x_{22} + c_{23}x_{23} + c_{31}x_{31}.$$

Припустимо, що план змінено, обчислимо вартість зміненого плану:

$$W^{*} = \begin{bmatrix} c_{11} & c_{12} & \widetilde{c_{13}} \\ \widetilde{c_{21}} & c_{22} & c_{23} \\ c_{31} & \widetilde{c_{32}} & \widetilde{c_{33}} \end{bmatrix} \cdot \begin{bmatrix} 0 & x_{21}^{*} & 0 \\ x_{12}^{*} & 0 & 0 \\ x_{13}^{*} & x_{23}^{*} & x_{33}^{*} \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} =$$

$$= c_{12}x_{12}^{*} + \widetilde{c_{31}}x_{31}^{*} + \widetilde{c_{21}}x_{21}^{*} + c_{23}x_{23}^{*} + \widetilde{c_{33}}x_{33}^{*} \geq$$

$$\geq \widetilde{c_{12}}x_{12}^{*} + \widetilde{c_{31}}x_{31}^{*} + \widetilde{c_{21}}x_{21}^{*} + \widetilde{c_{23}}x_{23}^{*} + \widetilde{c_{33}}x_{33}^{*}.$$

Доведена теорема виконується і для виродженого плану. Таким чином доведено, що ознакою оптимальності плану перевезень є виконання двох умов:

1. $\widetilde{c_{ij}} = c_{ij}$ для всіх базисних віконець;

2. $\widetilde{c_{ij}} \le c_{ij}$ для всіх вільних віконець.

План для якого виконуються властивості 1 та 2 носить назву *потенціального*, а платежі, що йому відповідають (α_i, β_j) носять назву *потенціалів пунктів* A_i та B_j.

Використовуючи вищеозначені терміни, можливо переформулювати теорему коротше: будь-який потенціальний план є оптимальним.

10.2. Побудова потенціального плану

Для розв'язання транспортної задачі необхідно побудувати потенціальний план. Цей план можливо побудувати за рахунок послідовних наближень від початкового плану, знайденого, наприклад, методом північно-західного кута. Існує властивість платежів та псевдовартостей, яка полягає в наступному: якою не була б система платежів (α_i, β_j), що задовольняє умові $\widetilde{c_{ij}} = c_{ij}$ для всіх базисних віконець, для кожного вільного віконця ціна циклу перерахунку дорівнює різниці між вартістю c_{ij} та псевдовартістю $\widetilde{c_{ij}}$, що відноситься до даного віконця.

Приклад 10.1. Розглянемо транспортну задачу, не проставляючи запасів і заявок, наприклад, для випадку $m = 5$, $n = 6$ (табл.10.1)

Таблиця 10.1.

Приклад циклу перерахунку

ПВ	Пункти призначення					
	B_1	B_2	B_3	B_4	B_5	B_6
A_1	c_{11}	c_{12} −	c_{13}	c_{14}	c_{15} +	c_{16}
A_2	c_{21}	c_{22} +	c_{23} −	c_{24}	c_{25}	c_{26}
A_3	c_{31}	c_{32}	c_{33} +	c_{34}	c_{35} −	c_{36}
A_4	c_{41}	c_{42}	c_{43}	c_{44}	c_{45}	c_{46}
A_5	c_{51}	c_{52}	c_{53}	c_{54}	c_{55}	c_{56}

❍Знайдемо

$$r = n + m - 1 = 5 + 6 - 1 = 10.$$

В таблиці для прикладу заштриховано 10 базисних віконець . Обираємо будь-яке вільне віконце, наприклад, $(1, 5)$ і будуємо цикл перерахунку так, щоб додатна вершина перебувала у цьому віконці, а всі інші в базисних. Ціна визначеного циклу дорівнює:

$$V_{15} = c_{15} - c_{35} + c_{33} - c_{23} + c_{22} - c_{12}.$$

Зважаючи на те, що для всіх базисних клітинок псевдовартості дорівнюють вартостям:

$$V_{15} = c_{15} - (\alpha_3 + \beta_5) + (\alpha_3 + \beta_3) - (\alpha_2 + \beta_3) + (\alpha_2 + \beta_2) - (\alpha_1 + \beta_2) =$$
$$= c_{15} - (\alpha_1 + \beta_5) = c_{15} - \tilde{c}_{15}. \bullet$$

Висновок: при використанні методу потенціалів для розв'язання транспортної задачі немає необхідності знаходити цикли із від'ємною ціною, що є найбільш трудомісткою процедурою розподільчого методу.

Методика розв'язання транспортної задачі методом потенціалів полягає в наступному:

1. Виконати перше наближення до оптимального плану, наприклад, за допомогою метода північно-західного кута. Перевірити задачу на невиродженість, тобто обчислити кількість базисних клітинок, перевезення в яких повинні бути додатними. Якщо перше наближення продемонструвало виродженість транспортної задачі (кількість вільних віконець з нульовими перевезеннями перевищує $m \cdot n - (m + n - 1)$), необхідно застосувати метод збурення заявок та запасів і перейти до не виродженої задачі.
2. Визначити для плану першого наближення платежі, виходячи із умови, що в будь-якому базисному віконці псевдовартості дорівнюють вартостям $\tilde{c}_{ij} = \alpha_i + \beta_j = c_{ij}$. Всього записаних вище рівнянь буде $r = m + n - 1$, а невідомих, що входять до їх складу $\alpha_i \left(i = \overline{1,m}\right)$ та $\beta_j \left(j = \overline{1,n}\right)$ - $m + n$. Тобто одну з невідомих можливо для простоти покласти нульовою, а всі інші обчислити так, щоб виконувалось записане вище рівняння.
3. Обчислити псевдовартості для всіх вільних віконець. Якщо виявляється, що псевдовартості вільних віконець не перевищують вартості цих віконець, то оптимальний план знайдено.
4. Якщо хоча б у одному з вільних віконець псевдовартість перевищує вартість $\tilde{c}_{ij} > c_{ij}$, то потрібно покращити план перевезень за рахунок переміщення перевезень за циклом, який відповідає обраному вільному віконцю із від'ємною різницею $0 > \tilde{c}_{ij} - c_{ij}$.
5. Знову обчислити платежі та псевдовартості. Перерахунки завершуються лише тоді, коли буде знайдено оптимальний план.

Зауваження 10.2. Поняття платежів, вартостей, псевдоплатежів, псевдовартостей, мають наглядну економічну інтерпретацію: компанії групи A, які зберігають товари, та компанії групи B, що накопичують та реалізують товари, разом фінансують перевезення. Оптимальним є той план, коли компанії A та B за ці перевезення не переплачують.

10.3. Приклади розв‘язання транспортної задачі методом потенціалів

Приклад 10.2. Розв’язати транспортну задачу, задану таблицею 10.2.

Таблиця 10.2

Вихідні дані

ПВ	Пункти призначення					Запаси	Платежі
	B_1	B_2	B_3	B_4	B_5	a_i	α_i
A_1	**10** 10 **17** -	**8** 8 **8 Ц1** +	6 9	3 6	4 5	25	$\alpha_1 = 0$
A_2	8 5 +	**6** 6 **13** -	**4** 4 **19**	3 3	2 8	32	$\alpha_2 = -2$
A_3	9 9	7 7	**5** 5 **22**	**4** 4 **14**	**3** 3 **4**	40	$\alpha_3 = -1$
A_4	14 14	12 10	10 8	9 8	**8** 8 **20**	20	$\alpha_4 = 4$
b_j	17	21	41	14	24	117	
Платежі β_j	10	8	6	5	4		

❍ Знайдемо методом північно-західного кута перше наближення до оптимального плану (див. табл. 10.2). Кількість ненульових віконець дорівнює 8, що співпадає із контрольним числом

$$r = m + n - 1 = 4 + 5 - 1 = 8.$$

Висновок – транспортна задача невироджена.

Додамо до транспортної таблиці рядок та стовпець платежів відповідно β_j та α_i. Псевдовартості запишемо зліва зверху у кожному віконці.

Враховуючи, що для базисних віконець $\widetilde{c_{ij}} = \alpha_i + \beta_j$, знаходимо $\widetilde{c_{11}} = c_{11} = 10 = 0 + \beta_1$ (за припущенням $\alpha_1 = 0$), тобто $\beta_1 = 10$. Аналогічно знаходимо $\beta_2 = 8$. Далі $\alpha_2 + \beta_2 = 6 \to \alpha_2 = -2$ тощо. Виконаємо перерахунок базисного віконця за циклом Ц1 (див. табл. 10.2). Перенесемо за цим циклом 13 одиниць вантажу (більше неможливо, інакше у віконці $(2{,}2)$ виникне від’ємне число і план не буде допустимим) і зменшимо вартість плану на $13 \cdot 3 = 39$.

Не всі псевдовартості у вільних віконцях (табл. 10.3) задовольняють умові $\tilde{c}_{ij} \le c_{ij}$. Тому, план, наведений у таблиці першої ітерації не є оптимальним. Спробуємо покращити цей план шляхом переведення у базисні змінні одну з вільних змінних, для якої у відповідному віконці $\tilde{c}_{ij} > c_{ij}$, наприклад, у віконці $(1{,}4)$. Ціна циклу для віконця $(1{,}4)$ дорівнює $v_{14} = 6 - 8 = -2$.

Таблиця 10.3

Перша ітерація

ПВ	Пункти призначення					запаси	платежі
	B_1	B_2	B_3	B_4	B_5	a_i	α_i
A_1	**10** 10 **4** -	**8** 8 **21**	9 9	8 6 +	7 5	25	$\alpha_1 = 0$
A_2	**5** 5 **13** +	6 6	**4** **Ц2** 4 **19** -	3 3	2 8	32	$\alpha_2 = -5$
A_3	6 9	4 7	**5** 5 **22** +	**4** 4 **14** -	**3** 3 **4**	40	$\alpha_3 = -4$
A_4	11 14	9 10	10 8	9 8	**8** 8 **20**	20	$\alpha_4 = 1$
b_j	17	21	41	14	24	117	
β_j	10	8	9	8	7		

Перенесемо по циклу Ц2 (див. табл. 10.3) 4 одиниці і занесемо результат у таблицю другої ітерації (табл. 10.4).

Таблиця 10.4

Друга ітерація

ПВ	Пункти призначення					Запаси	платежі
	B_1	B_2	B_3	B_4	B_5	a_i	α_i
A_1	8 10	**8** 8 **21**	7 9	**6** 6 **4**	5 5	25	$\alpha_1 = 0$
A_2	**5** 5 **17**	5 6	**4** 4 **15**	3 3	2 8	32	$\alpha_2 = -3$
A_3	6 9	6 7	**5** 5 **26** -	**4** 4 **10**	**3** 3 **4** +	40	$\alpha_3 = -2$
A_4	11 14	11 10	10 8 +	9 **Ц3** 8	**8** 8 **20** -	20	$\alpha_4 = 3$
b_j	17	21	41	14	24	117	
β_j	8	8	7	6	5		

План другої ітерації теж не оптимальний. Обираємо віконце $(4{,}3)$ і переносимо 20 одиниць перевезень по циклу Ц3. Остаточно отримаємо таблицю третьої ітерації (табл. 10.5), яка за всіма ознаками надає оптимальний план перевезень (усі псевдовартості не перевищують відповідних вартостей).

Значення показника ефективності дорівнює:
$$W = 21\cdot 8 + 4\cdot 6 + 17\cdot 5 + 15\cdot 4 + 6\cdot 5 + 10\cdot 4 + 24\cdot 3 + 20\cdot 8 = 639.$$

Таблиця 10.5

Третя ітерація

ПВ	Пункти призначення					Запаси	платежі
	B_1	B_2	B_3	B_4	B_5	a_i	α_i
A_1	*8* *10*	***8*** *8* ***21***	*7* *9*	***6*** *6* ***4***	*5* *5*	25	$\alpha_1 = 0$
A_2	***5*** *5* ***17***	*5* *6*	***4*** *4* ***15***	*3* *3*	*2* *8*	32	$\alpha_2 = -3$
A_3	*6* *9*	*6* *7*	***5*** *5* ***6***	***4*** *4* ***10***	***3*** *3* ***24***	40	$\alpha_3 = -2$
A_4	*9* *14*	*9* *10*	***8*** *8* ***20***	*7* *8*	*6* *8*	20	$\alpha_4 = 1$
b_j	17	21	41	14	24	117	
β_j	8	8	7	6	5		

Таким чином, фірма, яка володіє усіма складами A та всіма пунктами реалізації B, але не включає до свого складу перевізників, має заплатити мінімальну суму $W_{min} = 639$. Платежі α_i та β_j мають умовний зміст і використані як математична абстракція.●

Приклад 10.3. Розв'язати методом потенціалів вироджену транспортну задачу. Умови транспортної задачі задані вихідною транспортною таблицею 10.6, $m = 3$, $n = 3$.

Таблиця 10.6

Вихідні дані

ПВ	Пункти призначення			запаси	платежі
	B_1	B_2	B_3	a_i	α_i
A_1	***6*** *6* ***20*** -	***4*** *4* **ε** +	*3* *2*	$20 + \varepsilon$	$\alpha_1 = 0$
A_2	*7* *3* +	***5*** **Ц1** *5* ***25-ε*** -	***4*** *4* ***2ε***	$25 + \varepsilon$	$\alpha_2 = 1$
A_3	*6* *3*	*4* *6*	***3*** *3* ***30-2ε***	$30 - 2\varepsilon$	$\alpha_3 = 0$
b_j	20	25	30	75	
β_j	6	4	3		

❍ Перевіримо: $r = m + n - 1 = 3 + 3 - 1 = 5 > 3-$ транспортна задача вироджена. Скористаємось методом збурення запасів та заявок (див. табл. 10.6).

Перерахунок за циклом Ц1 дозволяє заповнити таблицю першої ітерації (табл. 10.7), а за циклом Ц2 – таблицю другої ітерації (табл. 10.8).

Таблиця 10.7

Перша ітерація

ПВ	Пункти призначення			Запаси	Платежі
	B_1	B_2	B_3	a_i	α_i
A_1	2 6	**4** 4 **20+ε** -	3 2 **Ц2** +	$20+\varepsilon$	$\alpha_1=0$
A_2	**3** 3 **20**	**5** 5 **5-ε** +	**4** 4 **2ε** -	$25+\varepsilon$	$\alpha_2=1$
A_3	2 3	4 6	**3** 3 **30-2ε**	$30-2\varepsilon$	$\alpha_3=0$
b_j	20	25	30	75	
β_j	2	4	3		

Таблиця 10.8

Друга ітерація

ПВ	Пункти призначення			Запаси	Платежі
	B_1	B_2	B_3	a_i	α_i
A_1	2 6	**4** 4 **20+ε**	**2** 2 **2ε**	$20+\varepsilon$	$\alpha_1=0$
A_2	**3** 3 **20**	**5** 5 **5+ε**	3 4	$25+\varepsilon$	$\alpha_2=1$
A_3	3 3	5 6	**3** 3 **30-2ε**	$30-2\varepsilon$	$\alpha_3=1$
b_j	20	25	30	75	
β_j	2	4	2		

$$W_{\min}\Big|_{\varepsilon=0} = 20\cdot 4 + 20\cdot 3 + 5\cdot 5 + 30\cdot 3 = 255$$

Підкреслимо, що у виродженій транспортній задачі при виборі Ц1 або Ц2 із іншими вільними віконцями остаточний, тобто оптимальний, план зміниться, але мінімальне значення показника ефективності залишиться незмінним: $W = 255$.●

Висновок: особливістю вироджених транспортних задач є те, що однакові мінімальні значення показника ефективності можуть досягатися на декількох різних планах. Тобто, для "вироджених" транспортних задач характерною ознакою є неоднозначність розв’язку.

Лекція 11. Деякі окремі випадки транспортних задач

11.1. Транспортні задачі із неправильним балансом

11.1.1. Транспортна задача із надмірністю запасів

Досі розглядались транспортні задачі із виконанням умови: сума запасів дорівнює сумі заявок

$$\sum_{i=1}^{m} a_i = \sum_{j=1}^{n} b_j. \tag{11.1}$$

Ця класична транспортна задача має назву *транспортної задачі з правильним балансом*. Порушення умови правильного балансу перетворює класичну транспортну задачу в *транспортну задачу із неправильним балансом*. Порушення балансу поділяють на 2 типи:

1. Сума запасів у пунктах відправлення перевищує суму заявок

$$\sum_{i=1}^{m} a_i > \sum_{j=1}^{n} b_j, \tag{11.2}$$

тобто має місце *транспортна задача із надмірністю запасів*.

2. Сума заявок перевищує наявні запаси

$$\sum_{i=1}^{m} a_i < \sum_{j=1}^{n} b_j \tag{11.3}$$

тобто маємо *транспортну задачу із надмірністю заявок*.

У випадку надмірності запасу, постановка транспортної задачі набуває вигляду: знайти такий план перевезень $\left(x_{ij}\right)$, при якому усі заявки будуть виконані, а загальна вартість перевезень набуває мінімального значення:

$$W = \sum_{i=1}^{m}\sum_{j=1}^{n} c_{ij} x_{ij} \to \min_{x_{ij} \geq 0 \left(i=\overline{1,m}\ \ j=\overline{1,n}\right)} \tag{11.4}$$

при виконанні системи обмежень

$$\begin{cases} \sum_{j=1}^{n} x_{ij} < a_i \ \left(i = \overline{1, m_1}\right), \\ \sum_{j=1}^{n} x_{ij} = a_i \ \left(i = \overline{m_1 + 1, m}\right), \\ \sum_{j=1}^{m} x_{ij} = b_j \ \left(j = \overline{1, n}\right), \end{cases} \tag{11.5}$$

та умов (11.2).

Як бачимо, частина або всі рівняння-обмеження перетворюються в нерівності-обмеження. Розв’язок цієї транспортної задачі можливо отримати або класичним симплекс-методом або за допомогою методів розв'язання транспортних задач, розглянутих вище (розподільчий метод, метод потенціалів).

Перед тим, як скористатись відомими методами розв'язання транспортної задачі,зведемо транспортну задачу із надмірністю запасів, до класичної транспортної задачі із правильним балансом.

Для цього використаємо уявний пункт призначення (фіктивний, хибний) $B_ф$, якому належить заявка, що якраз дорівнює перевищенню запасів над заявками:

$$b_ф = \sum_{i=1}^{m} a_i - \sum_{j=1}^{n} b_j .$$

Покладемо вартості перевезень від всіх пунктів відправлення до фіктивного пункту призначення $B_ф$ рівними 0, тобто $c_{iф} = 0 \left(i = \overline{1,m}\right)$.

За фізичним змістом $x_{iф}$ – означає кількість вантажу, який залишається невідправленим з i-го пункту відправлення.

Висновок: за рахунок фіктивного пункту призначення $B_ф$ із заявкою $b_ф$, всі нерівності-обмеження у транспортній задачі із надмірністю запасів перетворено у рівняння-обмеження, тобто задача зведена до класичної транспортної задачі із правильним балансом.

11.1.2. Транспортна задача із надмірністю запасів

У цьому випадку виконується нерівність (11.3):

$$\sum_{i=1}^{m} a_i < \sum_{j=1}^{n} b_j ,$$

тобто запасів недостатньо для задоволення усіх заявок. Необхідно скласти план таких перевезень, при якому усі запаси будуть вивезені і при цьому вартість перевезень буде мінімальною. Для розв'язання введемо фіктивний пункт відправлення $A_ф$, із фіктивним уявним запасом $a_ф$,який дорівнює саме тій кількості вантажу, якої не вистачає для задоволення усіх заявок

$$a_ф = \sum_{j=1}^{n} b_j - \sum_{i=1}^{m} a_i .$$

При цьому покладемо, що усі вартості перевезень із фіктивного пункту відправлення, до будь якого пункту призначення, дорівнюють 0:

$$c_{Фj} = 0 \left(j = \overline{1,n}\right).$$

Складові плану перевезень $x_{Ф_j}$ $(j = \overline{1,n})$ показують величину недопостачання вантажу у відповідний пункт призначення.

Висновок: за допомогою використання фіктивного пункту відправлення, транспортну задачу, з надмірністю заявок,зведено до транспортної задачі із правильним балансом.

Зауваження 11.1. Окрім прийому зведення транспортної задачі із неправильним балансом до транспортної задачі із правильним балансом за допомогою фіктивних пунктів відправлень та пунктів прийому, можливо використати умови «нормування», наприклад домножити усі заявки на коефіцієнт

$$k = \frac{\sum_{i=1}^{m} a_i}{\sum_{j=1}^{n} b_j}$$

для транспортної задачі із надмірністю заявок або на коефіцієнт

$$k = \frac{\sum_{j=1}^{n} b_j}{\sum_{i=1}^{m} a_i}$$

для транспортних задач із надмірністю запасів.

Приклад 11.1. Розв’язати транспортну задачу із надмірністю запасів (табл. 11.1).

Таблиця 11.1

Вихідні дані

ПВ	Пункти призначення				запаси	платежі
	B_1	B_2	B_3	$B_ф$	a_i	α_i
A_1	*5* 5 ***18***	*7* 7 ***21*** -	*6* 6 ***11*** +	*1* 0	50	$\alpha_1 = 0$
A_2	*4* 6	*6* 6	*5* 5 ***22*** Ц1 -	*0* 0 ***18*** +	40	$\alpha_2 = -1$
A_3	*4* 8	*6* 4 +	*5* 5	*0* 0 ***20*** -	20	$\alpha_3 = -1$
b_j	18	21	33	38	110	
β_j	5	7	6	1		

❍Перевіряємо балансове співвідношення

$$\left.\begin{aligned} a_1 + a_2 + a_3 &= 50 + 40 + 20 = 110, \\ b_1 + b_2 + b_3 &= 18 + 21 + 23 = 72, \end{aligned}\right\} \Rightarrow \sum_{i=1}^{3} a_i > \sum_{j=1}^{3} b_j \,.$$

Має місце надмірність запасів.

Використовуємо фіктивний пункт призначення $B_ф$ із заявкою $b_ф = 110 - 72 = 38$. В результаті використання $B_ф$ вдалося перейти до транспортної задачі із правильним балансом.

Розв'язуємо класичну транспортну задачу методом потенціалів. Перевіряємо задачу на невиродженість:

$$r = n + m - 1 = 4 + 3 - 1 = 6 .$$

Кількість ненульових базисних змінних, отриманих методом північно-західного кута дорівнює 6 і співпадає із r. Тобто транспортна задача є невиродженою. Виконаємо заміну базисних і вільних віконець за циклами Ц1 (див. табл. 11.1) , Ц2 (табл. 11.2) та Ц3 (табл. 11.3). Заповнимо відповідно таблиці: першого (див. табл. 11.2), другого (див. табл. 11.3) та третього (табл. 11.4) наближення.

Таблиця 11.2

Перше наближення

ПВ	Пункти призначення				Запаси	Платежі
	B_1	B_2	B_3	$B_ф$	a_i	α_i
A_1	**5** 5 **18**	**7** 7 **1**	**6** 6 **31** -	1 0 +	50	$\alpha_1 = 0$
A_2	4 6	6 6	**5** **Ц2** 5 **2** +	**0** 0 **38** -	40	$\alpha_2 = -1$
A_3	2 8	**4** 4 **20**	3 5	-2 0	20	$\alpha_3 = -3$
b_j	18	21	33	38	110	
β_j	5	7	6	1		

Таблиця 11.3

Друге наближення

ПВ	Пункти призначення				Запаси	Платежі
	B_1	B_2	B_3	$B_ф$	a_i	α_i
A_1	**5** 5 **18**	**7** 7 **1** -	5 6	**0** 0 **31** +	50	$\alpha_1 = 0$
A_2	5 6	7 + 6	**5** Ц3 5 **33**	**0** - 0 **7**	40	$\alpha_2 = 0$
A_3	2 8	**4** 4 **20**	2 5	-3 0	20	$\alpha_3 = -3$
b_j	18	21	33	38	110	
β_j	5	7	5	0		

Висновок:

1) Третє наближення дає оптимальний план, із якого випливає, що $W_{min} = 341$.

2) У пункті відправлення A_1 залишилось невідправлених 32, у A_2 – 6, а в A_3 – 0.

Таблиця 11.4

Третє наближення

ПВ	Пункти призначення				Запаси	Платежі
	B_1	B_2	B_3	$B_ф$	a_i	α_i
A_1	**5** 5 **18**	6 7	5 6	**0** 0 **32**	50	$\alpha_1 = 0$
A_2	5 6	**6** 6 **1**	**5** 5 **33**	**0** 0 **6**	40	$\alpha_2 = 0$
A_3	3 8	**4** 4 **20**	3 5	-2 0	20	$\alpha_3 = -2$
b_j	18	21	33	38	110	
β_j	5	6	5	0		

●

11.2. Розв'язок транспортної задачі за критерієм часу

Досі транспортна задача ставилась як задача, у якій необхідно було мінімізувати вартість перевезень. Але в багатьох випадках практики важливу роль відіграє тривалість часу T, протягом якого усі перевезення буде завершено. Так, наприклад: при ліквідації наслідків надзвичайних ситуацій необхідно мінімізувати як час надходження інформації про стихійне лихо, так і час до надання допомоги; при багатопроцесорній обробці інформації необхідно розподілити обчислювальні ресурси, щоб мінімізувати тривалість обчислень.

Найкращим слід вважати такий план перевезень $\left(x_{ij}\right)$, при якому час закінчення усіх перевезень мінімізується

$$T \to \min_{\substack{x_{ij} \geq 0 \\ i=\overline{1,m},\ j=\overline{1,n}}} . \tag{11.6}$$

Транспортна задача, у якій оптимальним вважається план із мінімальним часом усіх перевезень, носить назву *транспортна задача за критерієм часу*.

Математична постановка транспортної задачі за критерієм часу стосовно балансових обмежень виконується так само, як і класична транспортна задача, із тією різницею, що критерій, тобто показник ефективності, має вигляд:

$$T = \max t_i \to \min_{x_{ij} > 0} \tag{11.7}$$

і виконуються балансові обмеження:

$$\sum_{i=1}^{m} x_{ij} = b_j \ (j = \overline{1,n}), \tag{11.8}$$

$$\sum_{j=1}^{n} x_{ij} = a_i \ (i = \overline{1,m}). \quad (11.9)$$

Зауваження 11.2. Використання строгого обмеження $x_{ij} > 0$ показує, що максимальний інтервал часу обирається не взагалі, а лише з тих віконець транспортної таблиці, в яких перевезення строго більше нуля (строго додатні, тобто реально виконуються). Для розв'язку транспортної задачі за критерієм часу теж зручно використовувати транспортну таблицю, у якій в правому верхньому куті буде записано t_{ij} – час перевезення вантажу із пункту відправлення i в пункт призначення j (замість c_{ij}).

Загалом транспортна задача за критерієм часу не відноситься до задач лінійного програмування тому, що показник ефективності T є нелінійною функцією змінних x_{ij}. Але зручна форма транспортної таблиці, яка використовується для розв'язку класичної транспортної задачі, дозволяє виконати обчислення без порушень балансових обмежень, тобто побудувати алгоритм, який ґрунтується на використанні допустимого плану.

Метод обчислень, що мінімізує час перевезень T, побудований із використанням транспортної таблиці і технології циклічних перенесень, які дозволяють не порушити балансові обмеження, отримав назву *методу заборонених віконець*. "Забороненими" вважаються віконця без перевезень і при цьому із найбільшим часом перевезень t_{ij}. "Заборона" вказує, із якого допустимого віконця потрібно "виштовхнути" перевезення в інше віконце із меншим часом перевезень. Покращення плану перевезень припиняється тоді, коли найбільший час перевезень вже зменшити неможливо. Ознака: всі віконця заблоковані перевезеннями або заборонені до перевезень.

Приклад 11.2. Вихідні дані транспортної задачі представлені у таблиці 11.5. Необхідно мінімізувати час перевезень.

❍ 1.Викреслюємо віконця із координатами $(1{,}1)$, $(2{,}5)$, $(4{,}1)$ та $(4{,}5)$, як віконця із значним часом перевезень.

2. Будь-яким способом заповнюємо транспортну таблицю допустимим планом.

1) Якщо транспортну таблицю заповнити будь-яким планом не звертаючи уваги на баланс записів та заявок, але лише виконувати умову невід'ємності перевезень, то отримаємо план перевезень.

2)Якщо транспортну таблицю, додатково до невід'ємності, заповнити такими значеннями x_{ij}, які враховують умови виконання балансу запасів та заявок, то отримаємо допустимий план перевезень.
3) Якщо вимоги п.2. доповнити вимогою додатності $r = m + n - 1$ базових перевезень та нульовими перевезеннями в інших віконцях таблиці, то отримаємо опорний план перевезення.

Таблиця 11.5

Вихідні дані

ПВ	Пункти призначення					Запаси
	B_1	B_2	B_3	B_4	B_5	a_i
A_1	*10*	*8*	*5* **25**	*6*	*7*	25
A_2	*5* ***21*** -	*6* Ц1 +	*6* ***2***	*6* ***11***	*9*	34
A_3	*4* +	*8* ***14*** -	*7* ***13***	*8*	*5* ***15***	42
A_4	*11*	*4* ***23***	*5*	*8*	*9*	23
b_j	21	37	40	11	15	124

Підкреслимо ще раз, що *оптимальний план перевезень для транспортної задачі за критерієм вартості* досягається лише на опорному плані. Для транспортної задачі за критерієм часу оптимальний план перевезень не обов'язково досягається на опорному, він може досягатися на будь-якому допустимому плані перевезень. Виконаємо перенесення перевезень за циклами Ц1 (табл. 11.6) та Ц2 (табл. 11.7).

Таблиця 11.6

Перша ітерація

ПВ	Пункти призначення					Запаси
	B_1	B_2	B_3	B_4	B_5	a_i
A_1	*10*	*8*	*5* ***25*** -	*6*	*5* +	25
A_2	*7* ***7*** -	*6* ***14***	*6* ***2*** +	*6* Ц2 ***11***	*9*	34
A_3	*4* ***14*** +	*8*	*7* ***13***	*8*	*5* ***15*** -	42
A_4	*11*	*4* ***23*** -	*5* +	*8*	*9*	23
b_j	21	37	40	11	15	124

Таблиця 11.7

Друга ітерація

ПВ	Пункти призначення					запаси
	B_1	B_2	B_3	B_4	B_5	a_i
A_1	*10*	*8*	*5* ***18***	*6*	*5* ***7***	25
A_2	*5*	*6* ***14***	*6* ***9***	*6* ***11***	*9*	34
A_3	*4* ***21***	*8*	***5*** ***13***	*8*	*5* ***8***	42
A_4	*11*	*4* ***23***	*5*	*8*	*9*	23
b_j	21	37	40	11	15	124

Отже, $T_{min} = 6$. ●

Приклад 11.3. Скласти план перевезень, якщо вихідні дані транспортної задачі задані таблицею 11.8.

Таблиця 11.8

Вихідні дані

ПВ	Пункти призначення					Запаси
	B_1	B_2	B_3	B_4	B_5	a_i
A_1	1 15	3 5	4	2	5	20
A_2	5	2 15	3 15	1	2	30
A_3	8	7	6 10	10	4 25	35
A_4	3	4	3	6 25	5 15	40
b_j	15	20	25	25	40	125

Остаточний результат

Таблиця 11.9

ПВ	Пункти призначення					Запаси
	B_1	B_2	B_3	B_4	B_5	a_i
A_1	1 10	3	4	2 10	5	20
A_2	5	2 10	3	1 15	2 5	30
A_3	8	7	6	10	4 35	35
A_4	3 5	4 10	3 25	6	5	40
b_j	15	20	25	25	40	125

Отже, $T_{min} = 4$. ●

Лекція 12. Післяоптимізаційний аналіз задачі лінійного програмування

12.1. Ідея аналізу параметричної чуттєвості

Післяоптимізаційний аналіз або, як ще кажуть, аналіз чуттєвості полягає у зв'язуванні впливу структурних, параметричних та структурно-параметричних змін у математичній моделі (математичній постановці) задачі на отриманий оптимальний розв‘язок для тієї постановки задачі лінійного програмування, яка вважається вихідною.

Розглянемо приклад 5.1, в якому було з'ясовано, що для отримання максимального прибутку необхідно випустити продукцію типу A (супутники зв'язку) $\hat{x}_1 = 17\frac{1}{7} \approx 17$ та типу B (навігаційні супутники) $\hat{x}_2 = 23\frac{4}{7} \approx 23$. Оптимальний розв'язок було отримано за умови, що вартість виробів A та B складає відповідно 40 та 50 умовних одиниць. У зв'язку зі змінами, що відбуваються в світовій економіці, керівнику фірми (особі, що приймає рішення) важливо знати, як вплине зміна вартості продукції A та B (питомий прибуток) на запланований випуск продукції $\hat{x}_{1,2}$ (рис.12.1).

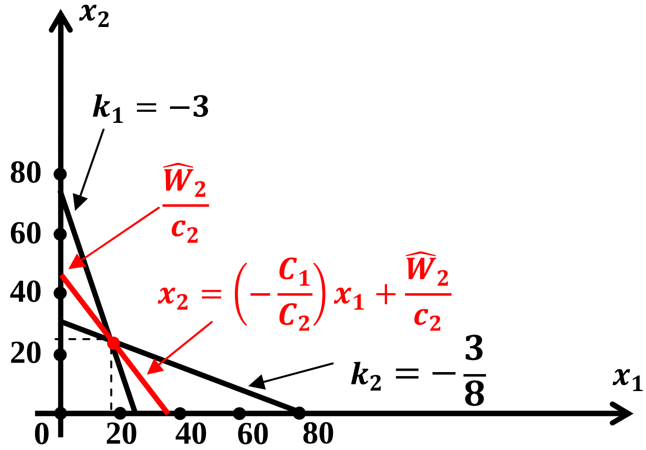


Рис. 12.1. Графічна ілюстрація алгоритму визначення меж зміни співвідношення між c_1 та c_2 у виразі для обчислення показника ефективності, при яких оптимальний розв‘язок $(\hat{x}_1, \hat{x}_2)$ залишається незмінним.

Як бачимо, при умові

$$-3 < -\frac{c_1}{c_2} < -\frac{3}{8} \Leftrightarrow 3 > \frac{c_1}{c_2} > \frac{3}{8} \quad \Leftrightarrow \quad 3 > \frac{c_1}{c_2} > 0,375$$

оптимальний план випуску продукції не зміниться, тому що «основна пряма»

$$W = c_1 x_1 + c_2 x_2$$

проходитиме через кутову точку із координатами $\left(\hat{x}_1 = 17\frac{1}{7},\ \hat{x}_2 = 23\frac{4}{7}\right)$, що забезпечує максимальне значення показника ефективності. При $\frac{c_1}{c_2} = \frac{3}{8}$ або $\frac{c_1}{c_2} = 3$ має місце невизначеність.

Якщо $\frac{c_1}{c_2} < \frac{3}{8}$, то оптимальною точкою буде $(0{,}30)$, а якщо $\frac{c_1}{c_2} > 3$, то оптимальною точкою буде $(25{,}0)$, що з фізичної точки зору означає відповідно недоцільність випуску супутників зв'язку або навігаційних супутників. При зміні співвідношення між c_1 та c_2 відбувається зміна оптимального значення показника ефективності.

Зрозуміло, що взагалі зміна будь-якого параметра математичної моделі задачі лінійного програмування (об'єм ресурсів та їх вартість), норми витрат ресурсів на виготовлення одиниці продукції впливатиме на оптимальне рішення і значення показника ефективності. Кількісне дослідження цих змін і виконується в процесі аналізу чутливості математичної моделі задачі лінійного програмування. Наведений простий приклад аналізу чутливості оптимального рішення задачі лінійного програмування та значення показника ефективності до зміни параметрів математичної моделі має ілюстративний характер і для використання у багатовимірному випадку ($n \geq 3$, де n– кількість змінних) стає громіздким.

Для дослідження впливу властивостей задачі лінійного програмування на її оптимальне за вихідною постановкою рішення в загальному багатовимірному випадку використовують спеціальним чином переформульовану вихідну задачу лінійного програмування, яка отримала назву *двоїста (спряжена) задача лінійного програмування*, або використовують так зване *параметричне програмування*.

12.2. Ідея фізичного змісту побудови математичної моделі двоїстої задачі лінійного програмування

У прикладі 5.1 було поставлено за мету отримати максимальний прибуток від виробництва виробів A та B, але нічого не було сказано про вартість ресурсу, який використовувався для створення цих виробів, тобто нічого не було сказано про витрати, які необхідно мінімізувати, але таким чином, щоб не зменшити питомий прибуток (собівартість виробленої продукції).

Припустимо, що одиниця ресурсу C, D, E коштує відповідно u_1, u_2, u_3. Тоді вартість виробів можливо обчислити за виразами відповідно $24u_1 + 8u_2 + 3u_3$ і $8u_1 + 8u_2 + 8u_3$.

Задача лінійного програмування із змінними $u_{1,2,3}$ формулюється так: якою повинна бути вартість кожного окремого ресурсу для того, щоб не зменшуючи вартості одиниці виробу, досягти мінімуму сумарної вартості ресурсів. Математична постановка задачі в цьому випадку набуває вигляду:

$$W = 600u_1 + 480u_2 + 240u_3 \to \min_{u_{1,2,3} \geq 0}$$

за умов

$$\begin{cases} 24u_1 + 8u_2 + 3u_3 \geq 40, \\ 8u_1 + 8u_2 + 8u_3 \geq 50. \end{cases}$$

Записана задача лінійного програмування має назву двоїстої по відношенню до задачі прикладу 5.1.

