МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ

НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ

«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ»

МАТЕМАТИЧНІ МЕТОДИ ОПТИМІЗАЦІЇ

МЕТОДИЧНІ ВКАЗІВКИ

ДО ВИКОНАННЯ ПРАКТИЧНИХ ЗАНЯТЬ

для магістрів спеціальності

«Обладнання лісового комплексу»

Рекомендовано Вченою радою інженерно-хімічного факультету

Київ

НТУУ «КПІ»

2016

Математичні методи оптимізації. : Метод. вказівки до виконання практичних занять для магістрів спеціальності «Обладнання лісового комплексу», / Уклад.: Л. Р. Ладієва – К. : НТУУ «КПІ», 2016. – 62 с.

Гриф надано Вченою радою ІХФ
(Протокол № 3 від 25 квітня 2016 р.)

Навчальне завдання

Математичні методи оптимізації
Методичні вказівки до виконання практичних занять для магістрів спеціальності «Обладнання лісового комплексу»

Укладач: Ладієва Леся Ростиславівна, канд. техн. наук, доц.

Відповідальний
редактор А. І. Жученко, докт.техн. наук, проф.

Рецензент А.Р. Степанюк, к.т.н.,доц.

Авторська редакція

Вступ

Методичні вказівки присвячені питанням практичного використання методів статичної оптимізації для практичних робіт. Основна увага приділяється методам і алгоритмам, що використовуються при проектуванні, керуванні і аналізі функціонування технологічних об'єктів. Розглядаються методи дослідження функцій класичного аналізу, методи лінійного програмування, як розділу дослідження операцій , які знайшли застосування при вирішенні задач, що пов'язані з ефективним використанням обмежених ресурсів. Методи нелінійного програмування орієнтовані на вирішення задач з обмеженнями.

Мета методичних вказівок для практичних робіт допомогти студентам освоїти розділи статичної оптимізації і застосувати знання на прикладах, представлених у формалізованому вигляді.

1. Постановка задач оптимізації. Приклади моделей процесів.

1.1. Теплообмінники типу «змішування — змішування»

Математична модель

Як приклади математичних моделей теплообмінних апаратів нижче проаналізовані моделі теплообмінників простіших типів, в яких здійснюється передача тепла між двома потоками — теплоносієм і хладоагентом. У всіх математичних описах передбачається, що рух потоків теплоносія і хладоагента характеризується простими гідродинамічними моделями «ідеальне змішування» і «ідеальне витіснення». Крім того, допускається, що коефіцієнт теплопередачі через стінку, що розділяє теплоносій і хладоагент, є постійною заданою величиною, яка не залежить від їх об'ємних витрат. Останнє допущення, строго кажучи, неточне [1], проте воно прийняте в майбутньому для спрощення математичних викладень при вирішенні завдань оптимізації.

Теплообмінник типу «змішування — змішування» (рис. 1.1). Математичний опис теплообмінника в даному випадку задають системою рівнянь типу

$$\frac{d(Vc_pT)}{dt} = \upsilon^{(0)} c_{p0} T^{(0)} - \upsilon c_p T + V q_T \,,$$

де V - об’єм зони ідеального змішування, м^3; C_p - теплоємність потоку теплоносія, $\frac{Дж}{м^3 k}$; V - витрати теплоносія, $\frac{кг}{с}$;

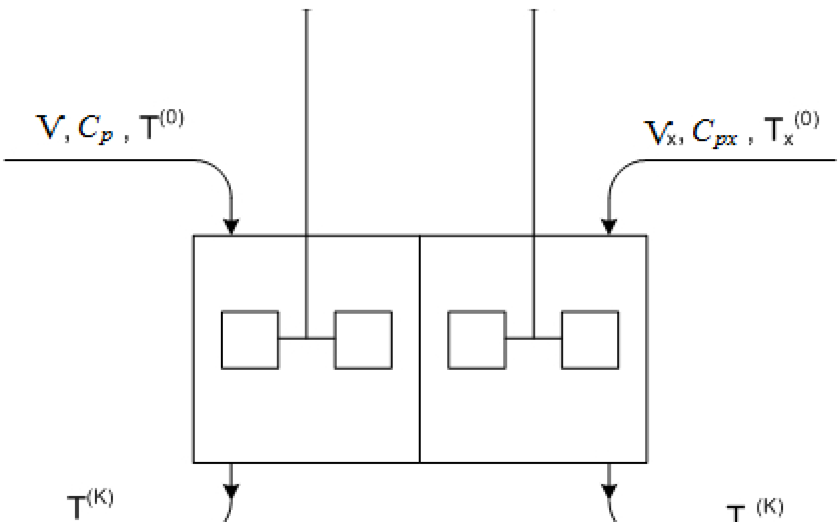

Рис. 1.1. Схематичне зображення теплообмінника типу „змішування — змішування“.

що відносяться до обох теплоносіїв. Інтенсивність джерела тепла при цьому визначається співвідношенням

$$Q_T = F_T K_T \left(T^* - T\right),$$

де K_T - коефіцієнт теплопередачі, $\frac{Вт}{м^2 K}$; F_T - поверхня теплообміну, м^2.

Стаціонарний режим теплообмінника можна описати нестаціонарними рівняннями, в яких похідні за часом вважаються рівними нулю. В результаті буде справедлива наступна система рівнянь:

$$\upsilon c_p \left(T^{(0)} - T^{(k)}\right) - K_T F \left(T^{(k)} - T_x^{(k)}\right) = 0$$

$$\upsilon_x c_{px} \left(T_x^{(0)} - T_x^{(k)}\right) + K_T F \left(T^{(k)} - T_x^{(k)}\right) = 0$$

яка може бути вирішена відносно будь-яких двох параметрів, що входять в ці рівняння. Зокрема, для вихідних температур теплоносія $T^{(k)}$ і хладоагента $T_x^{(k)}$ можна отримати вирази:

$$T^{(k)} = T^{(0)} \frac{1 + K_T F\left(\dfrac{1}{\upsilon_x c_{px}} + \dfrac{T_x^{(0)}}{T^{(0)}} \cdot \dfrac{1}{\upsilon c_p}\right)}{1 + K_T F\left(\dfrac{1}{\upsilon_x c_{px}} + \dfrac{1}{\upsilon c_p}\right)}$$

$$T_x^{(k)} = T_x^{(0)} \frac{1 + K_T F\left(\dfrac{T^{(0)}}{T_x^{(0)}} \cdot \dfrac{1}{\upsilon_x c_{px}} + \dfrac{1}{\upsilon c_p}\right)}{1 + K_T F\left(\dfrac{1}{\upsilon_x c_{px}} + \dfrac{1}{\upsilon c_p}\right)}$$

При розрахунках теплообмінних апаратів зазвичай задають теплове навантаження на теплообмінник Q, тобто кількість тепла, яке потрібно передати від теплоносія хладоагенту в одиницю часу. Для певної витрати теплоносія, відомої його теплоємності і заданої вхідної температури це по суті еквівалентний завданню необхідної вихідної температури теплоносія, оскільки

$$Q = \upsilon c_p \left(T^{(0)} - T^{(k)}\right)$$

Завдання розрахунку теплообмінника ставиться як завдання визначення поверхні теплообміну F. В даному випадку обчислюють також вихідну температуру хладоагента $T_x^{(k)}$, яку можна знайти із загального балансу тепла, не залежного від типа теплообмінника

$$\upsilon c_p \left(T^{(0)} - T^{(k)}\right) + \upsilon_x c_{px} \left(T_x^{(0)} - T_x^{(k)}\right) = 0$$

Звідки

$$T_x^{(k)} = T_x^{(0)} + \left(T^{(0)} - T^{(k)}\right)\frac{\upsilon c_p}{\upsilon_x c_{px}}$$

Підставляючи в рівняння або вираження для вхідної температури хладоагента і вирішуючи отримане рівняння відносно поверхні теплообміну, знайдемо:

$$F = \frac{\upsilon_x c_{px}}{K_T\left(\nu \dfrac{\upsilon_x c_{px}}{\upsilon c_p} - 1\right)}$$

Де

$$\nu = \frac{T^{(k)} - T_x^{(0)}}{T^{(0)} - T^{(k)}}$$

Знаменник виразу може перетворюватися на нуль або набувати негативних значень. При цьому поверхня теплообміну, визначена вказаним рівнянням, дорівнює нескінченності або від’ємна, що в обох випадках означає неможливість створення теплообмінника на задане теплове навантаження при прийнятих параметрах хладоагента. Тому умова додатності знаменника у виразі можна розглядати як умова фізичної реалізації теплообмінника при заданих параметрах теплоносія і хладоагента:

$$\nu \frac{\upsilon_x c_{px}}{\upsilon c_p} > 1.$$

Оптимізація теплообмінних апаратів

Основним питанням, яке доводиться розглядати при проектуванні теплообмінних апаратів, є вибір поверхності теплообміну F і відповідного нею навантаження по хладоагенту υ_x для заданого теплового навантаження на теплообмінник Q. У зв'язку з цим для економічної оцінки ефективності теплообмінника заданої конструкції може бути використаний критерій оптимальності, об'єднуючий параметри F і υ_x, котрий може бути записаний у вигляді співвідношення:

$$R = s_x \upsilon_x + s_F F$$

у якому s_x — вартість одиниці об'єму хладоагента; s_F — вартість одиниці поверхні теплообміну, що обчислюється з врахуванням амортизації теплообмінника.

Критерій є сумарними витратами на експлуатацію теплообмінника в одиницю часу. Природньо, що зв'язок цих витрат з величиною поверхні теплообміну має значно складніший характер, проте в першому наближенні її все ж можна застосовувати у формі виразу для ілюстрації загальної методики оптимізації теплообмінних апаратів.

Щоб отримати можливість використовувати критерій оптимальності для вибору оптимальної поверхні теплообміну $F_{опт}$ і оптимальної витрати хладоагента $\upsilon_{x,опт}$, необхідно знайти зв'язок між F і υ_x яку дають рівняння математичного опису теплообмінника. Зрозуміло, що на вибір оптимальних значень F і υ_x при цьому робить вплив тип використаного теплообмінного апарату. Тому при рішенні задачі оптимального проектування потрібно також розглянути можливі варіанти реалізації теплообміну, тобто по суті оцінити економічну ефективність від використання того або іншого варіанту теплообмінника.

Для різних теплообмінників може бути отримана залежність поверхні теплообмена F від витрати хладоагента для заданого теплового навантаження на теплообмінник, що характеризується параметрами теплоносія— температурою, теплоємністю і витратою, тобто

$F=F(V_X)$.

Скориставшись даною залежністю в загальному вигляді, з виразу можна визначити необхідні умови оптимальності. Для цього продиференціюємо критерій по v_x і прирівняємо отриманий вираз нулю. В результаті отримаємо рівняння

$$s_x + s_F \frac{dF}{d\upsilon_x} = 0$$

якому повинні задовольняти оптимальні значення F і υ_x.

Оскільки в даному випадку використовується залежність, фактично визначенню підлягає лише значення $\upsilon_{x,onm}$, вживане потім для розрахунку оптимальної поверхні теплообміну.

Розглянемо деякі частинні завдання оптимізації теплообмінників, математичні описи, яких наведені вище.

Теплообмінник типу «змішування — змішування». Для теплообмінника цього типу була отримана залежність поверхні теплообміну від навантаження по хладоагенту у формі виразу. В результаті диференціювання цього виразу по витраті хладоагента знаходимо:

$$\frac{dF}{d\upsilon_x} = -\frac{c_{px}}{K_T} \frac{1}{\left(v \frac{\upsilon_x c_{px}}{\upsilon c_{px}} - 1 \right)^2}$$

Підстановка виразу в рівняння дає:

$$\left(v \frac{\upsilon_x c_{px}}{\upsilon c_p} - 1 \right)^2 - \frac{s_p c_{px}}{s_x K_T} = 0$$

Рівнянням є умова перетворення на нуль похідної від критерію оптимальності. Вирішення цього рівняння дозволяє отримати два значення навантаження по хладоагенту:

$$\upsilon_{x1} = \frac{\upsilon c_p}{v c_{px}} \left[1 - \left(\frac{s_F c_{px}}{s_x K_T} \right)^{0,5} \right]$$

$$\upsilon_{x2} = \frac{\upsilon c_p}{v c_{px}} \left[1 + \left(\frac{s_F c_{px}}{s_x K_T} \right)^{0,5} \right]$$

Крім того, згідно з правилами дослідження функції однієї змінної, оптимальне значення слід шукати також в тих точках, де похідна від критерію оптимальності не існує. У нашому випадку є така точка, в якій похідна $dF/d\upsilon_x$ не існує і яка відповідає перетворенню на нуль знаменника виразу:

$$\upsilon_{x3} = \frac{\upsilon c_p}{\nu c_{px}}$$

Таким чином, є сукупність трьох значень υ_x, кожне з яких має бути перевірене на екстремум функції R. Аналіз достатніх умов в даному випадку доцільно провести, виходячи з фізичного сенсу вирішуваного завдання. З цією метою розглянемо вираз для поверхні теплообміну:

$$F = \frac{\upsilon_x c_{px}}{K_T\left(\nu \frac{\upsilon_x c_{px}}{\upsilon c_p} - 1\right)}$$

Очевидно, що фізичний сенс може мати лише додатнє значення величини поверхні F. Це відповідає виконанню умови:

$$\upsilon_x > \frac{\upsilon c_p}{\nu c_{px}}$$

Умові задовольняє лише значення υ_{x1} визначуване формулою. Залишається перевірити, чи при даному значенні υ_{x1} досягатиметься мінімальне значення критерію оптимальності.

1.2. Реактори ідеального змішування

Математична модель

Математичний опис даного реактора можна отримати із загальних рівнянь гідродинаміки потоку для випадку ідеального змішення

$$\frac{d(Vx_j)}{dt} = \upsilon(x_i^{(0)}) - \upsilon x_i + Vq_i$$

$$\frac{d(Vc_pT)}{dt} = \upsilon(c_{p0}T^{(0)}) - \upsilon c_pT + Vq_T\,,$$

де $x_i^{(0)}$ - вхідні концентрації реагентів, $T^{(0)}$ - температура суміші на вході в реактор, V - об’єм реагента, q - тепловий ефект стадії реакції.

якщо підставити в них відповідні вирази для інтенсивності джерел маси і тепла. Інтенсивність джерел маси в цьому випадку дорівнює швидкостям утворення реагентів. Вважаючи, що в процесі хімічного перетворення число моделей реагуючих речовин не змінюється, знаходять наступні рівняння для ключових компонентів реакції:

$$V_r \frac{dx_i}{dt} = \upsilon\left(x_i^{(0)} - x_i\right) + V_r \omega_i \quad i = 1,\ldots,m\,,$$

де w_i - швидкість утворення продукту.

За наявності теплового ефекту реакції система рівнянь має бути доповнена співвідношенням, що визначає характер зміни температури в зоні реакції, яке можна отримати з рівняння:

$$\frac{d\left(c_p T\right)}{dt} = \upsilon\left(c_{p0} T^{(0)}\right) - \upsilon c_p T + V q_T$$

У загальному випадку можливий теплообмін реагуючої суміші із зовнішнім теплоносієм. Щоб врахувати його, потрібно використовувати вираження для джерела тепла у формі рівняння

$$q_T = Q_r + Q_T$$

тобто з включенням у нього теплового ефекту реакції і теплопередачі:

$$V_r c_p \frac{dT}{dt} = \upsilon c_p \left(T^{(0)} - T\right) + V_r Q_r + K_T F\left(T_x - T\right)$$

При записі виразу передбачається, що теплоємність c_p реагуючої суміші не змінюється в процесі хімічної реакції.

