

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
Національний Університет «Запорізька політехніка»

КОНСПЕКТ ЛЕКЦІЙ
З дисципліни

**«ОПТИМІЗАЦІЯ ТЕХНІЧНИХ РІШЕНЬ В
МАШИНОБУДУВАННІ ТА ЧИСЕЛЬНІ МАТЕМАТИЧНІ
МЕТОДИ»**

Для студентів спеціальностей:
133 Галузеве машинобудування
131 Прикладна механіка
всіх форм навчання

Конспект лекцій з дисципліни «Оптимізація технічних рішень в машинобудуванні та чисельні математичні методи» для студентів спеціальностей: 133 Галузеве машинобудування, 131 Прикладна механіка всіх форм навчання / Укл.: М. В. Фролов. – Запоріжжя: НУ «Запорізька політехніка», 2023. – 90 с.

Укладач: М. В. Фролов, доц., к.т.н.

Рецензент: В. В. Циганов, доц., д.т.н.

Відповідальний за випуск: В. П. Загородній

Затверджено на засіданні
кафедри Металорізальних
верстатів та інструментів
Протокол № 12 від 16/06/23

Рекомендовано
до видання НМК
Машинобудівного
факультету. Протокол № 2
від 14/09/23

Зміст

1. ПОСТАНОВКА ЗАДАЧІ ОПТИМІЗАЦІЇ.....	4
2. ЗАГАЛЬНІ ВІДОМОСТІ ПРО ОПТИМІЗАЦІЙНІ МОДЕЛІ, ЇХ СТРУКТУРУ ТА ПРИНЦИПИ ПОБУДОВИ	6
3. ЗАГАЛЬНІ ВІДОМОСТІ ПРО КРИТЕРІЇ ОПТИМАЛЬНОСТІ ТА ЦІЛЬОВІ ФУНКЦІЇ.....	10
4. ДЕЯКІ ТИПОВІ КРИТЕРІЇ ОПТИМАЛЬНОСТІ ТА ЦІЛЬОВІ ФУНКЦІЇ ПРИ МЕХАНІЧНІЙ ОБРОБЦІ	14
5. ОДНОМІРНА БЕЗУМОВНА ОПТИМІЗАЦІЯ	16
5.1 Загальні положення.....	16
5.2. Методи звуження інтервалу невизначеності.....	18
5.3. Методи з використанням похідних: метод Ньютона-Рафсона	28
6. МІНІМУМ ФУНКЦІЇ БАГАТЬОХ ЗМІННИХ	31
6.1 Рельєф функції.....	31
6.2 Метод покоординатного спуску: метод Гауса	33
7. УМОВНА ОПТИМІЗАЦІЯ.....	36
7.1. Загальні положення.....	36
7.2. Задачі з обмеженнями у вигляді рівностей. Множники Лагранжа.....	38
7.3. Задачі з обмеженнями у вигляді нерівностей.	40
7.4. Методи штрафних функцій. Нежорсткі обмеження	43
8. ДЕЯКІ ТИПОВІ ОБМЕЖЕННЯ ПРИ МЕХАНІЧНІЙ ОБРОБЦІ	51
9. ЗАДАЧІ ЛІНІЙНОГО ПРОГРАМУВАННЯ	55
9.1 Загальні положення.....	55
9.2 Задача про використання сировини.....	56
9.3 Задача про використання продуктивності обладнання.....	61
9.4 Транспортна задача.....	65
10. Оптимізація ЗА експериментальНИМИ дослідженняМИ	67
10.1. Загальні положення. Повний факторний експеримент	67
10.2. Ортогональні центральні композиційні плани.....	74
10.3. Планування експерименту та оптимізація за концепцією Тагучі.....	80
ЛІТЕРАТУРА	90

1. ПОСТАНОВКА ЗАДАЧІ ОПТИМІЗАЦІЇ

В машинобудуванні як і в будь-якій області техніки а також науці взагалі, економіці та в інших сферах діяльності часто зустрічаються задачі, що мають численні рішення. Якщо ці рішення порівнюються, то це робиться на основі певного критерію та вибирається те, яке найкращим чином відповідає цьому критерію. Таке рішення називається **оптимальним**, задача його визначення на заданій множині рішень називається задачею оптимізації, а критерій – критерієм оптимальності.

Необхідність в оптимізації виникає при виконанні найбільш складних процедур при проектуванні та експлуатації технічних об'єктів. При цьому ця множина рішень може складатися як з характеристик об'єкту так і з умов його експлуатації.

Рішення задач оптимізації за допомогою обчислювальної техніки має деякі характерні особливості:

• **Постановка задачі** – передбачає добре формалізований опис множини рішень та призначення актуального критерію. При постановці задачі оптимізації необхідні наступні передумови:

1. Наявність об'єкта оптимізації і мети оптимізації. При цьому формулювання кожного завдання оптимізації повинно вимагати екстремального значення лише однієї величини, тобто одночасно системі не повинно приписуватися два і більше критерію оптимізації, так як практично завжди екстремум одного критерію не відповідає екстремуму іншого. Типовий приклад неправильної постановки завдання оптимізації: «Отримати максимальну продуктивність при мінімальній собівартості». Помилка полягає в тому, що ставиться завдання пошуку оптимуму двох величин, що суперечать один одному за своєю суттю. Правильною постановкою завдання може бути:

а) отримати максимальну продуктивність при заданій собівартості.

б) отримати мінімальну собівартість при заданій продуктивності

У першому випадку критерій оптимізації - продуктивність, а у другому - собівартість. В загальному випадку для двох критеріїв повинні бути встановлені свої пріоритети, про що буде сказано нижче.