12.3. Загальна постановка і правила побудови двоїстої задачі

Кожній задачі лінійного програмування можливо поставити у відповідність задачу, яку називають двоїстою до неї. Припустимо, що загальна задача лінійного програмування (вихідна, або, як ще кажуть, пряма задача) задана у вигляді:

$$W = \sum_{j=1}^{n} c_j x_j \to \max_{x_j \geq 0 \left(j=\overline{1,n}\right)} \tag{12.1}$$

за умови виконання обмежень

$$\sum_{j=1}^{n} a_{ij} x_j \leq b_i \ \left(i = \overline{1,m}\right). \tag{12.2}$$

Двоїста (спряжена до неї) задача має вигляд:

$$\tilde{W} = \sum_{i=1}^{m} b_i u_i \to \min_{u_i \geq 0 \left(j=\overline{1,m}\right)}, \tag{12.3}$$

$$\sum_{i=1}^{m} a_{ij} u_i \geq c_j \ \left(j = \overline{1,n}\right). \tag{12.4}$$

Задача (12.3), (12.4) – двоїста до задачі (12.1), (12.2) – будується за наступними правилами:

1. Виконується впорядкування вихідної задачі до виду (12.3), (12.4), тобто якщо цільова функція максимізується, то нерівності-обмеження повинні бути приведені до вигляду «$\leq$», а якщо мінімізується, то до вигляду «$\geq$». Досягти виконання потрібної орієнтації знаку обмежень можливо множенням обох його частин на $\left(-1\right)$.

2. Якщо вихідна задача є задачею максимізації, то двоїста буде задачею мінімізації. При цьому вектор, який утворено із коефіцієнтів при невідомих в показнику ефективності вихідної задачі співпадає із вектором констант в

правих частинах системи обмежень двоїстої задачі, і, навпаки, коефіцієнти при невідомих в показнику ефективності двоїстої задачі є відповідними правими частинами системи обмежень вихідної задачі.

3. Кожній змінній u_i двоїстої задачі відповідає i-те обмеження вихідної задачі і навпаки, кожній змінній x_j вихідної задачі відповідає j-те обмеження двоїстої задачі.

4. Матриця коефіцієнтів двоїстої задачі може бути отримана транспонуванням матриці $A = \left\| a_{ij} \right\|_{m \times n}$, складеної із коефіцієнтів при невідомих в системі обмежень вихідної задачі.

Задачі (12.1), (12.2) та (12.3), (12.4) утворюють симетричну пару взаємно двоїстих задач. Якщо використати позначення:

$C = \left(c_1, c_2, \ldots, c_n\right)$,

$B = \left(b_1, b_2, \ldots, b_m\right)^T$,

$A = \left\| a_{ij} \right\|_{\substack{i=\overline{1,m} \\ j=\overline{1,n}}}$ – прямокутна матриця розміром $m \times n$,

$X = \left(x_1, x_2, \ldots, x_n\right)^T$,

$U = \left(u_1, u_2, \ldots, u_m\right)^T$,

$Q_k = \left(0, \ldots, 0\right)_k^T$ – вектор-стовпець нулів розмірності k;

то пряму і двоїсту задачі лінійного програмування можливо записати у вигляді:

$$\begin{cases} W = C \cdot X \to \max\limits_{X \ge o_n} \\ A \cdot X \le B \end{cases} \text{ – пряма,}$$

$$\begin{cases} \tilde{W} = B^T \cdot U \to \min\limits_{U \ge o_m} \\ A^T \cdot U \ge C^T \end{cases} \text{ – двоїста.}$$

Розглянемо двоїсту задачу лінійного програмування (12.3), (12.4) як вихідну задачу лінійного програмування і, скориставшись правилами переходу до двоїстої задачі, отримаємо пряму задачу лінійного програмування (10.1), (10.2):

$$\begin{cases} -\tilde{W} = -B^T \cdot U \to \max\limits_{U \ge o_m}, \\ -A^T \cdot U \ge -C^T \end{cases}.$$

І далі, після перетворення, можемо записати:

$$\begin{cases} -W = -C \cdot X \to \min\limits_{\mathrm{X} \ge \mathrm{o_n}}, \\ -A \cdot X \ge -B; \end{cases} \Leftrightarrow \begin{cases} W = C \cdot X \to \max\limits_{X \ge o_n}, \\ A \cdot X \le B. \end{cases}$$

Висновок:

1) Двоїста задача лінійного програмування до двоїстої задачі лінійного програмування є прямою задачею.

2) В системі пряма-двоїста обидві задачі лінійного програмування рівноправні. Кожну з них можливо розглядати як пряму, і тоді інша вважається двоїстою до неї.

Приклад 12.1. Задана задача лінійного програмування:

$$\begin{cases} W = x_1 - 4x_2 - 3x_3 \to \min\limits_{0 \le x_{1,2,3}}; \\ 3x_1 + 4x_2 + x_3 \le 7, \\ x_1 + 2x_2 + x_3 = 6, \\ x_3 \ge 4. \end{cases}$$

Необхідно побудувати двоїсту задачу лінійного програмування.

❍Приведемо задану задачу лінійного програмування до вигляду (12.1), (12.2).

$$\begin{cases} -W = -x_1 + 4x_2 + 3x_3 \to \max\limits_{x_{1,2,3} \ge 0}, \\ 3x_1 + 4x_2 + x_3 \le 7, \\ x_1 + 2x_2 + x_3 \le 6, \\ -x_1 - 2x_2 - x_3 \le -6, \\ -x_3 \le -4, \end{cases} \Leftrightarrow$$

$$\Leftrightarrow \begin{cases} -W = \begin{pmatrix} -1 & 4 & 3 \end{pmatrix} \cdot \begin{pmatrix} x_1 & x_2 & x_3 \end{pmatrix}^T \to \max\limits_{x_{1,2,3} \ge 0}, \\ \begin{bmatrix} 3 & 4 & 1 \\ 1 & 2 & 1 \\ -1 & -2 & -1 \\ 0 & 0 & -1 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \le \begin{bmatrix} 7 \\ 6 \\ -6 \\ -4 \end{bmatrix} \end{cases}$$

Двоїста ЗЛП набуває вигляду:

$$\tilde{W} = \begin{bmatrix} 7 & 6 & -6 & -4 \end{bmatrix} \cdot \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{bmatrix} \to \min\limits_{u_{1,2,3,4} \ge 0}.$$

$$\begin{bmatrix} 3 & 1 & -1 & 0 \\ 4 & 2 & -2 & 0 \\ 1 & 1 & -1 & -1 \end{bmatrix} \cdot \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{bmatrix} \geq \begin{bmatrix} -1 \\ 4 \\ 3 \end{bmatrix} \Leftrightarrow \begin{cases} \tilde{W} = 7u_1 + 6u_2 - 6u_3 - 4u_4 \to \min\limits_{u_{1,2,3,4} \geq 0}; \\ 3u_1 + u_2 - u_3 \geq -1, \\ 4u_1 + 2u_2 - 2u_3 \geq 4, \\ u_1 + u_2 - u_3 - 4u_4 \geq 3. \end{cases} \bullet$$

Приклад 12.2. Пряма задача лінійного програмування задана у вигляді:

$$\begin{cases} W = 6x_1 + 10x \to \max\limits_{x_1 \geq 0}, \\ 5x_1 + 3x \geq 10, \\ x_1 - x \leq 4. \end{cases}$$

Побудувати двоїсту задачу.

❍Приведемо пряму задачу до стандартного вигляду. Для цього позначимо

$$x = x_2 - x_3 \in \mathbb{R},\ x_{2,3} \geq 0$$

Тоді:

$$\begin{cases} W = 6x_1 + 10x_2 - 10x_3 \to \max\limits_{x_{1,2,3} \geq 0}; \\ -5x_1 - 3x_2 + 3x_3 \leq -10, \\ x_1 - x_2 + x_3 \leq 4, \end{cases} \Leftrightarrow \begin{cases} W = \begin{bmatrix} 6 & 10 & -10 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \to \max\limits_{x_{1,2,3} \geq 0}; \\ \begin{bmatrix} -5 & -3 & 3 \\ 1 & -1 & 1 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \leq \begin{bmatrix} -10 \\ 4 \end{bmatrix}. \end{cases}$$

Двоїста задача набуває вигляду:

$$\begin{cases} W = \begin{bmatrix} -10 & 4 \end{bmatrix} \cdot \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \to \min\limits_{u_{1,2} \geq 0}, \\ \begin{bmatrix} -5 & 1 \\ -3 & -1 \\ 3 & 1 \end{bmatrix} \cdot \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \geq \begin{bmatrix} 6 \\ 10 \\ -10 \end{bmatrix}, \end{cases} \Leftrightarrow \begin{cases} \tilde{W} = -10u_1 + 4u_2 \to \min\limits_{u_{1,2} \geq 0}, \\ -5u_1 + u_2 \geq 6, \\ -3u_1 - u_2 \geq 10, \\ 3u_1 + u_2 \geq -10, \end{cases} \Leftrightarrow$$

$$\Leftrightarrow \begin{cases} \tilde{W} = -10u_1 + 4u_2 \to \min\limits_{u_{1,2} \geq 0}, \\ -5u_1 + u_2 \geq 6, \\ -3u_1 - u_2 = 10. \end{cases} \bullet$$

Приклад 12.3. Побудувати задачу двоїсту до заданої задачі лінійного програмування

$$\begin{cases} W = x_1 + x_2 + x_3 \to \min\limits_{x_{123} \ge 0}, \\ x_1 + 2x_2 + x_3 \ge 9, \\ 2x_1 + x_3 \ge 4. \end{cases}$$

❍ Маємо:

$$\begin{cases} W = \begin{bmatrix}111\end{bmatrix} \cdot \begin{bmatrix}x_1 \, x_2 \, x_3\end{bmatrix}^T \to \min\limits_{x_{123} \ge 0}, \\ \begin{bmatrix} 1 & 2 & 1 \\ 2 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \ge \begin{bmatrix} 9 \\ 4 \end{bmatrix}. \end{cases}$$

$$\Rightarrow \begin{cases} \tilde{W} = \begin{bmatrix}9 & 4\end{bmatrix} \cdot \begin{bmatrix}u_1 \, u_2\end{bmatrix}^T, \\ \begin{bmatrix} 1 & 2 \\ 2 & 0 \\ 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \le \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \end{cases} \Leftrightarrow \begin{cases} \tilde{W} = 9u_1 + 4u_2 \to \max\limits_{x_{123} \ge 0}; \\ u_1 + 2u_2 \le 1, \\ 2u_1 \le 1, \\ u_1 + u_2 \le 1. \end{cases}$$ ●

12.4. Основні теореми двоїстості

Теорема 12.1. Якщо одна із двоїстих задач має оптимальний розв’язок $\hat{X} = \left[\hat{x}_1, \hat{x}_2, \ldots, \hat{x}_n\right]^T$, то і інша має оптимальний розв’язок $\hat{U} = \left[\hat{u}_1, \hat{u}_2, \ldots, \hat{u}_m\right]^T$, при цьому екстремальні значення показника ефективності обох задач співпадають $\sum\limits_{j=1}^{n} c_j \hat{x}_j = \sum\limits_{i=1}^{m} b_i \hat{u}_i$.

Зауваження 12.1. Якщо показник ефективності однієї із задач необмеженний, то інша задача не має розв’язків.

Наслідок 12.1.

1. Для існування розв’язку однієї з двоїстих задач, необхідно і достатньо, щоб інша мала хоча б один розв’язок;
2. Для того щоб $X = \left[x_1, x_2, \ldots, x_n\right]^T$ та $U = \left[u_1, u_2, \ldots, u_m\right]^T$ були оптимальними розв'язками пари двоїстих задач, необхідно і достатньо, щоб виконувалась рівність:

$$\sum_{j=1}^{n} c_j x_j = \sum_{i=1}^{m} b_i u_i.$$

Теорема 12.2. Якщо будь яка змінна $\hat{x}_j \left(j = \overline{1,n}\right)$ оптимального розв'язку прямої задачі додатна, то j-те обмеження двоїстої задачі на оптимальному розв'язку перетворюється у рівність. Якщо оптимальне рівняння прямої задачі перетворює будь яке i-те обмеження у строгу нерівність, то в оптимальному розв'язку двоїстої задачі змінна $\hat{u}_i$ дорівнює 0.

Зауваження 12.2. Теорема 12.2 ще має назву «умова доповняльної нежорсткості» і формально записується у вигляді:

$$\begin{cases} \left(\sum_{i=1}^{m} a_{ij}\hat{u}_i - c_j\right) \cdot \hat{x}_j = 0, \ j = \overline{1,n}, \\ \left(\sum_{j=1}^{n} a_{ij}\hat{x}_i - b_i\right) \cdot \hat{u}_i = 0, \ i = \overline{1,m}. \end{cases}$$

Теорема 12.3. В оптимальному розв'язку двоїстої задачі значення змінних чисельно дорівнюють частинним похідним $\dfrac{\partial W_{max}}{\partial b_i}$, для прямої задачі:

$$\hat{u}_i = \frac{\partial W_{\max}}{\partial b_i}, \ i = \overline{1,m},$$

де ΔW_{max} – зміна максимального значення показника ефективності W, при одному і тому самому двоїстому плані, викликано малою зміною b_i прямої задачі.

►Покажемо, що при зміненні правих частин $b_i \left(i = \overline{1,m}\right)$ обмежень прямої задачі невідомі двоїстої задачі можливо інтерпретувати як оцінки впливу цих змінних на оптимальне значення показника ефективності прямої задачі. Позначимо

$$\Delta B = \left[\Delta b_1, \ldots, \Delta b_m\right]^T,$$

де Δb_i – приріст i-ої правої частини прямої задачі.

Розглянемо дві двоїсті задачі:

1) Пряма:

$$\begin{cases} W = C \cdot X_1 \to \max\limits_{X_1 \geq 0_n}, \\ A \cdot X_1 \leq B \end{cases}$$

та двоїста:

$$\begin{cases} \tilde{W} = B^T \cdot U_1 \to \min\limits_{U_1 \geq 0_m}, \\ A^T \cdot U_1 \geq C^T. \end{cases}$$

2) Пряма:

$$\begin{cases} W = C \cdot \mathrm{X}_2 \to \max\limits_{\mathrm{X}_1 \geq 0_\mathrm{n}}, \\ A \cdot \mathrm{X}_2 \leq \left(\mathrm{B} + \Delta \mathrm{B}\right) \end{cases}$$

де

$$\Delta B = \begin{bmatrix} \Delta b_1 \\ \vdots \\ \Delta b_m \end{bmatrix} \to 0 \Leftrightarrow \begin{bmatrix} \Delta b_1 \to 0 \\ \vdots \\ \Delta b_m \to 0 \end{bmatrix}$$

та двоїста:

$$\begin{cases} \tilde{W} = \left(B + \Delta B\right)^T \cdot U_2 \to \min\limits_{U_2 \geq 0_m}, \\ A^T \cdot U_2 \geq C^T. \end{cases}$$

Вимога, яка полягає в тому, що $\Delta B \to 0$, означає, що заміна B на $B + \Delta B$ у двоїстій задачі не призводить до зміни її оптимального розв'язку $\hat{U}_2 = \hat{U}_1.$. Тоді, відповідно до першої теореми двоїстості, можна записати для першої пари задач:

$$W_{\max_1} = \hat{W}_{\min_1} \Leftrightarrow C\hat{X}_1 = B^T \hat{U}_1$$

та для другої пари задач:

$$\mathrm{W}_{\max_2} = \hat{W}_{\min_2} \Leftrightarrow \mathrm{C}\hat{\mathrm{X}}_2 = \left(\mathrm{B} + \Delta \mathrm{B}\right)^\mathrm{T} \hat{U}_1.$$

Обчислимо приріст показника ефективності, прямої задачі

$$\Delta W_{\max} = W_{\max_2} - W_{\max_1} = \left(B + \Delta B\right)^T \cdot \hat{U}_1 - B^T \hat{U}_1 =$$

$$= \left(B^T + \Delta B^T - B^T\right)\hat{U}_1 = \Delta B^T \hat{U}_1 = \left[\Delta b_1, \ldots, \Delta b_m\right] \cdot \begin{bmatrix} \hat{u}_1 \\ \vdots \\ \hat{u}_m \end{bmatrix} = \sum_{i=1}^{m} \Delta b_i \cdot \hat{u}_i. \blacktriangleleft$$

Із доведеної теореми 12.3 випливає, що розв'язок двоїстої задачі u_i (двоїста оцінка) кількісно дорівнює приросту максимального значення показника ефективності прямої задачі при зміні правої частини i-го обмеження прямої задачі на одиницю.

Лекція 13. Аналіз розв’язку задачі лінійного програмування

13.1. Аналіз розв’язку задач лінійного програмування на основі двоїстих оцінок

Аналіз розв’язку задачі лінійного програмування

$$W=\sum_{j=1}^{n} c_j x_j \to \max_{x_j\ge 0\ \left(j=\overline{1,n}\right)} \tag{13.1}$$

за умови виконання обмежень

$$\sum_{j=1}^{n} a_{ij} x_j \le b_i \quad \left(i=\overline{1,m}\right) \tag{13.2}$$

базується на дослідженні впливу варіації параметрів c_j, a_{ij}, b_i ($i=\overline{1,m}$, $j=\overline{1,n}$) на цей розв’язок. Такий аналіз можливо виконати із використанням наступних прийомів та способів:

1) аналізу внутрішньої структури кожного із отриманих розв’язків із використанням властивостей двоїстих оцінок;

2) співставлення варіантів розв’язків, отриманих шляхом розв’язання прямої задачі, при різних значеннях її параметрів.

Перший напрямок використовує теореми двоїстості, другий напрямок пов’язаний із так званою *параметричною оптимізацією* (*параметричне програмування*), яка базується, здебільшого, на комп’ютерному обчислювальному експерименті.

Аналіз розв’язків задачі лінійного програмування на основі двоїстих оцінок

Двоїсті оцінки слід розглядати як:

1) ***Показники дефіцитності (коштовності) ресурсів і продукції.***

Як випливає із теореми 12.2. $\hat{u}_i$ є оцінкою i-го ресурсу. Чим більше значення оцінки, тим вище дефіцитність ресурсу. Для недефіцитного ресурсу $\hat{u}_i=0$.

2) ***Показник впливу обмежень на значення показника ефективності.***

В теоремі 12.3 було показано, що $\hat{u}_i=\dfrac{\partial W_{\max}}{\partial b_i}$, $i=\overline{1,m}$. При незначному прирості Δb_i, оцінка $\hat{u}_i$ є точною мірою впливу обмеження на показник ефективності. Для практики важливо знайти граничні значення правих частин системи обмежень-нерівностей прямої задачі (нижньої та верхньої границь), при яких величини оцінок $\hat{u}_i$ $\left(i=\overline{1,m}\right)$ залишаються незмінними.

3) ***Показники доцільності виробництва окремих видів продукції.***

Ця властивість випливає із теореми 12.2. Головний зміст цієї властивості полягає в тому, що в оптимальний план виробництва доцільно включати лише виробництво тієї j-ої продукції, для якої витрати на придбання і використання сировини не перевищують прибуток:

$$\sum_{i=1}^{m} a_{ij}\hat{u}_i < c_j \left(j = \overline{1,n}\right).$$

4) ***Інструмент співставлення сумарних умовних витрат та результатів.***

Ця властивість випливає із теореми 12.1, в якій встановлено зв'язок між екстремальними значеннями показника ефективності прямої та двоїстої задач.

Перейдемо до детального розгляду чотирьох записаних вище властивостей. З економічної інтерпретації двоїстих задач випливає, що рівність показника ефективності при оптимальних планах означає, що оцінка всіх затрат виробництва повинна дорівнювати оцінці вартості продукту, який виготовлено.

Означення 13.1. *Допустимим інтервалом стійкості двоїстих оцінок по відношенню до i-го обмеження* називають інтервал вигляду

$$\left[b_i - \Delta b_i^{\text{н}},\ b_i + \Delta b_i^{\text{в}}\right],\ i = \overline{1,m},$$

де $\Delta b_i^{\text{н}}$ – нижня границя зменшення; $\Delta b_i^{\text{в}}$ – верхня границя збільшення.

Якщо

$$b_i + \Delta b_i \in \left[b_i - \Delta b_i^{\text{н}}, b_i + \Delta b_i^{\text{в}}\right],$$

де Δb_i – зміна i-го ресурсу, то вплив Δb_i на змінення величини прибутку, обчисленого в результаті оптимізації прямої задачі лінійного програмування, можливо оцінити за виразом:

$$\Delta \hat{W}_i = \hat{u}_i \cdot \Delta b_i \left(i = \overline{1,m}\right). \tag{13.3}$$

При цьому сумарний вплив змінення усіх Δb_i обчислюється як сума окремих впливів.

Обчислити $\Delta b_i^{\text{н}}$ та $\Delta b_i^{\text{в}}$ можливо за формулами:

$$\Delta b_i^{\text{н}} = \min_{j=\overline{1,m_1}} \left\{\frac{\hat{x}_j}{d_{ij}}\right\},\ d_{ij} > 0 \tag{13.4}$$

та

$$\Delta b_i^{\text{в}} = \left|\max_{j=\overline{1,m_1}} \left\{\frac{\hat{x}_j}{d_{ij}}\right\}\right|,\ d_{ij} < 0, \tag{13.5}$$

де $\left\|d_{ij}\right\| = \left\|a_{ij}\right\|^{-1}$, $i = \overline{1,m_1}$, $j = \overline{1,m_1}$, m_1 – кількість базисних змінних, a_{ij} – елемент матриці A.

$$A=\left[\begin{array}{ccc:ccc} a_{11} & \cdots & a_{1m_1} & a_{1\,m_1+1} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ a_{m_11} & \cdots & a_{m_1m_1} & a_{m_1\,m_1+1} & \cdots & a_{m_1n} \end{array}\right]=\left[\begin{array}{c:ccc} A_1 & a_{1\,m_1+1} & \cdots & a_{1n} \\ & \vdots & \ddots & \vdots \\ & a_{m_1\,m_1+1} & \cdots & a_{m_1n} \end{array}\right],$$

де $A_1=\begin{bmatrix} a_{11} & \cdots & a_{1m_1} \\ \vdots & \ddots & \vdots \\ a_{m_11} & \cdots & a_{m_1m_1} \end{bmatrix}$ – квадратна матриця, що складається із коефіцієнтів при базисних змінних ($m_1 \le m$, де m_1 – кількість лінійно незалежних обмежень в прямій задачі лінійного програмування)

Доцільність включення до плану випуску нових видів продукції оцінюється параметром:

$$\Delta_j=\sum_{i=1}^{m} a_{ij}\hat{u}_i-c_j \quad \left(j=\overline{1,n}\right). \tag{13.6}$$

Якщо $\Delta_j<0$, то даний вид продукції після його введення до плану покращує план.

Якщо $\Delta_j>0$, то включення продукту до плану випуску недоцільно.

Зрозуміло, що приріст прибутку за рахунок збільшення деякого ресурсу на величину Δb_i дорівнює

$$\Delta\hat{W}_i=\Delta b_i\cdot\hat{u}_i \tag{13.7}$$

і при цьому витрати на придбання ресурсу складають

$$\Delta V_i=\Delta b_i\cdot v_i, \tag{13.8}$$

де v_i – вартість одиниці i-го ресурсу.

Висновок.

Захід із придбанням ресурсу слід вважати доцільним, якщо $\Delta\hat{W}_i-\Delta V_i>0$.

Приклад 13.1. Для виготовлення 4-х видів продукції А, Б, В, Г використовується 3 види ресурсів R_1, R_2, R_3. Наявність ресурсу, норми її витрат на виготовлення одниці продукції, питомий прибуток наведені в таблиці 13.1.

Таблиця 13.1.

Вихідні дані прямої задачі лінійного програмування

Вид ресурсу	Наявність ресурсу	Норми на виготовлення витрат одиниці продукції			
		А	Б	В	Г
R_1	240	2	1	1	3
R_2	60	1	0	2	1
R_3	300	1	2	1	0
Питомий прибуток, c_j	–	4	2	3	5

Завдання полягає в тому, щоб розв'язати пряму задачу лінійного програмування і виконати післяоптимізаційний аналіз.

Основні етапи виконання завдання полягають у наступному:

1. Знайти оптимальний розв'язок прямої та двоїстої задач $\hat{X}, \hat{U}$.

2. Визначити зміни максимального прибутку при зміні ресурсу:

$$R_1 \text{ на } \Delta b_1 = -10,$$

$$R_2 \text{ на } \Delta b_2 = 60,$$

$$R_3 \text{ на } \Delta b_3 = 30.$$

3. Оцінити доцільність введення у план випуску п'ятого виду продукції Д, якщо норми витрат ресурсів R_1, R_2, R_3 на одиницю випуску складатимуть відповідно 2, 4, 2, а питомий прибуток – $C_5 = 15$.

4. Оцінити доцільність придбання 100 одиниць ресурсу R_3 за ціною $V_3 = 0,5$.

❍**1.** 1) Постановка та розв'язок прямої та двоїстої задач лінійного програмування

$$\begin{cases} W = 4x_1 + 2x_2 + 3x_3 + 5x_4 = \begin{bmatrix} 4 & 2 & 3 & 5 \end{bmatrix} \cdot X \rightarrow \max\limits_{x_{1,2,3,4} \geq 0}, \\ \begin{bmatrix} 2 & 1 & 1 & 3 \\ 1 & 0 & 2 & 1 \\ 1 & 2 & 1 & 0 \end{bmatrix} \cdot X \leq \begin{bmatrix} 240 \\ 60 \\ 300 \end{bmatrix}, \end{cases}$$

де $X = \left[x_1, x_2, x_3, x_4\right]^T$.

2) Двоїста задача лінійного програмування

$$\begin{cases} \tilde{W} = \begin{bmatrix} 240 & 60 & 300 \end{bmatrix} \cdot U \to \min\limits_{u_{1,2,3} \geq 0}, \\ \begin{bmatrix} 2 & 1 & 1 \\ 1 & 0 & 2 \\ 1 & 2 & 1 \\ 3 & 1 & 0 \end{bmatrix} \cdot U \geq \begin{bmatrix} 4 \\ 2 \\ 3 \\ 5 \end{bmatrix}, \end{cases}$$

де $U = \left[u_1, u_2, u_3\right]^T$.

3) Результат обчислення оптимального рішення за допомогою функцій Mathlab набуває вигляду:

- Прямої задачі лінійного програмування

$\hat{X} = \left[\hat{x}_1, \hat{x}_2, \hat{x}_3, \hat{x}_4\right]^T = \begin{bmatrix} 60 & 120 & 0 & 0 \end{bmatrix}^T$,

$\hat{W} = 480$.

- Двоїстої задачі лінійного програмування

$\hat{U} = \left[\hat{u}_1,\ \hat{u}_2,\ \hat{u}_3\right]^{\mathrm{T}} = \begin{bmatrix} 1{,}6 & 0{,}6 & 0{,}2 \end{bmatrix}^{\mathrm{T}}$,

$\hat{W} = 480$.

Аналіз результатів розв’язку прямої задачі лінійного програмування:

а) найбільш дефіцитним є ресурс R_1, $\hat{u}_1 = 1{,}6$, найменш дефіцитним є ресурс R_3, $\hat{u}_3 = 0{,}2$;

б) обчислимо зміну максимального прибутку при зміні ресурсу.

1. Знайдемо інтервали стійкості двоїстих оцінок, тобто межі, в яких вони точно вимірюють вплив зміни обмежень на показник ефективності прямої задачі лінійного програмування.

1.1. Визначення інтервалу стійкості оцінки по відношенню до обмеження b_1, пов’язаного з першим ресурсом.

Базисними невідомими, що входять до оптимального плану є x_1, x_2, x_3, тому матриця A_1 набуває вигляду:

$$A_1 = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 0 & 2 \\ 1 & 2 & 1 \end{bmatrix} = \left\| a_{ij} \right\|_{3\times 3}.$$

Тоді $A_1^{-1} = \left\| a_{ij} \right\|_{3\times 3}^{-1} = \left\| d_{ij} \right\|_{3\times 3}$:

$$A_1^{-1} = \begin{bmatrix} 0,8 & -0,2 & -0,4 \\ -0,2 & -0,2 & 0,6 \\ -0,4 & 0,6 & 0,2 \end{bmatrix}.$$

Знаходимо

$$\Delta b_1^{\text{н}} = \min_{j=1,2,3} \left\{ \frac{60}{0,8} \right\} = 75\,,\ d_{1j} > 0\,;$$

$$\Delta b_1^{\text{в}} = \left| \max_{j=1,2,3} \left\{ \frac{120}{-0,2}; \frac{0}{-0,4} \right\} \right| = 0\,,\ d_{1j} < 0\,.$$

Остаточно інтервал стійкості оцінок по відношенню до першого обмеження набуває вигляду:

$$\left[b_1 - \Delta b_1^{\text{н}}, b_1 + \Delta b_1^{\text{в}} \right] = \left[240 - 75, 240 + 0 \right] = \left[165, 240 \right].$$

1.2. Визначаємо інтервал стійкості по відношенню до ресурсу 2:

$$\Delta b_2^{\text{н}} = \min_{j=1,2,3} \left\{ \frac{0}{0,6} \right\} = 0\,,\ d_{2j} > 0;$$

$$\Delta b_2^{\text{в}} = \left| \max_{j=1,2,3} \left\{ \frac{60}{-0,2}; \frac{120}{-0,2} \right\} \right| = 300,\ d_{2j} < 0.$$

Остаточно:

$$\left[b_2 - \Delta b_2^{\text{н}}, b_2 + \Delta b_2^{\text{в}} \right] = \left[60 - 0, 60 + 300 \right] = \left[60, 360 \right].$$

1.3. Визначаємо інтервал стійкості по відношенню до ресурсу 3:

$$\Delta b_3^{\text{н}} = \min_{j=1,2,3} \left\{ \frac{120}{0,6}; \frac{0}{0,2} \right\} = 0\,,\ d_{3j} > 0\,;$$

$$\Delta b_3^{\text{в}} = \left| \max_{j=1,2,3} \left\{ \frac{60}{-0,4} \right\} \right| = 150\,,\ d_{3j} < 0\,.$$

Остаточно:

$$\left[b_3 - \Delta b_3^{\text{н}}, b_3 + \Delta b_3^{\text{в}} \right] = \left[300 - 0, 300 + 150 \right] = \left[300; 450 \right].$$

Перевіримо, чи перебуває ресурс в межах стійкості після його зміни згідно п.2. завдання.

2. $R_1 : b_1 = 240 - 10 = 230 \in \left[165, 240 \right],$

$R_2 : b_2 = 60 + 60 = 120 \in \left[60, 360 \right],$

$R_3 : b_3 = 300 + 30 = 330 \in \left[300, 450 \right].$

Висновки:

1) Після зміни весь ресурс перебуває в межах стійкості

2) Вплив зміни окремого ресурсу на максимальний прибуток, тобто $\Delta \hat{W}_i \ \left(i = 1,2,3 \right)$ можливо обчислити за формулою

$$\underline{\Delta \hat{W}_i = \hat{u}_i \cdot \Delta b_i \ \left(i = 1,2,3\right)},$$

$$\begin{aligned} \Delta \hat{W}_1 &= \hat{u}_1 \cdot \Delta b_1 = 1,6 \cdot \left(-10\right) = -16, \\ \Delta \hat{W}_2 &= \hat{u}_2 \cdot \Delta b_2 = 0,6 \cdot 60 = 36, \\ \Delta \hat{W}_3 &= \hat{u}_3 \cdot \Delta b_3 = 0,2 \cdot 30 = 6. \end{aligned} \quad ,$$

Сумарний вплив знаходимо як суму окремих впливів:

$$\Delta \hat{W} = -16 + 36 + 6 = 26.$$

3. Оцінимо доцільність введення у план випуску п’ятого виду продукції Д. Для цього обчислимо характеристику Δ_5:

$$\Delta_5 = \sum_{i=1}^{3} a_{i5}\hat{u}_i - C_5 = 2 \cdot 1,6 + 4 \cdot 0,6 + 2 \cdot 0,2 - 15 = -9 < 0$$

Висновок: прибуток перевищує витрати, тому введення у план випуску п’ятого виду продукції є вигідним.

4. Приріст ресурсу R_3 на величину $\Delta b_3 = 100$

$$b_3 = 300 + 100 = 400 \in \left[300, 450\right]$$

не виводить цей ресурс за межі стійкості, тому

$$\Delta \hat{W}_3 = \hat{u}_3 \cdot \Delta b_3 = 0,2 \cdot 100 = 20,$$

але витрати на придбання 100 одиниць ресурсу R_3 складатимуть:

$$\Delta V_3 = \Delta b_3 \cdot V_3 = 100 \cdot 0,5 = 50 > 20,$$

тобто величина додаткового прибутку менше за витрати.

Висновок: закупка ресурсу R_3 на умовах п.4 недоцільна.●

13.2. Аналіз розв’язку задачі лінійного програмування на основі використання параметричного програмування

Дослідження змінення розв’язку задачі лінійного програмування залежно від змінення параметрів математичної моделі задачі лінійного програмування є предметом *параметричного програмування* в задачі лінійного програмування. Постановка загальної задачі параметричного програмування в задачі лінійного програмування формулюється таким чином: знайти на відрізку $\left[\alpha;\beta\right]$ скінчену кількість інтервалів $I_q \subset \left[\alpha;\beta\right]$, $q = \overline{1,Q}$, $Q \in \mathbb{N}$, що містять значення параметра $t \in I_q$ $\left(q = \overline{1,Q}\right)$, для яких оптимальний розв’язок задачі лінійного програмування виду

$$W = \sum_{j=1}^{n} \left(c'_j + c''_j t\right) \cdot x_j \to \max_{x_j \ge 0 \left(j = \overline{1,n}\right)} \tag{13.9}$$

за умови виконання обмежень

$$\sum_{j=1}^{n}\left(a_{ij}^{'}+a_{ij}^{''}t\right)\cdot x_j \le b_i^{'}+b_i^{''}t,\ i=\overline{1,m} \tag{13.10}$$

досягається в одній і тій самій кутовій точці (вершині) області допустимих розв'язків.

Економічний зміст задачі параметричного програмування в задачі лінійного програмування полягає в тому, що з плином часу, який можливо інтерпретувати як параметр t, змінюються як вартість $c_j\left(t\right)$ одиниці j - того виробленого товару (вартість на момент виготовлення плюс вартість, пов'язана із строком збереження), так і запаси $b_i\left(t\right)$ i - того ресурсу (відбувається вихід із ладу за рахунок старіння та перебування довгий час у неробочому стані) та витрат ресурсу $a_{ij}\left(t\right)$ на виготовлення одиниці товару (наприклад, пов'язані з модернізацією виробництва).

Дамо геометричну інтерпретацію задачі параметричного програмування в задачі лінійного програмування. Розглянемо залежність від часу t лише коефіцієнтів показника ефективності. Задача параметричного програмування в задачі лінійного програмування набуває вигляду: для кожного $t\in\left[\alpha;\beta\right]$ знайти свій вектор змінних, який максимізує показник ефективності:

$$W_t=\sum_{j=1}^{n}\left(c_j^{'}+c_j^{''}t\right)\cdot x_j=C\left(t\right)\cdot X\to \max_{x_j\ge 0\left(j=\overline{1,n}\right)} \tag{13.11}$$

при виконанні умов:

$$\sum_{j=1}^{n}a_{ij}x_j\le b_i^{'},\ i=\overline{1,m}, \tag{13.12}$$

де $C\left(t\right)=\left[c_1^{'}+c_1^{''}t,\ldots,c_n^{'}+c_n^{''}t\right],\ X=\left[x_1,\ldots,x_n\right]^T$.

Припустимо, що область допустимих розв'язків є опуклим багатогранником. Рівнянню $C\left(t\right)\cdot X=0$ відповідає сім'я гіперплощин, що проходять через початок координат.

Якщо параметру t надати деяке значення $t=t_1\in\left[\alpha;\beta\right]$, то гіперплощина займе фіксоване положення. Переміщення гіперплощини $C\left(t_1\right)\cdot X=W_{t_1}$ паралельно фіксованому положенню у бік збільшення W_{t_1} призведе до попадання цієї площини, наприклад, у вершину A області допустимих розв'язків (рис. 13.1).

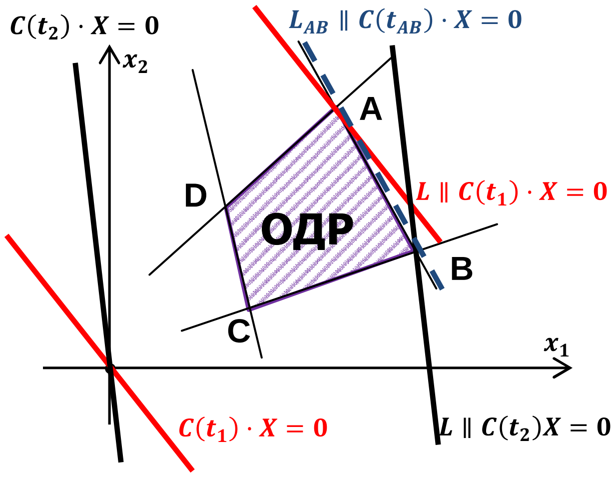


Рис. 13.1. Геометрична інтерпретація задачі параметричного програмування в задачі лінійного програмування

Змінення t в межах $t \in \left(t_{AB}, t_{AD}\right)$ залишає незмінним оптимальний розв'язок $\hat{X} = X_A$.