Спільне вирішення системи рівнянь, доповнене при необхідності стехіометричними співвідношеннями для неключових реагентів, визначає при зроблених вище припущеннях поведінку реактора ідеального змішення в разі нестаціонарних режимів. Для того, щоб отримати систему рівнянь, що характеризує стаціонарні режими реактора, досить в рівняннях покласти похідні за часом рівними нулю.

В результаті система кінцевих рівнянь запишеться у вигляді

$$\upsilon\left(x_i^{(0)} - x_i\right) + V_r \omega_i = 0 \quad i = 1,\ldots,m$$

$$\upsilon c_p \left(T^{(0)} - T\right) + V_r Q_r + K_T F\left(T_x - T\right) = 0$$

звідки для стаціонарного режиму можна знайти значення концентрацій реагентів і температури в реакторі при заданих значеннях вхідних концентрацій реагентів $x_i^{(0)}$ температури суміші $T^{(0)}$, що подається в реактор, і температури теплоносія T_x.

Система рівнянь може бути переписана в зручнішому вигляді, якщо ввести позначення для середнього часу перебування реагентів в апараті:

$$\tau = \frac{V_r}{\upsilon}$$

з урахуванням якого рівняння наберуть вигляду:

$$x_i^{(0)} - x_i + \tau\omega_i = 0$$

$$T^{(0)} - T + \tau\frac{Q_r}{c_p} + \frac{K_T F}{\upsilon c_p}(T_x - T) = 0$$

Нижче розглянутий ряд прикладів побудови математичного опису реактора ідеального змішення для різних типів хімічних реакцій, що проводяться в ізотермічних умовах.

Приклад.

Для реакції типа

$$A \rightarrow P$$

швидкість якої виражається рівнянням

$$r = kx_A$$

знайти склад суміші на виході реактора ідеального змішення.

Розв'язок.

Для опису ізотермічного режиму апарату в даному випадку досить мати лише одне рівняння, що характеризує концентрацію одного з реагентів, наприклад, речовини *A*. Концентрацію другого реагенту при цьому обчислюємо із співвідношення:

$$x_p = x_p^{(0)} + (x_A^{(0)} - x_A)$$

З врахуванням виразу для швидкості стадії знаходимо швидкість утворення реагенту *A*:

$$\omega_A = -kx_A$$

Скориставшись одним з рівнянь системи, отримуємо

$$x_A^{(0)} - x_A - \tau kx_A = 0$$

звідки легко розраховуємо концентрацію реагенту *A* для заданого значення

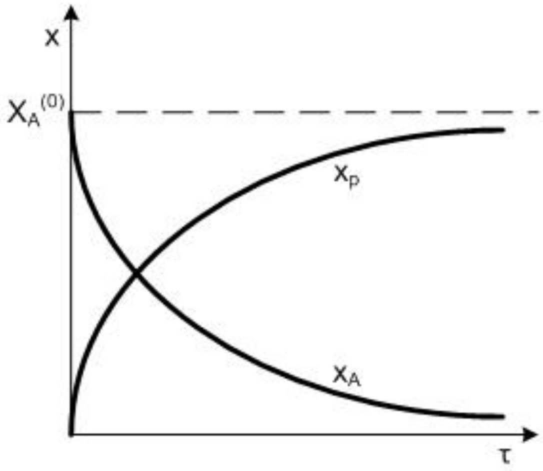

Рис. 1.2. Зміна концентрацій в реакторі ідеального змішення залежно від часу перебування τ для реакції $A \rightarrow P$.

середнього часу перебування τ при заданій константі швидкості реакції k, значення якої визначається температурою реагуючої суміші:

$$x_A = \frac{x_A^{(0)}}{1 + \tau k}$$

На рис. 1.2 показаний характер зміни концентрацій x_A і x_p залежно від часу перебування τ.

Оптимізація реакторів ідеального змішення.

Під оптимальними умовами розуміються оптимальний час перебування реагентів в реакторі $\tau_{опт}$ і оптимальна температура реакції $T_{опт}$, що забезпечують максимальне або мінімальне (залежно від постановки завдання) значення заданого критерію R.

Приклад.

Для реакції першого порядку визначити оптимальні умови, при яких собівартість продукту P буде мінімальною. При обчисленні собівартості можна скористатися виразом

$$R = s_{пр} = \frac{1}{\upsilon x_p}\left(s_A \upsilon x_A^{(0)} + s_V V + s_Д\right)$$

де s_A — вартість одиниці сировини, що витрачається; s_V — вартість одиниці об'єму реактора, що обчислюється з обліком його амортизації; $s_Д$ — вартість додаткового устаткування реактора, що обчислюється з врахуванням амортизації; υ — навантаження реактора по вихідній сировині; x_p — концентрація отримуваного продукту після реакції.

Приймається, що навантаження на реактор υ задана. При цьому визначення оптимального часу перебування реагентів в реакторі фактично зводиться до знаходження його оптимального об'єму $V_{опт}$.

Розв'язок

Для даного випадку концентрація продукту реакції на виході реактора може бути визначена із співвідношень:

$$x_p = x_p^{(0)} + \left(x_A^{(0)} - x_A\right)$$

$$x_A = \frac{x_A^{(0)}}{1 + \tau k}$$

звідки знаходимо:

$$x_p = \frac{\tau k x_A^{(0)}}{1 + \tau k}$$

Підстановка співвідношення у вираження для критерію оптимальності дає:

$$R = \left(s_A + \frac{s_Д}{\upsilon x_A^{(0)}}\right)\left(\frac{1}{\tau k} + 1\right) + \frac{s_V}{k x_A^{(0)}}\left(\tau k + 1\right)$$

З врахуванням отриманої залежності критерій оптимальності *R* може розглядатися як функція двох змінних-часу перебування τ і температури *T*, від якої залежить значення константи швидкості хімічної реакції *k*. Для спрощення замість змінної *T* в даному випадку можна використати змінну *k*, оскільки передбачається, що між ними існує однозначна залежність, визначена рівнянням Ареніуса.

$$k_j(T) = k_{j\infty} \exp\left(-\frac{E_j}{R_Г T}\right)$$

Необхідні умови екстремуму функції *R* можуть бути отримані диференціюванням функції *R* по обом змінним і прирівнюванням похідних нулю. В результаті знайдемо:

$$\frac{\partial R}{\partial \tau} = -\left(s_A + \frac{s_Д}{\upsilon x_A^{(0)}}\right)\frac{1}{\tau^2 k} + \frac{s_V}{x_A^{(0)}} = 0$$

$$\frac{\partial R}{\partial k} = -\left(s_A + \frac{s_Д}{\upsilon x_A} \right) \frac{1}{\tau k^2} + \frac{s_V}{k^2 x_A} = 0$$

Легко бачити, що останнє рівняння не може задовільняти жодним кінцевим значенням τ і k. Єдина можливість його виконання буде в тому випадку, якщо прийняти для константи швидкості реакції нескінченно велике значення. При цьому для кінцевих значень k похідна $\partial R/\partial k$ завжди залишається від'ємною величиною, що характеризує можливість досягнення мінімального значення критерію R лише при нескінченно великій температурі реакції. Оскільки в реальних умовах температура реакції завжди обмежена певними технологічними межами, отримані результати означають, що наближення до оптимальних умов можна досягти лише проведенням процесу при максимально можливій температурі. Рівняння, що залишається, в цьому випадку дозволяє знайти оптимальне значення часу перебування, відповідне заданому значенню температури процесу:

$$\tau_{опт} = \left(\frac{s_A \upsilon x_A + s_Д}{s_V \upsilon k} \right)^{0,5}$$

Підстановка співвідношення у вираз для критерію оптимальності дозволяє визначити його мінімальне значення для заданого максимально можливого значення температури T:

$$R_{опт} = \frac{s_V}{k x_A} \left[1 + k \left(\frac{s_A \upsilon x_A + s_Д}{s_V \upsilon k} \right)^{0,5} \right]^2$$

За допомогою формул можна отримати також кількісну і якісну оцінки оптимальних умов проведення реакції достатній загального вираження критерію оптимальності.

2. Методи дослідження функцій класичного аналізу

Методи дослідження функцій класичного аналізу - це найвідоміші методи розв'язання порівняно нескладних оптимальних задач. Це передусім:

1) метод пошуку безумовних екстремумів;

2) метод невизначених множників Лагранжа.

Методи дослідження функцій класичного аналізу в основному застосовують тоді, коли відомо аналітичний вигляд залежності функції, що оптимізується f від незалежних змінних x;. Це дозволяє знайти також в аналітичному вигляді похідні оптимізованої функції, використання яких і формулюють потрібні та достатні умови існування екстремуму.

2.1. Метод пошуку безумовних екстремумів

Цей метод можна застосовувати для розв'язання задачі оптимізації, якщо немає обмежень. Він оснований на знаходженні першої похідної цільової функції і прирівнюванні її до нуля. При цьому можна розв'язувати багатопараметричну задачу, тобто задачу з кількома варійованими параметрами. Для розв'язання такої задачі треба знайти перші похідні за кожним з варійовних параметрів, і всі ці похідні прирівняти до нуля. При цьому потрібно розв'язати систему кінцевих рівнянь, найчастіше нелінійних.

Додаткові ускладнення виникають унаслідок того, що розв'язок задачі пошуку оптимуму дає лише необхідні, але не достатні умови. Тому всі розв'язки цієї системи належить перевіряти на достовірність. Крім того, методи дослідження функцій класичного аналізу не дозволяють відразу виділити глобальний екстремум. Тому після перевірки розв'язку оптимальної задачі на достовірність виконують пошук серед розв'язків глобального екстремуму.

2.2. Матриця Гессе

Необхідними умовами наявності в точці X^* локального мінімуму є виконання рівності

$$\nabla f(X^*) = 0,$$

і матриця $\nabla^2 f(X^*)$ - додатно напіввизначена.

Достатніми умовами того, що X^* - точка ізольованого (строгого) локального мінімуму, є

$$\nabla f(X^*) = 0,$$

і матриця $\nabla^2 f(X^*)$ - додатно визначена.

Зазвичай доводиться обмежуватись знаходженням локального мінімуму. Але якщо можна показати, що $\Delta X^T \nabla^2 f(X^{(k)}) \Delta X \geq 0$ для всіх X, то $f(X)$ називають опуклою функцією, а локальний мінімум виявляється глобальним.

Якщо в стаціонарній точці $\overline{X}$ матриця Гессе $\nabla^2 f(\overline{X})$ - від'ємно визначена, то це точка максимуму, якщо $\nabla^2 f(\overline{X})$ - не визначена, то це сідлова точка.

У матричній формі можна записати матрицю Гессе - квадратну матрицю других частинних похідних $f(X)$, узятих у точці $X^{(k)}$:

$$\nabla^2 f\left(X^{(k)}\right) = H\left(X^{(k)}\right) = \begin{bmatrix} \dfrac{\partial^2 f\left(X^{(k)}\right)}{\partial x_1^2} & \cdots & \dfrac{\partial^2 f\left(X^{(k)}\right)}{\partial x_1 \partial x_n} \\ \cdots & \cdots & \cdots \\ \dfrac{\partial^2 f\left(X^{(k)}\right)}{\partial x_n \partial x_1} & \cdots & \dfrac{\partial^2 f\left(X^{(k)}\right)}{\partial x_n^2} \end{bmatrix}.$$

Для випадку двох змінних

$$H\left(X^{(k)}\right) = \begin{bmatrix} \dfrac{\partial^2 f\left(X^{(k)}\right)}{\partial x_1^2} & \dfrac{\partial^2 f\left(X^{(k)}\right)}{\partial x_1 \partial x_2} \\ \dfrac{\partial^2 f\left(X^{(k)}\right)}{\partial x_2 \partial x_1} & \dfrac{\partial^2 f\left(X^{(k)}\right)}{\partial x_2^2} \end{bmatrix}.$$

П р и к л а д.

Розглянемо функцію

$$f(x) = 2x_1^2 + 4x_1 x_2^3 - 10 x_1 x_2 + x_2^2,$$

лінії рівня якої зображено на рис.2.3. Потрібно класифікувати точку $\bar{x} = [0,0]^T$.

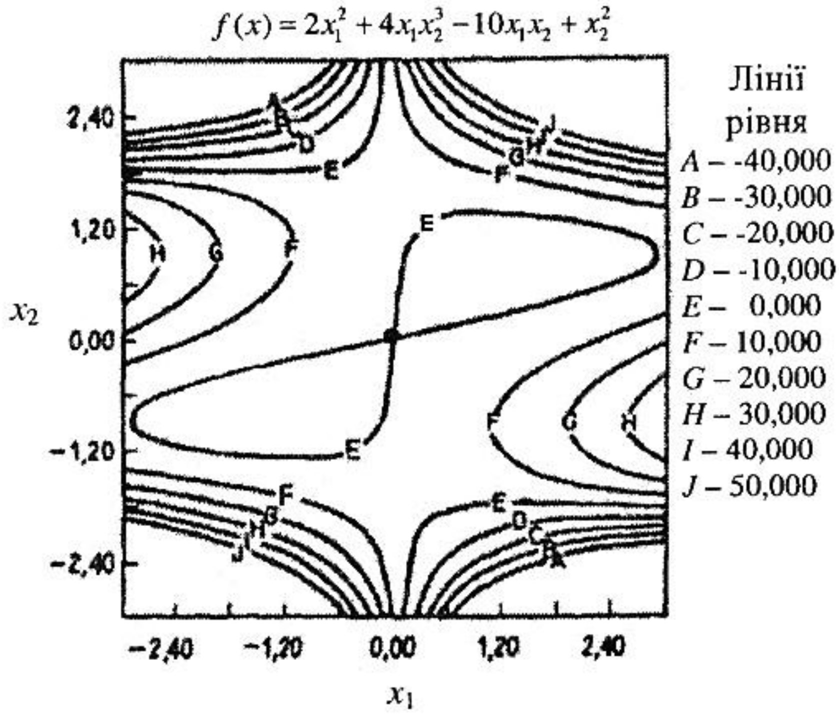

Рис. 2.3. Лінії рівня нелінійної функції двох змінних

Розв'язання:

$$\frac{\partial f}{\partial x_1} = 4x_1 + 4x_2^3 - 10x_2;$$

$$\frac{\partial f}{\partial x_2} = 12x_1x_2^2 - 10x_1 + 2x_2;$$

$$\nabla f(\bar{x}) = [0,0]^T.$$

Тоді точка $\bar{x}$ - стаціонарна:

$$\frac{\partial^2 f}{\partial x_1^2} = 4;$$

$$\frac{\partial^2 f}{\partial x_2^2} = 24x_1x_2 + 2 = +2;$$

$$\frac{\partial^2 f}{\partial x_1 x_2} = 12x_2^2 - 10 = -10.$$

Тоді

$$\nabla^2 f(\bar{x}) = H_f(\bar{x}) = \begin{bmatrix} +4 & -10 \\ -10 & +2 \end{bmatrix}.$$

Матриця $\nabla^2 f(\bar{x})$ - невизначена. Тому $\bar{x}$ являє собою сідлову точку, що і зображено на рис. 2.3.

Завдання 2.1.

Визначити і класифікувати стаціонарні точки функції.

2.1.

$$f(x) = x_1^2 - x_1 x_2 + x_2^2 - 2x_1 - 2x_1 + 3x_2 - 4;$$

2.2.

$$f(x) = 2x_1^3 + 4x_1 x_2^2 - 10x_1 x_2 + x_2^2;$$

2.3.

$$f(x) = x_1^2 - x_1 x_2^2 + 4x_2^2 - 2x_1 - 6;$$

2.4.

$$f(x) = 2x_1^2 - 6x_1 x_2^2 - 10x_1^2 x_2 + x_2^2 - 4.$$

2.3. Метод множників Лагранжа

1. Лінійна система описується рівнянням

$$f(X,U) = AX + BU + C = 0$$

І потрібно мінімізувати функцію вартості

$$I(X,U) = \frac{1}{2} X^T Q X + \frac{1}{2} U^T R U$$

Записати функцію Лагранжа і необхідні умови оптимальності та оптимальне керування.

Відповідь:

Функція Лагранжа формується наступним чином:

$$L = \frac{1}{2} U^T R U + \frac{1}{2} X^T Q X + \lambda^T (AX + BU + C)$$

Необхідні умови оптимальності

$$\frac{\partial L}{\partial X} = QX + A^T \lambda = 0$$

$$\frac{\partial L}{\partial U} = RU + B^T \lambda = 0$$

Оптимальне керування:

$$U^* = -R^{-1} B^T \lambda$$

2. Знайти оптимальне керування U використовуючи метод множників Лагранжа

$$I = X^2 + U^2 \to \min$$

При $XU = 1$

Відповідь:

Функція Лагранжа

$$L = x^2 + U^2 + \lambda(xU - 1)$$

Необхідні умови оптимальності

$$\frac{\partial L}{\partial x} = 2x + \lambda U = 0$$

$$\frac{\partial L}{\partial U} = 2U + \lambda x = 0$$

Розв'язуємо систему рівнянь методом підстановки

$$\lambda = -\frac{2x}{U}$$

$$2U - \frac{2x^2}{U} = 0 \quad U = x$$

З обмеженням $U^2 = 1;\ U^* = \pm\sqrt{1}$

3. Побудувати функцію Лагранжа і необхідні умови оптимальності для задачі умовної оптимізації

$$Z = -x_1 + x_2 + 2x_1^2$$

$$3x_1 + 4x_2 = 12$$

$$-x_1 + 2x_2 = 6$$

Відповідь:

Функція Лагранжа

$$L = -x_1 + x_2 + 2x_1^2 + \lambda_1(3x_1 + 4x_2 - 12) + \lambda_2(-x_1 + 2x_2 - 6)$$

Необхідні умови оптимальності

$$\frac{\partial L}{\partial x_1} = -1 + 4x_1 + 3\lambda_1 - \lambda_2 = 0$$

$$\frac{\partial L}{\partial x_2} = 1 + 4\lambda_1 + 2\lambda_2 = 0$$

$$\frac{\partial L}{\partial \lambda_1} = 3x_1 + 4x_2 - 12 = 0$$

$$\frac{\partial L}{\partial \lambda_2} = -x_1 + 2x_2 - 6 = 0$$

4. Перевірити чи має функція $F(X) = x_1x_2 + 2x_1^2 - 4x^2$ сідлову точку на площині X.

Відповідь:

Необхідні умови оптимальності

Градієнт функції $\nabla F(X) = \begin{bmatrix} x_2 + 4x_1 \\ x_1 - 8x_2 \end{bmatrix} = 0$

Матриця Гессе $\nabla^2 F(X) = \begin{bmatrix} 4 & 1 \\ 1 & 8 \end{bmatrix}$

Матриця Гессе не визначена і точка $X = (0 \quad 0)^T$ – сідлова точка.

5. Класифікувати точку $\overline{X} = (0 \quad 0)^T$ для функції

$f(X) = 2x_1 + 4x_1x_2^3 - 10x_1 + x_2^2$

Відповідь:

Градієнт функції $\nabla f(X) = \begin{bmatrix} 4x_1 + 4x_2^3 - 10x_2 \\ 12x_1x_2^2 - 10x_1 + 2x_2 \end{bmatrix} = 0$

Тобто т. X – стаціонарна.

Матриця Гессе $\nabla^2 f(X) = \begin{bmatrix} 4 & 12x_2^2 - 10 \\ 12x_2^2 - 10 & 24x_1x_2 + 2 \end{bmatrix} = 0$

$\nabla^2 f(\overline{X}) = \begin{bmatrix} 4 & -10 \\ -10 & 2 \end{bmatrix} \geq 0$

Матриця $\nabla^2 f(\overline{X})$ – невизначена і $\overline{X}$ – сідлова точка.

Приклад.

Знайти точки умовного екстремуму функції

$$Z = x_1 + x_2 + 2x_1^2$$
$$3x_1 + 4x_2 = 12$$
$$-x_1 + 2x_2 = 6$$

Побудуємо функцію Лагранжа

$$F = -x_1 + x_2 + 2x_1^2 + \lambda_1(3x_1 + 4x_2 - 12) + \lambda_2(-x_1 + 2x_2 - 6)$$

і знайдемо її похідні за $x_1, x_2, \lambda_1, \lambda_2$.

Прирівнявши здобуті вирази до нуля, дістанемо систему рівнянь

$$\begin{cases} \dfrac{\partial F}{\partial X_1} = -1 + 4X_1 + 3\lambda_1 - \lambda_2 = 0 \\ \dfrac{\partial F}{\partial X_2} = 1 + 4\lambda_1 + \lambda_2 = 0 \\ \dfrac{\partial F}{\partial \lambda_1} = 3X_1 + 4X_2 - 12 = 0 \\ \dfrac{\partial F}{\partial \lambda_2} = X_1 + 2X_2 - 6 = 0 \end{cases}$$

Розв'язавши цю систему методом Жордана-Гауса, дістанемо $x_1 = 0$; $x_2 = 3$.

Приклад.

Виробник бляшаних консервних банок бажає максимізувати деяку партію банок із заданою площею жесті.

Об’єм $V(r,l) = \pi r^2 l.$

Площа $A(r,l) = 2\pi r^2 + 2\pi r^2 l = A_0.$

Розглянемо задачу, використовуючи множник Лагранжа. Спочатку сформуємо функцією

$$V' = (r,l) = V(r,l) + \lambda \left(A(r,l) - A_0\right),$$

де λ - множник Лагранжа. Цей вираз можна записати через параметри консервної банки так:

$$V'(r,l) = \pi r^2 l + \lambda \left(2\pi r^2 + 2\pi rl - A_0\right).$$

Візьмемо першу частинну похідну за кожною змінною і прирівняємо до нуля. Отримуємо:

$$\frac{\partial V'(r,l)}{\partial l} = \pi r^2 + \lambda 2\pi r = 0, \;\; r = -2\lambda;$$

$$\frac{\partial V'(r,l)}{\partial r} = 2\pi rl + \lambda \left(4\pi r + 2\pi l\right) = 0, \;\; l = 2r.$$

Розрахуємо тепер λ з урахуванням заданого обмеження $A(r,l) = A_0$ чи $A_0 = 2\pi r^2 + 2\pi rl$. Одержуємо

$$A_0 = 2\pi\left(4\lambda^2\right) + 2\pi\left(-2\lambda\right)\left(-4\lambda\right).$$

Тоді $\lambda = \pm\sqrt{A_0 / 24\pi}.$

Отже, маємо $r = 2\sqrt{A_0 / 24\pi}, \;\; l = 4\sqrt{A_0 / 24\pi}.$

Завдання 2.2.

Використовуючи метод множників Лагранжа, знайти точки умовного екстремуму функції Z.

2.5.

$$Z = x_1^2 + 2x_2^2 + 3x_3^2;$$
$$x_1 + 2x_2 + x_3 = 8.$$

2.6.

$$Z = 2x_1 \cdot x_2 + x_2^2;$$
$$2x_1 + 4x_2 = 8.$$

2.7.

$$Z = x_1^2 + 2x_1 + x_2^2 - 5x_2;$$
$$x_1 + 3x_2 = 6.$$

2.8.

$$Z = 3x_1^2 + 2x_1 + x_2^2 - 2x_2;$$
$$2x_1 + 3x_2 = 5;$$
$$x_1 + 4x_2 = 7.$$

3. Лінійне програмування

3.1. Геометрична інтерпретація задач лінійного програмування

Приклад.

Розв’язати графічно таку задачу лінійного програмування (рис.2.1.):

$$Z = x_1 + 2x_2 (\max)$$

$$2x_1 + 4x_2 \ge 8$$

$$2x_1 - 3x_2 \le 6$$

$$-3x_1 + 2x_2 \ge 9$$

$$x_1 \ge 0, x_2 \ge 0$$

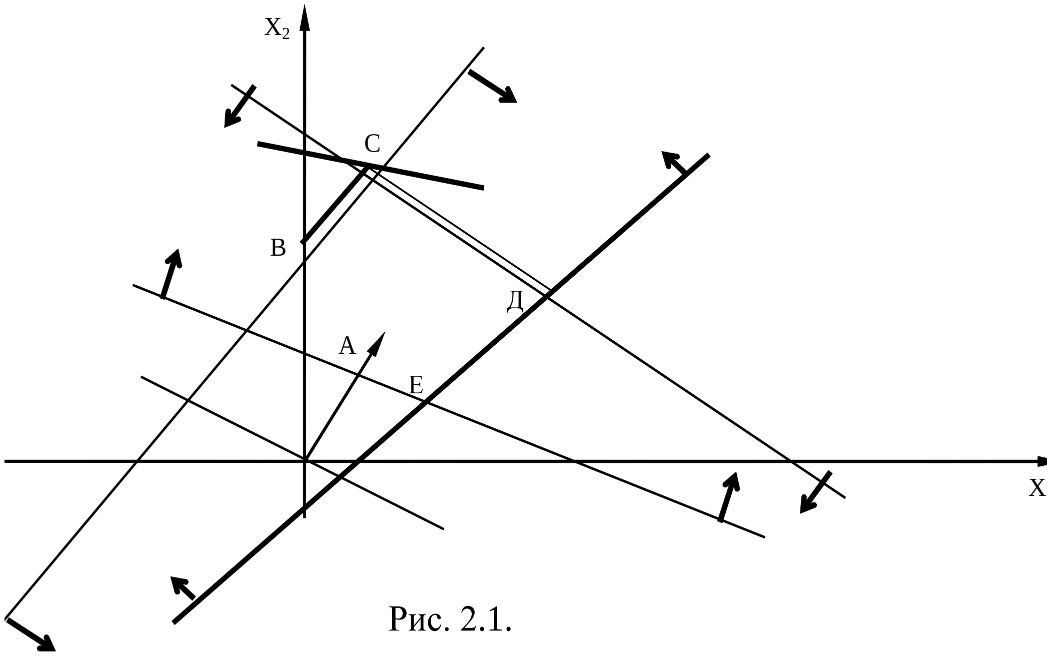

Рис. 2.1.

Завдання 3.1.

Розв’язати графічно такі задачі лінійного програмування:

3.1. $Z = 7x_1 + 6x_2 (\max)$

$2x_1 + 5x_2 \geq 10$

$5x_1 + 2x_2 \geq 10$

$x_1 \leq 6$

$x_2 \leq 5$

$x_1 \geq 0,\ x_2 \geq 0$

3.2. $Z = 5x_1 + x_2 (\max)$

$x_1 + x_2 \geq 2$

$x_1 - x_2 \leq 2$

$4x_1 - 8x_2 \geq 16$

$x_1 \geq 1$

$0 \leq x_2 \leq 3$

3.3. $Z = 5x_1 - 3x_2 (\min)$

$3x_1 + 2x_2 \geq 6$

$2x_1 - 3x_2 \geq -6$

$x_1 + x_2 \leq 4$

$4x_1 + 7x_2 \leq 28$

$x_1 \geq 0,\ x_2 \geq 0$

Використовуючи геометричну інтерпретацію задачі лінійного програмування в задачах 3.4.-3.6., визначити область змінювання параметрів, для яких:

а) задача несумісна;

б) область необмежена;

в) задача має розв’язок;

г) область розв’язування зображена точкою;

д) якщо задача несумісна чи область необмежена, то змінити умови, щоб задача мала розв’язок.

3.4. $Z = x_1 + x_2 (\max)$

$x_1 + ax_2 \le 1$

$ax_1 - x_2 \le 1$

$x_1, x_2 \ge 0$

3.5. $Z = x_1 - x_2 (\min)$

$x_1 \le 1$

$-ax_1 + x_2 \ge 4$

$x_1, x_2 \ge 0$

3.6. $Z = 2x_1 + 3x_2 (\min)$

$x_1 + x_2 \le 1$

$x_1 + bx_2 \le 1$

$x_1 - x_2 \ge c$

$x_1, x_2 \ge 0$

3.2. Симплекс метод розв'язування задач лінійного програмування

Розглянемо приклад:

$$Z = 5x_1 + 2x_2 - 6x_3 (\max)$$

$$x_1 + 2x_2 + x_3 = 8 \qquad \textbf{(1)}$$

$$3x_1 + 5x_2 \le 15 \qquad \textbf{(2)}$$

$$6x_1 + x_2 \ge 12 \qquad \textbf{(3)}$$

$$x_i \ge 0, \ j=1.3$$

Зведемо систему обмежень до стандартного вигляду. Для цього в обмеження (2) і (3) введемо додаткові змінні. У цільову функцію додаткові змінні входять з нульовими коефіцієнтами.

$$Z = 5x_1 + 2x_2 - 6x_3 + 0x_4 + 0x_5 (\max)$$

$$x_1 + 2x_2 + x_3 = 8$$

$$3x_1 + 5x_2 + x_4 = 15$$

$$6x_1 + x_2 - x_5 = 12$$

$$x_i \ge 0, \ j=1.3$$

Запишемо матрицю A і вектор B, утворені коефіцієнтами при змінних:

$$\overline{A}_1 = \begin{bmatrix} 1 \\ 3 \\ 6 \end{bmatrix} \quad \overline{A}_2 = \begin{bmatrix} 2 \\ 5 \\ 1 \end{bmatrix} \quad \overline{A}_3 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \quad \overline{A}_4 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \quad \overline{A}_5 = \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix} \quad \overline{B} = \begin{bmatrix} 8 \\ 15 \\ 12 \end{bmatrix}$$

У системі немає одиничного базиса, тому для його утворення не вистачає вектора $\overline{A}_6$:

$$\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

Щоб дістати його, введемо штучну змінну в третє рівняння. В цільову функцію x_6 увійде з коефіцієнтом -M, тому що цільова функція максимізуватиметься. Дістанемо:

$$Z = 5x_1 + 2x_2 - 6x_3 + 0x_4 + 0x_5 (\max)$$

$$x_1 + 2x_2 + x_3 = 8$$

$$3x_1 + 5x_2 + x_4 = 15$$

$$6x_1 + x_2 - x_5 + x_6 = 12$$

$$x_i \geq 0$$

За допомогою М-метода (метода великих штрафів) побудуємо першу симплексну таблицю (таб.3.1.) і будемо далі розв’язувати задачу в табличній формі до здобуття оптимального плану:

Таблиця 3.1.