Якщо

$$\begin{cases} t \in \left(-\infty, t_{AB}\right) \bigcap \left[\alpha;\beta\right] \\ t_{AB} > \alpha \end{cases}$$

тобто $t \in \left[\alpha, t_{AB}\right)$, то $\hat{X} = X_B$ і т.д.

При $t = t_{AB}$ задача лінійного програмування стає виродженою. Остаточний результат розв'язання задачі параметричного програмування в задачі лінійного програмування призведе до розбиття відрізка $\left[\alpha;\beta\right]$ на кінцеве число інтервалів, в яких зміна t не призводить до зміни оптимального розв‘язку, але в послідовно розташованих інтервалах ці розв‘язки будуть різними. Методика розв'язання задачі параметричного програмування в задачі лінійного програмування із використанням стандартних процедур лінійного програмування складається з 2 етапів:

I етап – обчислення оптимального розв‘язку задачі лінійного програмування.

Параметру $t \in \left[\alpha;\beta\right]$ надається фіксоване значення. Задача параметричного програмування перетворюється в задачу лінійного програмування із фіксованими параметрами, яку розв'язують, наприклад, симплекс-методом.

II етап – обчислення інтервалу сталості оптимального розв‘язку.

На цьому етапі знаходять інтервал $\left(t_{\text{н}}, t_{\text{в}}\right) \subset \left[\alpha;\beta\right]$, для якого оптимальний розв‘язок, отриманий за допомогою процедур етапу 1 не змінюється. Знайдений інтервал виключають із $\left[\alpha;\beta\right]$, задають нове фіксоване значення для інтервалу і повторюють обчислення у відповідності з першим етапом. Розв'язок припиняється лише тоді, коли буде виконано розбиття всього відрізку $[\alpha;\beta]$ на інтервали сталості оптимального розв'язку.

Приклад 13.2. Визначити інтервали зміни параметра $t \in [0,8]$ і знайти значення змінних x_1, x_2, при яких лінійна функція

$$W_t = 4x_1 + (2+t)x_2$$

досягає максимуму в одній і тій самій вершині області допустимих розв'язків при виконанні системи обмежень:

$$\begin{cases} 2x_1 - 5x_2 \le 10, \\ x_1 + x_2 \ge 5, \\ -x_1 + x_2 \le 4, \\ 4x_1 + 5x_2 \le 40, \quad x_{1,2} \ge 0. \end{cases}$$

❍ 1. Побудова області допустимих розв'язків (рис.13.2):

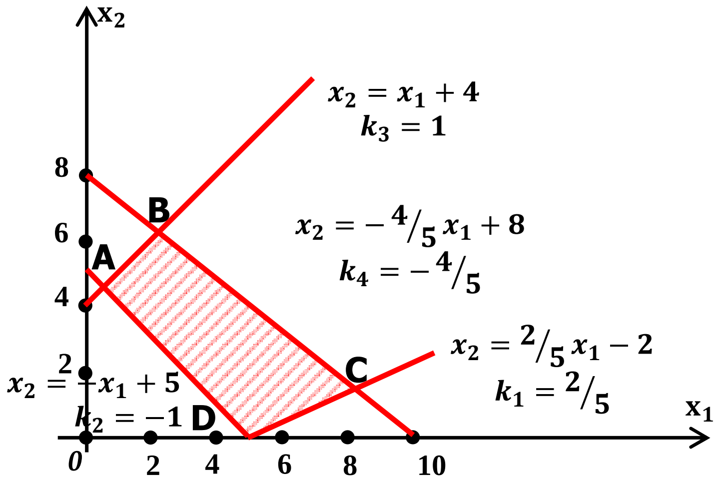


Рис. 13.2. Графічна ілюстрація побудови області допустимих розв'язків:

2. Побудова головної прямої (рис.13.3):

$$\begin{cases} W_t = 4x_1 + (2+t)\cdot x_2 \\ t \in [0,8] \end{cases} \Leftrightarrow \begin{cases} x_2 = \left(-\dfrac{4}{t+2}\right)x_1 + \left(\dfrac{W}{t+2}\right) = k(t)\cdot x_1 + b(t) \\ t \in [0,8] \end{cases},$$

$$k(t) = -\frac{4}{t+2},\ b(t) = \frac{W_t}{t+2}.$$

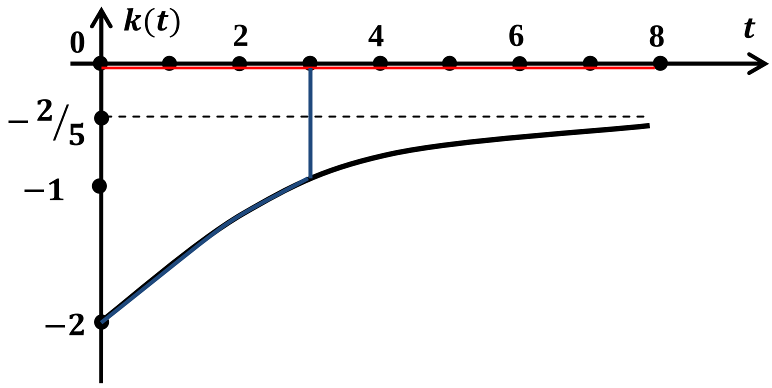


Рис. 13.1. Графік зміни кутового коефіцієнту головної прямої в залежності від зміни параметру t

Якщо $t \in [0,3)$, то $k(t) \in \left[-2,-\frac{4}{5}\right)$, $\hat{X} = X_C = \begin{bmatrix} x_{1_C} \\ x_{2_C} \end{bmatrix}$.

Знайдемо координати точки C:

$$\begin{cases} x_2 = -\frac{4}{5}x_1 + 8 \\ x_2 = \frac{2}{5}x_1 - 2 \end{cases} \Leftrightarrow \begin{cases} x_2 = \frac{25}{3} \\ x_2 = \frac{4}{3} \end{cases} \Leftrightarrow \hat{X}_C = \begin{bmatrix} 25/3 \\ 4/3 \end{bmatrix}$$

Якщо $t \in (3,8]$, то $k(t) \in \left(-\frac{4}{5},-\frac{2}{5}\right]$, $\hat{X} = X_B = \begin{bmatrix} x_{1_B} \\ x_{2_B} \end{bmatrix}$.

Знайдемо координати точки B:

$$\hat{X}_{\text{B}} = \begin{bmatrix} 20/9 \\ 56/9 \end{bmatrix}.$$

3. Відповідь:

1) $t \in [0, 3) \to \hat{X}_C = \left[\frac{25}{3}; \frac{4}{3}\right]^{\text{T}}$;

2) $t \in (3, 8] \to \hat{X}_B = \left[\frac{20}{9}; \frac{56}{9}\right]^{\text{T}}$;

3) $t = 3$, безліч розв'язків: $\hat{X} = \left[x_1 \in \left[\frac{20}{9}; \frac{25}{3}\right], x_2 = -\frac{4}{5}x_1 + 8\right]^T$. ●

Приклад 13.3. Розв'язати графічним методом задачу параметричного програмування в задачі лінійного програмування

$$\begin{cases} W_t = tx_1 + (1+t)x_2 \to \max\limits_{\substack{x_{1,2}\ge 0 \\ t\in[1;7]}}; \\ -3x_1 + 4x_2 \le 11, \\ 4x_1 + x_2 \le 8. \end{cases}$$

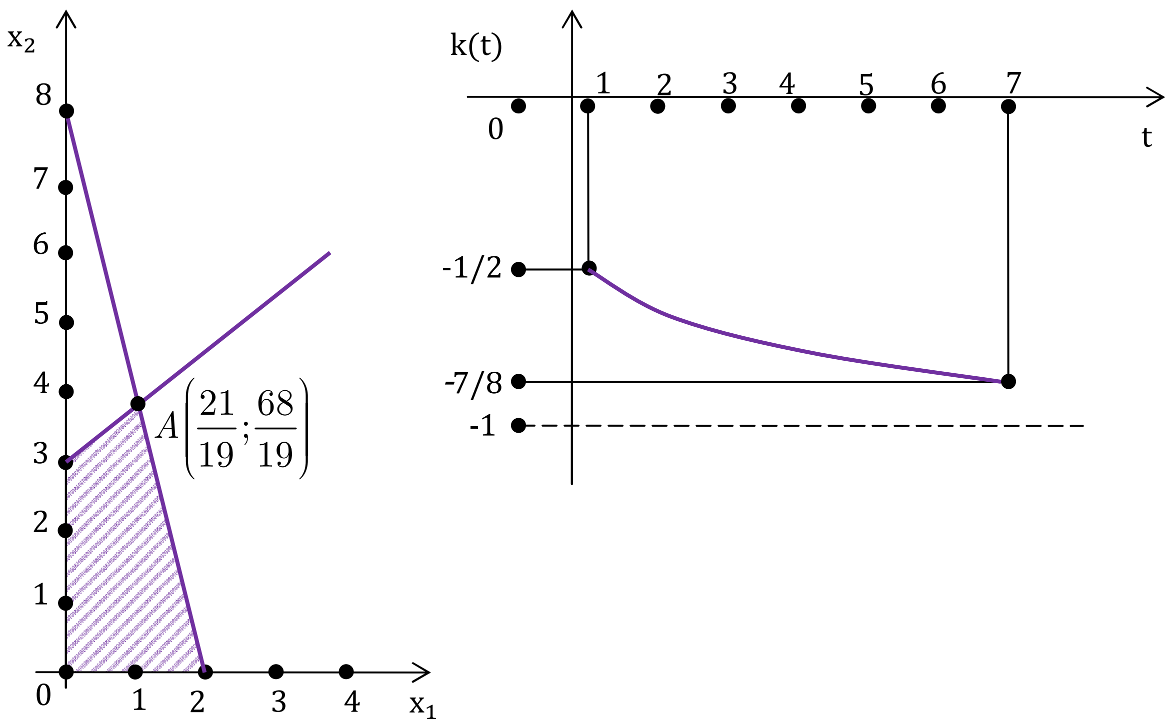


Рис. 13.4. Графічне розв'язання задачі параметричного програмування

Відповідь:

$$t \in [1;7] \;\to\; \hat{X}_A = \left[\frac{21}{19}; \frac{68}{19}\right]^{\mathrm{T}} \bullet.$$

Лекція 14. Цілочисельні задачі лінійної оптимізації

14.1. Постановка цілочисельних задач лінійної оптимізації

При розв'язанні більшості телекомунікаційних задач необхідно, щоб величини (кількість каналів зв'язку, окремих агрегатів обладнання, обслуговуючого персоналу, абонентів, запитів та відповідей і т.д.) обчислювались у цілих, невід'ємних числах.

Означення 14.1. *Задачі цілочисельного програмування* — різновид задач, що припускають, що шукані значення повинні бути цілими числами.

Задачі цілочисельної оптимізації можуть бути, як лінійними, так і нелінійними. В даній лекції розглядаються цілочисельні задачі, саме лінійної оптимізації, або, як ще кажуть, цілочисельні задачі лінійного програмування.

В цих задачах і показник ефективності і обмеження є лінійними функціями своїх аргументів, але всі, або частина цих аргументів можуть бути лише цілими та невід'ємними. Якщо вимога невід'ємної цілочисельності стосується лише частинних змінних, то така задача носить назву «частково цілочисельна».

Математична постановка цілочисельних задачі лінійного програмування має вигляд:

$$\begin{cases} W = \sum_{j=1}^{n} c_j x_j \to \max\limits_{x_j \in \mathbb{N}_0 \left(j=\overline{1,n}\right)} (\min) \\ \sum_{j=1}^{n} a_{ij} x_j \left\{\leq = \geq\right\} b_i, \; i = \overline{1,m} \\ \mathbb{N}_0 = \left\{0,1,2,3,\ldots\right\} - \text{розширений ряд натуральних чисел} \end{cases} \tag{14.1}$$

14.2. Метод Гоморі розв'язання цілочисельних задач лінійного програмування

Метод Гоморі будується на застосуванні симплекс методу і методу відтинання.

Головна ідея методу Гоморі полягає в наступному.

На першому етапі знаходиться оптимальний розв'язок задачі, яка була отримана із вихідної цілочисельної задачі лінійного програмування, за умови, що вимога цілочисельності не враховувалась. Якщо отримане рішення є цілочисельним, то мета цілочисельної задачі лінійного програмування досягнута і обчислення припиняється. Якщо розв‘язок задачі лінійного програмування не є цілочисельним, то його не вважають розв'язком задачі лінійного програмування і відбувається перехід до другого етапу.

На другому етапі до умов задачі додаються обмеження, які відтинають від області допустимих розв‘язків отриманий нецілочисельний розв'язок, але не відтинають жодного цілочисельного розв'язоку, що належить області допустимих розв‘язків. Після цього відбувається повернення до першого етапу і знову виконується розв'язання задачі лінійного програмування симплекс методом, але із урахуванням нового обмеження. Якщо новий розв'язок теж не є цілочисельним, то відбувається перехід до другого етапу і до складу обмеження включається ще одне додаткове обмеження і виконується повернення до першого етапу.

Вихід з двохетапної процедури відбувається тоді, коли буде знайдено цілочисельний розв’язок, що належить області допустимих розв‘язків.

Для з’ясування специфічних особливостей методу Гоморі, розглянемо його геометричний образ (рис.14.1).

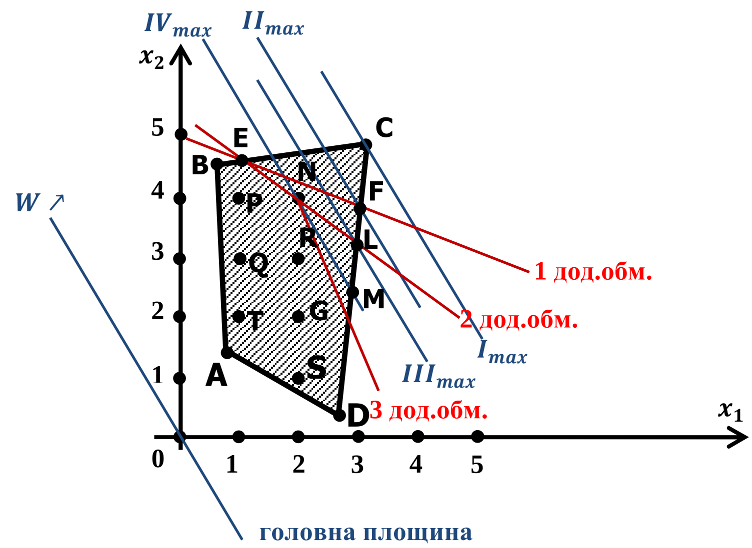


Рис. 14.1. Геометричнй образ ідеї методу Гоморі

Означення 14.2. *Послабленою цілочисельною задачею лінійного програмування* називають задачу, у якій знято обмеження на цілочисельність.

Для послабленої цілочисельної задачі лінійного програмування (або просто задачі лінійного програмування) максимум досягається в точці C, але x_2, не є цілим. Точка С(дробовий розв’язок) відсікається додатковим обмеженням EF (перше додаткове обмеження), зберігаючи область допустимих рішень цілочисельної задачі лінійного програмування (точки P, N, Q, R, T, G, S).

Будуємо додаткове обмеження 1 та переходимо до послабленої задачі що досягає максимуму в точці F, і при цьому оптимальне значення x_1 буде дробовим. Будуємо додаткове обмеження 2 (ENL) та переходимо до максимуму показника ефективності задачі лінійного програмування в точку L, в якій знову оптимальні x_1, x_2 – дробові. Будуємо додаткове обмеження 3, яке в послабленій задачі лінійного програмування дає кутову точку N, в якій оптимальні x_1, x_2 є цілочисельними і додатними. Вважаємо, що цілочисельна задача лінійного програмування розв’язана.

Як бачимо, додаткові обмеження кожного разу відтинали від області допустимих розв‘язків послабленої цілочисельної задачі лінійного програмування (тобто області допустимих розв‘язків задачі лінійного програмування) дробові розв’язки і зберігали область допустимих розв‘язків цілочисельної задачі лінійного програмування.

Критеріәм побудови "правильного" додаткового обмеження є те, що воно повинно відтинати оптимальну нецілочисельну точку і залишати незайманими усі цілочисельні точки вихідної цілочисельної задачі лінійного програмування.

Приклад 14.1. Розв‘язати цілочисельну задачу лінійного програмування методом Гоморі:

$$\begin{cases} W = 2x_1 + 4x_2 + 3x_3 \to \max\limits_{x_{1,2,3} \in \mathbb{N}_0}; \\ 2x_1 + 3x_2 + x_3 \le 20, \\ 9x_1 + 7x_2 + 10x_3 \le 40. \end{cases}$$

❍ **Перший етап.** Сформулюємо послаблену цілочисельну задачу лінійного програмування у вигляді основної задачі лінійного програмування:

$$\begin{cases} -W = -2x_1 - 4x_2 - 3x_3 + 0x_4 + 0x_5 \to \min\limits_{x_{1,2,3,4,5} \ge 0} \\ 2x_1 + 3x_2 + x_3 + x_4 + 0x_5 = 20, \\ 9x_1 + 7x_2 + 10x_3 + 0x_4 + x_5 = 40. \end{cases}$$

За допомогою функції **linprog** системи комп’ютерної математики MATLAB знайдемо оптимальний розв’язок:

```
>> f=[-2;-4;-3;0;0]
f =        -2
          -4
          -3
           0
           0
>> A=[2  3  1    1  0
      9  7  10   0  1]
A = 2    3    1    1    0
    9    7    10   0    1
>> b=[20;40]
b =20
   40
>> lb=[0;0;0;0;0]
lb = 0
     0
     0
     0
     0
>> [X,W]=linprog(f,[],[],A,b,lb,[])
Optimization terminated successfully.
X = 0.0000
    5.7143
    0.0000
    2.8571
```

0.0000

W = -22.8571

Висновок: серед базисних змінних x_1, x_2, x_3 є дробовий розв'язок x_2. Тому цей розв‘язок не можна вважати оптимальним розв'язком вихідної цілочисельної задачі лінійного програмування.

Другий етап. Виконаємо побудову першого додаткового обмеження:

$$x_2 = \frac{40}{7} - \left(\frac{9}{7}x_1 + \frac{10}{7}x_3 + \frac{1}{7}x_5\right) \Leftrightarrow$$

$$\Leftrightarrow x_2 = \left[\frac{40}{7}\right] + \left\{\frac{40}{7}\right\} - \left(\left(\left[\frac{9}{7}\right] + \left\{\frac{9}{7}\right\}\right)x_1 + \left(\left[\frac{10}{7}\right] + \left\{\frac{10}{7}\right\}\right)x_3 + \left(\left[\frac{1}{7}\right] + \left\{\frac{1}{7}\right\}\right)x_5\right) \Leftrightarrow$$

$$\Leftrightarrow -\left\{\frac{40}{7}\right\} + \left\{\frac{9}{7}\right\}x_1 + \left\{\frac{10}{7}\right\}x_3 + \left\{\frac{1}{7}\right\}x_5 = \left[\frac{40}{7}\right] - \left(\left[\frac{9}{7}\right]x_1 + \left[\frac{10}{7}\right]x_3 + \left[\frac{1}{7}\right]x_5\right) - x_2$$

Якщо $x_{1,2,3}$ та $x_{4,5}$ цілі, то і права частина записаного вище рівняння – ціле число. Тоді і ліва частина останнього рівняння теж буде цілим числом: $-\frac{5}{7} + \frac{2}{7}x_1 + \frac{3}{7}x_3 + \frac{1}{7}x_5 = x_6$, яке задовольняє нерівності $x_6 \ge -\frac{5}{7}$, x_6 – ціле невід'ємне число, $x_6 \in \{0,1,2,\ldots\}$. Нагадаємо, що $x_{1,2,3,4,5}$ невід'ємні.

Записане означає, що додаткое обмеження у вигляді нової змінної x_6, не обежено нічим, окрім загальної вимоги до невід'ємності та цілочисельності. З цього випливає, що нове обмеження може вплинути на дробовий розв'язок задачі лінійного програмування, залишивши цілочисельні незмінними.

Сформулюємо на прикладі x_6 правило побудови додаткового обмеження:

$$\left(\frac{9}{7} - \left[\frac{9}{7}\right]\right)x_1 + \left(\frac{10}{7} - \left[\frac{10}{7}\right]\right)x_3 + 0x_4 + \left(\frac{1}{7} - \left[\frac{1}{7}\right]\right)x_5 - x_6 = \frac{40}{7} - \left[\frac{40}{7}\right].$$

Послаблену цілочисельну задачу лінійного програмування із додатковим обмеженням розвязуеимо із використанням функції **linprog:**

```
>> f=[-2;-4;-3;0;0;0]
f = -2
    -4
    -3
     0
     0
     0
>>  A=[2  3  1    1   0  0
       9  7  10   0   1  0
       2/7 0 3/7  0  1/7 -1]
A =  2.0000    3.0000    1.0000    1.0000         0         0
     9.0000    7.0000   10.0000         0    1.0000         0
     0.2857         0    0.4286         0    0.1429   -1.0000
```

```
>> b=[20;40;5/7]
b = 20.0000
     40.0000
      0.7143
>> lb=[0;0;0;0;0;0]
lb =  0
      0
      0
      0
      0
      0
>> [X,W]=linprog(f,[],[],A,b,lb,[])
Optimization terminated successfully.
X =
   0.0000
   5.0000
   0.0000
   5.0000
   5.0000
   0.0000
W = -20.0000.
```

Висновок: $x_1 = 0$, $x_2 = 5$, $x_3 = 0$, $W = 20$ - оптимальний цілочисельний розв'язок цілочисельної задачі лінійного програмування.●

14.3. Метод гілок та границь розв'язання цілочисельної задачі лінійного програмування

Ідея методу та прийоми його застосування полягають в тому, що спочатку в області допустимих розв‘язків системи обмежень знаходять оптимальний розв‘язок послабленої цілочисельної задачі лінійного програмування із використанням, наприклад, симплекс-методу. Для визначеності покладемо, що відбувається пошук максимуму показника ефективності. Якщо в отриманому розв'язку деякі базисні змінні є дробами, то обираємо будь-яку із цих змінних і будуємо два обмеження: в одному обмежена величина змінної менше або дорівнює розв'язку задачі лінійного програмування округленому до найближчого цілого числа знизу; а, в іншому, обмежена величина змінної більше або дорівнює розв'язку задачі лінійного програмування округленому до найближчого цілого числа зверху. Наприклад, при побудові додаткового обмеження за деякою i-ою змінною $x_i = \frac{9}{2} \in \left[4;5\right]$, перше обмеження набуває вигляду $x_i \le 4$, а друге обмеження набуває вигляду $5 \le x_i$.

Першим та другим додатковим обмеженням виключаємо з області допустимих розв‘язків вихідної цілочисельної задачі лінійного програмування проміжок $\left(4;5\right)$ із дробовими значеннями невідомої x_i. Цей проміжок поділяє область допустимих розв‘язків на дві частини ОДР$_1$ та ОДР$_2$, де нова ОДР$_1$ отримана завдяки врахуванню в обмеженні вихідної задачі додаткового обмеження $x_i \le 4$, а ОДР$_2$ – врахуванням обмеження $5 \le x_i$.

В результаті такого розподілу області допустимих значень отримаємо дві нові задачі (підзадачі) лінійного програмування. Якщо після їх розв’язання отримаємо значеня невідомих, які знову не є цілочисельними, то обираємо для продовження обчислень, що наближає до цілочисельного розв’язку ту підзадачу, у якої значення показника ефективності більше і знову повторюємо процедуру формування додаткових обмежень і пов’язаних з ними підзадач. Графічнообразне сприйняття процесу наближення до цілочисельного оптимального розв‘язку і дало назву методу "гілок і границь".

Приклад 14.2. Для ОДР$_1$ цілочисельної задачі лінійного програмування (ЦЗЛП) необхідно знайти максимум показника ефективності W_{max}.

❍ Побудуємо схему.

Рис. 14.2. Приклад структури методу гілок та границь при умові, що всі змінні окрім x_i на кожній ітерації залишаються цілочисельними та невід’ємними

Процедура розгалуження припиняється тоді, коли буде знайдено цілочисельний розв’язок. Границями вважаються значення показника ефективності задачі і підзадач кожної гілки. На кожній ітерації подальшому розгалуженню (розподіленню на нові підзадачі) підлягає та гілка, на якій значення показника ефективності виявилося більшим. Тому окремі підзадачі (гілки, за якими отримано менші значення показника ефективності) можливо відкинути, але можливі випадки з поверненням, коли у подальших ітераціях відбулося зменшення показника ефективності нижче раніше виявленого рівня на гілках, що вважалися неперспективними. Оскільки множина усіх розв’язків цілочисельної задачі лінійного програмування скінченна, то після скінченної кількості ітерацій, оптимальний розв’язок буде знайдено. ●

Приклад 14.3. Знайти максимум показника ефективності методом гілок та границь із використанням функції **linprog** системи комп’ютерної математики MATLAB:

$$\begin{cases} W = 2x_1 + 4x_2 + 3x_3 \to \max\limits_{x_{1,2,3} \in \mathbb{N}_0}; \\ 2x_1 + 3x_2 + x_3 \le 20, \\ 9x_1 + 7x_2 + 10x_3 \le 40. \end{cases}$$

❍ **Метод гілок і границь.**

Вихідна послаблена ЦЗЛП, **представлена як симетрична** ЗЛП у форматі

```
>> f=[-2;-4;-3]
>> A=[2   3   1
        9   7   10]
>> b=[20;40]
>> lb=[0;0;0]
>> [X,W]=linprog(f,A,b,[],[],lb,[])
Optimization terminated successfully.
X =   0.0000
        5.7143
        0.0000
W = -22.8571
```

Гілка 1

```
>> A=[2   3   1
         9   7   10
          0   1   0]
>> b=[20;40;5]
>> [X,W]=linprog(f,A,b,[],[],lb,[])
Optimization terminated successfully.
X = 0.0000
      5.0000
      0.5000
W = -21.5000
```

Гілка 1.1

```
>> Aeq=[0   0   1]
>> beq=[0]
>> [X,W]=linprog(f,A,b,Aeq,beq,lb,[])
```

```
Optimization terminated successfully.
X = 0.5556
      5.0000
        0
W = -21.1111
```

Гілка 1.1.1

```
>> Aeq=[1   0   0
              0   0   1]
>> beq=[0;0]
>> [X,W]=linprog(f,A,b,Aeq,beq,lb,[])
Optimization terminated successfully.
X =   0
        5.0000
        0
```

W = -20.0000

Гілка 2

```
>> A=[2   3   1
          9   7   10
          0   -1   0]
>> b=[20;40;-6]
>> [X,W]=linprog(f,A,b,[],[],lb,[])
Exiting: One or more of the residuals, duality gap, or total relative error
        has stalled:
        the primal appears to be infeasible (and the dual unbounded).
        (The dual residual < TolFun=1.00e-008.)
X =  0.0001
       5.8937
       0.0006
W =  -23.5768
```

Висновок: отримано оптимальний цілочисельний розв'язок

$x_1 = 0$, $x_2 = 5$, $x_3 = 0$, $W = 20$. ●

Лекція 15. Нелінійне програмування

15.1. Класичні умови екстремуму задачі нелінійного програмування

Загальна задача нелінійного програмування полягає в знаходженні екстремуму показника ефективності:

$$W = \left(x_1, \ldots, x_n\right) \to \underset{x_j \in \mathbb{R},\ \ j=\overline{1,n}}{extrem} \tag{15.1}$$

на множині планів задачі, яка задається системою умов

$$g_i\left(x_1,\ldots,x_n\right)\left\{\leq=\geq\right\}b_i\ \left(i=\overline{1,m}\right), \quad (15.2)$$

де умова дискретності змінних та їх знаку теж враховується за допомогою функцій $g_i\left(x_1,\ldots,x_n\right)$. Підкреслимо, що функції $g_i\left(x_1,\ldots,x_n\right)$ $\left(i=\overline{1,m}\right)$ та показник ефективності $W\left(x_1,\ldots,x_n\right)$ усі або частково є нелінійними.

До специфічних властивостей задачі нелінійного програмування слід віднести:

1) багатоекстремальність;

2) оптимальний план (план, на якому досягається, наприклад, глобальний максимум) може бути як внутрішньою, так і граничною точкою області допустимих розв'язків;

3) область допустимих розв'язків(ОДР) може не бути опуклою та зв'язаною, тобто, складається з декількох частин, що не перетинаються (рис.15.1);

4) показник ефективності та обмеження можуть мати кутові точки та розриви.

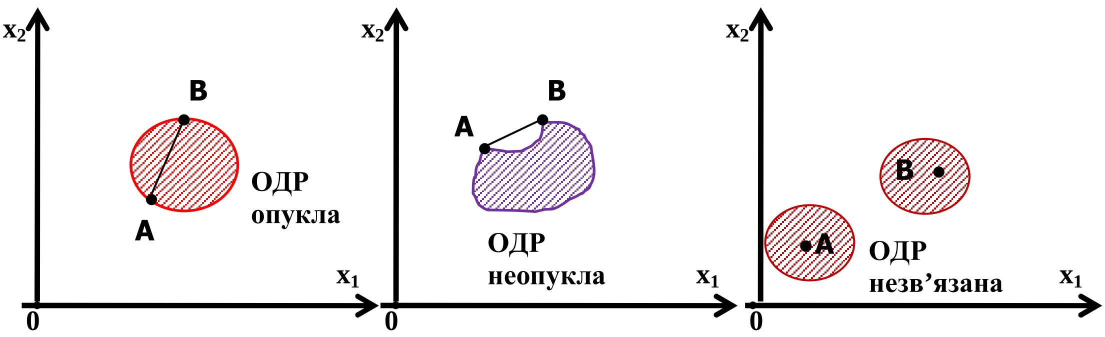


Рис.15.1. Графічне пояснення щодо особливостей ОДР

Класична детермінована задача пошуку екстремуму, тобто максимуму або мінімуму деякої неперервної та кусково-диференційованої функції n змінних ставиться наступним чином: знайти всі значення вектора $X^T=\left[x_1,\ldots,x_n\right]$, при яких функція $W\left(X\right)$ досягає екстремуму, при виконанні ОДР G , яка задається за допомогою обмежень у формі рівнянь,тобто у обмеженнях (15.2) існує лише m обмежень-рівностей, і математично записується у вигляді:

$$W\left(X\right)\to \underset{X\in G\subset\mathbb{R}^n}{extrem}.$$

Різницю $n-m$ називають *кількістю степенів свободи* задачі нелінійного програмування, а саму задачу нелінійного програмування із умовами (обмеженнями) називають *задачею нелінійного програмування на умовний екстремум*. Розглянемо застосування методу прямої підстановки та методу множників Лагранжа для розв'язання класичної задачі нелінійного програмування на умовний екстремум.

15.1.1. Метод прямої підстановки

Припустимо, що

$$X^T = \left[X_k^T, X_m^T\right],$$

де $X_k^T = \left[x_1, \ldots, x_k\right]$ – k-вимірний вектор вільних змінних, $X_m^T = \left[x_{k+1}, \ldots, x_n\right]$ – m-вимірний вектор базисних змінних($k+m=n$).

Тоді розв'язуючи у явному вигляді рівняння-обмеження відносно базисних змінних $x_{k+i} = \varphi_i$, $i = \overline{i,m}$ і підставляючи їх у вираз показника ефективності, отримаємо класичну безумовну задачу нелінійного програмування:

$$W\left(x_1, \ldots, x_k, \varphi_{k+1}\left(x_1, \ldots, x_k\right), \ldots, \varphi_n\left(x_1, \ldots, x_k\right)\right) \to \underset{\left[x_1, \ldots, x_k\right] \in \mathbb{R}^k}{extrem}. \quad (15.3)$$

Якщо $k=1$, то показник ефективності буде залежати лише від однієї змінної. Припустимо, що необхідно знайти максимум показника ефективності W_{max}. За визначенням, в точці $\hat{x}_1$ досягається строгий локальний максимум, якщо $W\left(\hat{x}_1\right) > W\left(\hat{x}_1 + \Delta x_1\right)$ $\forall \Delta x_1 \to 0$.

Знайдемо диференціальні умови існування $\hat{x}_1$ – точки локального максимуму. Розкладемо $W\left(x_1\right) = W\left(\hat{x}_1 + \Delta x_1\right)$ за умови $\Delta x_1 \to 0$ в ряд Тейлора в околі точки $\hat{x}_1$:

$$\Delta W\left(\hat{x}_1\right) = W\left(\hat{x}_1 + \Delta x_1\right) - W\left(\hat{x}_1\right) =$$

$$= \left.\frac{dW\left(x_1\right)}{dx_1}\right|_{x_1=\hat{x}_1} \cdot \Delta x_1 + \frac{1}{2!}\left.\frac{d^2W\left(x_1\right)}{dx_1^2}\right|_{x_1=\hat{x}_1} \cdot \Delta x_1^2 + 0\left(\Delta x_1\right) < 0, \quad \forall \Delta x_1 \to 0.$$

Із записаного виразу отримуємо достатні умови локального максимуму функції однієї змінної:

$$\begin{cases} \left.\dfrac{dW\left(x_1\right)}{dx_1}\right|_{x_1=\hat{x}_1} = 0, \\ \left.\dfrac{d^2W\left(x_1\right)}{dx_1^2}\right|_{x_1=\hat{x}_1} < 0. \end{cases} \quad (15.4)$$

Приклад 15.1. На космічній платформі із апаратурою системи космічного зв'язку розташовано контейнер із ядерною енергетичною установкою циліндричної форми. З метою максимізації маси ядерного пального при заданій масі контейнера необхідно виготовити контейнер максимального об'єму при заданій площі поверхні (рис.15.2).

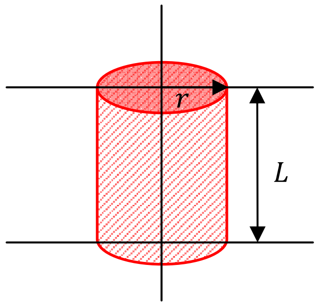


Рис. 15.2. Форма контейнера

❍Математична модель задачі:

$$\begin{cases} W(r,L) = \pi r^2 L; \\ S_0 = 2\pi r^2 + 2\pi r L. \end{cases}$$

Математична постановка задачі:

$$\begin{cases} W(r,L) = \pi r^2 L \to \max\limits_{r,L}; \\ S_0 = 2\pi r^2 + 2\pi r L \ . \end{cases}$$

Розв'язання методом прямої підстановки:

$$L = \frac{1}{2\pi r}\left(S_0 - 2\pi r^2\right),$$

$$W(r) = \frac{S_0}{2} r_0 - \pi r^3.$$

Необхідна умова екстремуму:

$$\frac{dW(r)}{dr} = 0.$$

Обчислюємо похідну:

$$\frac{S_0}{2} - 3\pi r^2 = 0 \Leftrightarrow r_{\max} = \sqrt{\frac{S_0}{6\pi}}.$$

За фізичним змістом мінімальний об'єм $V_{min} = 0$ при $r = 0$ або $L = 0$, тобто знайдене значення $r_{\max}$ відповідає максимальному об'єму. Використаємо достатні умови для формального підтвердження цього факту:

$$\frac{d^2W(r)}{dr^2} = -6\pi r < 0, \quad r > 0.$$

Отже

$$\begin{cases} r_{max} = \hat{r} = \sqrt{\dfrac{S_0}{6\pi}}; \\ L_{max} = \sqrt{\dfrac{2S_0}{3\pi}} \ . \end{cases}$$ ●

Приклад 15.2. Знайти екстремум функції $W(x_1,\ x_2)$:

$$W\left(x_1,x_2\right)=\frac{1}{\left(x_1-1\right)^2+\left(x_2-1\right)^2+1}.$$

❍

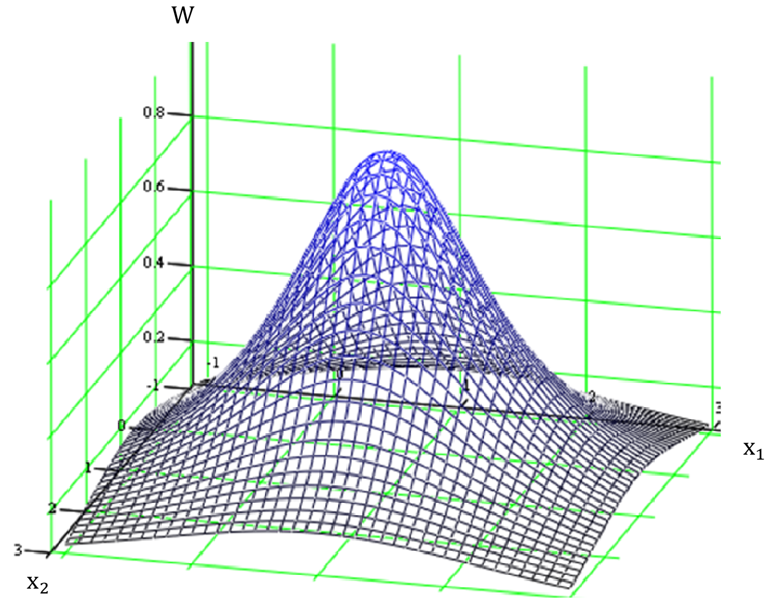


Рис. 15.3. Графічна ілюстрація щодо визначення типу екстремуму функції $W(x_1,\ x_2)$

Необхідна умова екстремуму:

$$\frac{\partial W\left(x_1,x_2\right)}{\partial x_1}=\frac{2\left(x_1-1\right)}{\left(x_1-1\right)^2+\left(x_2-1\right)^2+1}=0,$$

$$\frac{\partial W\left(x_1,x_2\right)}{\partial x_2}=\frac{2\left(x_2-1\right)}{\left(x_1-1\right)^2+\left(x_2-1\right)^2+1}=0.$$

За геометричним змістом задачі зрозуміло, що в точці екстремуму $\hat{x}_1=\hat{x}_2=1$ функція $W(x_1,\ x_2)$ приймає найбільше значення $W\left(\hat{x}_1,\hat{x}_2\right)=1$.