$C_{баз}$	**Базис**	**A_0**	**5**	**2**	**-6**	**0**	**0**	**-M**	
			x_1	**x_2**	**x_3**	**x_4**	**x_5**	**x_6**	
-6	**x_3**	**8**	**1**	**2**	**1**	**0**	**0**	**0**	**8**
0	**x_4**	**15**	**3**	**5**	**0**	**1**	**0**	**0**	**5 min {8,5,2}=2**
-M	**x_6**	**12**	**6**	**1**	**0**	**0**	**-1**	**0**	**2**
Zj-Cj	**0**	**-48**	**-29**	**-14**	**0**	**0**	**0**	**0**	**x_1 = (0,0,8,15,12)**
		-12M	**-M**	**-M**	**0**	**0**	**M**	**0**	**Z_1 = -48, -12M**
-6	**x_3**	**6**	**0**	**11/6**	**1**	**0**	**1/6**		**36/11**
0	**x_4**	**9**	**0**	**9/2**	**0**	**1**	**1/2**		**2**
5	**x_1**	**2**	**1**	**1/6**	**0**	**0**	**-1/6**		**12**

Zj-Cj	**0**	**-26**	**0**	**-73/6**	**0**	**0**	**-11/6**		$\mathbf{x_2 = (2,0,6,9,0)}$
									$\mathbf{Z_2 = -26}$
-6	$\mathbf{x_3}$	**7/3**	**0**	**0**	**1**	**-11/27**	**-1/27**		**-9/7**
2	$\mathbf{x_2}$	**2**	**0**	**1**	**0**	**2/9**	**1/9**		**18**
5	$\mathbf{x_1}$	**5/3**	**1**	**0**	**0**	**-1/27**	**-5/27**		**-9**
Zj-Cj	**0**	**-5/3**	**0**	**0**	**0**	**73/27**	**-20/27**		$\mathbf{x_3 = (5/3,2,7/3,0)}$
									$\mathbf{Z_3 = -5/3}$
-6	$\mathbf{x_3}$	**3**	**0**	**1/3**	**1**	**-1/3**	**0**		
0	$\mathbf{x_5}$	**18**	**0**	**9**	**0**	**2**	**1**		
5	$\mathbf{x_1}$	**5**	**1**	**5/3**	**0**	**1/3**	**0**		$\mathbf{x_0 = (5,0,3,0,18)}$
Zj-Cj	**0**	**7**	**0**	**13/3**	**0**	**11/3**	**0**		$\mathbf{Z_{max} = 7}$

Усі Zj-Cj ≥0; отже, план оптимальний.

Відповідь: $x_{опт} = (5,0,3,0,18)$; $Z_{max} = 7$.

Використаємо двоетапний симплекс-метод для розв’язку прикладу:

$$-3x_1 - 4x_2 = Z \to \min$$

$$x_1 \ge 10$$

$$x_2 \ge 5$$

$$x_1 + x_2 \le 20$$

$$-x_1 + 4x_4 \le 20$$

$$x_i \ge 0$$

Введемо додаткові змінні:

$x_1 - x_3 = 10$

$x_2 \ -x_4 = 5$

$x_1 + x_2 \ +x_5 = 20$

$-x_1 + 4x_2 \ +x_6 = 20$

Змінимо два перші обмеження, увівши штучні зміннu x_7 і x_8 і мінімізувати штучну цільову функцію $W = x_7 + x_8$. У канонічній формі отримаємо:

$-x_1 - x_2 + x_3 + x_4 = W - 15$. Розв`язок за допомогою двоетапного симплекс-методу показано у табл.3.2

Таблиця 3.2.

Ітерація	Базис	Значення	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
0	x_7	**10**	**1**	**0**	**-1**	**0**	**0**	**0**	**1**	**0**
	x_8	**5**	**0**	**1**	**0**	**-1**	**0**	**0**	**0**	**1**
	x_5	**20**	**1**	**1**	**0**	**0**	**1**	**0**	**0**	**0**
	x_6	**20**	**-1**	**4**	**0**	**0**	**0**	**1**	**0**	**0**
	-Z	**0**	**-3**	**-4**	**0**	**0**	**0**	**0**	**0**	**0**
	-W	**15**	**-1**	**-1**	**1**	**1**	**0**	**0**	**0**	**0**
1	x_1	**10**	**1**	**0**	**-1**	**0**	**0**	**0**	**1**	**0**
	x_8	**5**	**0**	**1**	**0**	**-1**	**0**	**0**	**0**	**1**
	x_5	**10**	**0**	**1**	**1**	**0**	**1**	**0**	**-1**	**0**
	x_6	**30**	**0**	**4**	**-1**	**0**	**0**	**1**	**1**	**0**
	-Z	**30**	**0**	**-4**	**-3**	**0**	**0**	**0**	**3**	**0**
	-W	**-5**	**0**	**-1**	**0**	**1**	**0**	**0**	**1**	**0**
2	x_1	**10**	**1**	**0**	**-1**	**0**	**0**	**0**	**1**	**0**
	x_2	**5**	**0**	**1**	**0**	**-1**	**0**	**0**	**0**	**1**
	x_5	**5**	**0**	**0**	**1**	**1**	**1**	**0**	**-1**	**-1**
	x_6	**10**	**0**	**0**	**-1**	**4**	**0**	**1**	**1**	**-4**
	-Z	**50**	**0**	**0**	**-3**	**-4**	**0**	**0**	**3**	**4**
	-W	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**1**	**1**
3	x_1	**10**	**1**	**0**	**-1**	**0**	**0**	**0**		
	x_2	**7.5**	**0**	**1**	**-0.25**	**0**	**0**	**0.25**		
	x_5	**2.5**	**0**	**0**	**1.25**	**0**	**1**	**-0.25**		
	x_4	**2.5**	**0**	**0**	**-0.25**	**1**	**0**	**0.25**		
	-Z	**60**	**0**	**0**	**-4**	**0**	**0**	**1**		
4	x_1	**12**	**1**	**0**	**0**	**0**	**0.8**	**-0.2**		
	x_2	**6**	**0**	**1**	**0**	**0**	**0.2**	**0.2**		
	x_3	**2**	**0**	**0**	**1**	**0**	**0.8**	**-0.2**		
	x_4	**3**	**0**	**0**	**0**	**1**	**0.2**	**0.2**		
	-Z	**68**	**0**	**0**	**0**	**0**	**3.2**	**0.2**		

Завдання 3.2.

Розв’язати двоетапним симплекс методом такі задачі:

3.7. $Z = 6x_1 + 5x_2 + x_3(\max)$

$x_1 + 4x_2 + 2x_3 = 8$

$2x_1 + x_2 + x_3 \ge 10$

$x_1 + 3x_2 + x_3 = 12$

$x_j \ge 0$

3.8. $Z = 2x_1 - 3x_2 + x_3(\min)$

$x_1 + 3x_2 + 8x_3 = 24$

$2x_1 + x_2 + x_3 \ge 18$

$x_1 + 3x_2 + 2x_3 = 18$

$x_j \ge 0$

3.9. $Z = -x_1 + x_2 + x_3(\max)$

$4x_1 + 5x_2 + x_3 \le 24$

$-6x_1 - 8x_2 \ge -24$

$x_1 + 5x_2 + x_3 \le 15$

$x_j \ge 0$

3.3. Двоїста задача

Приклад.

Нехай задана початкова задача

$Z = 8x_1 + 6x_2 + 9x_3(\max)$

$x_1 + 2x_2 + x_3 = 12$

$3x_1 + x_2 + 2x_4 = 18$

$x_j \ge 0$

Треба знайти розв’язок як заданої задачі, так і двоїстої для неї. Будуємо двоїсту задачу

$F = 12y_1 + 18y_2(\min)$

$y_1 + 3y_2 \ge 8$

$2y_1 + y_2 \ge 6$

$y_1 + 2y_2 \ge 9$

$x_j \ge 0$

y_1, y_2 - ,будь які значення

У двоїстої задачі (несиметричній) відсутні умови невід’ємності змінних. Тому її розв’язувати безпосередньо симплексним методом не можна.

Розв’яжемо початкову задачу і за її розв’язком знайдемо розв’язок двоїстої задачі (табл. 3.3.).

Таблиця 3.3.

Базис	**Значення**	**x_1**	**x_2**	**x_3**	**x_4**	**x_5**	
x_4	**12**	**1**	**2**	**1**	**1**	**0**	**12**
x_5	**18**	**3**	**1**	**2**	**0**	**1**	**6**
-Zj	**0**	**-8**	**-6**	**-9**	**0**	**0**	
-W	**-30**	**-4**	**-3**	**-3**	**0**	**0**	
x_4	**6**	**0**	**5/3**	**1/3**	**1**	**-1/3**	**18/5**
x_1	**6**	**1**	**1/3**	**2/3**	**0**	**1/3**	**18**
-Zj	**48**	**0**	**-10/3**	**1/3**	**0**	**8/3**	
-W	**-6**	**0**	**-5/3**	**3/3**	**0**	**1/3**	
x_2	**18/5**	**0**	**1**	**1/5**	**3/5**	**-1/5**	**18**
x_1	**24/5**	**1**	**0**	**3/5**	**-1/5**	**2/5**	**8**
-Zj	**60**	**0**	**0**	**-3**	**2**	**2**	
-W	**0**	**0**	**0**	**0**			
x_2	**2**	**-1/3**	**1**	**0**	**2/3**	**-1/3**	
x_3	**8**	**5/3**	**0**	**1**	**-1/3**	**2/3**	
-Zj	**84**	**5**	**0**	**0**	**1**	**4**	

Знайдемо розв'язок двоїстої задачі за формулою $y_{опт} = C_{баз} * D^{-1}$,

де $C_{баз}$=(6;9);

$$D = \begin{pmatrix} 2/3 & -1/3 \\ -1/3 & 2/3 \end{pmatrix}$$

$$Y_0 = (6,9)*\begin{pmatrix} 2/3 & -1/3 \\ -1/3 & 2/3 \end{pmatrix} = (1;4)$$

$F_{min} = 12*1+18*4 = 84$

Перевірка $Z_{max} = F_{min} = 84$.

Приклад.

На основі графічного аналізу двоїстої задачі до початкової дослідити розв'язок обох задач і в разі його присутності знайти оптимальний розв'язок, використовуючи другу теорему двоїстості.

Початкова задача:

$$F = y_1 + y_2 \quad (\max)$$

$$-3y_1 + 2y_2 \le 1$$
$$2y_1 + 2y_2 \le 13$$
$$3y_1 - y_2 \le 12$$
$$y_1 \ge 0, \; y_2 \ge 0$$

Розв'яжемо графічно двоїсту задачу (рис. 2.2.)

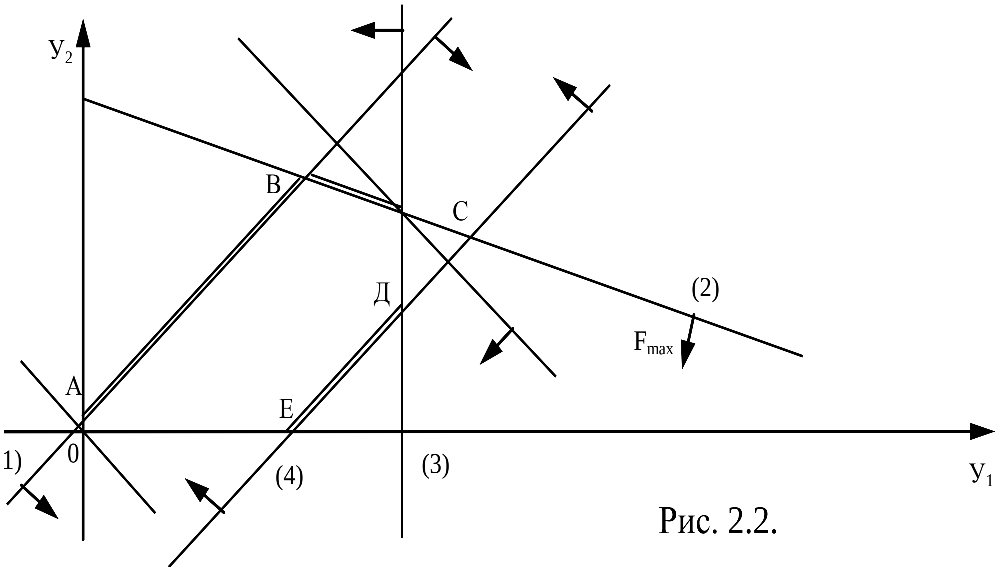

Рис. 2.2.

Побудуємо багатокутник розв'язку. Кожне обмеження є півплощиною. Перетинання відповідних півплощин з умовою невід'ємності змінних $y_j \ge 0$ утворить багатокутник розв'язання ABCDEO.

Визначимо крайню кутову точку, в якій пряма цільової функції буде опорною. Такою точкою буде точка С. Координати точки С визначають, розв'язуючи систему рівнянь для прямих, які перетинаються в точці С (прямі 2 і 3)

$$y_1 + 2y_2 = 14$$
$$2y_1 + y_2 = 13$$

Розв'язок системи визначить оптимальний план двоїстої задачі

$$y_1 = 4;\ y_2 = 5;\ y_0 = (4;5);\ Fmax = 9.$$

Згідно з першою теоремою двоїстості матиме розв'язок і початкова задача, причому Zmin = Fmax.

Для визначення оптимального плану початкової задачі використаємо другу теорему двоїстості. Якщо підставимо оптимальний розв'язок двоїстої задачі в першу і четверту нерівності системи обмеження цієї задачі, то побачимо, що ці нерівності виконуються як строгі, тобто

$$-12 + 10 = -2 < 1, \quad 12 - 10 \le 12.$$

Отже, відповідні змінні x_1 і x_4 початкової задачі повинні дорівнювати нулю (згідно з теоремою двоїстості). Обмеження початкової задачі перетворюються в рівняння, тому що відповідні їм змінні

$$y_1 = 4 > 0; \quad \text{і} \quad y_2 = 5 > 0.$$

Тоді з системи рівнянь початкової задачі

$$\begin{aligned} x_2 + 2x_3 &= 1 \\ 2x_2 + \ \ x_3 &= 1 \end{aligned}$$

Отже, $x_0=(0,1/3,1/3,0)$ Zmin = 9

При цьому Zmin = Fmax = 9, що підтверджує правильність розв'язку.

Завдання 3.3.

Побудувати двоїсту задачу до початкової. Розв'язати одну з пари задач і знайденим розв'язком дістати розв'язок двоїстої задачі:

3.10. $Z = 6x_1 + 7x_2 + 8x_3(\max)$

$$\begin{aligned} 4x_1 + 4x_2 + x_3 &= 12 \\ x_1 + 2x_2 + 2x_3 &\ge 12 \\ x_1 + 5x_2 + x_3 &\ge 20 \\ x_j &\ge 0 \end{aligned}$$

3.11. $Z = 5x_1 + x_2 + 4x_3(\max)$

$$\begin{aligned} x_1 + 4x_2 + x_3 &= 8 \\ 2x_1 + 4x_2 + x_3 &= 8 \\ x_1 + 5x_2 + 6x_3 &= 30 \\ x_j &\ge 0 \end{aligned}$$

3.12. $Z = 5x_1 + 6x_2 + x_3(\min)$

$$\begin{aligned} x_1 + 4x_2 + 3x_3 &\le 12 \\ 2x_1 + x_2 + x_3 &\ge 12 \\ x_1 + x_2 + 3x_3 &\ge 6 \\ x_j &\ge 0 \end{aligned}$$

Алгоритм подвійного симплекс-метода був виведений без звернення до двоїстості. Однак двоїстість дозволяє поглянути на процедуру по іншому.

Розглянемо задачу:

Знайти такі $x_1 \ge 0$, $x_2 \ge 0$, $x_3 \ge 0$, що

$$\begin{aligned} x_1 + 3x_3 &\ge 3 \\ x_2 + 2x_3 &\ge 5 \end{aligned}$$

та функцію $4x_1 + 6x_2 + 18x_3 = Z$ має мінімальне значення.

Оскільки коефіцієнти у виразі для функції Z додатні, можливо уникнути введення штучних змінних та вирішити задачу з використанням подвійного симплекс-методу.

Симплекс-множники доривнюють:

$$\pi^{T} = -(B^{-1})^{T} * C_{в} = -\frac{1}{3}\begin{bmatrix} 2 & -1 \\ -3 & 0 \end{bmatrix}\begin{bmatrix} 6 \\ 18 \end{bmatrix} = -\begin{bmatrix} 2/3 & 1/3 \\ -1 & 0 \end{bmatrix}\begin{bmatrix} 6 \\ 8 \end{bmatrix} = \begin{bmatrix} 2 \\ 6 \end{bmatrix}$$

$$B^{-1}A = \begin{bmatrix} 2/3 & -1 \\ -1/3 & 0 \end{bmatrix}\begin{bmatrix} -1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} -2/3 & 2/3 & -1 \\ 1/3 & -1/3 & 0 \end{bmatrix}$$

Приведемо таблицю послідовних обчислень подвійним симплекс методом:

Таблиця 3.4.