Як видно з графіку(див. рис.15.3), функція $W(x_1,\ x_2)$ має єдиний екстремум, що є глобальним максимумом:

$$\hat{x}_1=\hat{x}_2=1,$$

$$W\left(\hat{x}_1,\hat{x}_2\right)=1.$$ ●

Покладемо $k>1,\ m<n,\ k+m=n$, ($k=n-m-$ кількість степенів свободи) $k,n,m\in\mathbb{N}$. За означенням, $W\left(X_k\right)$ має в точці X_k строгий локальний максимум, якщо $W\left(X_k\right)>W\left(\hat{X}_k+\Delta X_k\right)\ \forall\Delta x_i\to 0\ \left(i=\overline{1,k}\right), X_k^T=\left[x_1,\ldots,x_k\right]$,

$\Delta X_k^T = \left[\Delta x_1, \ldots, \Delta x_k\right].$

Розкладання в ряд Тейлора скалярної функції векторного аргументу відносно X_k дає можливість обчислити приріст показника ефективності в малому околі точки максимуму цього показника ефективності, тобто $\hat{X}_k$:

$$\Delta W\left(\hat{X}_k\right) = W\left(\hat{X}_k + \Delta \hat{X}_k\right) - W\left(\hat{X}_k\right) =$$

$$= \left[\frac{\partial W\left(X_k\right)}{\partial X_k}\right]_{X_k = \hat{X}_k} \cdot \Delta X_k + \frac{1}{2!} \Delta X_k^T \left[\frac{\partial}{\partial X_k}\left(\frac{\partial W\left(X_k\right)}{\partial X_k}\right)^T\right]_{X_k = \hat{X}_k} \cdot \Delta X_k + o\left(\left\|\Delta X_k\right\|\right)$$

$$\forall \Delta X_k \to 0,$$

де $o\left(\left\|\Delta X_k\right\|\right)$ - доданки вищого порядку малості порівняно із

$\left\|\Delta X_k\right\| = \sqrt{\Delta x_1^2 + \Delta x_2^2 + \ldots + \Delta x_k^2}.$

Виходячи з означення локального максимуму, маємо:

$$\Delta W\left(\hat{X}_k\right) < 0, \ \forall \Delta X_k \to 0,$$

$$\left.\left(\frac{\partial W\left(X_k\right)}{\partial X_k}\right)^T\right|_{X_k = \hat{X}_k} = \left[\frac{\partial W}{\partial X_1}, \ldots, \frac{\partial W}{\partial X_k}\right] = 0.$$

тобто $\left.\dfrac{\partial W\left(X_k\right)}{\partial X_k}\right|_{X_k = \hat{X}_k}$ – градієнт показника ефективності в точці локального максимуму.

Квадратична форма

$$\frac{1}{2!} \Delta X_k^T \left[\frac{\partial}{\partial X_k}\left(\frac{\partial W\left(X_k\right)}{\partial X_k}\right)^T\right]_{X_k = \hat{X}_k} \cdot \Delta X_k < 0, \ \forall \Delta X_k \to 0, \tag{15.5}$$

побудована із використанням квадратичної симетричної матриці

$$\left[\frac{\partial}{\partial X_k}\left(\frac{\partial W\left(X_k\right)}{\partial X_k}\right)^T\right] = \begin{bmatrix} \dfrac{\partial^2 W}{\partial X_1^2} & \dfrac{\partial^2 W}{\partial X_1 \partial X_2} & \cdots & \dfrac{\partial^2 W}{\partial X_1 \partial X_k} \\ \dfrac{\partial^2 W}{\partial X_2 \partial X_1} & \dfrac{\partial^2 W}{\partial X_2^2} & \cdots & \dfrac{\partial^2 W}{\partial X_2 \partial X_k} \\ \vdots & \vdots & \ddots & \vdots \\ \dfrac{\partial^2 W}{\partial X_k \partial X_1} & \dfrac{\partial^2 W}{\partial X_k \partial X_2} & \cdots & \dfrac{\partial^2 W}{\partial X_k^2} \end{bmatrix}, \tag{15.6}$$

має назву *матриця Гессе*. Ця квадратична форма у випадку локального максимуму повинна буди від'ємно-визначеною, тобто:

$$\Delta X_k^T \left[\frac{\partial}{\partial X_k}\left(\frac{\partial W\left(X_k\right)}{\partial X_k}\right)\right]_{X_k=\hat{X}_k} \cdot \Delta X_k < 0 .$$

Квадратична форма буде від'ємно-визначеною, коли матриця Гессе задовольняє спеціальній умові. Цю матрицю Гессе, яка задовольняє умові від'ємної визначеності квадратичної форми, теж називають від'ємно-визначеною.

Ознакою від'ємної визначеності квадратичної форми, побудованої із використанням квадратної симетричної матриці

$$Q = \begin{bmatrix} q_{11} & \dots & q_{1n} \\ \dots & \dots & \dots \\ q_{n1} & \dots & q_{nn} \end{bmatrix}$$

є : усі кутові діагональні визначники матриці Q непарного порядку повинні бути від'ємними, а парного – додатними, тобто,

$$q_{11} < 0,$$

$$\begin{vmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{vmatrix} > 0,$$

$$\begin{vmatrix} q_{11} & q_{12} & q_{13} \\ q_{21} & q_{22} & q_{23} \\ q_{31} & q_{32} & q_{33} \end{vmatrix} < 0 \text{ тощо.}$$

Якщо показник ефективності в точці $\hat{X}_k$ має мінімум, то достатня умова мінімуму має вигляд:

$$\begin{cases} \left.\dfrac{\partial W\left(X_k\right)}{\partial X_k}\right|_{X_k=\hat{X}_k} = 0, \\ \Delta_k X_k^T \left.\left[\dfrac{\partial}{\partial X_k}\left(\dfrac{\partial W\left(X_k\right)}{\partial X_k}\right)^T\right]\right|_{X_k=\hat{X}_k} > 0. \end{cases} \tag{15.7}$$

Ознакою додатної визначеності квадратичної форми є додатність усіх кутових визначників квадратичної симетричної матриці Q, на якій ця форма побудована.

Приклад 15.3. Перевірити знаковизначеність квадратичної форми, заданої симетричною матрицею

$$Q=\begin{bmatrix}1 & 1 & 1\\ 1 & 2 & 1\\ 1 & 1 & 3\end{bmatrix},\ X_k \in \mathbb{R}^3.$$

❍ $q_1 = 1 > 0$,

$$q_2 = \begin{vmatrix}1 & 1\\ 1 & 2\end{vmatrix} = 2-1 = 1 > 0,$$

$$q_3 = \begin{vmatrix}1 & 1 & 1\\ 1 & 2 & 1\\ 1 & 1 & 3\end{vmatrix} = 6+1+1-2-1-3 = 2 > 0.$$

$$\left[\Delta x_1, \Delta x_2, \Delta x_3\right]\cdot\begin{bmatrix}1 & 1 & 1\\ 1 & 2 & 1\\ 1 & 1 & 3\end{bmatrix}\cdot\begin{bmatrix}\Delta x_1\\ \Delta x_2\\ \Delta x_3\end{bmatrix} > 0\ \forall\left[\Delta x_1, \Delta x_2, \Delta x_3\right]^T \in \mathbb{R}^3 \setminus \left[0,0,0\right]^T.$$

Висновок: матриця Q додатно-визначена.

Якщо задану матрицю Q помножити на -1, то отримаємо :

$$\Delta X_3^T\cdot\begin{bmatrix}-1 & -1 & -1\\ -1 & -2 & -1\\ -1 & -1 & -3\end{bmatrix}\cdot\Delta X_3 < 0,\ \text{де}\ \Delta X_3^T = \left[\Delta x_1, \Delta x_2, \Delta x_3\right]^T,$$

тобто отримана форма є від'ємно-визначеною.

Проілюструємо ознаку від'ємної визначеності квадратної симетричної матриці $-Q$:

$$-Q = \begin{bmatrix}-1 & -1 & -1\\ -1 & -2 & -1\\ -1 & -1 & -3\end{bmatrix},$$

$$q_1 = -1 < 0,\ q_2 = \begin{vmatrix}-1 & -1\\ -1 & -2\end{vmatrix} = 1 > 0,\ q_3 = \begin{vmatrix}-1 & -1 & -1\\ -1 & -2 & -1\\ -1 & -1 & -3\end{vmatrix} = -2 < 0. \bullet$$

Зауваження 15.2. Окрема умова $\dfrac{\partial W\left(X_k\right)}{\partial X_k} = 0$ має назву *необхідної умови екстремуму першого порядку*.

Умова

$$\begin{cases} \dfrac{\partial W\left(X_k\right)}{\partial X_k} = 0, \\ \dfrac{\partial}{\partial X_k}\left(\dfrac{\partial W\left(X_k\right)}{\partial X_k}\right)^T \le 0 \end{cases}$$

має назву *необхідної умови максимуму другого порядку*. Матрицю Гессе в цьому випадку називають *від'ємно-напіввизначеною*.

Необхідна умова мінімуму другого порядку набуває вигляду:

$$\begin{cases} \dfrac{\partial W\left(X_k\right)}{\partial X_k} = 0, \\ \dfrac{\partial}{\partial X_k}\left(\dfrac{\partial W\left(X_k\right)}{\partial X_k}\right)^T \ge 0. \end{cases}$$

Матрицю Гессе в цьому випадку називають *додатно-напіввизначеною*.

Як бачимо, для розв'язання задачі пошуку екстремуму скалярного показника ефективності із векторним аргументом на основі записаних вище достатніх умов необхідно, щоб $W(x_1 \dots x_k)$ мала частинні похідні до другого порядку включно за своїм аргументом, і при цьому існувала можливість отримання явного розв'язку системи обмежень, відносно базисних змінних. Але остання умова досить часто не виконується. Тому, за рахунок використання так званого принципу розширення вектору змінних та показника ефективності, було запропоновано спеціальний метод розв'язання класичної задачі нелінійного програмування на умовний екстремум, який отримав назву *метод множників Лагранжа* .

15.1.2. Метод множників Лагранжа

Метод множників Лагранжа полягає в тому, що класичну задачу нелінійного програмування на умовний екстремум з обмеженнями-рівностями замінюють спрощеною умовною задачею із розширеним вектором змінних $\left[X^T, \Lambda^T\right]^T$ де $X^T = [x_1, \dots, x_n]$– вихідній вектор змінних, $\Lambda^T = [\lambda_1, \dots, \lambda_m]$– вектор додаткових змінних, який називають *вектором множників Лагранжа*, а також із розширеним показником ефективності

$$L\left(X, \Lambda\right) = W\left(X\right) + \Lambda^T (B - G(X)). \tag{15.8}$$

Таким чином можна записати розширений показник ефективності

$$L\left(X, \Lambda\right) = W(x_1, \dots, x_k) + \sum_{i=1}^{m} \lambda_i (b_i - g_i(x_1, \dots, x_n)) \tag{15.9}$$

який називають *функцією Лагранжа* або *лагранжіаном*.

Необхідна умова екстремуму в розширеній задачі нелінійного програмування має вигляд:

$$\begin{cases} \dfrac{\partial L(X,\Lambda)}{\partial X} = \dfrac{\partial W(X)}{\partial X} + \left(-\dfrac{\partial G^T(X)}{\partial X} \right) \Lambda = 0 \Leftrightarrow \\ \Leftrightarrow \dfrac{\partial W(X)}{\partial x_j} - \sum_{i=1}^{m} \lambda_i \dfrac{\partial g_i(X)}{\partial x_j} = 0 \quad \left(j = \overline{1,n} \right); \\ \dfrac{\partial L(X,\Lambda)}{\partial \Lambda} = B - G(X) = 0 \Leftrightarrow b_i - g_i(X) = 0 \left(i = \overline{1,m} \right). \end{cases} \tag{15.10}$$

Після обчислення розв'язків записаної системи рівнянь знаходимо точки екстремумів. В залежності від умови задачі виділяємо точки, в яких показник ефективності $W(\hat{X})$ досягає максимуму або мінімуму і, в подальшому, виконуємо пошук глобального максимуму або мінімуму. В багатьох практичних випадках із фізичного змісту задачі зрозуміло, в якій точці досягається найбільше або найменше значення показника ефективності із урахуванням обмежень.

Приклад 15.4. Розв'яжемо приклад 15.1, використовуючи метод множників Лагранжа.

❍Запишемо функцію Лагранжа

$$L(r,l,\lambda) = \pi r^2 l + \lambda(S_0 - 2\pi r^2 - 2\pi r l).$$

Необхідна умова набуває вигляду :

$$\begin{cases} \dfrac{\partial L}{\partial r} = 2\pi r l + \lambda\left(-4\pi r - 2\pi l\right) = 0, & (1) \\ \dfrac{\partial L}{\partial l} = \pi r^2 + \lambda\left(-2\pi r\right) = 0, & (2) \\ \dfrac{\partial L}{\partial \lambda} = S_0 - 2\pi r^2 - 2\pi r l = 0; & (3) \end{cases}$$

Розв‘язання задачі (1) – (3):

$$(2) :\to \quad \lambda = \frac{r}{2},$$

$$(1) :\to rl - 2r^2 = 0 \to \left[\begin{array}{l} r = \dfrac{l}{2}; \\ r = 0 - \text{не задовольняє фізичному змісту.} \end{array} \right.$$

$$\left(3\right):\to S_0-6\pi r^2=0\to r_{max}=\sqrt{\frac{S_0}{6\pi}};$$

$$l_{max}=\sqrt{\frac{2S_0}{3\pi}};$$
$$\lambda_{max}=\frac{1}{2}\sqrt{\frac{S_0}{6\pi}}.\ \bullet$$

Розглянемо на прикладі двовимірної задачі лінійного програмування фізичний (економічний) зміст множників Лагранжа:

$$\begin{cases} W\left(x_1 x_2\right)\to max; \\ g\left(x_1 x_2\right)=b. \end{cases}$$

Припустимо, що умовний максимум досягається в точці $\hat{X}^T=[\hat{x}_1,\ \hat{x}_2]$. Розглядаючи залежність $\hat{X}$ від зміни параметра b, отримаємо залежності:

$$\hat{x}_1=\hat{x}_1(b), \hat{x}_2=\hat{x}_2(b),$$
$$W_{\max}=W\left(\hat{x}_1\left(b\right),\ \hat{x}_2\left(b\right)\right).$$

Знайдемо похідну від оптимального значення показника ефективності по параметру b:

$$\frac{\partial\hat{W}}{\partial b}=\frac{\partial W}{\partial\hat{x}_1}\cdot\frac{\partial\hat{x}_1}{\partial b}+\frac{\partial W}{\partial\hat{x}_2}\cdot\frac{\partial\hat{x}_2}{\partial b},$$

А також похідну по параметру b від обох частин обмеження-рівності:

$$\frac{\partial g}{\partial\hat{x}_1}\cdot\frac{\partial\hat{x}_1}{\partial b}+\frac{\partial g}{\partial\hat{x}_2}\cdot\frac{\partial\hat{x}_2}{\partial b}=1.$$

Виходячи з того, що в заданій задачі нелінійного лінійного програмування функція Лагранжа (15.9) має вигляд:

$$L\left(x_1,x_2,\lambda\right)=W\left(x_1,x_2\right)+\lambda\left(b-g\left(x_1,x_2\right)\right),$$

знаходимо частинні похідні в оптимальній точці:

$$\begin{cases} \dfrac{\partial L}{\partial\hat{x}_1}=\dfrac{\partial W}{\partial\hat{x}_1}-\hat{\lambda}\dfrac{\partial g}{\partial\hat{x}_1}=0, \\ \dfrac{\partial L}{\partial\hat{x}_2}=\dfrac{\partial W}{\partial\hat{x}_2}-\hat{\lambda}\dfrac{\partial g}{\partial\hat{x}_2}=0. \end{cases}$$

Підставляючи частинні похідні у вираз для обчислення похідної для оптимального значення показника ефективності, отримаємо:

$$\frac{\partial\hat{W}}{\partial b}=\frac{\hat{\lambda}\partial g}{\partial\hat{x}_1}\cdot\frac{\partial\hat{x}_1}{\partial b}+\frac{\hat{\lambda}\partial g}{\partial\hat{x}_2}\cdot\frac{\partial\hat{x}_2}{\partial b}=\hat{\lambda}\left(\frac{\partial g}{\partial\hat{x}_1}\cdot\frac{\partial\hat{x}_1}{\partial b}+\frac{\partial g}{\partial\hat{x}_2}\cdot\frac{\partial\hat{x}_2}{\partial b}\right)=\hat{\lambda}.$$

Узагальнюючи отриманий результат на випадок n-мірного вектора X та на випадок дії m обмежень $\left(m < n\right)$, отримаємо:

$$\left.\frac{\partial \hat{W}}{\partial b_i}\right|_{\substack{X=\hat{X}\\ \Lambda=\hat{\Lambda}}} = \hat{\lambda}_i \quad \left(i = \overline{1,m}\right). \tag{15.11}$$

Фізичний (економічний) зміст.

Вважаючи, що $W(X)$– прибуток і при цьому виконується максимізація показника ефективності, або $W(X)$– вартість (витрати) виробництва і при цьому виконується мінімізація показника ефективності, за умови, що b_i – об’єм i-го ресурсу, робимо висновок про те, що i-й множник Лагранжа показує на скільки зміниться максимальний прибуток або мінімальна вартість виробництва при зміні i-го ресурсу на 1.

Зауваження 15.3. В більшості задач, присвячених системним дослідженням складних технічних систем спеціально виділяють вектор стану, який можливо інтерпретувати як вектор базисних змінних.

Одноетапною процедурою прийняття рішення із обмеженнями у формі рівнянь називається результат розв’язання задачі нелінійного програмування, поставленої у вигляді:

$$\begin{cases} W(X,U) \to \max\limits_{X\in\mathbb{R}^n, U\in\mathbb{R}^m} (\min), \\ F\left(X,U\right) = 0. \end{cases} \tag{15.12}$$

де X– n-вимірний вектор, U– m-вимірний вектор, $W\left(X,U\right): R^n \times R^m \to R$– скалярний показник ефективності, $F\left(X,U\right): R^n \times R^m \to R^n$ $-n$- вимірна векторна функція, $F^T\left(X,U\right) = [f_1\left(X,U\right),\dots,f_n(X,U)]$.

Необхідні умови для пошуку оптимального розв’язку мають вигляд:

$$\begin{cases} \left.\dfrac{\partial L}{\partial X}\right|_{\substack{X=\hat{X}\\ U=\hat{U}\\ \Lambda=\hat{\Lambda}}} = \dfrac{\partial W\left(X,U\right)}{\partial X} + \left.\dfrac{\partial}{\partial X} F^T\left(X,U\right)\Lambda\right|_{\substack{X=\hat{X}\\ U=\hat{U}\\ \Lambda=\hat{\Lambda}}} = 0; \\ \left.\dfrac{\partial L}{\partial U}\right|_{\substack{X=\hat{X}\\ U=\hat{U}\\ \Lambda=\hat{\Lambda}}} = \dfrac{\partial W\left(X,U\right)}{\partial U} + \left.\dfrac{\partial}{\partial U} F^T\left(X,U\right)\Lambda\right|_{\substack{X=\hat{X}\\ U=\hat{U}\\ \Lambda=\hat{\Lambda}}} = 0; \\ F\left(\hat{X},\hat{U}\right) = 0, \end{cases} \tag{15.13}$$

де $L = L\left(X,U,\Lambda\right) = W\left(X,U\right) + \Lambda^T \cdot F\left(X,U\right)$, $\Lambda^T = [\lambda_1,\ldots,\lambda_m]$ – множник Лагранжа.

Приклад 15.5. Дано:

$$W\left(x_1,x_2\right) = \frac{1}{\left(x_1 - 1\right)^2 + \left(x_2 - 1\right)^2 + 1}.$$

1) Знайти точку екстремуму та із використанням матриці Гессе з'ясувати тип екстремуму.
2) Розв'язати задачу

$$\begin{cases} W\left(x_1,x_2\right) \to \max(\min); \\ x_1^2 + x_2^2 = 1 \end{cases}$$

застосовуючи метод множників Лагранжа.

❍1) Враховуючи результат розв'язання прикладу 15.2 маємо $\hat{x}_1 = \hat{x}_2 = 1$, $W\left(\hat{x}_1,\hat{x}_2\right) = 1$. Тоді в точці екстремуму матриця Гессе та її діагональні визначники набувають вигляду:

$$Q = \begin{bmatrix} -2 & 0 \\ 0 & -2 \end{bmatrix},\ q_1 = -2 < 0\ ,\ q_2 = \begin{vmatrix} -2 & 0 \\ 0 & -2 \end{vmatrix} = 4 > 0.$$

Отже, квадратична форма від'ємно визначена, тобто $(1;1)$ – точка максимуму.

2) Побудуємо лагранжіан:

$$L\left(x_1,x_2,\lambda\right) = W\left(x_1,x_2\right) + \lambda(1 - x_1^2 - x_2^2)$$

та скористаємося необхідною умовою екстремуму

$$\begin{cases} \dfrac{\partial L}{\partial x_1} = 0, \\ \dfrac{\partial L}{\partial x_2} = 0, \\ \dfrac{\partial L}{\partial \lambda} = 0. \end{cases}$$

В результаті отримаємо

$$\left[\begin{aligned} & x_{11} = x_{12} = -\frac{\sqrt{2}}{2}; \\ & x_{21} = x_{22} = \frac{\sqrt{2}}{2}. \end{aligned} \right.$$

За геометричним змістом задачі зрозуміло, що

$$\left[\begin{aligned} &x_{1\min} = x_{2\min} = -\frac{\sqrt{2}}{2};\\ &x_{1\max} = x_{2\max} = \frac{\sqrt{2}}{2}, \end{aligned}\right.$$

$$W_{min}\left(-\frac{\sqrt{2}}{2};-\frac{\sqrt{2}}{2}\right) = \frac{1}{\left(\frac{\sqrt{2}}{2}+1\right)^2 + \left(\frac{\sqrt{2}}{2}+1\right)^2 + 1},$$

$$W_{\max}\left(\frac{\sqrt{2}}{2};\frac{\sqrt{2}}{2}\right) = \frac{1}{\left(\frac{\sqrt{2}}{2}-1\right)^2 + \left(\frac{\sqrt{2}}{2}-1\right)^2 + 1}.\bullet$$

Лекція 16. Найпростіша задача нелінійного програмування в умовах невід'ємності змінних

16.1. Постановка задачі нелінійного програмування в умовах невід'ємності змінних

У випадку найпростішої задачі нелінійного програмування із усіх обмежень загальної задачі нелінійного програмування залишається лише вимога невід'ємності змінних:

$$W(X) \to \max_{X \geq 0}. \tag{16.1}$$

На прикладі скалярного показника ефективності, що залежить від двох змінних, розглянемо умови існування його максимуму при невід'ємності змінних.

Припустимо:

$$W = W\left(x_1, x_2\right) = \frac{1}{\left(x_1 - a\right)^2 + \left(x_2 - b\right)^2 + 1}.$$

Будемо позначати $x_{1\max}$, $x_{2\max}$ точки максимуму $W\left(x_1, x_2\right)$, коли обмежень немає, $x_{1,2} \in R$, $x_{1\max} = a$, $x_{2\max} = b$ (див. рис.16.1).

Будемо позначати $\hat{x}_1$, $\hat{x}_2$ координати точки максимуму $W\left(x_1, x_2\right)$ в умовах, коли $x_1 \geq 0$, $x_2 \geq 0$ (рис.16.1).

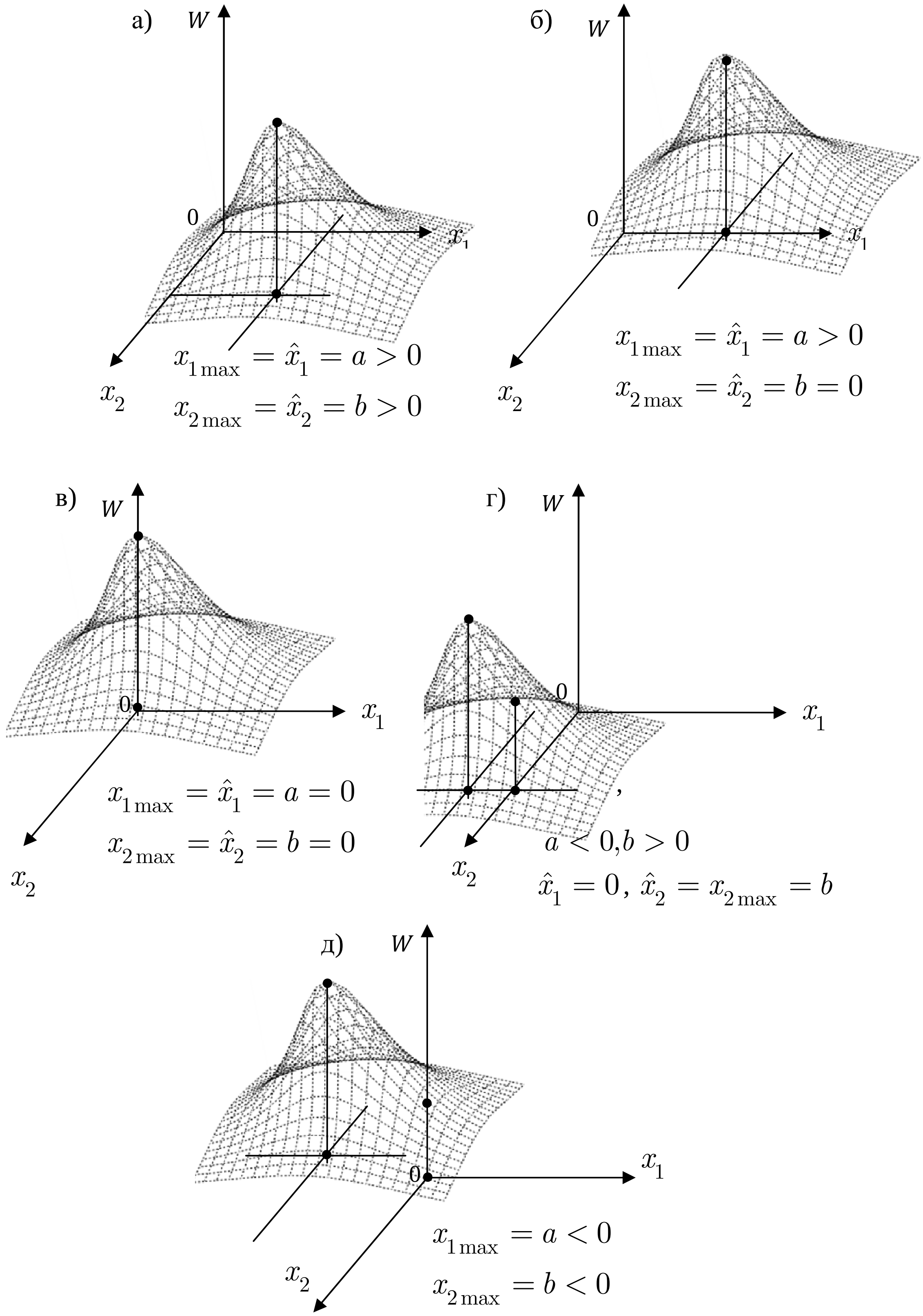


Рис. 16.1. Графічна ілюстрація стосовно пошуку необхідних умов існування найбільшого значення показника ефективності у точці $\hat{x}_{1,2}$, представленого у вигляді унімодальної двовимірної функції в умовах дії найпростіших обмежень $x_{1,2} \geq 0$: для а), б), в) виконується умова $\dfrac{\partial W}{\partial x_i} = 0, i = 1,2;$ для г) і д) відповідно $\dfrac{\partial W}{\partial x_1} < 0, \dfrac{\partial W}{\partial x_2} = 0;$ $\quad \dfrac{\partial W}{\partial x_1} < 0, \dfrac{\partial W}{\partial x_2} < 0.$

Висновок щодо необхідної умови існування максимуму:

$$\begin{cases} \left.\dfrac{\partial W}{\partial x_1}\right|_{x_1=\hat{x}_1} \le 0\ ; \\ \left.\dfrac{\partial W}{\partial x_2}\right|_{x_2=\hat{x}_2} \le 0; \\ \hat{x}_1 \ge 0; \\ \hat{x}_2 \ge 0; \\ \left.\dfrac{\partial W}{\partial x_1}\right|_{x_1=\hat{x}_1} \cdot \hat{x}_1 = 0; \\ \left.\dfrac{\partial W}{\partial x_2}\right|_{x_2=\hat{x}_2} \cdot \hat{x}_2 = 0, \end{cases} \rightarrow \begin{cases} \left.\dfrac{\partial W(X)}{\partial X}\right|_{X=\hat{X}} \le 0; \\ \hat{X} \ge 0\ ; \\ \left(\left.\dfrac{\partial W}{\partial X}\right|_{X=\hat{X}}\right)^T \cdot \hat{X} = 0, \end{cases}$$

де $X^T = \left[x_1, x_2\right]$.

Якщо замість $W\left(x_1, x_2\right)$ із одним максимумом використовується показник ефективності із декількома максимумами (як кажуть, багатомодова функція), то записана умова зберігається і використовується як необхідна умова перевірки існування локального максимуму в умовах найпростіших обмежень-нерівностей (всі координати невід'ємні).

Для n-вимірного випадку, тобто коли $X^T = \left[x_1, \ldots, x_n\right]$, можливо отримати строго аналітично ті ж самі співвідношення, що і для розглянутого двовимірного прикладу, як необхідну умову існування локального максимуму, якщо скористатись розкладанням у ряд Тейлора показника ефективності поблизу точки $\hat{X}$ і виконати вимогу

$$\begin{cases} \Delta W\left(\hat{X}\right) = W\left(\hat{X} + \Delta X\right) - W\left(\hat{X}\right) \le 0 \quad \forall \Delta X \to 0_n; \\ X \ge 0_n. \end{cases}$$

Ці необхідні вимоги існування локального максимуму в n-вимірному випадку запишемо у вигляді:

1) якщо $\hat{x}_j = 0$, то $\left.\dfrac{\partial W}{\partial x_j}\right|_{x=\hat{x}_j} \le 0;$

2) якщо $\hat{x}_j > 0$, то $\left.\dfrac{\partial W}{\partial x_j}\right|_{x=\hat{x}_j} = 0.$

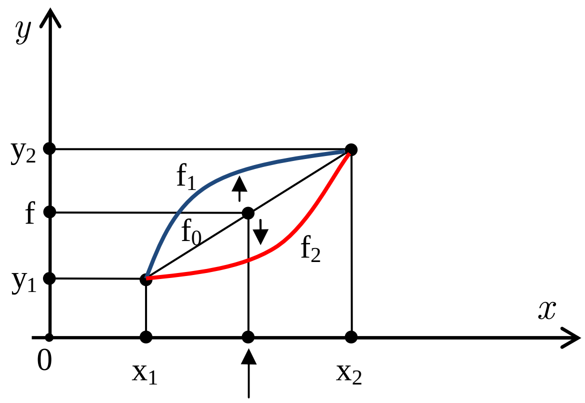
Зауваження 16.1. В умовах 1), 2) j може приймати будь-яке значення від 1 до n як окремо, так і по групі координат.

Зауваження 16.2. У випадку опуклої догори функції $W\left(X\right)$ для будь-якого $X \in \mathbb{R}^n$ отримаємо достатню умову існування глобального максимуму при $X \geq 0$ у тому ж самому вигляді, який було записано вище у пунктах 1) та 2).

Зауваження 16.3. Визначення опуклості функції (рис. 16.2):
Для векторного аргументу вважаємо функцію опуклою догори та донизу відповідно

$$f_1\left(\lambda X_1 + \left(1-\lambda\right)X_2\right) > \lambda f\left(X_1\right) + \left(1-\lambda\right)f\left(X_2\right),$$
$$f_2\left(\lambda X_1 + \left(1-\lambda\right)X_2\right) < \lambda f\left(X_1\right) + \left(1-\lambda\right)f\left(X_2\right),$$

де λ – скаляр.



$$\lambda x_1 + \left(1-\lambda\right)x_2$$

$f_0\left(\lambda x_1 + \left(1-\lambda\right)x_2\right) = \lambda f\left(x_1\right) + \left(1-\lambda\right)f\left(x_2\right)$ – пряма

$f_1\left(\lambda x_1 + \left(1-\lambda\right)x_2\right) > \lambda f\left(x_1\right) + \left(1-\lambda\right)f\left(x_2\right)$ – опукла догори

$f_2\left(\lambda x_1 + \left(1-\lambda\right)x_2\right) < \lambda f\left(x_1\right) + \left(1-\lambda\right)f\left(x_2\right)$ – опукла донизу

Рис. 16.2. Графічна ілюстрація стосовно визначення характеру опуклості двовимірної функції

Приклад 16.1. Розв’язати графічно:

$$1)\ \begin{cases} W\left(x_1\ \ x_2\right) = \left(x_1-5\right)^2 + \left(x_2-8\right)^2 \underset{x_{1,2}\geq 0}{\to \max\left(\min\right)}; \\ x_1 + 2x_2 \geq 4; \\ \dfrac{5}{2}x_1 + 2x_2 \leq 20, \end{cases}$$

$$2)\begin{cases} W\left(x_1\ \ x_2\right) = x_1^2 - 2x_1 + x_2^2 \to \max\limits_{x_{1,2}\geq 0}\left(\min\right); \\ x_1 \leq 10; \\ x_2 \leq 9; \\ x_1 + \dfrac{3}{2}x_2 \leq \dfrac{15}{2}; \\ x_1 \cdot x_2 \leq 6, \end{cases}$$

$$3)\begin{cases} W\left(x_1\ \ x_2\right) = x_2 - x_1^2 + 4x_1 \to \max\limits_{x_{1,2}\geq 0}; \\ x_1 \cdot x_2 \geq 6\ ; \\ x_1^2 + x_2^2 \leq 49; \\ x_2 \leq 6. \end{cases}$$

❍Методичні вказівки стосовно розв’язання задач:
1)

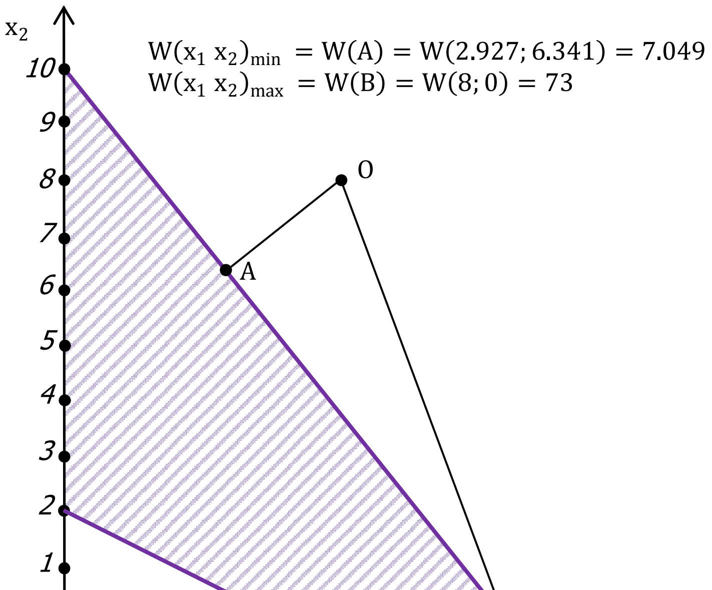


Рис. 16.3. Графічна ілюстрація щодо розв’язання задачі 1)

2)

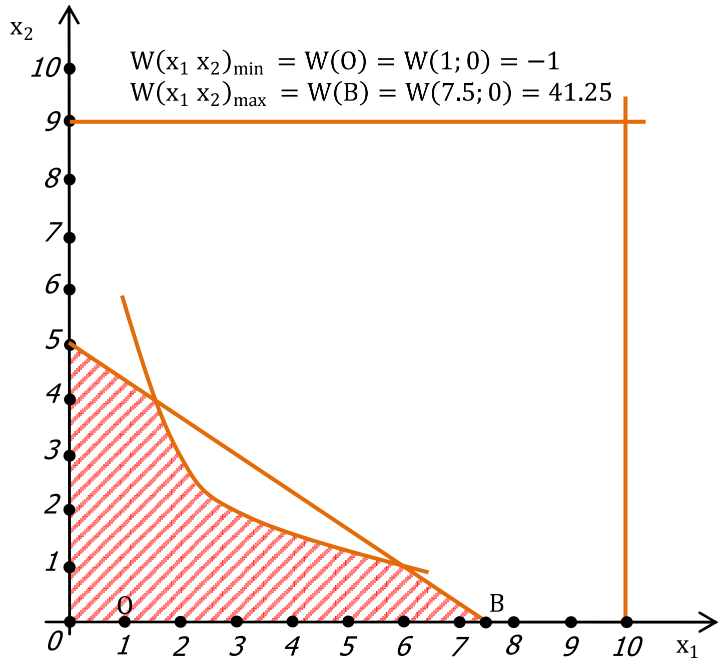


Рис.16.4. Графічне розв'язання задачі 2)

3)

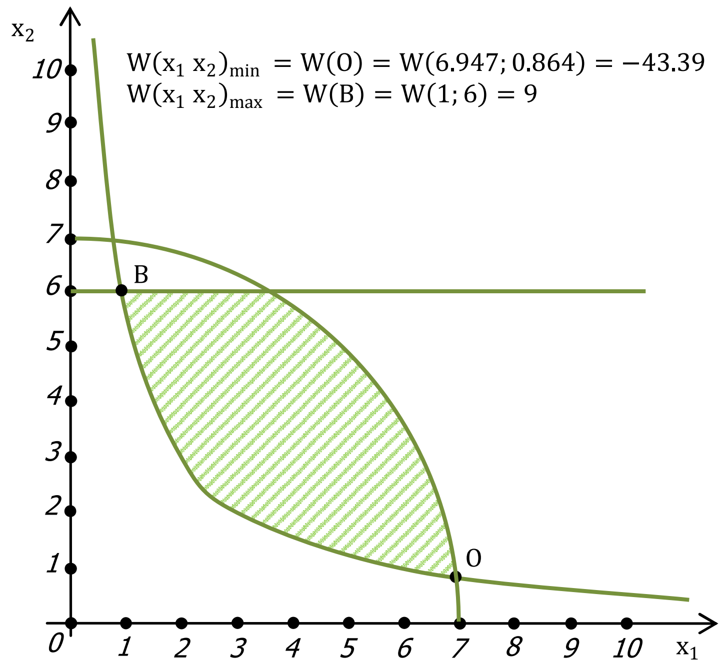


Рис. 16.5. Графічне розв'язання задачі 3)

Приклад 16.2. Телекомунікаційна компанія надає користувачам послуги супутникового зв'язку $(i = 1)$ та безпроводового доступу до інтернету $(i = 2)$.