Ітерація	**Базис**	**Значення**	x_1	x_2	x_3	x_4	x_5
0	x_4	**-3**	**-1**	**0**	**-3**	**1**	**0**
	x_5	**-5**	**0**	**-1***	**-2**	**0**	**1**
	-Z	**0**	**4**	**6**	**18**	**0**	**0**
1	x_4	**-3**	**-1**	**0**	**-3***	**1**	**0**
	x_5	**5**	**0**	**1**	**2**	**0**	**-1**
	-Z	**-30**	**4**	**0**	**6**	**0**	**6**
	x_2	**1**	**1/3**	**0**	**1**	**-1/3**	**0**
	x_3	**3**	**-2/3**	**1**	**0**	**2/3**	**-1**
	-Z	**-36**	**2**	**0**	**0**	**2**	**6**

Таким чином в оптимальному рішенні

$\mathbf{x_1}$=0, $\mathbf{x_2}$=3, $\mathbf{x_3}$=1 та $\mathbf{Z}_{min}$=36.

Симплекс-множники (коефіцієнти при нових змінних $\mathbf{x_4}$ та $\mathbf{x_5}$ кінцевому вигляді для функції **Z**) дорівнюють π_1=2 та π_2=6.

Розглянемо двоїсту задачу:

Знайти такі $\mathbf{y_1 \geq 0}$, $\mathbf{y_2 \geq 0}$, що

$\mathbf{y_1 \leq 4}$

$\mathbf{y_2 \leq 6}$

$\mathbf{y_1 + y_2 \leq 18}$

та функція $3\mathbf{y_1} + 5\mathbf{y_2} = \mathbf{W}$ має максимальне значення.

При звичайному підході до задачі мінімізуємо функцію

$W' = -3y_1 - 5y_2$

Приведемо таблицю послідовних обчислень:

Таблиця 3.5.

Ітерація	Базис	Значення	y_1	y_2	y_3	y_4	y_5
0	y_3	**4**	**1**	**0**	**1**	**0**	**0**
	y_4	**6**	**0**	**1***	**0**	**1**	**0**
	y_5	**18**	**3**	**2**	**0**	**0**	**1**
	$-W'$	**0**	**-3**	**-5**	**0**	**0**	**0**
1	y_3	**4**	**1**	**0**	**1**	**0**	**0**
	y_4	**6**	**0**	**1**	**0**	**1**	**0**
	y_5	**6**	**3***	**0**	**0**	**-2**	**1**
	$-W'$	**30**	**-3**	**0**	**0**	**5**	**0**
2	y_3	**2**	**0**	**0**	**1**	**2/3**	**-1/3**
	y_2	**6**	**0**	**1**	**0**	**1**	**0**
	y_1	**2**	**1**	**0**	**0**	**-1/3**	**1/3**
	$-W'$	**36**	**0**	**0**	**0**	**3**	**1**

Симплекс-множники (для цільової функції W') $\rho_1=0$, $\rho_2=3$, $\rho_3=1$, (коефіцієнти при нових змінних $\mathbf{y_3}$, $\mathbf{y_4}$, $\mathbf{y_5}$). Вони дають значення змінних прямої задачі. Двоїсті змінні $\mathbf{y_1=2}$ та $\mathbf{y_2=6}$ є симплекс-множниками прямої задачі. У проведених обчисленнях подвійний симплекс-метод для рішення прямої задачі та симплекс-метод для рішення зворотної задачі, за суттю, ідентичні.

Завдання 3.4.

Розв'язати пряму задачу і порівняти з рішенням двоїстої задачі.

3.13. $-2x_1 - 5x_2 + x_3 \geq 7$

$-5x_1 - 2x_2 + x_3 + x_4 \geq 6$

$x_j \geq 0 \quad j = \overline{1.4};$

$Z = -10x_1 - 10x_2 + 6x_3 - 5x_4 \text{ (min)}$

3.14. $-x_1 + 6x_2 + 2x_3 - x_4 \geq 2$

$x_1 + 7x_2 - 3x_3 - x_4 \geq -1$

$x_j \geq 0 \quad j = \overline{1.4};$

$Z = 3x_1 + 42x_2 + 6x_3 - 4x_4 \text{ (min)}$

3.15. $-2x_1 + 3x_2 + 3x_3 \leq 3$

$-x_1 - 2x_2 + 4x_3 \leq -2$

$x_j \geq 0 \quad j = \overline{1.3};$

$Z = -14x_1 - 9x_2 + 27x_3 \text{ (max)}$

3.16. $x_1 + 3x_2 - 5x_3 \geq 1$

$2x_1 - 5x_2 - 3x_3 \geq 2$

$x_j \geq 0 \quad j = \overline{1.3};$

$Z = 15x_1 + 8x_2 - 26x_3 \text{ (min)}$

3.4. Транспортна задача лінійного програмування.

Транспортною задачею лінійного програмування називають специфічну задачу пошуку найбільш економічного варіанту транспортування вантажу від кількох пунктів, де вантаж знаходився спочатку, до пунктів де його потребують споживачі. Транспортна задача це задача лінійного програмування, яку можна розв'язати звичайним симплексним методом, але її значно легше та ефективніше розв'язувати спеціальними алгоритмами, зокрема методом потенціалів.

Загальна постановка транспортної задачі має такий вигляд:

$$Z = \sum_{j=1}^{n} \sum_{i=1}^{m} C_{ij}\, x_{ij} \to \text{(min)};$$

$$\sum_{j=1}^{n} x_{ij} \leq a_{ij}; i = \overline{1,m}$$

$$\sum_{j=1}^{m} x_{ij} \geq b_j; j = \overline{1,n};\ x_{ij} \geq 0,$$

де $C_{i\,j}$ - коефіцієнт питомих транспортних витрат при перевезенні вантажу з пункту i до пункту j; $x_{i\,j}$ - кількість вантажу, який перевозиться з пункту (i = 1,m) до пункту j (j = 1,n); a_i - потужність поставщика i; b_i - потужність споживачі j.

У тому разі, коли $\sum_{i=1}^{m} a_i = \sum_{j=1}^{n} b_j$, транспортна задача є закритою.

Коли ж $\sum a_i \neq \sum b_j$ впроваджується фіктивний поставщик або споживач залежно від того, де сумарна потужність менша за розміром. До матриці коефіцієнтів питомих транспортних витрат додається стовпчик або рядок з нульовими питомими витратами $C_{i\,j}$ (залежно від того, де додається фіктивний контрагент).

Розв’язання транспортної задачі складається з двох основних етапів. На першому будується початковий опорний план, на другому - пошук оптимального плану. Перший опорний план можна будувати одним з чотирьох методів: північно-західного кута, мінімального елемента, подвійної переваги та апроксимації Фогеля. Пошук оптимального плану на другому етапі треба робити з допомогою методу потенціалів. План перевезень буде оптимальним, якщо знайдено систему з (n+m) чисел α_i, β_j, які відповідають умовам

$$\alpha_i + \beta_j = C_{i\,j} \text{ при } x_{i\,j} > 0$$
$$\alpha_i + \beta_j \quad C_{i\,j} \text{ для } x_{i\,j} = 0$$

або

$$\Delta_{i\,j} = C_{i\,j} = (\alpha_i + \beta_j) \quad 0$$

Приклад.

Розв’язати транспортну задачу перевезень однорідного вантажу.

- Потужності поставщиків: a_i = (40, 30, 20),
- потужності споживачів b_j = (40, 30, 40, 10).
- Матриця коефіцієнтів питомих транспортних витрат

$$C_{i\,j} = \begin{pmatrix} 4 & 1 & 2 & 1 \\ 5 & 2 & 4 & 2 \\ 1 & 3 & 3 & 3 \end{pmatrix}$$

Перший опорний план побудуємо методом подвійної переваги. Він матиме такий вигляд (табл. 3.6.).

Таблиця 3.6.

a_i	**Перевезення для b_i**				
	40	30	40	10	$\propto_i$
40	^{4}	30 ^{1}	+ ^{2}	-10 ^{1}	0
30	^{5}	^{2}	30 ^{4}	+0 ^{2}	1
20	20 ^{1}	^{3}	^{3}	^{3}	-2
30	^{0}	^{0}	10 ^{0}	^{0}	-3
β_j	3	1	3	1	

Для першого варіанту плану цільова функція

$$Z = 30\times1 + 10\times1 + 30\times4 + 0\times2 + 20\times1 + 20\times0 + 10\times0 = 180.$$

Для перевірки оптимального плану обчислимо значення

$$\Delta_{ij} = C_{ij} - (\alpha_i + \beta_j).$$

Серед них відберемо лише Δ_{ij}, що менше за нуль. У даному разі це всього одне значення (якщо їх було б кілька, то вибирали б максимальне за модулем). Таким чином,

$$\Delta_{1,3} = 2 - (3 + 0) = -1$$

Побудуємо цикл у клітинці (1,3). Потім за зайнятими маршрутами або клітинками побудуємо цикл перебудови плану (його вдосконалення). Перенумеруємо вершини циклу знаками “+” і “-”, починаючи з нової клітинки (1,3). Далі виберемо за маршрутом, позначеним знаком “-”, найменше число (10). Це означає, що x_{13} = 10.

Число 10 додамо до кількості вантажу в клітинках, позначених знаком ”+”, та віднімемо в клітинках, позначених знаком “-”. Після цього побачимо, що один маршрут ввійшов до плану, а один вийшов з нього. Таким чином, дістали новий план. Він має такий вигляд (табл. 2.7.).

Таблиця 3.7.

a_i	Перевезення для b_i				
	40	30	40	10	$\propto_i$
40	4	-30 1	+10 2	1	0
30	5	+ 2	-20 4	10 2	2
20	20 1	3	3	3	-1
30	20 0	0	10 0	0	-2
β_j	2	1	2	1	

Знов обчислимо значення Δ_{ij}, серед яких від'ємним буде те, яке відповідає маршруту (2,3):

$\Delta_{2,3} = 2 - 1 + 2 = -1$

Побудуємо цикл зміни плану. Він відображений на схемі транспортної задачі. Дістанемо оптимальний розв'язок, який має такий вигляд (табл. 2.8.).

Таблиця 3.8.

a_i	Перевезення для b_i				
	40	30	40	10	$\propto_i$
40	4	10 1	30 2	1	0
30	5	20 2	4	10 2	1
20	20 1	3 3	3	3	-1
30	20 0	0 0	10 0	0	-2
β_j	2	1	2	1	

$Z = 10 \times 1 + 30 \times 2 + 20 \times 2 + 10 \times 2 + 20 \times 1 + 20 \times 0 + 10 \times 0 = 150$.

Завдання 3.5.

Розв'язати транспортну задачу:

3.17. a = (20, 10, 15, 8);
b = (10, 17, 15, 6);

$$C = \begin{pmatrix} 10 & 11 & 12 & 16 \\ 21 & 15 & 13 & 18 \\ 17 & 10 & 11 & 12 \\ 14 & 13 & 9 & 8 \end{pmatrix}$$

3.18. a = (8, 7, 15, 15);
b = (6, 9, 20, 22);

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

3.19. a = (80, 40, 80, 80);
b = (90, 60, 20, 40);

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

4. Нелінійне програмування з обмеженнями у вигляді рівностей нерівностей. Умови КУНА - ТАККЕРА

Кун і Таккер побудували необхідні і достатні умови оптимальності для задач нелінійного програмування, виходячи з припущення про диференційованість функцій f, g_j і h_j.

Знайти сукупність векторів X^*, U^*, V^*, задовольняючих наступні умови:

$$\nabla L(X^*, U^*, V^*) \equiv \nabla f(X^*) - \sum_{j=1}^{m} U_j \nabla h_j(X^*) - \sum_{j=m+1}^{p} U_j^* \nabla g_j(X^*) = 0, \quad (4.1)$$

$$h_j(X^*) = 0, j = 1, ..., m, \quad (4.2)$$

$$g_j(X^*) \geq 0, j = m+1, ..., p, \quad (4.3)$$

$$U_j^* g_j(X^*) = 0, j = m+1, ..., p, \quad (4.4)$$

$$U_j^* \geq 0, j = m+1, ..., p, \quad (4.5)$$

де $L(X, U, V)$ - узагальнена функція Лагранжа; U, V - множники Лагранжа.

Достатньою умовою того, що обмеження - першого порядку, є те, що всі градієнти активних обмежень нерівностей ($g_j(X^*) = 0$) і обмежень-рівностей, взятих у деякій довільній точці X^*, лінійно незалежними. Умова лінійної незалежності являє собою деяку умову регулярності припустимої області, яка майже завжди виконується для задач оптимізації, які зустрічаються на практиці.

Перевірка умов лінійної незалежності являє складність, тому що потрібно, щоб оптимальне рішення задачі було відомо раніше.

Теорема 1.

Якщо функції $f(X), h_j(X)\ (j = 1,...,m)$,

$g_j(X)\ (j = m+1,...,p)$ диференційовані в точці X^* і якщо у точці X^* обмеження являються обмеженнями першого порядку, то необхідною умовою наявності в точці X^* локального мінімуму задачі нелінійного програмування полягає у тому, що існують множники Лагранжа U^* і V^* такі, що сукупність векторів X^*, U^*, V^* задовільняє співвідношенням (4.1) - (4.5).

Умови Куна - Таккера стають достатніми умовами наявності глобального мінімуму, якщо цільова функція опукла, обмеження - нерівності утворенні вгнутими функціями, а обмеження - рівності мають лінійні функції. Розглянемо слідуючу задачу нелінійного програмування.

Приклад 4.1

Мінімізувати $f(X) = X_1^2 - X_2$ при обмеженнях

$g_1(X) = X_1 - 1 \geq 0,\quad g_2(X) = 26 - X_1^2 - X_2^2 \geq 0,\quad h_1(X) = X_1 + X_2 - 6 = 0$.

За допомогою теореми доведемо, що рішення $X^* = [\;\; 5]^T$ - оптимальне. Маємо

$$\nabla f(X) = \begin{bmatrix} 2X_1 \\ -1 \end{bmatrix} \quad \text{і} \quad H_f(X) = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}.$$

Тому що матриця $H_f(X)$ - додатньо напіввизначена при всіх X, функція $f(X)$ виявляється опуклою. Перше обмеження - нерівність має лінійну функцію $g_1(X)$, яка одночасно являється і опуклою, і вгнутою. Для того щоб показати, що функція $g_2(X)$ - вгнута, обчислимо

$$\nabla g_2(X) = \begin{bmatrix} -2X_1 \\ -2X_2 \end{bmatrix} \quad \text{і} \quad H_{g_2}(X) = \begin{bmatrix} -2 & 0 \\ 0 & -2 \end{bmatrix}.$$

Оскільки матриця $H_{g_2}(X)$ від′ємно визначена, функція $g_2(X)$ - увігнута. Функція $h_1(X)$ входить у лінійне обмеження - рівність. Отже, всі умови теореми виконані; якщо покажемо, що $X^* = [1 \quad 5]^T$ - точка Куна - Таккера, встановимо оптимальність рішення X^*.

Наприклад, умова Куна - Таккера має вигляд:

$$2X_1 - U_1 + 2X_1U_2 - V_1 = 0, \tag{4.6}$$

$$-1 + 2X_2U_2 - V_1 = 0, \tag{4.7}$$

$$X_1 - 1 \geq 0, \tag{4.8}$$

$$26 - X_1^2 - X_2^2 \geq 0, \tag{4.9}$$

$$X_1 + X_2 - 6 = 0, \tag{4.10}$$

$$U_1(X_1 - 1) = 0, \tag{4.11}$$

$$U_2(26 - X_1^2 - X_2^2) = 0 \tag{4.12}$$

$$U_1, U_2 \geq 0. \tag{4.13}$$

Точка $X^* = [1 \quad 5]^T$ задовільняє обмеженням (4.8) - (4.10) і, отже, - припустима. Рівняння (4.6) - (4.7) приймають наступний вигляд:

$$2 - U_1 + 2U_2 - V_1 = 0; \qquad -1 + 10U_2 - V_1 = 0.$$

Припустивши $V_1 = 0$, отримаємо $U_2 = 0{,}1$; $U_1 = 2{,}2$. Рішення $X^* = [1 \quad 5]^T$, $U^* = [2{,}2 \quad 0{,}1]$ і $V^* = 0$ задовільняють умовам Куна - Таккера.

Щоб врахувати нелінійний характер функції в задачах , Мак - Кормік сформулював необхідні і достатні умови оптимальності другого порядку для задач з двічі диференційованими функціями.

Точка Куна - Таккера в задачі нелінійного програмування визначається вектором (X^*, U^*, V^*) задовільняючим співвідношенням (4.1) - (4.5).

Точка X^* являє собою точку локального мінімуму в задачі нелінійного програмування, якщо X^* відповідає припущеному рішенню і $f(X^*) \leq f(\overline{X})$ для всіх припущених $\overline{X}$ з деякої малої околиці $\delta(X^*)$ точки X^*

Точка X^* являє собою точку строгого (єдиного або ізольованого) локального мінімуму, якщо X^* відповідає припущеному рішенню задачі і $f(X^*) < f(\overline{X})$ для всіх припустимих $\overline{X} \neq X^*$ і деякої малої околиці $\delta(X^*)$ точки X^*

Теорема 2. Необхідні умови оптимальності другого порядку.

Нехай $f(X)$, $h_1(X),...,h_m(X)$, $g_{m+1}(X),...,g_p(X)$ - двічи диференційовані функції, X^* - припустиме рішення нелінійної задачи, а множина активних в точці X^* обмежень визначається множиною

$I = \{j \mid g_j(X^*) = 0\}$. Припустимо, що градієнт $\nabla g_j(X^*)$ при $j = m+1,...,p$ і $\nabla h_j(X^*)$ при $j = 1,...,m$ - лінійно незалежні. Тоді необхідні умови того, щоб точка X^* була точкою локального мінімума в задачі нелінійного програмування, формулюються наступним чином:

1) існує такий вектор (U^*, V^*) що вектор (X^*, U^*, V^*) визначає точку Куна - Таккера;

2) для будь-якого нульового вектора Y, задовільняючого системі рівнянь

$$\nabla g_j(X^*)Y = 0, \quad j = m+1,...,p, \tag{4.14}$$

$$\nabla h_j(X^*)Y = 0, \quad j = 1,...,m, \tag{4.15}$$

виконується нерівність

$$Y^T H_L(X^*, U^*, V^*)Y \geq 0, \tag{4.16}$$

при

$$L(X;U,V) = f(X) - \sum_{j=1}^{m} V_j h_j(X) - \sum_{j=m+1}^{p} U_j g_j(X), \tag{4.17}$$

а $\mathrm{H}_L\left(X^*,U^*,V^*\right)$ - матриця Гессе, утворена значеннями других частинних похідних L по X у точці $\left(X^*,U^*,V^*\right)$.

Достатні умови Куна - Таккера виконуються при дуже жорстких припущеннях. У зв′язку з цим розглянемо достатні умови другого порядку, формулювання яких не потребує припущень про опуклість та вгнутість відповідних функцій і про лінійність обмежень - рівностей.

Теорема 3. Достатні умови оптимальності другого порядку.

Достатні умови того, щоб точка X^* була точкою строгого локального оптимуму в задачі нелінійного програмування, де $f\left(X\right)$, $g_j\left(X\right), j=m+1,...,p, h_j\left(X\right), j=1,...,m$ - двічі диференційовані функції, формулюються наступним чином:

1) існує такий вектор $\left(U^*,V^*\right)$, що вектор $\left(X^*,U^*,V^*\right)$ визначає точку Куна - Таккера;

2) для кожного ненульового вектора Y, задовільняючого системі рівнянь і нерівностей

$$\nabla g_j\left(X^*\right)Y=0,\ j\in \mathrm{I}_1=\left\{j \middle| g_j\left(X^*\right)=0, U_j^*>0\right\} \tag{4.18}$$

$$\nabla g_j\left(X^*\right)Y=0,\ j\in \mathrm{I}_2=\left\{j \middle| g_j\left(X^*\right)=0, U_j^*=0\right\} \tag{4.19}$$

$$\nabla h_j\left(X^*\right)Y=0,\quad j=1,...,m,\quad Y\neq 0, \tag{4.20}$$

виконується нерівність

$$Y^T\mathrm{H}_L\left(X^*,U^*,V^*\right)Y>0. \tag{4.21}$$

Множина $\mathrm{I}=\mathrm{I}_1\bigcup\mathrm{I}_2$ має номера всіх активних у точці X^* обмежень.

Порівнюючи теореми 2 і 3, нескладно побачити, що достатні умови лише дечим відрізняються від необхідних умов. Основна різниця полягає у тому, що рівняння (4.21) не повинні виконуватися для всіх активних обмежень, а нерівність (4.23) перетворюється у строгу нерівність.

Приклад 4.2

Мінімізувати $f(X) = (X_1 - 1)^2 + X_2^2$ при обмеженні

$g_1(X) = -X_1 + \left(\dfrac{X_2^2}{5}\right) \geq 0$ [13].

Припустимо, потрібно перевірити , чи буде точка $X^* = [1 \quad 0]^T$ точкою локального мінімуму. Запишемо умови Куна - Таккера для данної задачи:

$$2(X_1 - 1) + U_1 = 0;\ 2X_2 - \frac{2}{5}X_2 U_1 = 0;$$

$$U_1\left[-X_1 + \left(\frac{X_2^2}{5}\right)\right] = 0;\ U_1 \geq 0.$$

Точка $X^* = [1 \quad 0]^T$, $U_1^* = 2$ задовільняє умовам Куна - Таккера. Необхідні умови наявності мінімума виконуються, однак не можна стверджувати, що X^* - точка локального мінімума, оскільки функція $g_1(X)$ в суперечь припущень теореми 1 (про достатність умов Куна - Таккера) являється опуклою.

Використовуючи достатні умови другого порядку, знаходимо матрицю

$$H_L(X^*; U^*) = \begin{bmatrix} 2 & 0 \\ 0 & 1{,}2 \end{bmatrix}.$$

Вектор $Y = [Y_1 \quad Y_2]^T$, задовільняючий умовам (4.25) та (4.26)

$\nabla g_1(X^*)Y = 0$, або $[-1 \quad 0]\begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix} = 0$ необхідно розглядати у вигляді $[0 \quad Y_2]^T$.

Нерівність (4.21) приймає вигляд

$$[0 \quad Y_2]\begin{bmatrix} 2 & 0 \\ 0 & 1{,}2 \end{bmatrix}\begin{bmatrix} 0 \\ Y_2 \end{bmatrix} = 1{,}2Y_2^2 > 0 \text{ при всіх } Y_2 \neq 0$$

Таким чином, за теоремою 3 $X^* = [1 \quad 0]^T$ - точка строгого локального мінімума.

Приклад 4.3.

Мінімізувати функцію $f(x) = (x_1 - 1)^2 + x_2^2$ при обмеженні

$g_1(X) = -x_1 + x_2^2 \geq 0$

Відповідь.

$$\nabla f(X) = \begin{bmatrix} \dfrac{\partial f}{\partial x_1} \\ \dfrac{\partial f}{\partial x_2} \end{bmatrix} = \begin{bmatrix} 2(x_1 - 1) \\ 2x_2 \end{bmatrix}$$

$$\nabla g_1(X) = \begin{bmatrix} \dfrac{\partial g_1}{\partial x_1} \\ \dfrac{\partial g}{x_2} \end{bmatrix} = \begin{bmatrix} -1 \\ 2x_2 \end{bmatrix}$$

Умови Куна-Таккера.

$$\frac{\partial L}{\partial x_1} = 2(x_1 - 1) + U_1 = 0$$

$$\frac{\partial L}{\partial x_2} = 2x_2 - 2x_2 U_1 = 0$$

$$U_1(-x_1 + x_2^2) = 0, \quad U_1 \geq 0$$

Достатні умови оптимальності першого порядку

$$\nabla^2 f(X) = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} > 0$$

$$\nabla^2 g_1(X) = \begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix} \geq 0$$

не виконуються і точка $[0 \quad 0]$ не є точкою глобального оптимуму.

Перевіримо умови оптимальності другого порядку

$$H_L(X, U) = \begin{bmatrix} 2 & 0 \\ 0 & 2 - 2U_1 \end{bmatrix}$$

$$H_L(X^*, U^*) = \begin{bmatrix} 2 & 0 \\ 0 & -4 \end{bmatrix}$$

Перевіримо виконання нерівності

$$y^T \begin{bmatrix} 2 & 0 \\ 0 & -4 \end{bmatrix} y \geq 0$$

для всіх у, що задовольняють

$\nabla^2 g_1(X^*)У = 0$ чи $\begin{bmatrix} -1 & 0 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = 0$

Маємо

$\begin{bmatrix} 0 & y_2 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & -4 \end{bmatrix} \begin{bmatrix} 0 \\ y_2 \end{bmatrix} = -4y_2^2 < 0$ для всіх $y_2 \neq 0$

Точка $X^* = [0 \quad 0]$ не є точкою локально мінімума.

Приклад 4.4

Мінімізувати $f(X) = (x_1 - 1)^2 + x_2^2$при обмеженні $g_1(X) = -x_1 + \frac{x_2^2}{5} \geq 0.$

Перевірити точку $X^* = [0 \quad 0]$ на наявність локального мінімума.

Відповідь. Відмітимо, що область $S = \{X | g_1(X) \geq 0\}$ не є вгнутою множиною

Умови Куна-Таккера

$$\frac{\partial L}{\partial x_1} = 2(x_1 - 1) + U_1 = 0$$

$$\frac{\partial L}{\partial x_2} = 2x_2 - \frac{2}{5} x_2 U_1 = 0$$

$$U_1 \left(-x_1 + \left(\frac{x_2^2}{5}\right) \right) = 0, \quad U_1 \geq 0$$

Точка $X^* = [0 \quad 0]$, $U_1 = 2$ задовольняє умови Куна-Таккера, але функція $g_1(X)$ є опуклою. Тому використаємо достатні умови оптимальності другого порядку

$$H_L(X^*, U^*) = \begin{bmatrix} 2 & 0 \\ 0 & 1{,}2 \end{bmatrix}$$

Перевіримо нерівність

$\begin{bmatrix} 0 & y_2 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 1{,}2 \end{bmatrix} \begin{bmatrix} 0 \\ y_2 \end{bmatrix} = 1{,}2y_2^2 > 0$ для всіх $y_2 \neq 0$

Таким чином $X^* = [0 \quad 0]$ – точка строгого локального мінімуму.

Приклад 4.5

Мінімізувати $f(X) = 1 - x^2$ при обмеженні$-1 \leq x \quad 3.$

Відповідь.

$g_1(X) = x + 1 \geq 0,$

$g_2(X) = 3 - x \geq 0,$

Умови Куна-Таккера

$L = 1 - x^2 - U_1(x+1) - U_2(3-x)$

$\frac{\partial L}{\partial x} = -2x - U_1 + U_2 = 0$

$U_1(x+1) = 0$

$U_2(3-x) = 0$

$U_1, U_2 \geq 0$

Для точки $x = 3$

$U_1 = 0$

$U_2 > 0, \quad U_2 = 6$

Тобто для точки $x = 3$ виконуються необхідні умови оптимальності першого порядку. Але $\frac{\partial^2 f}{\partial x^2} = -2$, тобто функція $f(X) = 1 - x^2$ є вгнутою, а не опуклою.

Точка $x = 3$ не є точкою мінімуму.

Приклад 4.6

Для функції $f(X) = x_1^2 - x_2$ і обмежень $x_1 - 1 \geq 0; 26 - x_1^2 - x_2^2 \geq 0,$

$x_1 + x_2 - 6 = 0$ записати умови Куна-Таккера і перевірити точку

$X^* = [1 \quad 5]^T$

ВІдповідь.

$L(X^*, U^*, V^*) = x_1^2 - x_2 - V(x_1 + x_2 - 6) - U_1(x_1 - 1) - U_2(26 - x_1^2 - x_2^2)$

Умови Куна-Таккера

$$\frac{\partial L}{\partial x_1} = 2x_1 - V - U_1 + 2U_2x_1 = 0$$

$$\frac{\partial L}{\partial x_2} = -1 - V + 2U_2x_2 = 0$$

$$x_1 - 1 \geq 0$$

$$26 - x_1^2 - x_2^2 = 0$$

$x_1 + x_2 - 6 = 0$

$$U_1(x_1 - 1) = 0$$

$$U_2(26 - x_1^2 - x_2^2) = 0$$

$U_1, U_2 \geq 0$

Підставивши координати т. X^* в умови додаткової нежорсткості (6)-(7), бачимо, що

$U_1, U_2 \geq 0$

Рівняння набувають вигляду

$2 - U_1 + U_2 - V_1 = 0,$

$-1 + U_2 - V_1 = 0$

Припустивши V_1=0, отримаємо $U_2 = 0{,}1; U_1 = 2{,}2$

Розвя'зок $X^* = [1 \quad 5]^T$ задовольняє умовам Куна-Таккера.

Приклад 4.7

Мінімізувати функцію $f(x) = 4(x_1 - 1)^2 + 3x_2^2$ при обмеженні

$g_1(X) = -x_1 + x_2^2 \geq 0$

і перевірити точку $X^* = [0 \quad 0]$ на наявність оптимуму.

Відповідь.

$$\nabla f(X) = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \end{bmatrix} = \begin{bmatrix} 8(x_1 - 1) \\ 6x_2 \end{bmatrix}$$

$$\nabla g_1(X) = \begin{bmatrix} \frac{\partial g_1}{\partial x_1} \\ \frac{\partial g}{x_2} \end{bmatrix} = \begin{bmatrix} -1 \\ 2x_2 \end{bmatrix}$$

Умови Куна-Таккера.

$$\frac{\partial L}{\partial x_1} = 8(x_1 - 1) + U_1 = 0$$

$$\frac{\partial L}{\partial x_2} = 6x_2 - 2x_2 U_1 = 0$$

$U_1(-x_1 + x_2^2) = 0, \quad U_1 \geq 0$

Достатні умови оптимальності першого порядку

$$\nabla^2 f(X) = \begin{bmatrix} 8 & 0 \\ 0 & 6 \end{bmatrix} > 0$$

$\nabla^2 g_1(X) = \begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix} \geq 0$

не виконуються і точка $[0 \quad 0]$ не є точкою глобального оптимуму.

Перевіримо умови оптимальності другого порядку

$H_L(X, U) = \begin{bmatrix} 8 & 0 \\ 0 & 6 - 2U_1 \end{bmatrix}$

$H_L(X^*, U^*) = \begin{bmatrix} 8 & 0 \\ 0 & -10 \end{bmatrix}$

Перевіримо виконання нерівності

$y^T \begin{bmatrix} 8 & 0 \\ 0 & -10 \end{bmatrix} \text{У} \geq 0$

для всіх у, що задовольняють

$\nabla^2 g_1(X^*)\text{У} = 0$ чи $[-1 \quad 0] \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = 0$

Маємо

$[0 \quad y_2] \begin{bmatrix} 8 & 0 \\ 0 & -10 \end{bmatrix} \begin{bmatrix} 0 \\ y_2 \end{bmatrix} = -10y_2^2 < 0$ для всіх $y_2 \neq 0$

Точка $X^* = [0 \quad 0]$ не є точкою локально мінімума.