Знижка витрат на одиницю послуги при зростанні кількості користувачів обчислюються за формулою:

$\Delta V_i = l_i \cdot x_i$,

де x_i – об'єм послуги i-го виду (кількість користувачів даної послуги);

V_i – витрати фірми на реалізацію i-го виду послуги при мінімальній сумарній кількості користувачів обох послуг;

l_i – коефіцієнти знижок питомих витрат при зростанні кількості користувачів.

Необхідно з'ясувати об'єм послуг кожного виду, при якому забезпечується мінімізація сумарних витрат при виконанні наступних обмежень:

– мінімальна сумарна кількість користувачів: 40;

– техніко-технологічній ресурс двох видів послуг складає відповідно: 160 одиниць; 210 одиниць;

– норми витрат першого ресурсу на обслуговування одного користувача супутникового зв`язку дорівнюють 2, а користувача безпроводового доступу до інтернету - $2,67$;

– норми витрат другого ресурсу складають відповідно 3 і 2.

Необхідно побудувати математичну модель задачі, виконати математичну постановку задачі і розв'язати її графічним методом, якщо:

$V_1 = 100$ у.о., $V_2 = 140$ у.о., $l_1 = l_2 = 1$ у.о. на одного користувача.

❍ 1. Математична модель:

$$\begin{cases} W\left(x_1,x_2\right) = \left(100 - x_1\right)x_1 + \left(140 - x_2\right)x_2; \\ x_1 + x_2 \geq 40; \\ 2x_1 + 2.67x_2 \leq 160; \\ 3x_1 + 2x_2 \leq 210; \\ x_{1,2} \geq 0. \end{cases}$$

2. Математична постановка:

$$W(x_1,x_2) \to \min_{x_{1,2} \in G},$$

$$\left(x_1 - 50\right)^2 + \left(x_2 - 70\right)^2 = 50^2 + 70^2 - W$$

де ОДР G задається обмеженнями-нерівностями:

$$\begin{cases} x_2 \geq -x_1 + 40; \\ x_2 \leq \dfrac{160 - 2x_1}{2,67}; \\ x_2 \leq -\dfrac{3}{2}x_1 + 105; \\ x_{1,2} \geq 0. \end{cases}$$

3. Графічний розв'язок (рис. 16.6).

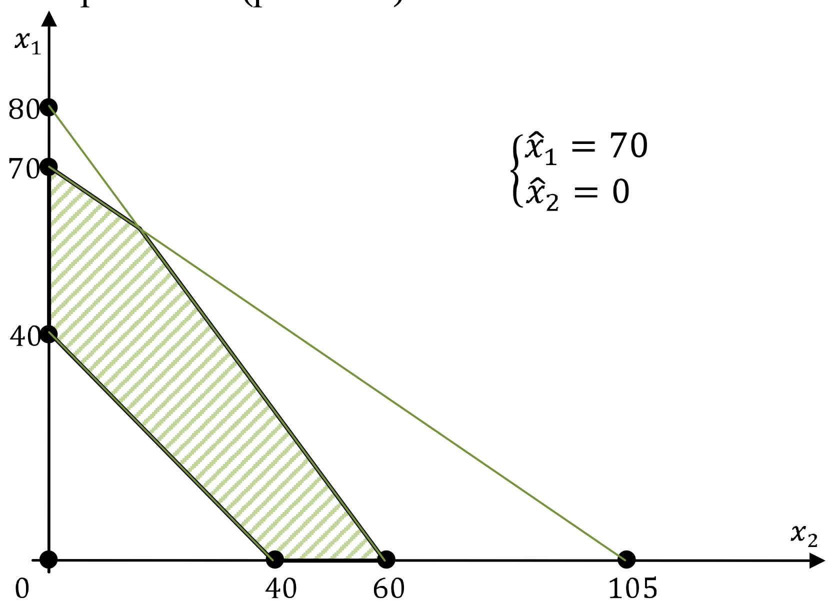


Рис. 16.6. Графічна ілюстрація побудови області допустимих значень

Як бачимо

$$\left(x_1 - 50\right)^2 + \left(x_2 - 70\right)^2 = 50^2 + 70^2 - W = R^2,$$

тобто графічним образом виразу, який пов'язаний із показником ефективності W, є коло радіусу $R^2 = 50^2 + 70^2 - W$.

Чим меньше W, тим більше R. Найбільше значення R досягається у точці

$$\begin{cases} \hat{x}_1 = 70, \\ \hat{x}_2 = 0. \end{cases}$$

Отже результат розв'язання сформульованої вище задачі нелінійного програмування в умовах дії обмежен-нерівностей, дозволяє зробити висновок про те, що надання послуг супутникового зв'язку економічно вигідніше, ніж безпроводового доступу до інтернету. Мінімальні витрати на обслуговування $W_{\min} = 2100$ у.о.●

Лекція 17. Задачі опуклого та квадратичного програмування

17.1. Умови та теорема Куна-Таккера

Задача нелінійного програмування у некласичній постановці, тобто у постанові, коли в якості умов-обмежень у задачі нелінійного програмування використовуються обмеження-нерівності, формулюється наступним чином:

$$\begin{cases} W(X) \to \max\limits_{x_i \ge 0, (i=\overline{1,n})}; \\ g_j\left(X\right) \le b_j \ (j = \overline{1,m}), \end{cases} \tag{17.1}$$

де $X^T = [x_1, \ldots, x_n]$.

В якості показника ефективності розглядають опуклу догори функцію для всіх $x_i \ge 0 \left(i = \overline{1,n}\right)$. Побудуємо функцію Лагранжа:

$$L\left(X,\Lambda\right) = W\left(X\right) + \Lambda^T \cdot \left(B - G\left(X\right)\right) \to \max\limits_{\substack{x_i \ge 0, \left(i=\overline{1,n}\right) \\ \lambda_j \ge 0, \left(j=\overline{1,m}\right)}}, \tag{17.2}$$

де $G^T\left(X\right) = \left[g_1\left(X\right), \ldots, g_m\left(X\right)\right]$, $B^T = \left[b_1, \ldots, b_m\right]$, $\Lambda^T = [\lambda_1, \ldots, \lambda_m]$.

Нам вдалося привести некласичну задачу нелінійного програмування до найпростішої некласичної задачі нелінійного програмування, тобто задачі нелінійного програмування, в якій на змінні накладена лише одна умова - умова їх невід`ємності.

Скористаємося результатом параграфу 16.1:

$$\begin{cases} \left.\dfrac{\partial L\left(X,\Lambda\right)}{\partial X}\right|_{X=\hat{X}} = \left.\left(\dfrac{\partial W\left(X\right)}{\partial X} - \dfrac{\partial G^T\left(X\right)}{\partial X} \cdot \Lambda\right)\right|_{X=\hat{X}} \le 0; \\ \hat{X} \ge 0; \\ \left.\left(\dfrac{\partial W\left(X\right)}{\partial X} - \dfrac{\partial G^T\left(X\right)}{\partial X} \cdot \Lambda\right)^T\right|_{X=\hat{X}} \cdot \hat{X} = 0; \\ \left.\dfrac{\partial L\left(X,\Lambda\right)}{\partial \Lambda}\right|_{\Lambda=\hat{\Lambda}} = B - G\left(\hat{X}\right) \ge 0; \\ \hat{\Lambda} \ge 0; \\ \left(B - G\left(\hat{X}\right)\right)^T \hat{\Lambda} = 0, \end{cases}$$

де $\dfrac{\partial G^T(X)}{\partial X} = \begin{bmatrix} \dfrac{\partial g_1}{\partial x_1} & \cdots & \dfrac{\partial g_n}{\partial x_1} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial g_1}{\partial x_n} & \cdots & \dfrac{\partial g_n}{\partial x_n} \end{bmatrix}.$

Записані умови носять назву умов Куна-Таккера і є необхідними і достатніми умовами існування глобального максимуму задачі нелінійного програмування із опуклим догори показником ефективності при умові виконання обмежень, які утворюють опуклу множину допустимих розв'язків.

Задачі такого типу складають спеціальний розділ задач нелінійного програмування, який називається *опуклим програмуванням*.

Важлива властивість функції Лагранжа полягає в тому, що:

$$L(X,\hat{\Lambda}) \le L(\hat{X},\hat{\Lambda}) \le L(\hat{X},\Lambda). \tag{17.3}$$

Це означає, що точка $(\hat{X},\hat{\Lambda})$ є *сідловою* точкою функції Лагранжа (рис.17.1).

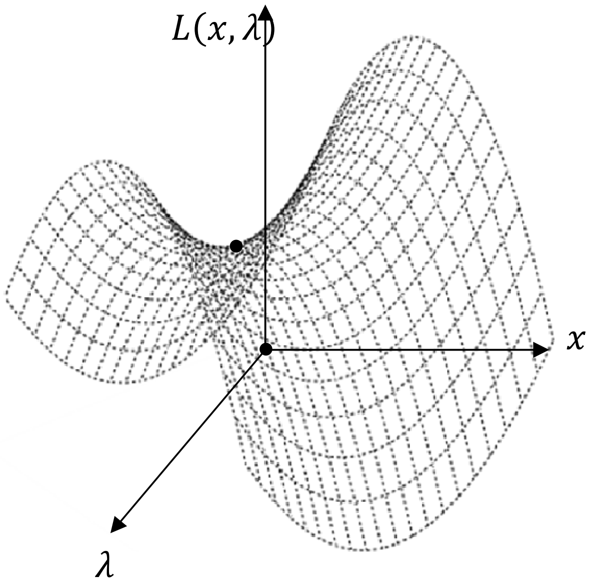


Рис. 17.1. Графічна ілюстрація стосовно існування сідлової точки для двовимірної функції Лагранжа

Тому умову Куна-Таккера можливо переписати у вигляді двоїстих задач:

$$\begin{cases} L\left(\hat{X},\Lambda\right) = \max\limits_{X\ge 0} L(X,\Lambda) \to \min\limits_{\Lambda\ge 0}; \\ \dfrac{\partial L}{\partial \Lambda} \ge 0, \end{cases} \tag{17.4}$$

та

$$\begin{cases} L\left(X,\hat{\Lambda}\right) = \min\limits_{\Lambda\ge 0} L(X,\Lambda) \to \max\limits_{X\ge 0}; \\ \dfrac{\partial L}{\partial X} \le 0. \end{cases} \tag{17.5}$$

Вперше властивість оптимальної точки $(\hat{X},\hat{\Lambda})$ як сідлової точки функції Лагранжа в задачах опуклого програмування, було доведено Куном і Таккером і сформульовано у формі теореми, яка в подальшому отримала назву теореми Куна-Таккера.

Теорема 17.1. Для того, щоб в задачі опуклого програмування деякий план $\hat{X}$ був точкою глобального максимуму необхідно і достатньо існування вектора множників Лагранжа таких, що точка $(\hat{X},\hat{\Lambda})$ буде сідловою точкою функції Лагранжа:

$$L\left(X,\Lambda\right)=W\left(X\right)+\Lambda^{T}(B-G(X)).$$

Зауваження 17.1. Умови Куна-Таккера ще записують у більш детальному вигляді:

$$\hat{x}_j>0 \;\to\; \left.\frac{\partial W}{\partial x_j}\right|_{X=\hat{X}}-\Lambda^T\cdot\left.\frac{\partial G}{\partial x_j}\right|_{X=\hat{X}}=0,$$

$$\left.\frac{\partial W}{\partial x_j}\right|_{X=\hat{X}}-\Lambda^T\cdot\left.\frac{\partial G}{\partial x_j}\right|_{X=\hat{X}}<0 \to \hat{x}_j=0,$$

$$\left.\frac{\partial W}{\partial x_j}\right|_{X=\hat{X}}-\Lambda^T\cdot\left.\frac{\partial G}{\partial x_j}\right|_{X=\hat{X}}=0 \to \;\hat{x}_j=0\,\text{або}\,\hat{x}_j>0,$$

$$\hat{x}_j=0 \to \left.\frac{\partial W}{\partial x_j}\right|_{X=\hat{X}}-\Lambda^T\cdot\left.\frac{\partial G}{\partial x_j}\right|_{X=\hat{X}}=0$$

або

$$\left.\frac{\partial W}{\partial x_j}\right|_{X=\hat{X}}-\Lambda^T\cdot\left.\frac{\partial G}{\partial x_j}\right|_{X=\hat{X}}<0$$

і далі, відносно змінної λ_i:

$$\lambda_i>0 \;\to\; g_i\left(\hat{X}\right)=b_i,$$

$$g_i\left(\hat{X}\right)<b_i \;\to\; \hat{\lambda}_i=0,$$

$$\hat{\lambda}_i=0 \to g_i\left(\hat{X}\right)=b_i$$

або

$$g_i\left(\hat{X}\right)<b_i,$$

$$g_i\left(\hat{X}\right) = 0 \rightarrow \lambda_i = 0$$
або
$$\lambda_i > 0 \left(i = \overline{1,m}\right).$$

Умови Куна-Таккера, записані у детальному вигляді носять назву умов доповнюючої нежорсткості.

17.2. Квадратичне програмування

Спеціальним класом задач нелінійного програмування є так звані задачі квадратичного програмування, в яких показник ефективності складається із лінійної та квадратичної функцій, і при цьому усі обмеження-нерівності – лінійні.

$$\begin{cases} W\left(X\right) = G \cdot X + \dfrac{1}{2} X^T \cdot D \cdot X \rightarrow \max\limits_{X \geq 0}; \\ AX \leq B, \end{cases} \tag{17.6}$$

де D – дійсна симетрична від'ємно-визначена матриця (тобто D складається із таких елементів, при яких $X^T \cdot D \cdot X < 0 \ \forall X \in \mathbb{R}^n \setminus 0_n$).

Лінійні обмеження-нерівності утворюють опуклу область допустимих розв'язків. Показник ефективності складається із суми квадратичної форми

$$X^T \cdot D \cdot X = \sum_{k=1}^{n}\sum_{j=1}^{n} \alpha_{k_j} x_n x_j < 0 \ \forall X \in R^n \setminus \{0_n\}$$

та лінійної форми:

$$C \cdot X = \sum_{j=1}^{n} c_j x_j$$

і є опуклою до гори функцією.

Це означає, що для розв'язання задачі квадратичного програмування можна застосувати методи та теореми опуклого програмування, зокрема теорему та умови Куна-Таккера. Функція Лагранжа для задачі квадратичного програмування має вигляд:

$$L\left(X,\Lambda\right) = GX + \frac{1}{2} X^T D X + \Lambda^T \left(B - AX\right), \tag{17.7}$$

що дає можливість отримати умови Куна-Такера:

$$\begin{cases} \left(\dfrac{\partial L\left(X,\Lambda\right)}{\partial X}\right)^T \cdot X\Bigg|_{X=\hat{X},\Lambda=\hat{\Lambda}} = \left(C + D\hat{X} - \hat{\Lambda}^T A\right)\hat{X}^T = 0; \\ \hat{X} \geq 0; \\ \dfrac{\partial L\left(X,\Lambda\right)}{\partial \Lambda}\Bigg|_{X=\hat{X},\Lambda=\hat{\Lambda}} = B - A\hat{X} \geq 0; \\ \dfrac{\partial L\left(X,\Lambda\right)}{\partial X}\Bigg|_{X=\hat{X},\Lambda=\hat{\Lambda}} = C + D\hat{X} - \hat{\Lambda}^T A \leq 0; \\ \Lambda^T \dfrac{\partial L\left(X,\Lambda\right)}{\partial \Lambda}\Bigg|_{X=\hat{X},\Lambda=\hat{\Lambda}} = \hat{\Lambda}^T\left(B - A\hat{X}\right) = 0; \\ \hat{\Lambda} \geq 0. \end{cases}$$

Розглянемо модифікацію щодо формулювання умов Куна-Таккера. Введемо вектори

$$V^T = \left(v_1, \ldots, v_n\right) \geq 0,$$

$$\Omega^T = (\omega_1, \ldots, \omega_m) \geq 0,$$

компоненти яких обираються, виходячи із умов:

1) якщо $\dfrac{\partial L}{\partial x_j} < 0$, то $v_j > 0$,

2) якщо $\dfrac{\partial L}{\partial x_j} = 0$, то $v_j = 0$;

3) якщо $\dfrac{\partial L}{\partial \lambda_i} < 0$, то $\omega_i > 0$,

4) якщо $\dfrac{\partial L}{\partial \lambda_i} = 0$, то $\omega_i = 0$ $\left(j = \overline{1,n}; i = \overline{1,m}\right)$.

Тоді умови Куна-Таккера можливо записати у вигляді:

$$\begin{cases} V^T \cdot \hat{X} = 0, \\ B - A\hat{X} - \Omega = 0, \\ C + D\hat{X} - \hat{\Lambda}^T A + V^T = 0, \\ \Lambda^T \cdot \Omega = 0, \\ \hat{X} \ge 0_n, \\ \hat{\Lambda} \ge 0_m, \\ V \ge 0_n, \\ \Omega \ge 0_m. \end{cases}$$

Теорема 17.2. *Необхідна та достатня умови оптимального розв'язку задачі квадратичного програмування.*

Вектор $\hat{X}$ є розв'язком задачі квадратичного програмування тоді і лише тоді, коли існують такі вектори:

- m-вимірний вектор $\hat{\Lambda} \ge 0$ та $\Omega \ge 0$,
- n-вимірний вектор $V \ge 0$,

при яких виконуються вищезазначені лінійні рівняння:

$$\begin{cases} B - A\hat{X} - \Omega = 0; \\ C + D\hat{X} - \hat{\Lambda}^T A + V^T = 0 \end{cases}$$

та нелінійні рівняння:

$$\begin{cases} V^T \cdot \hat{X} = 0; \\ \Lambda^T \cdot \Omega = 0, \end{cases}$$

що носять назву *умови доповнюючої нежорсткості*.

Для розв'язання записаної вище системи лінійних роівнянь можливо застосувати методи розв'язання задач лінійного програмування, наприклад, симплекс-метод .

Серед методів розв'язання задач квадратичного програмування найбільш відомим є метод Франка-Вульфа.

В якості прикладу аналітичного пошуку розв'язків задачі квадратичного програмування розглянемо задачу квадратичного програмування як одноетапну процедуру прийняття рішення в наступній постановці:

$$\begin{cases} W\left(X,U\right) = \frac{1}{2}U^T RU + \frac{1}{2}X^T QX = \frac{1}{2}\left\|U\right\|_R^2 + \frac{1}{2}\left\|X\right\|_Q^2 \to \min\limits_{U \in R^m}; \\ AX + BU + C = 0, \end{cases} \tag{17.8}$$

де A – невироджена матриця $n \times n$, B – матриця розміру $n \times m$; $X \in E^n$ – n-вимірний евклідів простір, $C \in E^n$, $0_n^T = [0_1, \dots, 0_n]$, R, Q – симетричні додатньо-визначені матриці розміру відповідно $m \times m$, $n \times n$.

Функція Лагранжа може бути записана у наступному вигляді:

$$L\left(X,U,\Lambda\right)=\frac{1}{2}U^TRU+\frac{1}{2}X^TQX+\Lambda^T(AX+BU+C). \qquad (17.9)$$

Для мінімізації показника ефективності необхідно, щоб:

$$\begin{cases}\left.\dfrac{\partial L}{\partial X}\right|_{X=\hat{X},U=\hat{U},\Lambda=\hat{\Lambda}}=Q\hat{X}+A^T\hat{\Lambda}=0;\\ \left.\dfrac{\partial L}{\partial U}\right|_{X=\hat{X},U=\hat{U},\Lambda=\hat{\Lambda}}=R\hat{U}+B^T\hat{\Lambda}=0;\\ A\hat{X}+B\hat{U}+C=0.\end{cases}$$

Із перших двох рівнянь знаходимо:

$$\hat{X}=-Q^{-1}A^T\Lambda,$$
$$\hat{U}=-R^{-1}B^T\Lambda$$

та підставляємо у третє рівняння (залежатиме лише від Λ):

$$-AQ^{-1}A^T\hat{\Lambda}-BR^{-1}B^T\hat{\Lambda}+C=0\Leftrightarrow\hat{\Lambda}=\left(AQ^{-1}A^T+BR^{-1}B^T\right)^{-1}C$$

Підставимо $\hat{\Lambda}$ у вирази для $\hat{X}$ і $\hat{U}$:

$$\hat{X}=-Q^{-1}A^T\left(AQ^{-1}A^T+BR^{-1}B^T\right)^{-1}\cdot C,$$

$$\hat{U}=-R^{-1}B^T\left(AQ^{-1}A^T+BR^{-1}B^T\right)^{-1}\cdot C.$$

Достатня умова того, що в точці $\left(\hat{X},\hat{U}\right)$ показник ефективності досягає мінімуму, полягає в тому, що друга варіація показника ефективності поблизу цієї точки більша 0:

$$\delta^2W\left(X,U\right)\Big|_{\substack{X=\hat{X}\\ U=\hat{U}}}>0,$$

де

$$\delta^2W\left(X,U\right)\Big|_{\substack{X=\hat{X}\\ U=\hat{U}}}=\frac{1}{2!}\left[\Delta X^T,\Delta U^T\right]\cdot\begin{bmatrix}Q & 0\\ 0 & R\end{bmatrix}\cdot\begin{bmatrix}\Delta X\\ \Delta U\end{bmatrix}\quad \text{при} \quad \begin{matrix}\Delta X\to 0\\ \Delta U\to 0\end{matrix}.$$

Враховуючи той факт, що:

$$\begin{cases}A\cdot\Delta X+B\cdot\Delta U=0.\\ \Delta X\to 0\\ \Delta U\to 0\end{cases}\quad\to\quad \Delta\mathrm{X}=-\mathrm{A}^{-1}\mathrm{B}\Delta\mathrm{U},$$

отримаємо:

$$\delta^2 W\left(X,U\right)\Big|_{\substack{X=\hat{X}\\U=\hat{U}}} = \frac{1}{2!}\Delta U^T\cdot R\cdot\Delta U + \frac{1}{2}\left(-A^{-1}\cdot B\cdot\Delta U\right)^T\cdot Q\left(A^{-1}\cdot B\cdot\Delta U\right) =$$

$$= \frac{1}{2}\Delta U^T\left(R + B^T\cdot A^{-T}\cdot Q\cdot A^{-1}\cdot B\right)\Delta U > 0.$$

Тобто, матриця в дужках повинна задовольняти ознаці додатньої визначеності.

Лекція 18. Огляд основних підходів до побудови чисельних методів розв'язання задач нелінійного програмування

18.1. Модельно-тестовий приклад

В якості модельно-тестового прикладу, який будемо використовувати для демонстрації змісту прийомів та способів, що складають зміст більшості чисельних методів розв'язання задач нелінійного програмування, розглянемо показник ефективності вигляду:

$$W\left(x_1, x_2\right)\Big|_{\substack{x_1=a\\x_2=b}} = a^2 + ab + b^2 \to \min_{\substack{a\in[-3;3]\\b\in[-3;3]}} \tag{18.1}$$

Використаємо необхідні і достатні умови пошуку екстремуму функції двох змінних і знайдемо при яких значеннях параметрів a та b показник ефективності досягає мінімуму:

1) необхідна умова:

$$\begin{cases}\dfrac{\partial W\left(a,b\right)}{\partial a} = 0,\\ \dfrac{\partial W\left(a,b\right)}{\partial b} = 0,\end{cases} \Rightarrow \begin{cases}2a + b = 0,\\ 2b + a = 0,\end{cases} \Rightarrow \begin{cases}\hat{a} = 0,\\ \hat{b} = 0.\end{cases} \tag{18.2}$$

2) достатня умова існування мінімуму в точці $\left(0;0\right)$: матриця Гессе повинна бути додатно-визначеною:

$$\begin{bmatrix}\dfrac{\partial^2 W}{\partial a^2} & \dfrac{\partial^2 W}{\partial a\partial b}\\ \dfrac{\partial^2 W}{\partial a\partial b} & \dfrac{\partial^2 W}{\partial b^2}\end{bmatrix} = \begin{bmatrix}2 & 1\\ 1 & 2\end{bmatrix} = \begin{bmatrix}q_{11} & q_{12}\\ q_{21} & q_{22}\end{bmatrix} > 0, \tag{18.3}$$

$\Delta_1 = q_{11} = 2 > 0$,

$\Delta_2 = 2\cdot 2 - 1\cdot 1 = 3 > 0$, де $\Delta_{1,2}$ визначники, що розташовані на головній діагоналі матриці Гессе.

Висновок:

1) Підтверджено, що матриця Гессе додатно-визначена.

2) Показник ефективності досягає мінімуму і його значення дорівнює 0:

$$\begin{cases}\hat{a} = 0,\\ \hat{b} = 0,\end{cases} \quad \hat{W}\left(\hat{a},\hat{b}\right) = 0.$$

18.2. Метод сканування (метод повного перебору)

Зазвичай, цей метод використовують для вибору початкових умов для роботи чисельного алгоритму або в задачах цілочисельного програмування для уточнення цілочисельних значень змінних, за якими розв'язується задача, розташованих поблизу екстремуму показника ефективності, який було знайдено в умовах послабленої цілочисельної задачі.

Для реалізації методу сканування необхідно виконати наступні дії:

1. Задати крок дискретизації змінних, за якими виконується пошук екстремуму.
2. Побудувати область дискретизації змінних, за якими виконується пошук екстремуму.

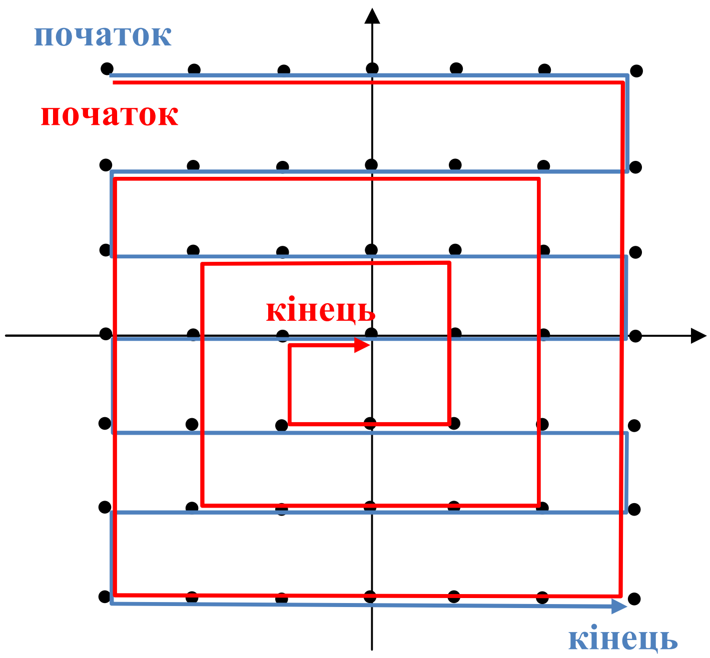


Рис. 18.1. Приклад побудови області дискретизації значень змінних, за якими виконується оптимізація показника ефективності

$$W = a^2 + ab + b^2 \to \min \forall a,b \in \left[-3;3\right]$$

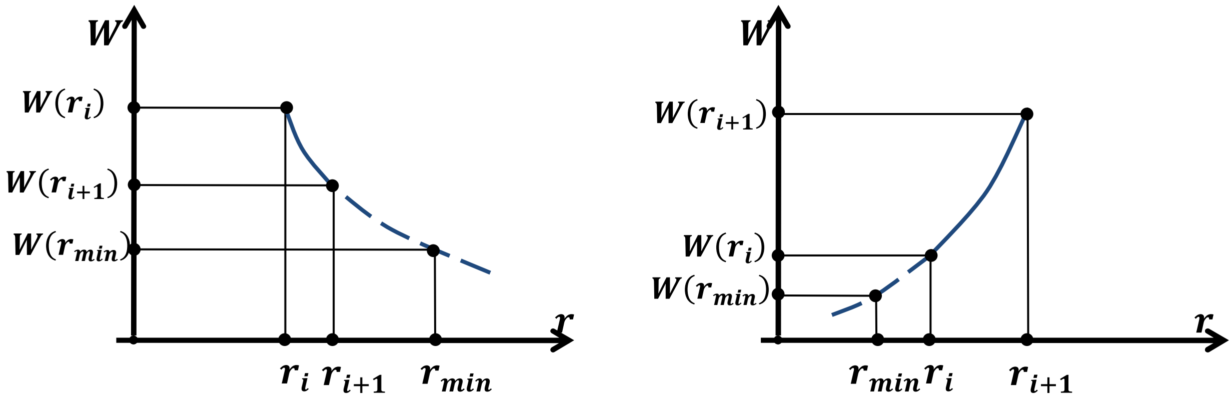
із вибором кроку дискретизації за параметрами $\varepsilon_a = 1,\ \varepsilon_b = 1$

З метою не пропустити жодної точки використовуються різні способи їх обходу (способи сканування вузлів). Наприклад, на рис.18.1 показано рядковий(синій колір) та спіральний(червоний колір) способи сканування.

18.3. Метод Гаусса-Зейделя (метод покоординатного спуску)

Якщо показник ефективності W залежить від m змінних, то в цьому методі на кожному етапі(циклі) пошуку оптимальних значень параметрів припускають, що лише одна змінна впливає на значення показника ефективності, а інші залишаються незмінними (сталими). Після того, як буде знайдено екстремальне значення показника ефективності за виділеною змінною, цю змінну вважають сталою із значенням, яке було знайдено при пошуку екстремуму показника ефективності за цією змінною. Покоординатний пошук екстремуму припиняється за відповідною умовою, яка може полягати в тому, що при повному виконанні етапів(циклів) покоординатної зміни за всіма змінними значення показника майже не змінюється.

Умови наближення до мінімуму показника ефективності при чисельній реалізації методу Гаусса-Зейделя представлені на рис.18.2.



а) якщо $r_{i+1} > r_i$ і при цьому $W(r_{i+1}) < W(r_i)$, то точка мінімуму розташована праворуч від r_{i+1}

б) якщо $r_{i+1} > r_i$ і при цьому $W(r_{i+1}) > W(r_i)$, то точка мінімуму розташована ліворуч від r_i

Рис. 18.2. Графічна ілюстрація стосовно основної ідеї усіх методів скалярної оптимізації (тобто пошуку екстремуму функції однієї змінної):
$\boldsymbol{r}$ – узагальнене позначення того параметру, від якого залежить показник ефективності і який на даному етапі(циклі) методу по координатного спуску вважається змінним

Існує велика кількість методів пошуку екстремуму функції однієї змінної, наприклад:

1) аналітичний (класичний) метод;
2) групи чисельних методів:
   - метод ділення відрізків навпіл;
   - метод золотого перерізу;
   - симетричні методи;
   - оптимальні методи;
   - метод ламаних;
   - метод покриттів;
   - метод дотичних;
   - метод парабол тощо.

18.3.1. Класичний метод мінімізації функції однієї змінної при його застосуванні в методі Гаусса-Зейделя

Показник ефективності має вигляд:

$$W = a^2 + ab + b^2 \rightarrow \min,$$

тобто вважається відомою залежність показника ефективності від змінних, за якими відбувається оптимізація.

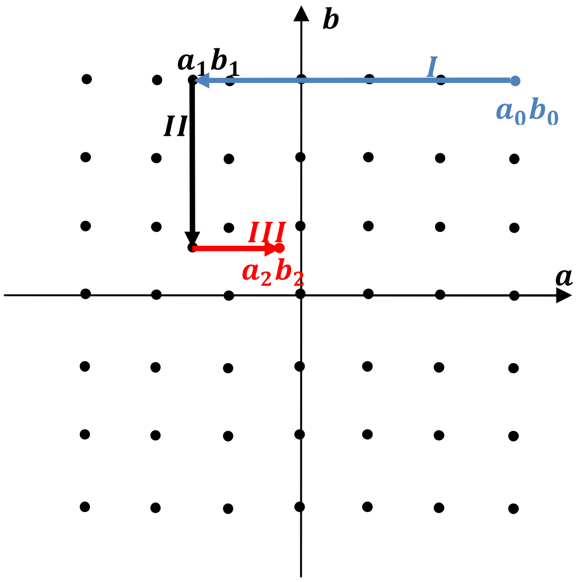


Рис. 18.3. Графічна ілюстрація застосування класичного методу пошуку екстремуму функції однієї змінної в методі Гаусса-Зейделя

I етап (цикл) (рис. 18.3)

$$\begin{cases} a_0 = 3, \\ b_0 = 3, \end{cases}$$

$$\begin{cases} b_0 = 3 = \text{const}, \\ a = \text{var}, \end{cases}$$

$$W\left(a, 3\right) = a^2 + 3a + 9,$$

$$\frac{dW}{da} = 2a + 3 = 0 \rightarrow a_1 = -\frac{3}{2}, b_1 = 3.$$

II етап (цикл)

$$\begin{cases} b_1 = \text{var}, \\ a_1 = -\dfrac{3}{2}, \end{cases}$$

$$W\left(-\frac{3}{2}, b\right) = \frac{9}{4} - \frac{3}{2}b + b^2,$$

$$\frac{dW}{db} = -\frac{3}{2} + 2b = 0 \rightarrow b_2 = \frac{3}{4} = 0{,}75,\ a_2 = -\frac{3}{2}.$$

III етап (цикл)

$$\begin{cases} b_2 = 0{,}75, \\ a_2 = \text{var}, \end{cases}$$

$$W\left(a,\frac{3}{4}\right)=a^2+\frac{3}{4}a+\frac{9}{16},$$

$$\frac{dW}{da}=2a+\frac{3}{4}=0\to a_3=-\frac{3}{8}$$

і т.д.

18.3.2. Метод екстраполяційного пошуку

Метод екстраполяційного пошуку базується на трьох припущеннях:

1) Припускається, що заздалегідь відоме бажане значення (область розташування бажаного значення показника ефективності):

$$W_б\in[W_0-\varepsilon_i;W_0+\varepsilon_i],$$

де ε_i– допустима помилка обчислення($\varepsilon_i>0$), W_0– певна величина, значення якої задають, базуючись на досвіді розв'язання подібних задач, i - номер етапу(циклу).

Але при цьому невідома точка екстремуму показника ефективності, тобто значення змінних, при яких досягається $W_б$.

2) Припускається, що значення показника ефективності може бути наближено представлене функціональною залежністю від змінних, за якими відбувається пошук екстремуму.
3) Значення показника ефективності при заданому векторі змінних, за якими відбувається оптимізація може бути виміряне або обчислене.

Приклад 18.1. Нехай залежність показника ефективності від змінного на даному циклі параметру є лінійною:

$$\frac{y-y_1}{y_2-y_1}=\frac{x-x_1}{x_2-x_1},$$

$$r_1=r^{(i)}-\delta_i\to W\left(r_1\right),$$

$$r_2=r^{(i)}+\delta_i\to W\left(r_2\right).$$

❍ Тоді можемо записати:

$$\left(\frac{W\left(r^{(i+1)}\right)-W\left(r_1\right)}{W\left(r_2\right)-W\left(r_1\right)}\right)=\frac{r^{(i+1)}-r_1}{r_2-r_1}\to r^{(i+1)}=r_i-\delta_i+2\delta_i\cdot\frac{W\left(r^{(i+1)}\right)-W\left(r_1\right)}{W\left(r_2\right)-W\left(r_1\right)},$$

де припускається, що $W\left(r^{(i+1)}\right)=W_б,\ \delta_i>0$.

Після обчислення $r^{(i+1)}$ виконується вимірювання або обчислення істинного значення показника ефективності $W_{іст}\left(r\right)$ в точці $r=r^{(i+1)}$.

Якщо $W_{\text{іст}}\left(r^{(i+1)}\right) \in [W_0 - \varepsilon_i; W_0 + \varepsilon_i]$, то вважаємо, що $r^{(i+1)} = r_{ext}$. Запам'ятовуємо це значення і переходимо до пошуку екстремуму показника ефективності за наступною координатою.

Якщо $W_{\text{іст}}\left(r^{(i+1)}\right) \notin [W_0 - \varepsilon_i; W_0 + \varepsilon_i]$, то повторюємо усі вищенаведені розрахунки при інших значеннях δ_i та ε_i.●

Приклад 18.2. Застосування методу екстраполяційного пошуку за умови квадратичної екстраполяції залежності показника ефективності від значень тієї змінної, за якою на даному циклі відбувається оптимізація.