5. Методи штрафних функцій

В основу методів штрафних функцій в області нелінійного програмування покладена ідея перебудови загальної нелінійної задачі) в послідовність задач без обмежень шляхом додавання до цільовій функції однієї або кількох функцій, які встановлюють обмеження для того, щоб обмеження, як такі, в задачі оптимізації не фігурували. В цьому випадку мінімізація може здійснюватись за допомогою більш простих алгоритмів. При використанні методів штрафних функцій отримується оптимальний ефект за рахунок постійного компромісу між необхідністю задоволення обмежень і процесом мінімізації f(x), який досягається шляхом присвоєння належної ваги цільової функції і функціям, які визначають обмеження. Методи штрафних функцій можна поділити на 2 класи: параметричні і непараметричні методи. Параметричні методи характеризуються наявністю одного чи кількох певним чином підібраних параметрів, які входять у склад штрафної функції в якості вагових коефіцієнтів. Параметричні методи розпадаються на 3 категорії: методи внутрішньої точки; методи зовнішньої точки; комбіновані методи. Штрафна функція не дозволяє вектору X надто відходити від межі припустимої області. Формально перебудова задачі нелінійного програмування, в задачу мінімізації обмежень проводиться шляхом переходу до задачі мінімізації:

$$P(X^{(к)}, P^{(к)}) = f(X^{(к)}) + \sum_{i=1}^{m} \rho_i^{(к)} H(h_i(X^{(к)})) + \sum_{i=m+1}^{p} \rho_i^{(к)} G(g_i(X^{(к)})) \qquad (5.1)$$

де $P(X^{(к)},\square^{(к)})$ - узагальнена приєднана функція чи штрафна функція; $\rho_i^{(к)} \geq 0$ - вагові коефіцієнти; К - кількість завершених етапів обчислювального оптимізаційного процесу. При будь-якому виборі функціоналів $H(h_i(X))$ і $G(g_i(X))$ потрібно, щоб:

$$\left.\begin{aligned} &\lim_{к \to \infty} \sum_{i=m+1}^{p} \rho_i^{(к)} G(g_i(X^{(к)})) = 0 \\ &\lim \sum_{i=1}^{m} \rho_i^{(к)} H(h_i(X^{(к)})) = 0 \\ &\lim \left| p(X^{(к)}, \rho^{(к)}) - f(X^{(к)}) \right| = 0 \end{aligned}\right\} \qquad (5.2)$$

В міру розвитку процесу оптимізованого пошуку вплив входящих в $P(X^{(к)},\rho^{(к)})$ функцій обмежень на значення даної приєднаної функції постійно слабне, а в межі повністю зникає. Значить, екстремум P(X) співпадає з екстремумом f(X).

5.1. Основні типи штрафів

Для обліку обмежень-рівностей часто використовують квадратичний штраф (Рис.5.1):

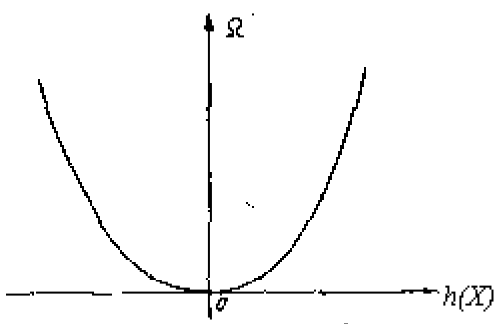

Рис.5.1. Квадратичний штраф

При мінімізації цей штраф запобігає відхиленню значення h(X) від нуля.

При врахуванні обмежень-нерівностей використовують різні типи штрафів. Простішим серед них є нескінченний бар'єр, показаний на рис.5.2.

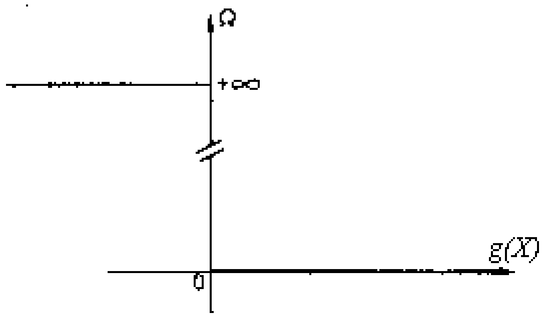

Рис.5.2. Нескінчений штраф

Відповідний вираз приймає нескінченно великі значення в неприпустимих точках і нульове значення - в припустимих. У даному випадку штрафна функція P(X,ρ) - розривна на межі припустимої області. В машинній реалізації нескінченних штрафів використовують велике додатне число. Наприклад, цей штраф використовується у формулі

$$\Omega = 10^{20}\sum_{j=g}\left|g_i(X)\right| \qquad (5.3)$$

де g - множина індексів порушень обмежень, $g_j(X)<0$ при $j\in g$. Іншим широко використовуваним типом штрафу є логарифмічний центр (рис.5.3).

$$\Omega = -\rho\ln\left[g(X)\right] \qquad (5.4)$$

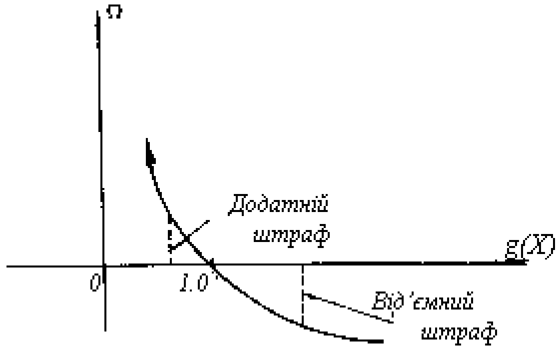

Рис.5.3. Логарифмічний штраф

Цей штраф додатній при всіх X таких, що $0<g(X)<1$ і від'ємний при $g(X)>1$. Логарифмічний штраф - бар'єрна функція, не визначена в неприпустимих точках (тобто для X таких, що g(X)<0). Ітераційний процес починається з припустимої точки початкової при додатному значенні ρ. Після рішення кожної підзадачі безумовної мінімізації параметр ρ зменшується і в межах прагне до нуля. Штраф, заданий оберненою функцією рис.5.4.

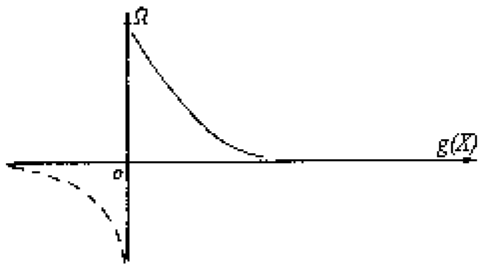

Рис.5.4. Штраф, заданий оберненою функцією

$$\Omega = \rho\left[\frac{1}{g(X)}\right] \qquad (5.5)$$

Штраф, заданий оберненою функцією, не має від'ємних значень в припустимій області. Цей штраф є бар'єром; можливі труднощі, які пов'язані з появою неприпустимих точок.

5.2 Метод поступової безумовної мінімізації

Алгоритм методу поступової безумовної мінімізації (МПБМ), розвинутий Фіакко і Мак-Корміком, який застосовується для рішення задачі нелінійного програмування вигляду, в якій f(X) і g(X) (i=m+1,..., ρ) можуть бути нелінійними функціями незалежних змінних, а $h_i(X)$ (i=1,...,m) повинні бути лінійними функціями незалежних змінних. При таких умовах гарантується збіжність послідовності проміжних рішень до оптимального рішення задачі нелінійного програмування.. В варіанті метода 1967 р. задачу перетворюють в послідовність задач без обмежень, використовуючи штраф, заданий оберненою функцією і шляхом побудови P- функції вигляду

$$P(X^{(к)}, r^{(к)}) = f(X^{(к)}) + (r^{(к)})^{-\frac{1}{2}}\sum_{i=1}^{m} h_i^2(X^{(к)}) + r^{(к)}\sum_{i=m+1}^{\rho}\frac{1}{g_i(X^{(к)})} \qquad (5.6)$$

де значення вагових коефіцієнтів r додатні і утворюють монотонно зменшуючу послідовність $\{r \mid r^{(0)}>r^{(1)}>...>0\}$.

У варіанті МПБМ 1970 р. застосовується логарифмічний штраф

$$P(X^{(\kappa)}, r^{(\kappa)}) = f(X^{(\kappa)}) + \frac{1}{r^{(\kappa)}}\sum_{i=1}^{m} h_i^2(X^{(\kappa)}) - r^{(\kappa)}\sum_{i=m+1}^{\rho} \ln g_i(X^{(\kappa)}) \qquad (5.7)$$

Як і у варіанті (5.6) тут використовується квадратичній штраф для урахування обмежень-рівностей. Процедура мінімізації функцій (5.6) і (5.7) починається з внутрішньої (або граничної) точки, тобто з точки $X^{(0)}$, в якій всі граничні умови у вигляді нерівностей задоволені. Після визначення $r^{(0)}$ точка $X^{(1)}$ розраховується мінімізацією $P(X, r^{(0)})$. Потім визначається $r^{(1)}$ і знаходиться $X^{(2)}$ і т.д. Швидкість збіжності залежить від початкового вибору $r^{(0)}$ і від способу редуцирування даного параметру. Одним з способів вибору початкового значення $r^{(0)}$, який було запропоновано Фіакко і Мак-Корміком, є $r^{(0)}=1$. При цьому вказано, що ефективність алгоритму не зміниться значно, якщо послідовність $r^{(1)}$, $r^{(2)}$,..., $r^{(\kappa)}$ індуціювати простим співвідношенням

$$r^{(\kappa)} = \frac{r^{(\kappa-1)}}{C}$$

де C>1 є константою (завжди вважають C=4).

Завдання

Приклад 5.1. Мінімізувати $f(X) = 100(X_2 - X_1^2)^2 + (1 - X_1)^2$ при обмеженнях:

$$g_1(X)=X_1+1\geq 0;$$

$$g_2(X)=1-X_2\geq 0;$$

$$g_3(X)=4X_2-X_1-1\geq 0;$$

$$g_4(X)=1-0,5X_1-X2\geq 0;$$

Записати для цієї задачі штрафні функції, використовуючи: логарифмічний штраф; штраф, який задано оберненою функцією.

Відповідь.

$$P(X,r) = 100(x_2 - x_1^2)^2 + (1 - x_2)^2 - r(\ln(x_1 + 1) + \ln(1 - x_2) + \\ +\ln(4x_1 - x_2 - 1) + \ln(1 - 0{,}5x_1 - x_2)) \qquad ;$$

$$P(X,r) = 100(x_2 - x_1^2)^2 + (1 - x_2)^2 + r\left[\frac{1}{x_1 + 1} + \frac{1}{1 - x_2} + \frac{1}{4x_1 - x_2 - 1} + \frac{1}{1 - 0{,}5x_1 - x_2}\right]$$

Спочатку

задається $r^{(0)} = 1$ і відбувається мінімізація штрафної функції ефективним методом нелінійного програмування. Потім штрафний параметр зменшується $r^{(k)} = \frac{r^{(k-1)}}{C}, \quad C = 4.$

Знову відбувається процес мінімізації штрафної функції. Ітераційний процес відбувається доти, поки не виконується умова збіжності.

Приклад 5.2. В задачі

$f(X) = 100(x_1 - 4)^2 + (x_2 - 4)^2 \rightarrow min$

При обмеженні $h(X) = x_1 + x_2 - 5 = 0$ ввести штрафну функцію використовуючи квадратичний штраф. Пояснити алгоритм реалізації методу послідовної безумовної мінімізації.

Відповідь . Штрафна функція

$P(X,r) = (x_1 - 4)^2 + (x_2 - 4)^2 + (\frac{1}{r})(x_1 + x_2 - 5)^2$, де r – штрафний параметр.

Спочатку задається $r^{(0)} = 1$ і відбувається мінімізація штрафної функції ефективним методом нелінійного програмування. Потім штрафний параметр зменшується $r^{(k)} = \frac{r^{(k-1)}}{C}, \quad C = 4.$

Знову відбувається процес мінімізації штрафної функції. Ітераційний процес відбувається доти, поки не виконується умова збіжності.

Завдання для самостійної роботи.

1. Мінімізувати функцію $f(X) = -X_1^2 - X_2^2$ при обмеженнях

$$X_1 \geq 0;$$

$$X_2 \geq 0;$$

$$X_1 + 2X_2 \leq 3.$$

2. Використовуючи МПБМ, мінімізувати функцію $f(X) = \frac{4}{X_1} + \frac{9}{X_2} + (X_1 + X_2)$ при обмеженнях

$$X_1 \geq 0;$$

$$X_2 \geq 0;$$

$$X_1 + 2X_2 \leq 4.$$

3. Мінімізувати функцію $f(X) = X_1^2 + X_2^2 + X_3^2$ при обмеженнях

$$X_1 + X_2 + X_3 \geq 3;$$

$$X_1 X_2 X_3 \geq 3;$$

$$X_1 > 0; X_2 \geq 0; X_3 \geq 0;$$

4. Мінімізувати функцію $f(X) = (X_1 - X_2)^2 + (X_2 - 1)^2$ при обмеженнях

$$y_1(X) = \frac{X_1^2}{4} + X_2^2 + 1 \geq 0,$$
$$h_2(X) = X_1 - 2X_2 + 1 = 0$$

а) побудувати функцію P(X, r) при r=0,04. Зобразити побудовану P-функцію графічно;

б) визначити напрямок оптимізаційного пошуку з початкової внутрішньої точки

$X^{(n)} = [0{,}748 \;\; 0{,}548]^T$;

в) знайти вектор $X^{(1)}$;

г) Чи може одна з точок $X^{(к)}$ виявитися зовні припустимої області.

6. Застосування градієнтних методів в однокрокових процедурах вибору рішень

Розглянемо задачу мінімізації функції вартості

$$J = F(X,U) \tag{6.1}$$

шляхом підбору значення вектора **u** при обмеженні у формі наступної рівності:

$$f(X,U) = 0 \tag{6.2}$$

Тут X – вектор стану, f – n-вектор, U – m-вектор, а F – скалярна функція. Сформульована проблема є задачею нелінійного програмування чи задачею однокрокового рішення. Можна ввести

$$L(X,\lambda,U) = F(X,U) + \lambda^T f(X,U) \tag{6.3}$$

Розрахуємо та прирівняємо нульовому вектору градієнт виразу (6.3) відносно вектора U:

$$\frac{\partial L}{\partial U} = \frac{\partial F(X,U)}{\partial U} + \left[\frac{\partial f^T(X,U)}{\partial U}\right]\lambda = 0 \tag{6.4}$$

Таким самим чином можна отримати співвідношення

$$\frac{\partial L}{\partial X} = \frac{\partial F(X,U)}{\partial X} + \frac{\partial f^T(X,U)}{\partial X}\lambda = 0 \tag{6.5}$$

Останні дві рівності є першою необхідною умовою мінімуму функції J. В доповнення до цього повинна виконуватись також і наступна умова:

матриця

$$\begin{bmatrix} \frac{\partial}{\partial X}\frac{\partial L}{\partial X} & \frac{\partial}{\partial U}\frac{\partial L}{\partial X} \\ \left[\frac{\partial}{\partial U}\frac{\partial L}{\partial X}\right]^T & \frac{\partial}{\partial U}\frac{\partial L}{\partial U} \end{bmatrix} \tag{6.6}$$

повинна бути невід’ємно визначеною впродовж “траєкторії” f(X,U)=0. В загальному випадку взагалі не вдається розв’язати систему нелінійних алгебраїчних рівнянь. Так як, при f(X,U)=0 маємо J=L, то поклавши ∂L/∂X=0, приблизно одержимо

$$\Delta J = \left[\frac{\partial L}{\partial U}\right]^T \Delta U \tag{6.7}$$

Якщо зараз необхідно забезпечити найбільшу зміну ΔJ значення функції J, то можна розрахувати градієнт ∂L/∂U та нове значення J знайти після введення приросту ΔU, напрямок якого протилежний одержаному градієнту:

$$\Delta U = -K\left[\frac{\partial L}{\partial U}\right] \tag{6.8}$$

В результаті маємо, що зміна значення функції визначається виразом

$$\Delta J = -K\left[\frac{\partial L}{\partial U}\right]^T\left[\frac{\partial L}{\partial U}\right] \qquad (6.9)$$

Ця процедура розрахунку починається з неоптимального вектора керування U^N. Для цього вектора керування треба знайти значення x^N вектора X, яке задовольняє рівності, тобто

$$f(X^N, U^N) = 0 \qquad (6.10)$$

Далі з (6.5) для забезпечення виконання умови $\partial J/\partial x^N=0$ знаходимо значення невизначеного множника

$$\lambda^N = -\left[\frac{\partial f^T(X^N, X^N)}{\partial X^N}\right]^{-1}\frac{\partial F(X^N, U^N)}{\partial X^N} \qquad (6.11)$$

Нарешті, на основі (6.3) розрахуємо градієнт

$$\frac{\partial L}{\partial U^N} = \frac{\partial f(X^N, U^N)}{\partial U^N} + \left[\frac{\partial f^T(X^N, U^N)}{\partial X^N}\right]\lambda^N \qquad (6.12)$$

Цей вектор буде нульовий тільки в тому випадку, коли виконується відмічена вище умова мінімуму. Одержане значення градієнта зараз дозволяє знайти за формулою (6.8) такий приріст вектора U, при якому забезпечується найшвидший спуск, а саме

$$\Delta U^N = -K\left[\frac{\partial L(X^N, \lambda^N, U^N)}{\partial U^N}\right] \qquad (6.13)$$

Нове значення вектора U, таким чином, визначається за формулою

$$U^{N+1} = U^N + \Delta U^N \qquad (6.14)$$

В свою чергу дане значення дозволяє знайти нові значення X^{N+1}, λ^{N+1} векторів X, λ і так далі. Ці розрахунки повторюються до тих пір, доки зміни значення функції вартості

$$\Delta J^N = -K\left[\frac{\partial L(X^N, \lambda^N, U^N)}{\partial U^N}\right]^T\left[\frac{\partial L(X^N, \lambda^N, U^N)}{\partial U^N}\right] \qquad (6.15)$$

на двох сусідніх кроках даного ітераційного процесу не буде менше деякої заданої малої величини.

Приклад 6.1

З метою ілюстрації послідовності розрахунків при використанні даного градієнтного методу розглянемо задачу мінімізації функції $J=X^2+U^2$ при обмеженні у формі рівності XU=1. Ця задача може бути легко вирішена аналітично. Вираз (6.3) тут має вигляд $L=X^2+U^2+\lambda(XU-1)$, а рівняння $\partial H/\partial U=\partial H/\partial X=0$ призводять до рішення $U=\pm1$, $X=\pm1$, $\lambda=-2$, $J_{min}=2$. Припустимо, що рішення цієї задачі невідоме. Виберемо початкове значення $U= U^0$. Згадані вище рівняння з (6.10) по (6.15) для даного прикладу приймають наступний вигляд

$$f(X^N,U^N)=0=X^NU^N-1,$$

$$\frac{\partial L}{\partial X}=2X+\lambda U=0; \qquad \lambda^N=-\left[\frac{\partial f^T}{\partial X^N}\right]^{-1}\frac{\partial F}{\partial X^N}=-\frac{2X^N}{U^N},$$

$$\frac{\partial L}{\partial U^N}=\frac{\partial F}{\partial U^N}+\left[\frac{\partial f^T}{\partial U^N}\right]\lambda^N=2U^N+X^N\lambda^N,$$

$$\Delta U^N=-K\frac{\partial L}{\partial U^N}=-K(2U^N+X^N\lambda^N)=-2K\left[U^N-\frac{1}{(U^N)^3}\right],$$

$$\Delta J^N=-K\left[\frac{\partial L}{\partial U^N}\right]^T\left[\frac{\partial L}{\partial U^N}\right]=-K(2U^N+X^N\lambda^N)^2,$$

$$U^{N+1}=U^N-\Delta U^N$$

Таким чином для даної задачі послідовні значення керування одержуються шляхом рішення різницевого рівняння

$$U^{N+1}=U^N-2K\left[U^N-\frac{1}{(U^N)^3}\right]$$

Послідовні етапи обчислень при використанні приведеної вище градієнтної процедури “першого порядку” полягають в наступному:

1. Обчислити чергове значення вектора U^i.
2. З рівняння $f(X^i, U^i)=0$ знаходимо X^i.
3. Вектор λ^i визначаємо за формулою

$$\lambda^i=-\left[\frac{\partial f^T(X^i,U^i)}{\partial X^i}\right]^{-1}\frac{\partial F(X^i,U^i)}{\partial X^i}$$

4. Знаходимо похідні

$$\frac{\partial L(X^i, U^i, \lambda^i)}{\partial U^i} = \frac{\partial F(X^i, U^i)}{\partial U^i} + \frac{\partial f^T(X^i, U^i)}{\partial X^i}\lambda^i$$

5. Обчислюємо чергове значення вектора

$$U^{i+1} = U^{(i)} - K^i\left[\frac{\partial L(X^i, U^i, \lambda^i)}{\partial U^i}\right]$$

6. Ця послідовність обчислень повторюється, якщо значення керування U змінюється від ітерації до ітерації і обчислення припиняються, коли значення U практично не змінюється.

Приклад 6.2.

1. Розв’язати задачу статичної оптимізації технологічного процесу за допомогою методу нелінійного програмування. Для технологічного процесу сформулювати задачу статичної оптимізації і вибрати критерій оптимальності.
2. Визначити обмеження, які накладаються на процес.
3. Вибрати математичну модель статики чи статичного режиму технологічного процесу.
4. Реалізувати один з методів нелінійного програмування для пошуку оптимуму задачі статичної оптимізації.

Біла сажа – неорганічна сполука, для виробництва якої використовують силікат натрію $\mathrm{Na_2O} \cdot mSiO_2$ (F1), хлористий кальцій $CaCl_2$ (F2), та хлореводневу кислоту (F3).

Отримання суспензії білої сажі відбувається в два етапи. На першому в реакторі 1 з розчинів силікату натрію та хлористого кальцію одержують напівпродукт – силікат кальцію $\mathrm{CaO} \cdot mSiO_2 \cdot nH_2O$, що поступає в реактор 3, в який через ротаметр 2 подається розчин хлороводневої кислоти HCl з напірного бачка 4. На другому етапі в реакторі 3 отримують суспензію білої сажі за рівнянням:

$\mathrm{CaO} \cdot mSiO_2 \cdot nH_2O + 2HCl = mSiO_2 \cdot nH_2O + CaCl_2 + H_2O$

Даний апарат є об’єктом моделювання. Його схему зображено на рис. 6.1.

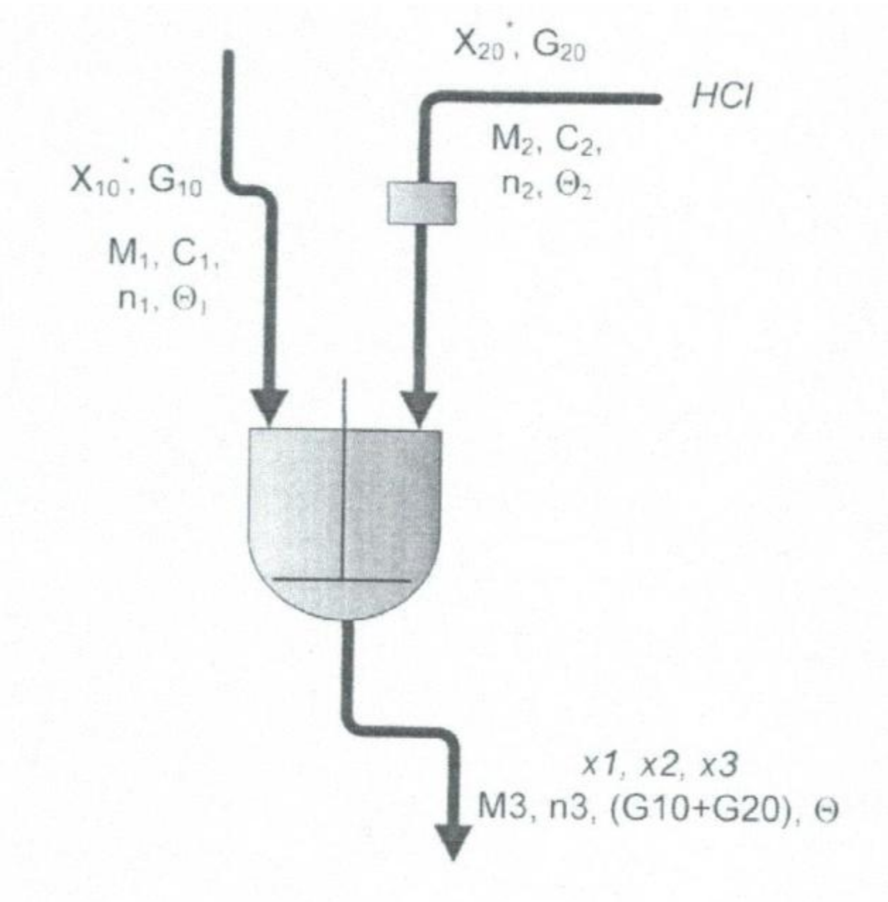

Рис.6.1 - Схема хімічного реактору

На цій схемі позначені такі технологічні параметри:

X_{10}^*, X_{20}^* - початкові концентрації компонентів 1 і 2 у потоках G_{10} і G_{20} відповідно; M_1, M_2, M_3 – молекулярні маси компонентів 1, 2 і 3; n_1, n_2, n_3 – стехіометричні коефіцієнти реакції; C_1, C_2, C_3, - питомі теплоємності потоків; Θ_1, Θ_2, Θ – температура потоків і в апараті; G_1, G_2, G_3 – кількість речовини 1 і 2, що прореагували і утворилося G_3.

Розробка математичної моделі статичного режиму об'єкта

Припущення для об'єкта:

1) В об'єкті розглянемо такі акумулюючи ємності, як елементарний об'єм потоку у реакторі;
2) Хімічні реакції в апараті присутні.
3) Апарат геометричний.
4) Об'єкт теплоізольований.
5) Витрати на входах дорівнюють витратам на виходах.
6) Об'єкт із зосередженими параметрами.

Статичний режим об'єкта представлений рівнянням

$$-x_3 \cdot (G_{10} + G_{20}) + k \cdot \left[\frac{G_{20} \cdot x_{10}}{G_{10} + G_{20}} - \frac{n_1 \cdot M_1}{n_3 \cdot M_3} \cdot x_3 \right] \cdot e^{-\frac{E}{R\Theta}} \cdot V \cdot \rho = 0 \qquad (6.16)$$

Виражаємо вихідну концентрацію, і отримуємо статичну модель реактора:

$$x_3 = \frac{k \cdot e^{-\frac{E}{R\Theta}} \cdot V \cdot \rho \cdot \ G_{20} \cdot x_{10}}{\left[G_{10} + G_{20} \ \ + k \cdot e^{-\frac{E}{R\Theta}} \cdot V \cdot \rho \cdot \frac{n_1 \cdot M_1}{n_3 \cdot M_3} \right] \cdot (G_{10} + G_{20})} \qquad (6.17)$$

Змінні: x^*_{10} - збурення. G_{10}– клерувальний вплив, x_3 - регульована величина.

Таблиця 6.1 – Таблиця значень основного статичного режиму.

Ємність	**Тип потоку**	**Речовина**	Технологічні параметри				
			Витрата(кількість речовини), **кг/с**	Температура, **°С**	Молекулярна маса, а.о.м.	Стехіометричний коефіцієнт	
Реактор	Вхід	Силікат кальцію	G_{10} =120…250	Θ_1 =50	M=212	n_1= 0.7	
		Хлороводнева кислота	G_{20} = 120	Θ_2 =50	M=72	n_2 = 1.2	
	Вихід	Силікат кальцію	(G_{10}+G_{20}) = 120…250+120	Θ =50	M=156	n_3= 2	
		Хлороводнева кислота					- - - - -
		Суспензія білої сажі					- - - - -

Формулювання критерію оптимальності для задачі статичної оптимізації

Сформудюємо критерій оптимальності, виходячи з того, що треба підтримувати задану концентрацію і при цьому мінімізувати витрату пари G_n.

Критерій оптимальності:

$$I = \frac{1}{2} \cdot q \cdot (x_3 - x_3^{зд})^2 + \frac{1}{2} \cdot r \cdot G_{10}^2 \to \min$$

Визначення обмежень, які накладаються на процес

Згідно з технологією, неможна змінювати температуру у будь-якому діапазоні в даному апараті, тому треба задати межі, в яких має знаходитися значення температури. Значення температури має лежати в діапазоні від 48 до 52 °C.

$$T_{\min} < T < T_{\max}$$

$$48 < T < 52$$

Перетворення задачі умовної оптимізації в безумовну

Маючи статичну модель, критерій оптимальності і обмеження перетворимо задачу умовної оптимізації в безумовну.

Функція Лангранжа:

$$L = \frac{1}{2} \cdot q \cdot (x_3 - x_3^{зд})^2 + \frac{1}{2} \cdot r \cdot G_{10}^2 + \lambda_1 (-x_3 \cdot (G_{10} + G_{20}) + k \cdot \left[\frac{G_{20} \cdot x_{10}}{G_{10} + G_{20}} - \frac{n_1 \cdot M_1}{n_3 \cdot M_3} \cdot x_3 \right] \cdot e^{-\frac{E}{R\Theta}} \cdot V \cdot \rho)$$

Необхідні умови оптимальності

Необхідними умовами оптимальності будуть:

$$\frac{\partial L}{\partial x_3} = (x_3 - x_3^{зд}) - \ G_{10} + G_{20} \ \cdot \lambda = 0;$$

$$\frac{\partial L}{\partial G_{10}} = r \cdot G_{10} - \lambda \cdot x_3 + \lambda (k_1 + k_2) = 0;$$

$$G_{10}^{(N+1)} = G_{10}^{N} - k \frac{\partial L}{\partial G_{10}};$$

Алгоритм пошуку

1. Задаємо значення G_{10}^{N}

2. Розв’язуємо математичну модель. Знаходимо X_3^N

3. Розраховуємо λ

4. Знаходимо $\dfrac{\partial L}{\partial G_{10}}$

5. Знаходимо G_{10}^{N+1} і повертаємось до пункту

6. Розраховуємо оптимальне значення керування G_{10}^{N}

Список літератури

1. Бояринов А.И., Кафаров В.В. Методы оптимизации в химической технологи. – М., Химия,1975. – 576с.
2. Акулич И.Л. Математическое программирование в примерах и задачах. - М.: Высшая школа, 1986.
3. Таха Т. Введение в исследование операций; В двух книгах. – М.:Мир, 1985. – 479 с.
4. Банди Б. Основы линейного программирования. – М. :Радио и связь, 1989. –176 с.
5. Ладієва Л.Р. Оптимальне керування системами.: Навчальний посібник. - К.НМЦ ВО, 2000- 187с.
6. Ладієва Л. Р. Оптимізація технологічних процесів. Навч.посіб. – К .: ІВЦ «Політехніка», 2004.-192с.