Нехай $W\left(r\right) = Ar^2 + Br + C, \quad A > 0.$

❍ Якщо на i-тому етапі(циклі) виконати обчислення (вимірювання) значення показника ефективності, а саме:

$$W\left(r^{(i)} - \delta_i\right) = W_1,$$

$$W\left(r^{(i)}\right) = W_2,$$

$$W\left(r^{(i)} + \delta_i\right) = W_3,$$

то отримаємо систему трьох лінійних рівнянь для обчислення трьох невідомих параметрів A, B, C:

$$\begin{cases} A\left(r^{(i)} - \delta_i\right)^2 + B\left(r^{(i)} - \delta_i\right) + C = W_1, \\ A\left(r^{(i)}\right)^2 + B\left(r^{(i)}\right) + C = W_2, \\ A\left(r^{(i)} + \delta_i\right)^2 + B\left(r^{(i)} + \delta_i\right) = W_3. \end{cases} \qquad \otimes$$

Отже,

$$\begin{cases} A = \dfrac{\Delta A}{\Delta}, \\ B = \dfrac{\Delta B}{\Delta}, \\ C = \dfrac{\Delta C}{\Delta}, \end{cases}$$

де $\Delta = \begin{vmatrix} \left(r^{(i)} - \delta_i\right)^2 & \left(r^{(i)} - \delta_i\right) & 1 \\ \left(r^{(i)}\right)^2 & \left(r^{(i)}\right) & 1 \\ \left(r^{(i)} + \delta_i\right)^2 & \left(r^{(i)} + \delta_i\right) & 1 \end{vmatrix}$, $\quad \Delta A = \begin{vmatrix} W_1 & \left(r^{(i)} - \delta_i\right) & 1 \\ W_2 & \left(r^{(i)}\right) & 1 \\ W_3 & \left(r^{(i)} + \delta_i\right) & 1 \end{vmatrix}$,

$$\Delta B = \begin{vmatrix} \left(r^{(i)} - \delta_i\right)^2 & W_1 & 1 \\ \left(r^{(i)}\right)^2 & W_2 & 1 \\ \left(r^{(i)} + \delta_i\right)^2 & W_3 & 1 \end{vmatrix}, \quad \Delta C = \begin{vmatrix} \left(r^{(i)} - \delta_i\right)^2 & \left(r^{(i)} - \delta_i\right) & W_1 \\ \left(r^{(i)}\right)^2 & \left(r^{(i)}\right) & W_2 \\ \left(r^{(i)} + \delta_i\right)^2 & \left(r^{(i)} + \delta_i\right) & W_3 \end{vmatrix}.$$

Мінімум квадратичної залежності досягається при $r_{ext} = -B \,/\, 2A$.●

Знайдене значення змінної r_{ext}, за якою відбувається пошук екстремуму дозволяє обчислювати (вимірювати) істинне значення показника ефективності: $W_{\text{іст}}\left(r_{ext}\right)$.

Якщо $W_{\text{іст}}\left(r_{ext}\right) \in \left[W_0 - \varepsilon_i, W_0 + \varepsilon_i\right]$, то вважаємо, що оптимізація за цією координатою завершена. Значення r_{ext} запам'ятовуємо і переходимо до пошуку екстремуму показника ефективності за наступною координатою. Якщо $W_{\text{іст}}\left(r_{ext}\right)$ не належить заданій області, то повторюємо наведені вище обчислення, змінивши δ_i або ε_i, або W_0.

18.4. Метод градієнту

При використанні методу градієнту на кожному етапі змінюються усі змінні, за якими відбувається пошук екстремуму. Припустимо, що таких змінних m. Показник ефективності має вигляд $W = W(x_1, \ldots, x_m)$. Градієнтом диференційовної функції називається вектор, проекція якого на вісі координат дорівнює відповідно $\dfrac{\partial W}{\partial x_1}, \ldots, \dfrac{\partial W}{\partial x_m}$.

В методі градієнту використовується його властивість, яка полягає в тому, що модуль градієнту зменшується при наближенні до точки екстремуму (рис. 18.4):

$$\left|\frac{\partial W}{\partial X}\right|_{X_i} = \left|\operatorname{grad} W\right|_i = \sqrt{\left(\frac{\partial W}{\partial x_1}\right)_i^2 + \ldots + \left(\frac{\partial W}{\partial x_m}\right)_i^2} \to 0 \ ,$$

$$X_i^T = [x_1^{(i)}, \ldots, x_m^{(i)}] \to X_{ext}.$$

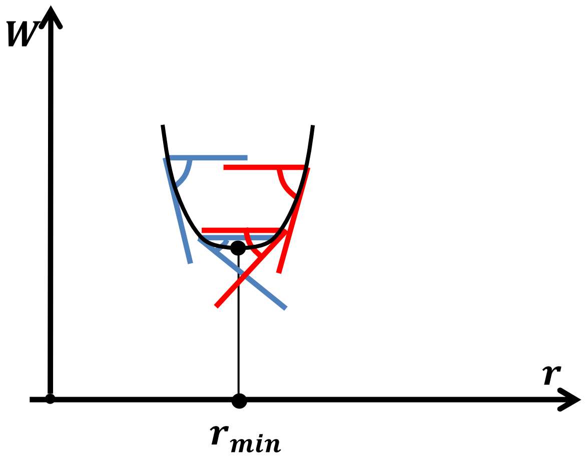


Рис.18.4. Графічна ілюстрація наближення модуля градієнту до "0", при наближенні аргументу до точки екстремуму показника ефективності

В загальному випадку, якщо невідома аналітична залежність показника ефективності від змінних, за якими відбувається оптимізація, значення градієнту в заданій точці X_i обчислюється наближено, тобто обчислюється субградієнт:

$$\begin{cases} \dfrac{\partial W}{\partial {x_1}^{(i)}} = \mathrm{grad}_1^{(i)} W \cong \dfrac{1}{2\varepsilon_1}\left[W\left(x_1^{(i)} + \varepsilon_1, x_2^{(i)} \dots x_m^{(i)}\right) - W\left(x_1^{(i)} - \varepsilon_1, x_2^{(i)} \dots x_m^{(i)}\right)\right], \\ \dots \\ \dfrac{\partial W}{\partial {x_m}^{(i)}} = \mathrm{grad}_m^{(i)} W \cong \dfrac{1}{2\varepsilon_m}\left[W\left(x_1^{(i)}, x_2^{(i)} \dots x_m^{(i)} + \varepsilon_m\right) - W\left(x_1^{(i)}, x_2^{(i)} \dots x_m^{(i)} - \varepsilon_m\right)\right], \end{cases}$$

де $\varepsilon_j > 0$ $(j = \overline{1,m})$ – пробний крок, який використовується для зміни j-ої змінної на i-ій ітерації.

Тобто в цілому:

$$\left(\overrightarrow{\mathrm{grad}}^{(i)} W\right)^T = \left(\mathrm{grad}_1^{(i)} W, \dots, \mathrm{grad}_m^{(i)} W\right).$$

Для обчислення нового (наступного) наближення змінних, за якими відбувається пошук екстремуму, використовуються вирази:

$$\begin{cases} x_1^{(i+1)} = x_1^{(i)} - \delta_1 \, \mathrm{grad}_1^{(i)} W, \\ \dots \\ x_m^{(i+1)} = x_m^{(i)} - \delta_m \, \mathrm{grad}_m^{(i)} W, \end{cases}$$

якщо відбувається пошук мінімального значення показника ефективності та

$$\begin{cases} x_1^{(i+1)} = x_1^{(i)} + \delta_1 \operatorname{grad}_1^{(i)} W, \\ \ldots \\ x_m^{(i+1)} = x_m^{(i)} + \delta_m \operatorname{grad}_m^{(i)} W, \end{cases}$$

якщо відбувається пошук максимального значення показника ефективності, де $\delta_j \ (j = \overline{1,m})$ – величина робочого кроку вздовж відповідної координати.

Таким чином, в методі градієнту процес пошуку екстремуму на кожній ітерації, тобто етапі(циклі)наближення до точки екстремуму, включає в себе дві обов'язкові операції (дії), які носять назву:

- пошук градієнту в поточній точці, коротка назва – пробний крок;
- пошук нового значення вектору змінних, коротка назва – робочий крок.

Графічна ілюстрація траєкторії наближення до мінімального значення показника ефективності при використанні методу градієнту представлена на рис.18.4.

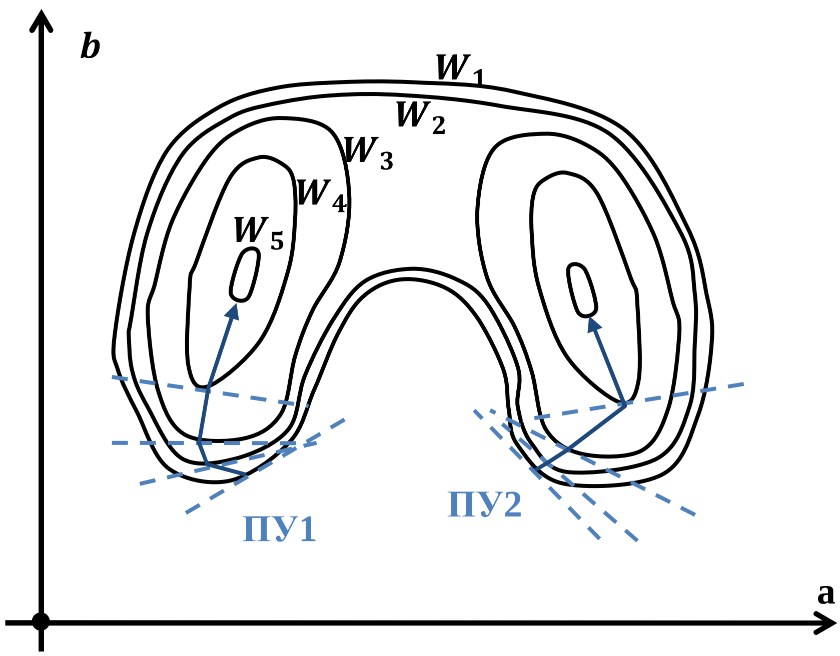


Рис.18.4. Графічна ілюстрація застосування методу градієнту
для пошуку екстремуму багатомодової функції: ПУ- початкова умова

Метод градієнту дозволяє знаходити локальний екстремум показника ефективності. Як бачимо (див. рис.18.4), результат пошуку локального екстремуму залежить від початкових умов.

Приклад 18.3. Припустимо, що показник ефективності має вигляд:
$W = a^2 + ab + b^2 \to \min$.
Розв'язати задачу методом градієнту (рис.18.5).

❍1. Перша ітерація. Початкові умови: $a_0 = b_0 = 3$.

1.1. Пробний крок:

$$\frac{\partial W}{\partial a} = 2a + b\Big|_{(a_0;b_0)} = 9,$$

$$\frac{\partial W}{\partial b} = 2b + a\Big|_{(a_0;b_0)} = 9.$$

1.2. Робочий крок:

$$\begin{cases} a_1 = a_0 - \delta \cdot \dfrac{\partial W}{\partial a} = -1.5; \\ b_1 = b_0 - \delta \cdot \dfrac{\partial W}{\partial b} = -1.5, \end{cases}$$

де $\delta = 1/2$.

2. Друга ітерація. Початкові умови: $a_1 = b_1 = -1.5$.

2.1. Пробний крок:

$$\frac{\partial W}{\partial a} = 2a + b\Big|_{(a_1;b_1)} = -4.5\ ,$$

$$\frac{\partial W}{\partial b} = 2b + a\Big|_{(a_1;b_1)} = -4.5\ .$$

2.2. Робочий крок:

$$\begin{cases} a_2 = a_1 - \delta \cdot \dfrac{\partial W}{\partial a} = 0.75; \\ b_2 = b_1 - \delta \cdot \dfrac{\partial W}{\partial b} = 0.75. \end{cases}$$

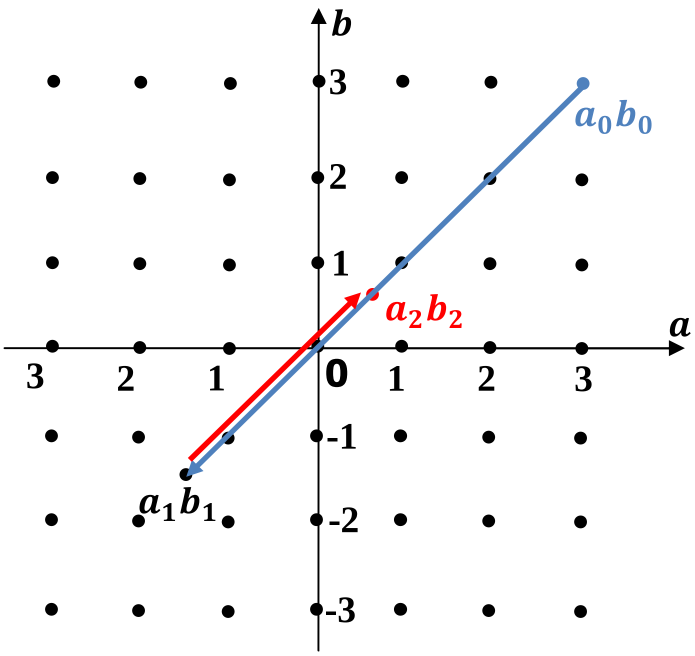


Рис. 18.51. Графічна ілюстрація використання методу градієнту для пошуку екстремуму показника ефективності

Збіжність методу градієнта залежить як від початкових умов, так і від величини робочого кроку, тобто для підвищення ефективності методу градієнту (пошук локального екстремуму відбувається за меншу кількість кроків), необхідно цей метод модифікувати, надавши йому деякі оптимальні властивості. ●

18.5. Метод найшвидшого спуску

Метод найшвидшого спуску – модифікація методу градієнту, суть якого полягає в тому, що після обчислення градієнту на i-ій ітерації, переміщення вздовж цього градієнту відбувається від $X^{(i)} \to X^{(i+1)}$ доти, доки зменшується значення показника ефективності, а не з фіксованим кроком, як це було в методі градієнту (рис. 18.6).

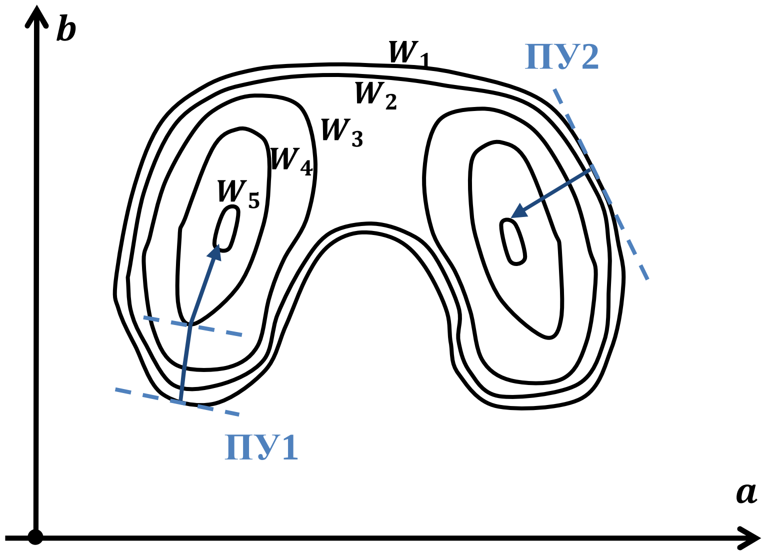


Рис.18.6. Графічна ілюстрація застосування методу найшвидшого спуску для пошуку екстремуму багатомодової функції

При використанні методу найскорішого спуску зменшується кількість ітерацій і навіть можливо досягти локального екстремуму за одну ітерацію при вдало обраних початкових умовах (приклад 18.4 та рис. 18.7).

Приклад 18.4. Припустимо, що показник ефективності має вигляд:
$W = a^2 + ab + b^2 \to \min$.
Розв‘язати задачу методом найшвидшого спуску.

❍ Початкові умови: $a_0 = b_0 = 3$.

1. Перша ітерація. Початкові умови: $a_0 = b_0 = 3$.

1.1. Пробний крок:

$$\frac{\partial W}{\partial a} = 2a + b\Big|_{(a_0;b_0)} = 9,$$

$$\frac{\partial W}{\partial b} = 2b + a\Big|_{(a_0;b_0)} = 9.$$

2) Робочий крок.

$$a_1 = a_0 - 9\delta = 3 - 9\delta;$$

$$b_1 = b_0 - 9\delta = 3 - 9\delta,$$

$$W(\delta) = 3(3 - 9\delta)^2,$$

$$\frac{\partial W}{\partial \delta} = -54(3 - 9\delta) = 0 \rightarrow \delta = \frac{1}{3}.$$

Остаточно маємо:

$$a_1 = a_0 - 9\delta = 3 - 9\delta = 3 - 9 \cdot \frac{1}{3} = 0;$$

$$b_1 = b_0 - 9\delta = 3 - 9\delta = 3 - 9 \cdot \frac{1}{3} = 0. \bullet$$

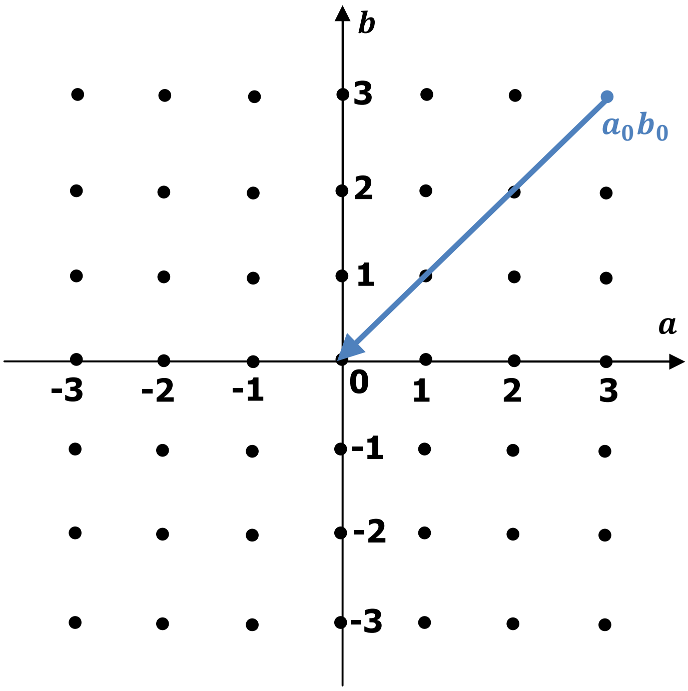


Рис.18.7. Графічна ілюстрація застосування методу найшвидшого спуску для пошуку мінімуму показника якості

18.6. Метод випадкового пошуку

В основу випадкового пошуку покладено метод випробувань і помилок, у відповідності з яким вдало знайдене значення керуючих параметрів зберігається, а невдале – відкидається.

Метод випадкового пошуку використовують для пошуку як глобальних, так і локальних екстремумів показника ефективності, причому як самостійно, так і комбіновано з методами регулярного (детермінованого) пошуку.

Класифікація методів випадкового пошуку:

1) локальний випадковий пошук з поверненням;
2) локальний випадковий пошук за найкращою пробою;
3) локальний випадковий пошук за статистичним градієнтом;

4) глобальний випадковий пошук із незалежним вибором щільності розподілу пробних кроків;

5) генетичні алгоритми та еволюційне програмування.

Розглянемо для прикладу основні елементи алгоритму локального випадкового пошуку (локального випадкового пошуку з поверненням, який на сьогодні є найбільш поширеним). Алгоритм полягає в наступному:

1-й крок.

У випадковому напрямку здійснюється зміна значень параметрів, за якими відбувається пошук екстремуму:

$$X^{(i+1)} = X^{(i)} + \Delta X^{(i+1)},$$

де $\Delta X^{(i+1)} = \delta E_{i+1}$;

$E_{i+1}^T = [e_1, \dots, e_n]_{i+1}$ – випадковий одиничний вектор $\left(e_1^2 + e_2^2 + \dots + e_n^2 = 1\right)$ із заданою щільністю ймовірності, реалізація значень компонент якого отримана для обчислення $\Delta X^{(i+1)}$;

δ – величина кроку за керуючими параметрами.

Якщо вимірювання (обчислення) показало, що показник ефективності у новій точці менше значення показника ефективності у попередній точці (відбувається пошук мінімуму), тобто $W\left(X^{(i+1)}\right) < W\left(X^{(i)}\right)$, то значення показника ефективності $W\left(X^{(i+1)}\right)$ та $X^{(i+1)}$ запам'ятовуються.

Якщо $W\left(X^{(i+1)}\right) \geq W\left(X^{(i)}\right)$, то виконується повернення до попередніх значень показника ефективності та керуючої змінної і знову обчислюється

$$X^{(i+2)} = X^{(i)} + \Delta X^{(i+2)},$$

де $\Delta X^{(i+2)} = \delta E_{i+2}$, $E_{i+2}^T = \left[e_1, \dots, e_n\right]_{i+2}$ - випадковий одиничний вектор реалізація значень компонент якого отримана для обчислення $\Delta X^{(i+2)}$.

Обчислення виконується доти, доки показник ефективності стане меншим, ніж його попереднє значення. Останнє значення вектора керуючих змінних та значення показника ефективності запам'ятовується.

2-й крок.

Виконується крок у випадковому напрямку від останнього значення вектора керуючих параметрів, яке було запам'ятовано наприкінці першого кроку.

В подальшому кроки 1 та 2 повторюються доти, доки не буде виконана умова зупинки алгоритму пошуку екстремального значення показника ефективності.

Лекція 19. Застосування системи комп'ютерної математики MATLAB для розв'язку задач квадратичного та нелінійного програмування

19.1. Квадратичне програмування: функція *quadprog*

Формат постановки задачі квадратичного програмування в MATLAB:

$$\begin{cases} W = \frac{1}{2} X^T \cdot H \cdot X + f \cdot X \to \min\limits_{X \in \mathbb{R}^n}, \\ A \cdot X \le b, \\ A_{eq} \cdot X = b_{eq}, \\ L_b \le X \le U_b. \end{cases} \tag{19.1}$$

Формат звернення до функції ***quadprog:***

[X,W,flag]= ***quadprog***(H , f , A , b , A_{eq} , b_{eq} , L_b , U_b**).**

Приклад 19.1. Розглянемо задачу рівня бізнес-управління телекомунікаційною компанією. Перед топ менеджерами компанії поставлено завдання щодо розподілу інноваційних ресурсів компанії на розвиток найприбутковіших технологій, тобто визначення так званого інноваційного портфелю. В силу невизначеності майбутнього запиту ринку цей портфель є ризикованим.

❍ Для зменшення ризику втрати наявних інноваційних ресурсів необхідно виконати їх розбиття, наприклад, на чотири частини x_1, x_2, x_3, x_4 для того, щоб розвивати. відповідно, чотири інноваційні технології: A_1, A_2, A_3, A_4 .

Позначимо

$$X^T = [x_1, x_2, x_3, x_4],$$

де $x_i \ (i = \overline{1,4})$ відносні витрати (частина від загальної інноваційної суми):

$$x_1 + x_2 + x_3 + x_4 = 1, \quad x_{1,2,3,4} \ge 0 \ .$$

Таким чином, інноваційний портфель визначається вектором X частинок від загальної суми інноваційних ресурсів. Ризик оцінюється як величина дисперсії доходу, що очікується. Вважаємо, що за результатом аналізу тенденцій розвитку ринку IT-технологій ідентифіковано вектор, що складається з математичних сподівань прибутковості чотирьох обраних для розвитку технологій

$$d^T = [d_1, d_2, d_3, d_4]$$

та матриця взаємних дисперсій V .

Тоді математичне сподівання прибутковості інноваційного портфелю можливо обчислити:

$$M_p = d^T \cdot X \ .$$

Припустимо, що:

$$d = [11.3; 13.2; 16.4; 17.4],$$

$$V = \begin{bmatrix} 102.0 & 27.1 & -52.3 & 66.5 \\ 27.1 & 148.8 & 42.1 & -66.4 \\ -52.3 & 42.1 & 246.5 & 56.9 \\ 66.5 & -66.4 & 56.9 & 272.3 \end{bmatrix}.$$

Дисперсія відхилення від прибутковості, що очікується ($M_{p0} = 15$) обчислюється за формулою:

$$D = X^T \cdot V \cdot X .$$

Вираз для обчислення дисперсії D використовують для оцінки ризику інноваційного портфеля. Цей ризик необхідно мінімізувати за рахунок оптимального вибору компонент вектора X (задача Марковица).

Для наведеного прикладу матриці, що відповідають загальній постановці задачі, набувають вигляду:

$H = V$;

f – відсутня, [ ];

A – відсутня, [ ];

B – відсутня, [ ];

$A_{eq} = \begin{bmatrix} d_1 & d_2 & d_3 & d_4 \\ 1 & 1 & 1 & 1 \end{bmatrix}$;

$b_{eq} = \begin{bmatrix} 15 \\ 1 \end{bmatrix}$;

$lb = [0, \ldots, 0]$;

Ub – відсутня, [ ].

Звернення до функції мінімізації квадратичного показника ефективності для нашого випадку:

$$\left[X, minW, flag\right] = quadprog\left(V, [], [], [], A_{eq}, b_{eq}, lb, []\right),$$

де *flag* – символ якості виконання оптимізації: якщо *flag* = цілому додатному числу, то операція оптимізації виконана успішно; якщо *flag* = 0, то виконана максимальна кількість ітерацій без покращення результату (ознака зациклювання алгоритму); якщо *flag* = цілому відємному числу, то досягти екстремуму не вдалося – операція пошуку екстремуму не віконана.

Командне віконце MATLAB:

```
>> V=[102.0   27.1   -52.3   66.5
       27.1  148.8    42.1  -66.4
      -52.3   42.1   246.5   56.9
       66.5  -66.4    56.9  272.3]
```

```
>> Aeq=[11.3   13.2   16.1   17.4
            1      1       1       1]

>> beq=[15;1]
>> lb=[0;0;0;0]

>> [X,W,flag]=quadprog(V,[],[],[],Aeq,beq,lb,[])
X =  0.0626
     0.4359
     0.1439
     0.3575
minW =  31.2725
flag = 1.
```
●

19.2. Нелінійне програмування: функція *fmincon*

Формат постановки задачі нелінійного програмування в MATLAB:

$$f(x) \to \min_{X \in \mathbb{R}^n},$$

$$\begin{cases} A \cdot X \le b, \\ A_{eq} \cdot X = b_{eq}, \\ lb \le X \le Ub, \\ C\left(X\right) \le 0\ , \\ C_{eq}\left(X\right) = 0. \end{cases} \tag{19.2}$$

Формат звернення до функції ***fmincon:***

$$\left[X, minW, flag\right] = fmincon(fun, X_0, A, b, A_{eq}, b_{eq}, lb, Ub, nonlcon, options, p_1, p_2),$$

де X – вектор змінних, за якими відбувається оптимізація (керуючі змінні); $minW$ – мінімальне значення показника ефективності; *flag* – символ якості виконання оптимізації(має той самій зміст, що і у випадку квадратичного програмування); fun – ім’я файл-функції, яка обчислює значення показника ефективності (може залежати від декількох параметрів p_1, p_2, значення яких передається за допомогою аргументів $fmincon$, які починаються з одинадцятої позиції у її вхідному списку; аргументом $fmincon$ є вектор X); вектори і матриці, які не використовуються, позначаються як []; X_0 – початкове значення вектора X; $nonlcon$ – ім’я файл-функції нелінійних обмежень (вхідним аргументом є вектор X, а вихідні аргументи – ліві частини обмежень); $options$ – управляюча структура, яка задає властивості обчислювального алгоритму.

Приклад 19.2.

$$W = 3x_1^2 + 2x_2^2 \to \min,$$
$$x_1^2 + x_2^2 \le 1.$$

❍ **Побудова файл-функції:File_New_M-file**

```
function fp1= myfunp1(x)
fp1=3*x(1)^2+2*x(2)^2;
```

Save as…myfunp1

Побудова файл-функції:File_New_M-file

```
function [c, ceq]=myconp1(x)
c(1)=x(1)^2+x(2)^2-1;
ceq=[];
```

Save as…myconp1

```
>> [x, fp1, flag]=fmincon(@myfunp1,[0.7 0.7],[],[],[],[],[],[],@myconp1)
x = 1.0e-004 * (0.1895   -0.0235)
fp1 = 1.0882e-009
```

flag = 1●

Приклад 19.3.

$$W = \left(x_1 - 5\right)^2 + \left(x_2 - 8\right)^2 \to \min\left(\max\right),$$
$$x_1 + 2x_2 \ge 4,$$
$$\frac{5}{2}x_1 + 2x_2 \le 20,$$
$$x_{1,2} \ge 0.$$

❍ **Побудова файл-функції:File_New_M-file**

```
function fp2min= myfunp2min(x)
fp2min=(x(1)-5)^2+(x(2)-8)^2;
```

Save as…myfunp2min

```
>> A=[-1   -2
  2.5   2]
>> b=[-4;20]
>> [x, fp2min, flag]=fmincon(@myfunp2min,[0 5],[A],[b],[],[],[0 0],[])
x =  2.9268    6.3415
fp2min =  7.0488
```

flag = 1

Побудова файл-функції:File_New_M-file

```
function fp2max= myfunp2max(x)
fp2max=-((x(1)-5)^2+(x(2)-8)^2);
```

Save as…myfunp2max

```
>>  [x, fp2max, flag]=fmincon(@myfunp2max,[0 5],[A],[b],[],[],[0 0],[])
x = 0     2
fp2max = -61
```

flag = 1●

Приклад 19.4.

$$W = x_1^2 - 2x_1 + x_2^2 \to \min\left(\max\right),$$
$$x_1 + \frac{3}{2}x_2 \le \frac{15}{2},$$
$$0 \le x_1 \le 10,$$
$$0 \le x_2 \le 9,$$
$$x_1x_2 \le 6.$$

❍**Побудова файл-функції:File_New_M-file**

function fp3min= myfunp3min(x)
fp3min=x(1)^2+x(2)^2-2*x(1);
Save as…myfunp3min
>> A=[1 1.5]

A = 1.0000   1.5000

>> b=[7.5]
b = 7.5000

Побудова файл-функції:File_New_M-file

function [c, ceq]=myconp3(x)
c(1)=x(1)*x(2)-6;
ceq=[];
Save as…myconp3

>> [x, fp3min, flag]=fmincon(@myfunp3min,[0 0],[A],[b],[],[],[0 0],[10 9], @myconp3)
x = 1.0000   0
fp3min = -1
flag = 1
Побудова файл-функції:File_New_M-file
function fp3max= myfunp3max(x)
fp3max=-(x(1)^2+x(2)^2-2*x(1));
Save as…myfunp3max
>> [x, fp3max, flag]=fmincon(@myfunp3max,[**0 0**],[A],[b],[],[],[0 0],[10 9], @myconp3)
x = 1.0e-007 *(0   0.1000)
fp3max = -1.0000e-016
flag = 1
>> [x, fp3max, flag]=fmincon(@myfunp3max,[**6 0**],[A],[b],[],[],[0 0],[10 9], @myconp3)
x = 7.5000   0
fp3max = -41.2500
flag = 1●

Приклад 19.5.

$$
\begin{aligned}
&W = -x_1^2 + 4x_1 + x_2 \rightarrow \min\left(\max\right),\\
&{x_1}^2 + {x_2}^2 \le 49,\\
&0 \le x_1,\\
&0 \le x_2 \le 6,\\
&x_1 x_2 \ge 6.
\end{aligned}
$$

❍ **Побудова файл-функції:File_New_M-file**

```
function fp4min= myfunp4min(x)
fp4min=-x(1)^2+4*x(1)+x(2);
```

Save as…myfunp4min

Побудова файл-функції:File_New_M-file

```
function [c, ceq]=myconp4(x)
c(1)=-x(1)*x(2)+6;
c(2)=x(1)^2+x(2)^2-49;
ceq=[];
```

Save as…myconp4

```
>> [x, fp4min, flag]=fmincon(@myfunp4min,[3 4],[],[],[],[],[0 0],[10 6], @myconp4)
x = 6.9465    0.8637
fp4min = -19.6042
```

flag = 1

Побудова файл-функції:File_New_M-file

```
function fp4max= myfunp4max(x)
fp4max=-(-x(1)^2+4*x(1)+x(2));
```

Save as…myfunp4max

```
>> [x, fp4max, flag]=fmincon(@myfunp4max,[3 4],[],[],[],[],[0 0],[10 6], @myconp4)
x = 2.0000    6.0000
```

fp4max = **-10.0000**

flag = 1●

Пошук глобального максимуму (мінімуму) полімодальної функції

Побудова файл-функції:File_New_M-file

```
function fP=myfunP(x)
fP=4*sin(2*pi*x(1))*cos(1.5*pi*x(2))*(1-x(1)^2)*x(2)*(1-x(2));
```

Save as…myfunP

```
>> [x1,x2]=meshgrid(-1:0.05:1,0:0.05:1);
>> fP=4.*sin(2.*pi.*x1).*cos(1.5*pi.*x2).*(1-x1.^2).*x2.*(1-x2);
```

```
>> mesh(x1,x2,fP)
```

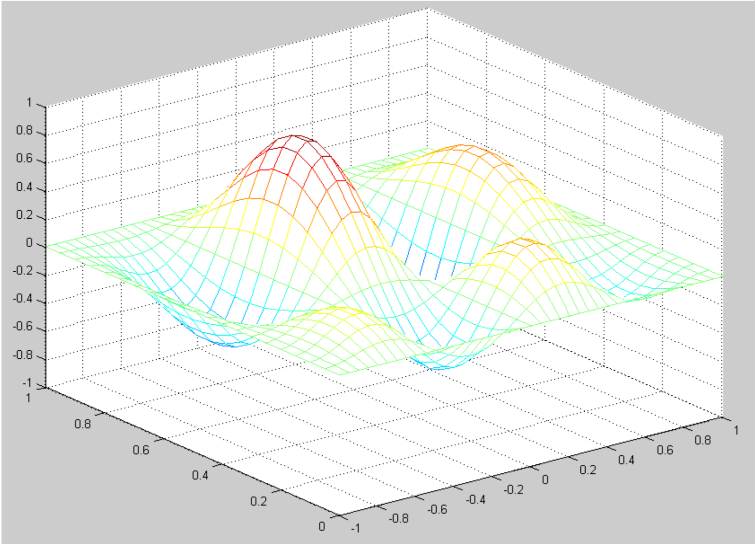


Рис. 19.1. Графічний образ полімодальної функції

```
>> [CMatr,h]=contour(x1,x2,fP);
>> clabel(CMatr,h)
>> grid on
```

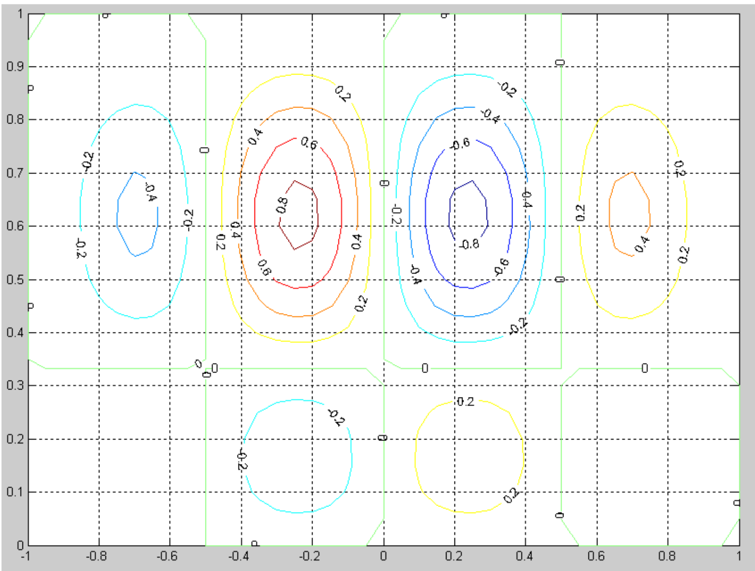


Рис. 19.2. Графічний образ ліній рівного рівня полімодальної функції

Початкові умови [–1; 0]

```
>> [X,fP,flag]=fmincon(@myfunP,[-1;0],[],[],[],[],[-1;0],[1;1])
X =   -1
         0
fP = 0
flag =  1
```

Початкові умови [0.2;0.6]

```
>> [X,fP,flag]=fmincon(@myfunP,[0.2;0.6],[],[],[],[],[-1;0],[1;1])
X =    0.2373
       0.6211
fP =  -0.8652
flag =   1
```

Початкові умови [-0.6;0.5]

```
>> [X,fP,flag]=fmincon(@myfunP,[-0.6;0.5],[],[],[],[],[-1;0],[1;1])
X =  -0.6874
      0.6211
fP =   -0.4481
flag =     1
```

Пошук глобального максимуму (мінімуму) функції Розенброка

Побудова файл-функції:File_New_M-file

```
function fR=myfunR(x)
fR=100*(x(2)-x(1)^2)^2+(1-x(1))^2;
```

Save as…myfunR

```
>> [x1,x2]=meshgrid(-2:0.05:2,-2:0.05:2);
>> fR=100.*(x2-x1.^2).^2.+(1-x1).^2;
>> mesh(x1,x2,fR)
```

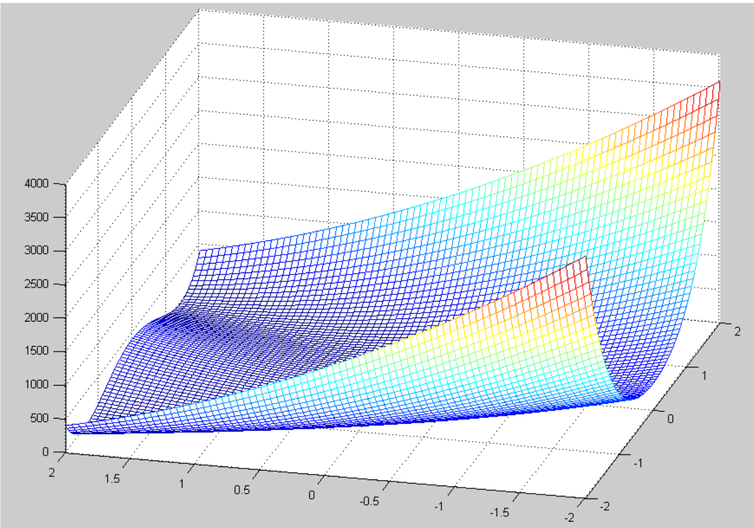


Рис. 19.3. Графічний образ функції Розенброка

```
>> [CMatr,h]=contour(x1,x2,fR);
>> clabel(CMatr,h)
>> grid on
```

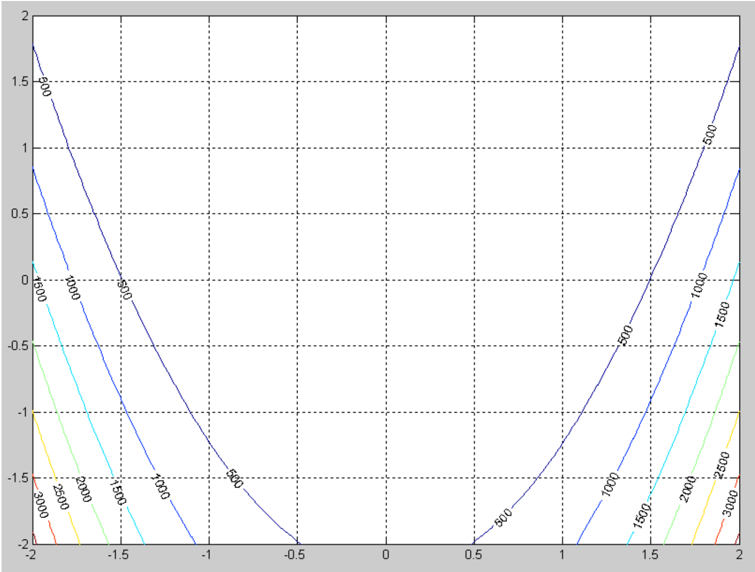


Рис. 19.4. Графічний образ ліній рівного рівня функції Розенброка

```
>> p=[0.5 2 10 50 100 150 200 250 300 350 400];
>> [CMatr,h]=contour(x1,x2,fR,p);
>> clabel(CMatr,h)
>> grid on
>> [X,fR,flag]=fmincon(@myfunR,[-1;-1],[],[],[],[],[-2;-2],[2;2])
X =    1.0000
       1.0001
fR =  9.7522e-009
flag =     1
```

Пошук глобального максимуму (мінімуму) функції Вуда

Побудова файл-функції:File_New_M-file

```
function fB=myfunB(x)
f1=100*(x(2)-x(1)^2)^2+(1-x(1))^2+90*(x(4)-x(3)^2)^2+(1-x(3))^2;
f2=10.1*((x(2)-1)^2+(x(4)-1)^2)+19.8*(x(2)-1)*(x(4)-1);
fB=f1+f2;
```

Save as…myfunB

Початкові умови [3;3;3;3]

```
>> [X,fB,flag]=fmincon(@myfunB,[3;3;3;3],[],[],[],[],[-3;-3;-3;-3],[3;3;3;3])
X =   1.0000
      1.0000
      1.0000
      1.0000
fB = 4.2696e-009
flag =    1
```

Початкові умови [-3;3;3;3]

```
>>  [X,fB,flag]=fmincon(@myfunB,[-3;3;3;3],[],[],[],[],[-3;-3;-3;-3],[3;3;3;3])
```

```
X =   1.0000
      1.0000
      1.0000
      1.0000
fB =  3.4980e-010
flag =    1.
```

Лекція 20. Нелінійне програмування з сепарабельними функціями. Дробово-лінійне програмування

20.1. Задачі нелінійного програмування з сепарабельними функціями

Означення 20.1. Функцію n змінних $f(x_1,x_2,\ldots,x_n)$ називають *сепарабельною*, якщо вона є сумою n функцій, кожна з яких залежить від однієї змінної, тобто:

$$f\left(x_1,x_2,\ldots,x_n\right) = \sum_{j=1}^{n} f_j(x_j).$$

Розглянемо задачу, в якій всі функції сепарабельні:

$$W(x_1,\ldots,x_n) = \sum_{j=1}^{n} f_j(x_j) \to \max, \qquad (20.1)$$

$$\sum_{j=1}^{n} g_{ij}\left(x_j\right)\left\{\le = \ge\right\} b_i, \quad i = \overline{1,m}, \qquad (20.2)$$

$$x_j \ge 0, \; j = \overline{1,n}. \qquad (20.3)$$

Оскільки всі лінійні функції є сепарабельними, то розглядувана задача має структуру, подібну до ЗЛП. Ця подібність використовувана для побудови методів розв'язування нелінійних задач із сепарабельними функціями.

Наближене розв'язування задачі (20.1) – (20.3) досить просте: функції f_j, g_{ij} замінюємо їх кусково-лінійними апроксимаціями, а вихідну задачу – наближеною задачею лінійного програмування, яка розв'язується симплекс методом. При цьому знаходимо її локальний максимум, який буде наближеним локальним максимумом вихідної задачі. Якщо ж задача (20.1) – (20.3) буде задачею опуклого програмування, то наближений локальний максимум буде одночасно і глобальним.

Заміну нелінійної функції її кусково-лінійною апроксимацією розглянемо на прикладі неперервної функції $f(x)$ однієї змінної. Позначимо інтервал допустимих значень змінної $[\alpha_1;\alpha_2]$. Розіб'ємо його за допомогою фіксованих точок $\alpha_1 = x_0, x_1, x_2 \ldots, x_{r-1}, x_r = \alpha_2$ і обчислимо значення функції $f(x)$ в точках розподілу $f\left(x_0\right), f\left(x_1\right), \ldots, f(x_r)$. Точки $(x_k, f(x_k)$ на кривій $f(x)$ сполучимо відрізками прямих ліній. В результаті отримуємо кусково-лінійну апроксимацію – функцію $\overline{f}(x)$, яка наближує функцію $f(x)$ (рис. 20.1).

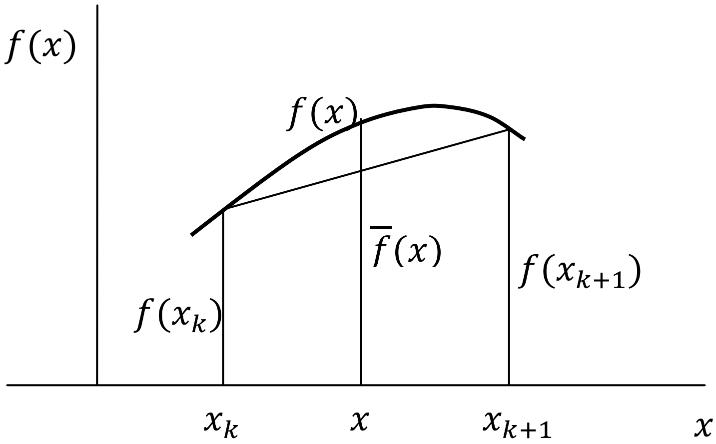


Рис. 20.1. Кусково-лінійна апроксимація нелінійної функції

На основі рівняння прямої, яка проходить через дві точки, можна записати:

$$\overline{f}\left(x\right) = f\left(x_k\right) + \frac{x - x_k}{x_{k+1} - x_k}\left[f\left(x_{k+1}\right) - f\left(x_k\right)\right] . \quad (20.4)$$

Очевидно, що для будь-якої точки x такої, що $x_k \le x \le x_{k+1}$, допустиме значення зображення:

$$x = \left(1 - \lambda\right)x_k + \lambda x_{k+1}, \qquad 0 \le \lambda \le 1.$$

$$x - x_k = \lambda\left(x_{k+1} - x_k\right) \quad (20.5)$$

Підставляючи (20.5) у (20.4), отримаємо:

$$\overline{f}\left(x\right) = \left(1 - \lambda\right)f\left(x_k\right) + \lambda f\left(x_{k+1}\right) . \quad (20.6)$$

Позначивши тепер в (20.6) $1 - \lambda = \lambda_k, \;\; \lambda = \lambda_{k+1}$, одержимо:

$$\overline{f}\left(x\right) = \lambda_k f\left(x_k\right) + \lambda_{k+1} f\left(x_{k+1}\right), \quad (20.7)$$

$$\lambda_k \ge 0, \;\; \lambda_{k+1} \ge 0, \;\; \lambda_k + \lambda_{k+1} = 1.$$

$$x = \lambda_k x_k + \lambda_{k+1} x_{k+1}. \quad (20.8)$$

Формули (20.7) і (20.8) можна узагальнити для зображення довільного $x \in [\alpha_1;\alpha_2]$ і відповідних довільних значень апроксимуючої функції, якщо прийняти таку додаткову умову: в узагальнених формулах відмінними від нуля (тобто додатними) можуть бути одне або два сусідніх λ_k.

Тоді узагальнені формули матимуть такий вигляд:

$$\overline{f}\left(x\right)=\sum_{k=0}^{r}\lambda_k f\left(x_k\right), \tag{20.9}$$

$$x=\sum_{k=0}^{r}\lambda_k x_k, \tag{20.10}$$

$$\lambda_k \geq 0\ , \quad \sum_{k=0}^{r}\lambda_k = 1 .$$

Користуючись виразами (20.9) і (20.10), замінимо для задачі (20.1) – (20.3) усі функції їх кусково-лінійними наближеннями:

$$\overline{f}_j\left(x\right)=\sum_{k=0}^{rj}\lambda_{kj} f_j\left(x_{kj}\right), \tag{20.11}$$

$$\overline{g}_{ij}=\sum_{k=0}^{rj}\lambda_{kj} g_{ij}\left(x_{kj}\right) \tag{20.12}$$

де $x_j=\sum_{k=0}^{r_j}\lambda_{kj}x_{kj},\ \lambda_{kj}\geq 0,\ \sum_{k=0}^{r_j}\lambda_{kj}=1$

для всіх значень k,j, причому для даного j не більше двох сусідніх за індексом k значень λ_{kj} можуть бути відмінними від нуля (додатними).

Підкреслимо, що при зображенні функцій $\overline{f}_j, \overline{g}_{ij}$ використовується одне й теж розбиття інтервалу $[\alpha_1;\alpha_2]$. При цьому точки x_{ij} вибираються так, щоб забезпечити необхідну точність апроксимації всіх функцій f_j, g_{ij}.

Таким чином, наближену задачу можемо записати так:

$$\overline{W}(x_1,\ldots,x_n)=\sum_{j=1}^{n}\sum_{k=0}^{r_j}\lambda_{kj} f_j\left(x_{kj}\right)\rightarrow \max, \tag{20.13}$$

$$\sum_{j=1}^{n}\sum_{k=0}^{r_j}\lambda_{kj} g_{ij}\left(x_{kj}\right)\left\{\leq=\geq\right\} b_i,\ i=\overline{1,m}, \tag{20.14}$$

$$\sum_{k=0}^{r_j}\lambda_{kj}=1,\quad j=\overline{1,n}\ ,\ \lambda_{kj}\geq 0,\quad k=\overline{0,r_j}. \tag{20.15}$$

Ця наближена задача називається задачею в λ-формі.

Змінними наближеної задачі є величини λ_{kj}, оскільки значення функцій $f_j\left(x_{kj}\right)$, $g_{ij}\left(x_{kj}\right)$ фіксовані (обчислені у вузлових точках – точках зламу апроксимованої ламаної лінії). Число змінних λ_{kj} легко підраховується, а саме:

$$N = \sum_{r=1}^{n} r_j + 1.$$

Розв'язуємо задачу (20.13)–(20.15) симплекс методом. Додержуємося такого правила: у базисі має бути не більше двох, і причому сусідніх, відмінних від нуля (додатних) змінних λ_{kj} для кожного $j = \overline{1,n}$.

Знайшовши розв'язок наближеної задачі в λ-формі, оптимальні значення $\hat{x}_j$ обчислюємо за формулою:

$$\hat{x}_j = \sum_{k=0}^{r_j} \lambda_{kj} x_{kj}.$$

Наближена задача, навіть для вихідної задачі невеликої розмірності, може бути дуже великих розмірів за рахунок різкого збільшення числа змінних. Тому потребують заміни через λ_{kj} лише змінні, відносно яких вихідна задача нелінійна, а решту можна розглядати як звичайні змінні лінійної задачі.

Теорема 20.1. Якщо функції $f_j(x_j)$ вгнуті, а функції $g_{ij}(x_j)$ опуклі або вгнуті, так що множина G допустимих розв'язків опукла, то, розв'язуючи наближену задачу (20.13)–(20.15) як задачу лінійного програмування, отримаємо оптимальний розв'язок наближеної задачі.

Приклад 20.1. Розв'язати наступну задачу методом кусково-лінійної апроксимації:

$$W(x_1, x_2) = 3x_1 + 2x_2 \to \max,$$

за умов:

$$4x_1^2 + x_2^2 \le 16,$$

$$x_1 \ge 0, x_2 \ge 0.$$

❍Цільову функцію й обмеження можна зобразити як суму двох функцій:

$$f_1\left(x_1\right) = 3x_1 \quad \text{і} \quad f_2\left(x_2\right) = 2x_2,$$

$$g_{11}\left(x_1\right) = 4x_1^2 \quad \text{і} \quad g_{12}\left(x_2\right) = x_2^2.$$

Кожна з яких є функцією однієї змінної. Отже, цільова функція й обмеження є сепарабельними.

Оскільки функції $f_1(x_1)$ і $f_2(x_2)$ лінійні, то апроксимуватимемо функції $g_{11}(x_1)$ і $g_{12}(x_2)$. Як видно з рис.20.2 змінна x_1 може приймати значення на інтервалі $[0;2]$, а змінна $x_2 \in [0;4]$. Розіб'ємо ці інтервали точками $x_{01} = 0,\ x_{11} = 1,\ x_{21} = 1.5,\ x_{31} = 2$ і $x_{02} = 0,\ x_{12} = 1,\ x_{22} = 2,\ x_{32} = 3,\ x_{42} = 4$. Обчислимо значення функцій в цих точках (табл. 20.1).

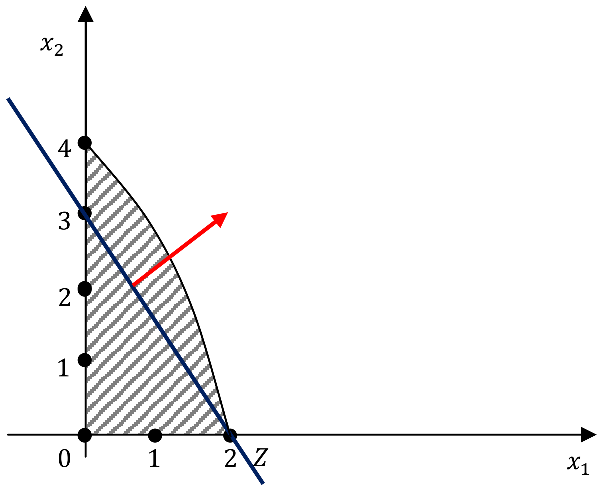


Рис. 20.2. Геометрична інтерпретація вихідної задачі

Таблиця 20.1.

Значення функцій в точках розбиття

K	X_{k1}	X_{k2}	$g_{11}(x_1)$	$g_{12}(x_2)$	$f_1(x_1)$	$f_2(x_2)$
0	0	0	0	0	0	0
1	2	1	4	1	3	2
2	1.5	2	9	4	4.5	4
3	2	3	16	9	6	6
4	-	4	-	16	-	8

Підставляючи отримані результати у вихідну задачу, маємо таку наближену задачу:

$$\bar{W} = 3\lambda_{11} + 4.5\lambda_{21} + 6\lambda_{31} + 2\lambda_{12} + 4\lambda_{22} + 6\lambda_{32} + 8\lambda_{42} \to \max,$$

$$4\lambda_{11} + 9\lambda_{21} + 16\lambda_{31} + \lambda_{12} + 4\lambda_{22} + 9\lambda_{32} + 16\lambda_{42} + x_3 = 16,$$

$$\lambda_{01} + \lambda_{11} + \lambda_{21} + \lambda_{31} = 1,$$

$$\lambda_{02} + \lambda_{12} + \lambda_{22} + \lambda_{32} + \lambda_{42} = 1,$$

$$\lambda_{kj} \ge 0,\ k = \overline{0,4}\ ,\ j = 1,2,\ \ x_3 \ge 0.$$

Розв'яжемо задачу симплекс методом. Оптимальний розв'язок: $\hat{x}_1 = 1.2,\ \ \hat{x}_2 = 3.2,\ \ W(\hat{x}_1, \hat{x}_2) = 10.$ ●

Зауваження 20.1. Для збільшення точності розв'язку задачі слід встановити деякий окіл точки, знайденого з наближеної задачі оптимального плану і знову розбити інтервал зміни окремих координат в цьому околі на певне число проміжків, скласти нову наближену задачу та розв'язати її. Ітерації можна повторювати до отримання необхідної точності розв'язку.

Зауваження 20.2. Деякі несепарабельні функції відносно даних змінних функції будуть сепарабельними відносно інших змінних. Знайшовши відповідну заміну змінних, можна застосувати описаний метод.

Наприклад, функція $W = x_1 x_2$ сепарабельна відносно змінних y_1, y_2, де $x_1 = y_1 + y_2,\ x_2 = y_1 - y_2$, звідки $W = y_1^2 - y_2^2$, або замість функції

$$W = \alpha_0 \prod_{j=1}^{n} x_j^{\alpha_j}$$

можна розглядати функцію

$$LW = \ln W \approx \ln \alpha_0 + \sum_{j=1}^{n} \alpha_j y_j, \quad y_j = \ln x_j .$$

20.2. Задачі дробово-лінійного (гіперболічного) програмування

Практичне значення задач зумовлюється тим, що значна кількість важливих, насамперед, економічних, показників є відносними, тобто вони визначаються як частка від ділення двох інших показників, наприклад, собівартість продукції, рівень рентабельності господарської діяльності тощо. Загальна дробово-лінійна задача полягає у відшуканні максимального значення функції

$$W = \sum_{j=1}^{n} c_j x_j \Big/ \sum_{j=1}^{n} d_j x_j \tag{20.16}$$

за умов:

$$\sum_{j=1}^{n} a_{ij} x_j = b_i, \quad i = \overline{1,m}, \tag{20.17}$$

$$x_j \geq 0, \quad j = \overline{1,n}, \tag{20.18}$$

де $c_j,\ d_j,\ a_{ij},\ b_i$ – деякі сталі числа.

Хоча вигляд цільової функції дає підстави віднести ці задачі до нелінійних, їх математична структура дозволяє розв'язувати ці задачі одним із варіантів симплекс методу.

Покладемо, що на множині планів задачі знаменник цільової функції не стає рівним нулю, тобто

$$\sum_{j=1}^{n} d_j x_j \neq 0 .$$

Крім того, він приймає лише додатні значення в допустимій області:

$$\sum_{j=1}^{n} d_j x_j > 0.$$

Якщо ця величина від'ємна, то знак мінус можна віднести до чисельника.

Теорема 20.2. Цільова функція задачі дробово-лінійного програмування досягає свого максимального значення в крайній точці многогранника розв'язків, тобто хоча б один з опорних планів буде розв'язком задачі (природно, за умов, що ця задача має оптимальний план).

Теорема 20.3. Якщо максимум дробово-лінійної функції досягається в кількох крайніх точках многогранника планів задачі, то кожна точка границі, що є опуклою лінійною комбінацією цих крайніх точок, є розв'язком задачі.

Відмінність геометрії задачі дробово-лінійного програмування від задач лінійного програмування полягає лише в інтерпретації гіперповерхні рівня цільової функції при зміні її значень. Рівнянням гіперповерхні рівня $W = const$, очевидно, буде рівняння гіперплощини, яка проходить через початок координат

$$\sum_{j=1}^{n} c_j x_j - W \sum_{j=1}^{n} d_j x_j = 0.$$

Змінюючи $Z \in (-\infty; +\infty)$, одержимо пучок гіперплощин. У випадку $n = 2$ центром симетрії пучка прямих буде початок координат, при $n \geq 3$ гіперплощини пучка симетричні відносно точкового підпростору (гіперсиметрії) розмірності $n - 2$. При зростанні (зменшенні) W гіперплощина обертається відносно гіперосі симетрії в одну сторону. Положення гіперплощини рівня, при яких вона дотикається до множини планів (у вершині чи по грані), відповідає екстремальним значенням цільової функції на множині планів задачі.

Отже, процес розв'язання дробово-лінійної задачі за допомогою симетричної інтерпретації має такі етапи:

1. Будуємо область допустимих розв'язків, яка визначається обмеженнями задачі.
2. Будуємо пряму рівня, рівняння якої одержуємо, якщо покласти значення цільової функції рівним деякому сталому числу. Пряма проходить через початок координат. Нехай також вона має спільні точки з многогранником розв'язків задачі.
3. Обертаючи її коло центру симетрії (початку координат), знаходимо крайню точку (вершину) або крайні точки, де цільова функція набуває максимального значення чи установлюємо необмеженість функції на множині планів задачі.
4. Визначаємо координати точки максимуму, якщо він існує.
5. Обчислюємо значення цільової функції в точці максимуму.

При розв'язанні конкретних задач можливі такі випадки.

1) Многогранник розв'язків обмежений. Максимум і мінімум досягається в його крайніх точках – вершинах.
2) Многогранник розв'язків необмежений, але існують крайні точки, в яких цільова функція задачі приймає відповідно максимальне і мінімальне значення (рис.20.3).
3) Многогранник розв'язків необмежений (многогранна множина), але один з екстремумів досягається. Наприклад, мінімум досягається в точці С. Функція W має асимптотичний максимум, тобто оптимального плану, як такого, не існує, бо точка знаходиться на нескінченності деякого променя-ребра множини планів (рис.20.4).
4) На многогранній множині планів задачі як максимум, так і мінімум мають асимптотичний характер.

В задачах з асимптотичним оптимумом цільової функції оптимальний план визначається наближено, як деяка досить віддалена від вершини точка ребра асимптотичного оптимуму, а саме $X^* = \hat{X}^{(r)} + \beta \overline{X}^{(r)}$, де $\overline{X}^{(r)}$ – напрямний вектор цього ребра, β– деяке досить велике число, таке, що $Z\left(\hat{X}^{(r)} + \beta \overline{X}^{(r)}\right) \cong \max Z$ з потрібною точністю.

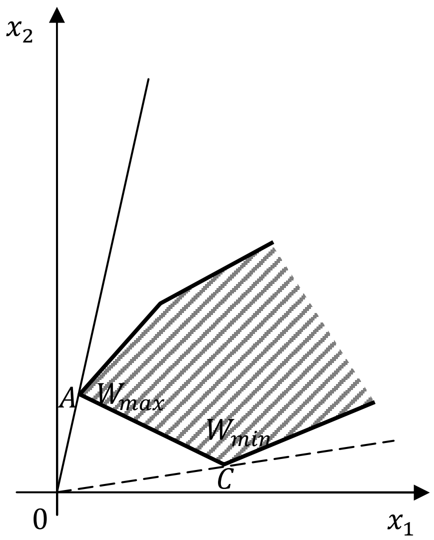


Рис. 20.3. Геометрична інтерпретація розв'язку у випадку необмеженої області допустимих рішень

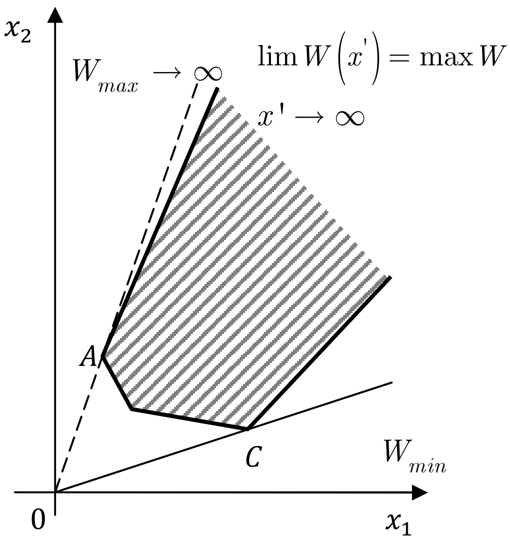


Рис. 20.4. Геометрична інтерпретація альтернативного максимуму

Зведемо сформульовану задачу дробово-лінійного програмування (20.16)–(20.18) до задачі лінійного програмування. Позначимо

$$y_0 = \frac{1}{\sum_{j=1}^{n} d_j x_j} \tag{20.19}$$

і введемо нові змінні

$$y_i = y_0 x_j,\ j = \overline{1,n}. \tag{20.20}$$

В нових змінних модель (20.16)–(20.18) набуває наступного вигляду.

Максимізувати

$$W = \sum_{j=1}^{n} c_j y_j \tag{20.21}$$

за умов

$$\sum_{j=1}^{n} a_{ij} y_j - b_i y_0 = 0,\ i = \overline{1,m}; \tag{20.22}$$

$$\sum_{j=1}^{n} d_j y_j = 1,\ y_0 \geq 0,\ y_j \geq 0,\ j = \overline{1,n}. \tag{20.23}$$

Отже, отримано модель задачі лінійного програмування, яка може бути розв'язана звичайним симплексним методом. З оптимального розв'язку $\hat{W} = (\hat{y}_1, \ldots, \hat{y}_n, \hat{y}_0)$ задачі (20.21)–(20.23), за умови $\hat{y}_0 > 0$, одержимо, за допомогою співвідношень (20.20), оптимальний розв'язок вихідної задачі (20.16) – (20.18).

Таким чином, процес розв'язання задачі дробово-лінійного програмування містить такі етапи:

1) Звести задачу (20.16)–(20.18) до задачі лінійного програмування (20.21)–(20.23).
2) Знайти розв'язок лінійної задачі (20.21)–(20.23).
3) Використовуючи співвідношення (20.20), визначити оптимальний план вихідної задачі дробово-лінійного програмування і знайти максимальне значення цільової функції.

Приклад 20.2. Знайти розв'язок задачі

$$W = \frac{-5x_1 + 2x_2}{3x_1 + 4x_2} \to \max,$$

$$x_1 + 4x_2 + x_3 = 16,$$

$$2x_1 + 3x_2 - x_4 = 12,$$

$$3x_1 - 2x_2 + x_5 = 18,$$

$$x_j \geq 0, \quad j = \overline{1,5}.$$

❍Позначимо $\left(3x_1 + 4x_2\right)^{-1}$ через y_0 і введемо нові змінні. Приходимо до такої задачі лінійного програмування: знайти максимум функції:

$$W = -5y_1 + 2y_2$$

за умов

$$y_1 + 4y_2 + y_3 - 16y_0 = 0,$$

$$
\begin{aligned}
&2y_1 + 3y_2 - y_4 - 12y_0 = 0,\\
&3y_1 - 2y_2 + y_5 - 18y_0 = 0,\\
&y_0 \ge 0, \quad y_j \le 0, \quad j = \overline{1,5}.
\end{aligned}
$$

Розв’язуємо задачу симплекс методом. Оптимальним планом задачі є

$$y_2 = \frac{1}{4}, \ \ y_5 = \frac{26}{16}, \ \ y_0 = \frac{1}{16}. \bullet$$

Лекція 21. Чисельні методи розв’язання багатовимірних задач нелінійного програмування за наявності обмежень

21.1. Загальна характеристика чисельних методів

Чисельні методи відіграють значну роль у прикладних дослідженнях. Це обумовлено рядом причин, серед яких головне місце займає різноманіття функцій $f(X)$ та $g_i(X), i = \overline{1,m}$, а також форм їх завдання. В окремих випадках навіть важко визначити, до якого класу відноситься та чи інші задача, а також чи існує для неї обґрунтований метод розв'язання.

Розроблено багато чисельних методів для задач як безумовної, так й умовної оптимізації. Природним є бажання вибрати для розв'язання конкретної задачі найкращий метод, який дозволяє отримати розв'язок із необхідною точністю з максимальним використанням потужності ЕОМ.

Якість чисельних методів характеризується багатьма факторами: швидкість збіжності, час виконання однієї ітерації, обсяг пам'яті ЕОМ, який є необхідним для реалізації методу, клас задач, що розв'язуються і т. ін. Розв'язувані задачі також дуже різноманітні: вони можуть мати велику чи малу вимірність, бути унімодальними чи багатоекстремальними тощо. Один і той же метод, що є ефективним для розв'язання задач одного типу, може бути зовсім непридатним для задач іншого типу.

Було запропоновано декілька функцій, які через свої властивості є тестовими для чисельних методів. Прикладом таких функцій є

- функція Розенброка:

$$Z = f\left(x_1, x_2\right) = 100\left(x_2 - x_1^*\right)^2 + \left(1 - x_1\right)^2, \ \hat{X} = \left(1,1\right); \tag{21.1}$$

- функція Пауелла:

$$\begin{aligned} &Z = f\left(x\right) = \left(x_1 + 10x_2\right)^2 + 5\left(x_3 - x_4\right)^2 + \left(x_1 - x_4\right)^4 + 10\left(x_1 - x_4\right)^4, \\ &\hat{X} \ = \left(0,0,0,0\right). \end{aligned} \tag{21.2}$$

Будь-яка оптимізаційна процедура має ефективно розв'язувати задачі (21.1), (21.2) та інші тестові задачі.

Методи оптимізації поділяють на класи залежно від інформації, що використовується. Якщо на кожній ітерації використовується лише значення функцій, то метод називається методом нульового порядку. Якщо, крім того, буде потрібним обчислення перших похідних функцій, то мають місце методи першого порядку, при необхідності додаткового обчислення других похідних — методи другого порядку.

21.2. Методи нульового порядку

Будемо вести мову про пошук мінімального значення функції n змінних.

У методах нульового порядку для визначення напряму спуску не потрібне обчислення похідних цільової функції. Напрям мінімізації повністю визначається послідовними обчисленнями значень функції. Зазначимо, що при розв'язанні задач безумовної оптимізації методи першого та другого порядку мають, як правило, більш високу швидкість збіжності, ніж методи нульового порядку. Але виникає проблема щодо обчислення перших та других похідних функцій багатьох змінних. В деяких випадках їх не можна отримати у вигляді аналітичних функцій. Визначення похідних за допомогою різних чисельних методів спричиняє помилки, які можуть обмежити застосування таких методів. Крім того, на практиці зустрічаються задачі, розв'язання яких можливе лише методами нульового порядку, наприклад, задачі мінімізації функцій із розривними першими похідними, Критерій оптимальності може задаватися не в явному вигляді, а системою рівнянь. У цьому випадку аналітичне або чисельне визначення похідних стає надто складним, а іноді й неможливим. Для розв'язання таких практичних задач оптимізації рекомендується застосувати методи нульового порядку.

21.2.1. Метод прямого пошуку (метод Хука-Дживса)

Пошук (наприклад мінімуму) складається з послідовності кроків дослідного пошуку навколо базисної точки, за яким у випадку успіху іде пошук за зразком.

Суть методу. Задається деяка початкова точка $X\left(0\right)$. Змінюючи компоненти вектора $X\left(0\right)$, обслідується окіл обраної точки, в результаті чого знаходиться напрям, в якому зменшується функція $Z = f(X)$. У вибраному напрямі здійснюється спуск доти, поки значення функції зменшується. Після того, як у даному напрямі не знаходиться точка з меншим значенням функції, зменшується величина кроку спуску. Якщо послідовні дроблення кроку не призводять до зменшення функції, від вибраного напряму спуску відмовляються і здійснюється нове обстеження околу і т. д.

Метод не потребує знання цільової функції в явному вигляді, дозволяє легко враховувати обмеження на окремі змінні, а також складні обмеження на допустиму область пошуку.

Недоліком методу є те, що у випадку сильно витягнутих, зігнутих ліній рівня цільової функції, він може не забезпечити рух до точки мінімуму.

Алгоритм методу прямого пошуку такий:

1. Задаються значення координат $x_j\left(0\right),\ \ j=\overline{1,n}$ початкової точки $X\left(0\right)$, вектором зміщення координат ΔX у процесі обстеження околу, найменшим допустимим значенням ε компонент ΔX .
2. Покладаючи, що $X\left(0\right)$є базисною точкою X^b, обчислюють $f\left(X^b\right)$.
3. Циклічно змінюють кожну координату $x_j^b,\ \ j=\overline{1,n}$, базисної точки X^b на величину $\Delta x_j, j=\overline{1,n}$, тобто $x_j\left(0\right)=x_j^b+\Delta x_j,\ x_j\left(0\right)=x_j^b-\Delta x_j$. При цьому обчислюються значення $f(X\left(k\right))$ і порівнюються з $f\left(X^b\right)$. Якщо $f\left(X\left(k\right)\right)<f\left(X^b\right)$, то відповідна координата $x_j,\ j=\overline{1,n}$ приймає нове значення, яке обчислюється за одним із наведених вище виразів. Такий крок вважають успішним. В протилежному випадку значення цієї координати залишається незмінним. Якщо після змінення останньої n-ї координати $f\left(X\left(k\right)\right)<f\left(X^b\right)$, то переходять до п.4. В протилежному випадку – до п.7.
4. Покладають, що $X\left(k\right)$ є новою базисною точкою X^b, обчислюють $f\left(X^6\right)$ і здійснюють пошук за зразком.
5. При пошуку за зразком використовується інформація. Отримана в процесі дослідження, і мінімізація функції завершується пошуком у напрямі, який задається зразком. Ця процедура здійснюється так: здійснюється спуск із точкою $X\left(k\right),\ x_j\left(k+1\right)=2x_j\left(k\right)-x_j^b,\ j=\overline{1,n}$, де x_j^b – координати попередньої базисної точки, обчислюється значення $f(X\left(k+1\right))$.
6. Як і в п.3, циклічно змінюють кожну координату точки $X\left(k+1\right)$, здійснюючи порівняння відповідних значень функції $f\left(X\right)$ із значенням $f(X\left(k+1\right))$, яке отримано в п.5. Після змінення останньої координати порівнюють відповідне значення функції $f(X\left(k+1\right))$ із значенням $f\left(X^6\right)$, отримано в п.4. Якщо $f\left(X\left(k+1\right)\right)<f\left(X^b\right)$, то переходять до п.4, в протилежному випадку – до п.7.
7. Порівнюють значення ΔX та ε. Якщо $\Delta X<\varepsilon$, то обчислення припиняють. Вважаючи $X\left(k+1\right)$– розв’язком. У протилежному випадку зменшують значення ΔX на $\Delta X\ /\ 2$ і переходять до п.3.

На практиці задовільним є зменшення кроку (кроків) у десять разів від початкової довжини.

21.2.2. Метод деформованого многогранника (метод Нелдера–Міда)

Симплексний метод полягає в тому, що для мінімізації функції n-змінних $Z = f(X)$ у n- вимірному просторі будується многогранник, який має $n+1$ вершину. Множина $(n+1)$-ї рівновіддаленої точки в n-вимірному просторі називається регулярним симплексом. Для випадку двох змінних регулярний симплекс буде рівнобічним трикутником, трьох змінних – тетраедром і т.д. Координати вершин початкового симплексу можна визначити за допомогою матриці R розміром $n \times (n+1)$:

$$R = \begin{vmatrix} 0 & r_1 & r_2 & \cdots & r_2 \\ 0 & r_2 & r_1 & \cdots & r_2 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & r_2 & r_2 & \cdots & r_1 \end{vmatrix}, \tag{21.3}$$

де

$$r_1 = \left[\frac{l}{n\sqrt{2}}\right]\left(\sqrt{n+1} + n - 1\right),$$

$$r_2 = \left[\frac{l}{n\sqrt{2}}\right]\left(\sqrt{n+1} - 1\right)$$

– параметр, який ототожнюється з віддалю між двома вершинами. Елемент r_{ij} матриці R дорівнює i-й координаті j-ї вершини.

Пошук мінімуму функції симплексним методом ведеться таким чином. Установлюються координати вершин симплексу і в кожній з них обчислюється значення функції $f(X)$. Визначається вершина з найбільшим значенням $f(X)$. Через цю вершину й центральну точку симплексу проводиться пряма, на якій з деяким віддаленням від центру установлюється нова вершина. Потім вершина з найбільшим значенням $f(X)$ вилучається, а новий симплекс будується на вершинах, що залишилися, та новій вершині. Цей метод, використовує регулярні симплекси, може зациклюватися, має певні труднощі в яристих ситуаціях.

Метод Нелдера-Міда усуває ці недоліки. Координати вершин початкового симплексу можуть бути призначені або обчислені за (21.3). На наступних етапах конфігурація симплексу змінюється. Симплекс втрачає регулярність. Тому цей метод має другу назву – метод деформованого многогранника.

Введемо такі позначення:

$$X\left(j,k\right) = \left(x_1\left(j,k\right),\ldots,x_i\left(j,k\right),\ldots,x_n\left(j,k\right)\right),$$

де $j = \overline{1,n+1}$, $k = 0,1-$ j-та вершина многогранника на k-му етапі пошуку;

$X\left(h,k\right)$– вершина, в якій значення функції максимальне, тобто:

$$f\left(X\left(h,k\right)\right) = \max\left\{f\left(X\left(1,k\right)\right),\ldots,f\left(X\left(n+1,k\right)\right)\right\},$$

$X(l,k)$ – вершина, в якій значення функції мінімальне, тобто:

$$f\left(X\left[l,k\right]\right)=\min\left\{f\left(X\left(1,k\right)\right),\ldots,f\left(X\left(n+1,k\right)\right)\right\};$$

$X(n+1,k)$ – «центр ваги» всіх вершин за виключенням $X(h,k)$.

Координати «центру ваги» обчислюються за формулою

$$x_i\left(n+2,k\right)=\frac{1}{n}\left(\sum_{j=1}^{n+1}x_j\left(j,k\right)-x_i\left(h,k\right)\right),\ i=\overline{1,n},$$

де i – номер координати.

У методі Нелдера-Міда над многогранником виконуються такі операції (рис.21.1):

1. *Відбиття*, тобто проектування $X[h,k]$ через центр ваги відповідно із співвідношенням:

$$X\left(n+3,k\right)=X\left(n+2,k\right)+\alpha\left(X\left(n+2,k\right)-X\left(n,k\right)\right),$$

де $\alpha>0$ - коефіцієнт відбиття.

2. *Розтяг* застосовується у тому випадку, коли відбиття є вдалим, тобто коли

$$f(X\left(n+3,k\right))\le f(X\left(l,k\right)).$$

При цьому вектор $X\left(n+3,k\right)-X\left(n+2,k\right)$ розтягується і координати нової вершини визначаються як:

$$X\left[n+4,k\right]=X\left[n+2,k\right]+\gamma\left(X\left[n+3,k\right]-X\left[n+2,k\right]\right),$$

де $\gamma>1$ – коефіцієнт розтягу.

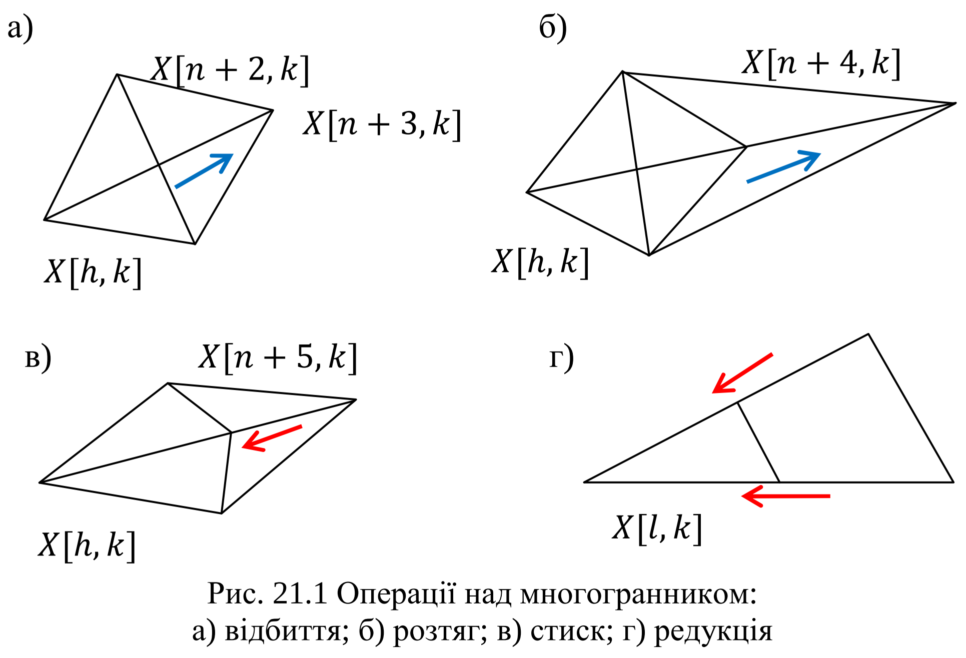


Рис. 21.1 Операції над многогранником:
а) відбиття; б) розтяг; в) стиск; г) редукція

3. *Стиск* виконується, якщо в результаті відбиття значення функції в точці $X[n+3,k]$, більше від значень у всіх інших вершинах симплексу за винятком вершини $X[h,k]$, тобто:

$$f\left(X\left[n+3,k\right]\right) > f\left(X\left[j,k\right]\right)$$

для всіх $j \neq h$.

Тоді вектор $X\left[h,k\right] - X[n+2,k]$ стискується так, що

$$X\left[n+5,k\right] = X\left[n+2,k\right] + \beta\left(X\left[h,k\right] - X\left[n+2,k\right]\right)$$

де $0 < \beta < 1$ – коефіцієнт стиску.

4. *Редукція*, тобто стиск симплексу у два рази відносно до вершини $X[l,k]$. Редукція застосовується у випадку, коли:

$$f\left(X\left[n+3,k\right]\right) > f(X[h,k])$$

виконується за формулою:

$$X\left[j,k\right] = X\left[l,k\right] + 0.5\left(X\left[j,k\right] - X\left[l,k\right]\right),\ j = \overline{1,n+1}.$$

Зауваження 21.1. Як правило, значення коефіцієнтів відбиття, розтягу та стиску треба приймати:

$$\alpha = 1; \quad 0.4 \le \beta \le 0.62; \quad 8 \le \gamma \le 3.0,$$

хоча Недлер і Мід рекомендують брати:

$$\alpha = 1; \quad \beta = 0.5; \quad \gamma = 2.$$

Умова припинення пошуку записується у вигляді:

$$\sigma = \left\{\frac{1}{n+1}\sum_{j=1}^{n+1}\left(f\left(X\left[j,k\right]\right) - f\left(X\left[n+2,k\right]\right)\right)^2\right\}^{0.5} \le \varepsilon \qquad (21.4)$$

де $\varepsilon > 0$ – мале число, яке визначає ε окіл пошуку екстремуму.

Алгоритм методу деформованого многогранника об’єднує такі кроки:

Крок 1. Установлення координат вершин початкового симплексу і обчислення в кожній із них функції $f(X)$.

Крок 2. Визначення вершин з найбільшим і найменшим значеннями $f(X)$. А також координат центру ваги симплексу: $X\left[h,k\right]$, $X\left[l,k\right]$, $x_i\left[n+2,k\right]$, $i = \overline{1,n}$.

Крок 3. Відбиття: обчислення координат відбитої вершини $x_i\left[n+2,k\right]$, $i = \overline{1,n}$ і значення функції $f(X[n+3,k])$.

Крок 4. Якщо $f\left(X\left[n+3,k\right]\right) > f(X[l,k])$, то обчислення продовжується з кроку 7.

Крок 5. Розтяг: обчислення координат $x_i\left[n+4,k\right]$, $i = \overline{1,n}$ і значення $f(X[n+4,k])$.

Крок 6. Приймається

$$X\left[h,k\right]=\begin{cases}X\left[n+4,k\right], \text{ якщо } f(X\left[n+4,k\right]<f(X[l,k]))\\ X\left[n+3,k\right], \text{ якщо } f(X\left[n+4,k\right]\geq f(X[l,k]))\end{cases}$$

і здійснюється перехід до кроку 12.

Крок 7. Перевірка умов $f\left(X\left[n+3,k\right]\right)>f(X[j,k])$ для всіх $j\neq k$. Якщо ця нерівність не виконується, то приймається $X\left[h,k\right]=X[n+3,k]$ і пошук продовжується з кроку 12.

Крок 8. Приймається

$$X\left[h,k\right]=\begin{cases}X\left[n+3,k\right], \text{ якщо } f(X\left[n+3,k\right]\geq f(X[h,k]))\\ X\left[h,k\right], \text{ якщо } f(X\left[n+3,k\right]<f(X[h,k]))\end{cases}.$$

Крок 9. Стиск: обчислення координат

$$x_i\left[n+5,k\right],\ i=\overline{1,n},\ f\left(X\left[n+5,k\right]\right).$$

Крок 10. Якщо $f(X[n+5,k])\leq f(X[h,k])$, то приймається $X\left[h,k\right]=X[n+5,k]$ і пошук проводиться з кроку 12.

Крок 11. Редукція.

Крок 12. Якщо умова припинення пошуку (21.4) не виконується, то індекс k збільшується на одиницю і пошук мінімуму повторюється з кроку 2.
Якщо умова (21.4) виконується, то усі значення функцій є дуже близькими одне до одного, тому вони, можливо, лежать поблизу точки мінімуму функцій $X[l,k]$. Виходячи з цього, такий критерій збіжності є розумним.

21.3. Прямі та непрямі методи умовної оптимізації

В загальному випадку методи поділяються на прямі та непрямі. Прямі методи оперують безпосередньо з вихідними задачами оптимізації і генерують послідовність точок $\{X[k]\}$ таких, що $f\left(X\left[k+1\right]\right)<f(X[k])$. В силу цього такі методи часто називають методами спуску. Математично перехід на певному k-кроці ($k=0,1,2,\ldots$) від точки $X[k]$ до точки $X[k+1]$ можна записати так:

$$X\left[k+1\right]=X\left[k\right]+\alpha_k P\left[k\right], \tag{21.5}$$

де $P[k]$– вектор, який визначає напрям спуску; α_k– довжина кроку вздовж даного напряму.

При цьому в одних алгоритмах прямих методів точки $X[k]$ вибираються так, щоб для них виконувалися всі обмеження задачі, в інших ці обмеження можуть порушуватися на деяких або всіх ітераціях.

Таким чином, в прямих методах при виборі напряму спуску обмеження, що визначають допустиму область G, враховують в явному вигляді.

Непрямі методи зводять вихідну задачу нелінійного програмування до послідовності задач безумовної оптимізації певних допоміжних функцій. При цьому обмеження вихідної задачі враховуються в явному вигляді.

Розглянемо деякі алгоритми прямих методів.

21.3.1. Прямі методи умовної оптимізації

21.3.1.1. Метод проекції градієнта

Нехай маємо таку загальну задачу: мінімізувати функцію $Z = f(X)$ за наявності m обмежень – нерівностей $g_i\left(X\right) \le 0,\ i = \overline{1,m}$ (обмеження $g_i(X) \le b_i$ можна записати так $g_i\left(X\right) - b_i \le 0$). За початкову вибирається деяка точка допустимої області G. Якщо $X[0]$ – внутрішня точка множини G, то розглядуваний метод є звичайним градієнтним методом:

$$X\left[k+1\right] = X\left[k\right] - \alpha_k \nabla f\left(X\left[k\right]\right),\ k = 0,1,2\ldots \tag{21.6}$$

де $\nabla f\left(X\left[k\right]\right)$– градієнт функції $f(X)$ в точці $X[k]$.

Після виходу на межу області G в якійсь точці $X[k]$ рух в напрямі антиградієнта – $\nabla f\left(X\left[k\right]\right)$ може вивести за межі допустимої множини (рис. 21.2). Тому антиградієнт проектується на лінійний многовид M, який апроксимує ділянку межі в околі точки $X[k]$. Рухаючись в напрямі проекції вектора $-\nabla f(X[k])$ на многовид M, відшукується нова точка $X[k+1]$, в якій $f(X[k+1]<f(X[k])$. $Xk+1$ приймається за вихідне наближення і процес продовжується.

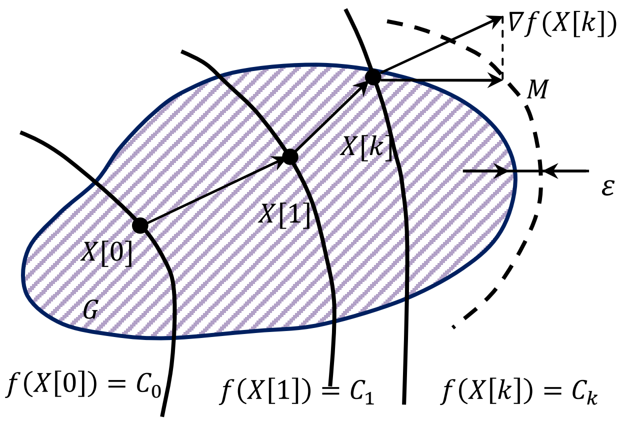


Рис. 21.2. Геометрична інтерпретація методу проекції градієнта

Розглянемо більш детальніше цю процедуру. В точці $X[k]$ частина обмежень – нерівностей задовольняються як рівності

$$g_i\left(X\right) = 0,\ \ i = \overline{1,l},\ \ l < m.$$

Такі обмеження називають активними. Позначимо через I множину індексів $i\ (1 \le i \le l)$ цих обмежень. Їх рівняння відповідають гіперповерхням, що утворюють межу області G в околі точки $X[k]$. В загальному випадку ця межа буде нелінійною. Обмеження $g_i\left(X\right)$, $i \in I$, апроксимуються гіперповерхням, які дотикаються до них у точці $X[k]$:

$$\sum_{j=1}^{n} \frac{\partial g_i\left(X\left[k\right]\right)}{\partial x_j}\left(x_j - x_j\left[k\right]\right) = 0,\ i \in I \tag{21.7}$$

Отримані гіперповерхні обмежують деякий многогранник M, який апроксимує допустиму область G в околі точки $X[k]$. Проекція $p\left[k\right]$ антиградієнта $-\nabla f(X[k])$ на многогранник M обчислюється за формулою:

$$P\left[k\right] = P[-\nabla f(X[k])] = 0,\ i \in I \tag{21.8}$$

де P – оператор ортогонального проектування, який визначається виразом:

$$P = E - A^T\left(AA^T\right)^{-1}A,$$

де E – одинична матриця розміру $n \times n$.

A – матриця розміру $l \times n$ – якобіан функції $g_i(X)$, який обчислений у точці $X[k]$

$$A = \left[\alpha_1,\ldots,\alpha_i,\ldots,\alpha_l\right]^T,$$

де $\alpha_i = \left(\dfrac{\partial g_i\left(X\left[k\right]\right)}{\partial x_1},\ldots,\dfrac{\partial g_i\left(X\left[k\right]\right)}{\partial n}\right),\quad i = \overline{1,l}.$

Далі здійснюється спуск у обраному напрямі

$$X\left[k+1\right] = X\left[k\right] + \alpha_k p[k].$$

Можна показати, що точка $X[k+1]$, буде розв’язком задачі мінімізації функції $f(X)$ в області G тоді і тільки тоді, коли $P\left[-\nabla f\left(X\left[k\right]\right)\right] = 0$, тобто

$$-\nabla f\left(X\left[k\right]\right) = \sum_{i=1}^{l} u_i\alpha_i \text{ та } U = \left(A^TA\right)^{-1}A^T\left(-f\left(X\left[k\right]\right)\right) \geq 0.$$

Ця умова означає, що антиградієнт $-\nabla f(X[k])$ цільової функції є лінійною комбінацією з невід’ємними коефіцієнтами градієнтів обмежень $g_i\left(X\right) = 0$.

Відповідно до викладеного вище алгоритм методу проекції градієнта складається з таких дій:

1. У точці $X[k]$ визначається напрям спуску $p[k]$.

Оскільки точка $X[k] \in G$ і відомі активні обмеження $g_i\left(X\right) = 0 i \in I$, то обчислюються $-\nabla f(X[k])$ і визначається проекція $P\left[-\nabla f\left(X\left[k\right]\right)\right]$. При цьому можливі два випадки:

а) $P\left[-\nabla f\left(X\left[k\right]\right)\right] \neq 0$, тоді за напрям спуску приймається отримана проекція,

б) $P\left[-\nabla f\left(X\left[k\right]\right)\right] = 0$, тобто

$$\nabla f\left(X\left[k\right]\right) = \sum_{i=1}^{l} u_i \alpha_i .$$

Якщо усі $u_i \geq 0, i \in I$, то точка $X[k]$ буде розв'язком задачі. Якщо ж деяка з компонент $u_q < 0$, то q-й градієнт вилучається з матриці A і породжується нова матриця проектування P, яка визначає новий напрям спуску.

1. Для визначення величини кроку α_k цільова функція мінімізується у напрямі $p\left[k\right]$ за умов виконання обмежень задачі із заданою точністю, що задається певним додатнім числом ε. Вважається, що точка X задовольняє умовам задачі із заданою точністю, якщо $g_i\left(X\right) = \varepsilon,\ i = \overline{1,m}$. Величина кроку α_k визначається розв'язком задачі вигляду:

$$f\left(X\left[k\right] + \alpha p\left[k\right]\right) \to \min ,$$

$$g_i\left(X\left[k\right] + \alpha p\left[k\right]\right) \leq \varepsilon, \quad i = \overline{1,m} .$$

2. Наступна точка обчислюється за формулою

$$X\left[k+1\right] = X\left[k\right] + \alpha p\left[k\right].$$

Ознакою збіжності є наближення до нуля векторів $p[k]$.

Розглянутий метод є аналогом градієнтних методів для розв'язання задач на безумовний екстремум, тому він також поволі збіжний.

21.3.1.2. Комплексний метод Бокса

Метод Бокса є модифікацією методу деформованого многогранника, що дозволяє врахувати обмеження нерівності. Задачі, що розв'язується, полягає в мінімізації функції

$$Z = f\left(X\right),$$

де $X = \left(x_1, \ldots, x_n\right)$ визначається явними обмеженнями

$$\alpha_j \leq x_j \leq b_j,\ j = \overline{1,n},$$

а також неявними обмеженнями

$$g_i\left(X\right) \le 0,\ i = \overline{1,m}.$$

Якщо цільова функція задачі опукла і функції $g_i\left(X\right)$ також опуклі, то задача має єдиний розв'язок. Значення α_j та b_j є нижньою та верхньою межами змінних.

Вибираємо q точок, які задовольняють обмеженням задачі, а також обчислимо в цих точках цільову функцію. Побудуємо многогранник, який має $q > n+1$ вершин (Бокс знайшов, що $q = 2n$). Цей многогранник називається комплексом.

Введемо такі позначення:

$$X\left[i,k\right] = \left(x_1\left[i,k\right],\ldots,x_j\left[i,k\right],\ldots,x_n\left[i,k\right]\right),$$

де $i = \overline{1,q}, \quad k = 0,1,2\ldots$– i-а вершина комплексу на k-му кроці,

$X[h,k]$– вершина, в якій значення цільової функції максимальне.

Тобто

$$f\left(X\left[h,k\right]\right) = \max\left\{X\left[1,k\right],\ldots,f\left(X\left[q,k\right]\right)\right\},$$

$X[k,k]$– центр ваги усіх вершин за виключенням $X\left[h,k\right]$.

Координати центру ваги обчислюються за формулою

$$x_j\left[l,k\right] = \frac{1}{q}\left(\sum_{i=1}^{q} x_j\left[i,k\right] - x_j\left[h,k\right]\right), \quad j = \overline{1,n}\,. \tag{21.9}$$

За першу вершину початкового комплексу вибирається деяка допустима точка $X[1,0]$. Координати інших $q-1$ вершин комплексу визначаються співвідношенням

$$x_j\left[i,0\right] = \alpha_j + r_j\left(b_j - \alpha_j\right) \tag{21.10}$$

для $j = \overline{1,n}, \quad i = 2,q, r_j$- псевдовипадкові числа, які рівномірно розподілені в проміжку$[0,1]$. Точки, вибрані з цим для даного j, будуть автоматично задовольняти обмеження $\alpha_j \le x_j \le b_j$, але обмеження $g_i(X) \le 0$ можуть бути порушені. В цьому випадку недопустима точка замінюється новою, такою, що лежить в середині відрізка. Який сполучає точку з центром ваги вибраних допустимих вершин, тобто формується точка

$$X\left[i,k\right] = \frac{X\left[i,k\right] + X\left[l,k\right]}{2} \tag{21.11}$$

Ця процедура може повторюватися доти, поки точка не стане допустимою. Якщо функції $g_i(X)$ опуклі, то, врешті-решт, обмеження виконуватимуться і комплекс утворюватиметься допустимими точками.

Розглянемо ітераційну процедуру комплексного методу, в якій здійснюється відшукання мінімуму переміщення за напрямом до мінімуму всередині області обмежень. Тут треба виконати такі кроки:

1. Знаходимо точку з найбільшим значенням функції $X[h,k]$ та центр ваги $(q-1)$ точок, що залишилися.
2. Як і в методі деформованого многогранника, замінюємо вершину $X\left[h,k\right]$ шляхом відображення відносно центру ваги $X[l,k]$, обчислюється за формулою

$$X\left[p,k\right]=\left(\alpha+1\right)X\left[l,k\right]+\alpha X\left[h,k\right],$$

де $\alpha>0$ – деяка константа, яку називають коефіцієнтом відображення. Найбільш задовільні результати значення $\alpha=1.3$.

3. Перевіряємо, чи буде отримана точка $X[p,k]$ допустимою.

а) якщо $X\left[p,k\right]$ не буде допустимою та не виконується обмеження для α_j, то покладемо $x_j\left[p,k\right]=\alpha_j+10^{-6}$, якщо ж не виконується обмеження для b_j, то покладемо

$$x_j\left[p,k\right]=b_j-10^{-6};$$

б) якщо обмеження не виконуються, то точку $X[p,k]$ пересувають на половину відстані між $X[p,k]$ та $X[l,k]$, тобто

$$X\left[p,k\right]\left(\text{нове}\right)=\frac{X\left[p,k\right]+X\left[l,k\right]}{2}.$$

Ця процедура може повторюватися декілька разів, поки не буде отримана допустима точка.

4. Якщо точка $X\left[p,k\right]$ допустима, то обчислюється значення функції в цій точці $f\left(X\left[p,k\right]\right)$ і порівнюється з найбільшим значення функції $f\left(X\left[h,k\right]\right)$. Якщо $f\left(X\left[p,k\right]\right)>f(X[h,k])$, то точка $X[p,k]$ зсувається до центру $X[l,k]$ на половину відстані між ними, тобто

$$X\left[p,k\right]\left(\text{нове}\right)=\frac{X\left[p,k\right]+X\left[l,k\right]}{2}$$

і процес повертається до кроку 3.

5. Якщо $f\left(X\left[p,k\right]\right)<f(X[h,k])$, то точка $X[h,k]$ замінюється точкою $X[p,k]$ і точки значення функцій комплексу знову упорядковуються.
6. Критерій завершення: обчислення закінчуються, коли значення цільової функції мало змінюється протягом п’яти послідовних ітерацій, тобто

$$\left|f\left(X\left[l,k+1\right]\right)-f\left(X\left[l,k\right]\right)\right|\le\varepsilon,\quad k=1,2,\ldots,5,$$

де ε – задана константа.

В цьому випадку середнє арифметичне п'яти останніх ітерацій вважають розв'язком задачі нелінійного програмування. В протилежному випадку треба повернутися до кроку 1 та повторити процедуру.

Комплексний метод можна застосовувати до широкого кола задач з обмеженнями. Він простий та надійний у роботі. На кожному кроці метод використовує тільки інформацію про значення цільової функції та функцій обмежень задачі. Якщо цільова функція опукла і, крім того, опукла область обмежень, то застосування методу буде успішним, хоча певні особливості задачі можуть вимагати певної модифікації критерію завершення.

21.3.2. Непрямі методи умовної оптимізації (методи штрафних функцій)

Методи штрафних функцій відносяться до непрямих методів розв'язування ЗНП. Методи полягають у зведенні задачі математичного програмування до однієї чи кількох задач безумовної оптимізації деяких спеціальним способом побудованих допоміжних функцій. У загальному вигляді допоміжну функцію можна записати так:

$$F\left(X,\alpha\right)=f\left(X\right)+\Phi\left(X,\alpha\right),$$

де $f(X)$ – мінімізуюча цільова функція задачі,

$\Phi\left(X,\alpha\right)$ – "штрафна" функція, параметр $\alpha>0$.

Точку безумовного мінімуму функції $F\left(X,\alpha\right)$ позначимо через $X(\alpha)$.

Залежно від вигляду $\Phi\left(X,\alpha\right)$ розрізняють методи внутрішніх або зовнішніх штрафних функцій, які інколи називають «бар'єрними» функціями.

21.3.2.1. Методи внутрішніх штрафних функцій

Методи внутрішніх штрафних функцій застосовуються при розв'язуванні ЗНП з обмеженнями – нерівностями типу $g_i\left(X\right)\le 0, i=\overline{1,m}$. Функції $\Phi\left(X,\alpha\right)$ підбирають так, щоб їх значення необмежено зростало при наближенні до межі допустимої області G (рис. 21.3.).

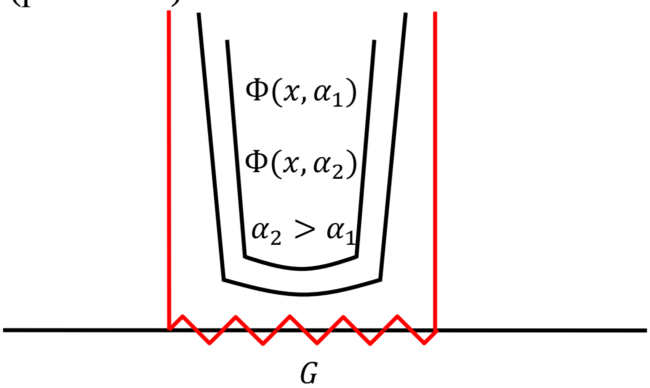


Рис. 21.3. Геометрична інтерпретація методу внутрішніх штрафних функцій

Наближення до межі "штрафується" різким збільшенням значення функції $F(X,\alpha)$. На межі G побудовано "бар'єр", який перешкоджає порушенню обмежень у процесі безумовної оптимізації $F(X,\alpha)$. Пошук мінімуму допоміжної функції треба починати із внутрішньої точки області G. Тоді в процесі оптимізації траєкторія спуску ніколи не вийде за межі допустимої області. Усі перелічені особливості функції $\Phi(X,\alpha)$ визначили найменування розглядуваної групи методів.

Отже, внутрішня штрафна функція $\Phi(X,\alpha)$ може бути визначена так:

$$\Phi\left(X,\alpha\right)=\begin{cases}\infty, & \text{якщо}\, X \overline{\in} G,\\ 0, & \text{якщо}\, X \in G,\ X \overline{\in} \partial G,\ \alpha \to 0,\\ \infty, & \text{якщо}\, X \in G,\ X \to \partial G.\end{cases} \quad (21.12)$$

де ∂G - межа області G.

Загальний вигляд внутрішньої штрафної функції

$$\Phi\left(X,\alpha\right)=\alpha\sum_{i=1}^{m}\varphi_i(g_i(X)), \quad (21.13)$$

де φ_i– неперервні диференційовані функції. Які визначаються обмеженнями-нерівностями вихідної ЗНП.

Допоміжна функція $F(X,\alpha)$ при цьому має вигляд:

$$F\left(X,\alpha\right)=f\left(X\right)+\alpha\sum_{i=1}^{m}\varphi_i(g_i(X)). \quad (21.14)$$

Вона визначена в області G і необмежено зростає, коли $g_i(X)\to 0$ для деякого i. Прикладами внутрішніх штрафних функцій для задач мінімізації є такі:

$$\Phi\left(X,\alpha\right)=\alpha\sum_{i=1}^{m}\frac{1}{g_i\left(X\right)};\ \Phi\left(X,\alpha\right)=-\alpha\sum_{i=1}^{m}\ln[g_i(X)].$$

Алгоритм методу такий. За початкову точку вибираємо довільну внутрішню точку області G. Задаємо деяку монотонно спадну збіжну до нуля послідовність $\left\{a_k\right\},k=1,2,\ldots$ додатних чисел. Для першого елемента α_i цієї послідовності розв’язуємо задачу безумовної мінімізації функції $F(X,\alpha_1)$, у результаті чого визначаємо точку $X(\alpha_1)$. Ця точка використовується за початкову для розв’язування задачі пошуку мінімуму функції $F(X,\alpha_2)$, де $\alpha_2\le\alpha_1$ і т.д. Отже, розв’язуємо послідовність задач безумовної мінімізації функції $F\left(X,\alpha_k\right),k=1,2,\ldots,$ причому розв’язок попередньої задачі $X\left(\alpha_k\right)$ використовуємо за початкову точку для пошуку наступного вектора $X(\alpha_{k+1})$. Послідовність точок $X(\alpha_k)$ збігається до оптимального розв’язку вихідної задачі – локального мінімуму X^*.

Обчислення припиняються, коли

$$\left|f\left(X\left[k\right]\right)-f\left(X\left[k+1\right]\right)\right|\le\varepsilon,$$

$$\left\|X\left[k\right]-X\left[k+1\right]\right\|\le\beta,$$

де ε,β – задані числа, які визначають точність обчислень.

Розглянутий метод внутрішніх штрафних функцій має такі властивості

1) $\lim\limits_{k\to\infty}\alpha_k\sum\limits_{i=1}^{m}\varphi_i\left(g_i\left(X\left[k\right]\right)\right)=0;$

2) $\lim\limits_{k\to\infty}f\left(X\left[k\right]\right)=f\left(X^*\right)$, $f\left(X\left[k\right]\right)$ – монотонно спадає;

3) $\lim\limits_{k\to\infty}F\left(X\left[k\right],\ \alpha_k\right)=f(X^*)$.

Ці властивості є вірними для задач, які мають неперервні функції і локальні мінімуми всередині області G.

21.3.2.2. Методи зовнішніх штрафних функцій

Функції $\Phi(X,\alpha)$ вибирають так, що їх значення дорівнює нулю всередині і на межі допустимої області G, а поза нею – додатні і зростають тим більше, чим сильніше порушуються обмеження (рис. 21.4.). Отже, тут «штрафується» віддалення від допустимої області. Зовнішня графічна функція $\Phi(X,\alpha)$ в загальному вигляді можна визначити так:

$$\Phi\left(X,\alpha\right)=\begin{cases}0, & \text{якщо}\, X\in G\ ,\\ \infty, & \text{якщо}\, X\,\overline{\in}\, G,\ \alpha\to\infty.\end{cases}$$

Пошук мінімуму допоміжної функції $F\left(X,\alpha\right)$ починається із довільної точки.

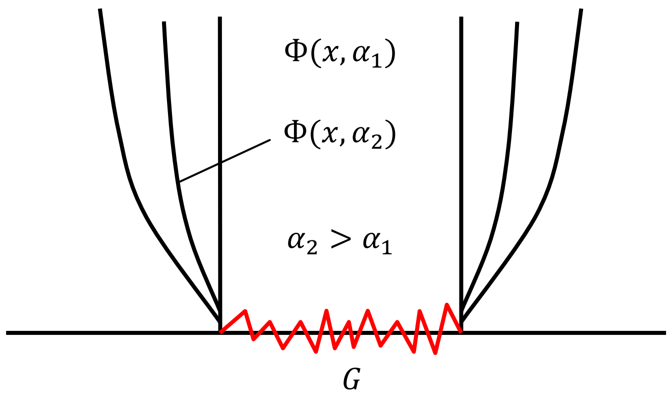


Рис. 21.4. Геометрична інтерпретація методу зовнішніх штрафних функцій

У більшості випадків вона є недопустимою, тому траєкторія спуску розміщується частково за допустимою областю. Якщо мінімум цільової функції знаходиться на межі допустимої області. То ця траєкторія повністю лежить за областю G. Перелічені особливості функції $\Phi(X,\alpha)$ визначили назву даної групи методів. Загальний вигляд зовнішньої функції

$$\Phi(X,\alpha) = \alpha\left(\sum_{i=1}^{m}\varphi_i\left(g_i\left(X\right)\right)\right) \qquad (21.15)$$

де φ_i – функція, яка визначається обмеженнями-нерівностями ЗНП

Допоміжна функція $F(X,\alpha)$ при цьому така:

$$F(X,\alpha) = f(X) + \alpha\left(\sum_{i=1}^{m}\varphi_i\left(g_i\left(X\right)\right)\right). \qquad (21.16)$$

Прикладами зовнішніх штрафних функцій можуть бути такі:

$$\Phi(X,\alpha) = \left\{\max\left[0, g_i,(X)\right]\right\}^2, \quad \Phi(X,\alpha) = \alpha\exp\left\{\frac{1}{\alpha}\sum_{i=1}^{m}\max\left[0, g_i\left(X\right)\right]\right\}.$$

Тут

$$\max\left[0, g_i,(X)\right] = \begin{cases} 0, & \text{якщо } g_i(X) \le 0, \\ g_i(X), & \text{якщо } g_i(X) > 0. \end{cases}$$

Алгоритм методу зовнішніх штрафних функцій формулюється також, як і алгоритм внутрішніх штрафних функцій і має аналогічні властивості. Однак в цьому випадку не потрібно, щоб точка $x[0] \in G$, а послідовність $\left\{\alpha_k\right\}, k = 1, 2, \ldots,$ додатних чисел була монотонно зростаючою.

Аналіз методів штрафних функцій дає такі висновки. Методи внутрішніх функцій ведуть пошук розв'язку, не виходячи за межі допустимої області. Це важливо, коли цільова функція чи обмеження не визначені за границями допустимої множини. Крім того, припиняючи обчислення, ми завжди отримуємо допустимий розв'язок. Але для визначення будь-якої початкової точки іноді треба розв'язати задачу, яка порівняна за складністю із вихідною задачею. Методи зовнішніх штрафних функцій забезпечують розв'язування із довільної точки. Загальним недоліком методів є складність допоміжної функції, яка може мати яристу структуру. Степінь яристості збільшується із збільшенням α. При великих значеннях α точність обчислень мінімуму $F(X,\alpha)$ значно зменшується через помилки округлення.

Список літератури

Основна література

1. Зайченко Ю.П. Исследование операций: Учебник. – 6 изд., перераб. и доп. – Киев: Издательский Дом «Слово»., 2003.- 688 с.
2. Зайченко О.Ю., Зайченко Ю. П. Дослідження операцій. Збірник задач. – К.: Видавничий Дім “Слово”, 2007.- 472 с.
3. Ларіонов Ю.І., Левикін В.М., Хажмурадов М.А. Дослідження операцій в інформаційних системах.-Харків.: Компанія СМІТ, 2005.-364 с.
4. Костевич Л.С. Математическое программирование: Информ. Технологии оптимальных решений: Учеб. Пособие.- Минск.: Новое знание, 2003.-424 с.
5. Томашевський В.М. Моделювання систем. Підручник. -К.: Видавнича група BHV, 2007.- 352 с.
6. Глоба Л.С. Математичні основи побудови інформаційно-телекомунікаційних систем.-К.: Норіта-плюс, 2007.-360 с.
7. Ільченко М.Ю.. Кравчук С.О. Сучасні телекомунікаційні системи.-К.: НВП «Видавництво «Наукова думка» НАН України», 2008.- 328 с.
8. Згуровський М.З., Панкратова Н.Д. Основи системного аналізу.-К.: Видавнича група BHV, 2007.-544с.
9. Бунин С.Г., Войтер А.П., Ильченко М.Е., Романюк В.А. Самоорганизующиеся радиосети сосверхширополосными радиосигналами. –К.: НПП «Издательство «Науква думка» НАН Украины», 2012. – 444 с.
10. Бакулин М.Г., Варукина Л.А., Крейнделин В.Б. Технология MIMO: принципы и алгоритмы. – М.: Горячая линия – Телеком, 2014. – 244 с.
11. Самсонов В.В. Алгоритми розв’язання задач оптимізації: Навчальний посібник. К.: НУХТ, 2014.-300 с.

Додаткова література

12. Таха Хемди А. Введение в исследование операций.-М.: издательский дом “Вильямс”, 2005.-912с.
13. ВолковИ.К., Загоруйко Е.А. Исследование операций: Учеб. для вузов.-М.: Изд-во МГТУ им.Н.Э. Баумана, 2002.- 436 с.
14. Вентцель Е.С. Исследование операций.-М.:”Советское радио”, 1972.-552с.
15. Дьяконов В.П. MATLAB 6.5 SP1/7 + Simulink 5/6. Основы применения.-М.: СОЛОН-Пресс, 2005.- 800 с.
16. Мэтьюз Д.Г.,Финк К.Д. Численные методы. Использование MATLAB.-М.: Издательский дом «Вильямс», 2001.- 720 с.
17. Корн Г., Корн Т. Справочник по математике для научных работников и инженеров. - М.:”Наука”, 1984.-832с.
18. Ануфриев И.Е.,Смирнов А.Б., Смирнова Е.Н. MATLAB 7.-СПб.: БХВ-Петербург, 2005.- 1104 с.
19. Черных И.В. Моделирование электротехнических устройств в MATLAB, SimPowerSystems & Simulink. – М.: ДМК Пресс. СПб.: Питер, 2008. – 288 с.

20. Герман-Галкин С.Г. Компьютерное моделирование полупрводникових систем в MATLAB 6.0: Учебное паособие. – СПб.: КОРОНА принт, 2001. -320 с.
21. Цисарь И.Ф. MATLAB Simulink. Компьютерное моделирование в экономике . – М.: СОЛОН-ПРЕССС, 2008.- 256 .
22. Дьяконов В.П. Simulink 5/6/7: Самоучитель. – М.: ДМК-Пресс, 2008.- 784 с.
23. Бенькович Е.С., Колесов Ю.Б., Сениченков Ю.Б. Практическое моделирование динамических систем.-СПб.: БХВ-Петербург, 2002. - 464 с.
24. Калюжний О.Я. Моделювання систем передачі сигналів в обчислювальному середовищі MATLAB-Simulink: Навч. Посібник. –К.: ІВЦ «Видавництво «Політехніка»», 2004. - 136 с.
25. http://www.uuooi.org/english/viewforum.php?f=289&topicdays=0&sort=0&order=0&sort2=0&order2=0&start=12