2. Наявність ресурсів оптимізації, під якими розуміють можливість вибору значень деяких параметрів об'єкту що

оптимізується. Об'єкт повинен мати певні ступені свободи – **керуючі впливи**.

3. Можливість кількісної оцінки величин що оптимізуються оскільки тільки в цьому випадку можна порівнювати ефекти від вибору тих чи інших дій, що управляють.

4. Урахування обмежень.

•**Математичний опис** – полягає в визначенні типу задачі оптимізації, виходячи з її постановки, а також у вивченні її специфічних особливостей та побудові математичної моделі, що пов'язує кількісні характеристики складових рішення з критерієм оптимальності. На цій основі вибирається, вивчається, а можливо і модифікується метод її чисельного вирішення

•**Складання алгоритму** – полягає в розробці загальної блок-схеми з її подальшою деталізацією програми оптимізації, що розробляються на основі мови програмування, та передбачає наявність інтерфейсу для відображення вихідних даних та результатів оптимізації, а також їх зберігання [2].

Таким чином постановку задачі оптимізації можна сформулювати наступним чином:

Дано – множина рішень або області їх можливих значень та критерій за допомогою якого ці рішення можна порівнювати

Треба – на заданій множині рішень використовуючи заданий критерій знайти найкраще (оптимальне) рішення.

Питання для самоперевірки

1. Яке рішення називається оптимальним?
2. З чого може складатися множина рішень при оптимізації технічного об'єкту?
3. Які особливості оптимізації існують при використанні обчислювальної техніки?
4. Що вимагає кожна задача оптимізації?
5. Які передумови необхідні для постановки задачі оптимізації?
6. В чому полягає математичний опис задачі оптимізації?
7. В чому полягає складання алгоритму оптимізації?

2.ЗАГАЛЬНІ ВІДОМОСТІ ПРО ОПТИМІЗАЦІЙНІ МОДЕЛІ, ЇХ СТРУКТУРУ ТА ПРИНЦИПИ ПОБУДОВИ

Математичне моделювання на основі принципу оптимізації передбачає можливість та необхідність цілеспрямованого регулювання умов функціонування об'єкту або системи з метою пошуку таких, у яких система буде функціонувати найкращим чином – оптимальних умов. Сукупність цих умов і буде представляти собою оптимальне рішення. Звідси маємо і складові елементи таких моделей до яких належать [1]:

Критерій оптимальності – прийнятий показник міри ефективності системи що досліджується. Максимальне або мінімальне значення критерію оптимальності і визначає умови у яких система працює найкращим чином тобто оптимальні умови. Треба наголосити, що зазвичай критеріями оптимальності не є фізичні величини, наприклад сила різання, температура, стійкість інструменту та ін., а показники тим чи іншим чином пов'язані з грошима: собівартість обробки, прибуток, якість, час виготовлення та ін., що і відображає міру ефективності системи. У якості прикладу критерію оптимальності візьмемо собівартість обробки C , що залежить від режимів різання:

$$C = f(t, S, V) \quad (2.1)$$

Незалежні змінні або фактори або керуючі фактори – величини, оптимальні значення яких треба знайти. У даному випадку це будуть режими різання t , S , V бо саме за їх допомогою можна цілеспрямовано змінювати умови функціонування системи. Сукупність знайдених оптимальних значень і буде тим самим оптимальним рішенням про яке було сказано в розділі 1.

Параметри – постійні величини моделі, що характеризують незмінні умови в яких модель була отримана. Відображаються у значеннях постійних коефіцієнтів моделі.

Цільова функція – функція, що пов'язує критерій оптимальності, фактори та параметри. Сама назва функції показує, що в неї є ціль. У нашому випадку це є мінімальна собівартість, тобто

$$C = f(t, S, V) \rightarrow \min \quad (2.2)$$

Звертаємо увагу, що тут перш за все цікавлять умови (режими різання) за яких собівартість буде мінімальною, а не значення собівартості як такої. Це означає, що з метою пошуку найкращих умов функціонування системи, цільова функція може спрощуватися і не включати елементи собівартості, що не залежать від режимів різання в даному випадку, або від факторів, якщо говорити взагалі.

Обмеження – області можливих значень факторів, у межах яких відбувається пошук оптимальних рішень. При вирішенні реальних виробничих задач обмеження існують завжди і вони можуть бути пов'язані з технічними можливостями обладнання (обмеження по швидкості різання, потужності верстату); з якістю обробленої поверхні (обмеження подачі); з умовами безпеки праці та ін. Обмеження можуть бути:

- **жорсткі** – вихід за межі яких неприпустимий при будь яких умовах. Це можуть бути обмеження пов'язані з фізичними константами, безпекою, технічними характеристиками обладнання тощо.
- **нежорсткі** - дозволяють вихід за їх межі, але треба розуміти що якість функціонування системи при цьому істотно погіршиться.

Як зазначалося вище, метою оптимізаційного моделювання є пошук умов в яких система функціонує найкращим образом згідно з прийнятим критерієм оптимальності. Процес отримання цільової функції не можна формалізувати – це результат аналізу процесу в якому використовується попередній досвід та здоровий глузд. Але для ефективного пошуку оптимальних умов, оптимізаційна модель у формі цільової функції повинна відповідати деяким основним принципам.

1. Принцип однозначності. Полягає в тому, що оптимізуватись повинна тільки одна цільова функція. Якщо ж за умовами задачі таких функцій повинно бути декілька, то вони все одно мають бути зведені до одної, наприклад за допомогою лінійної згортки:

$$F = \sum_{i=1}^k a_i f_i(X_1, X_2, \dots, X_n) \quad (2.3)$$

Тут a_i – масштабний або ваговий коефіцієнт, що відображає ступінь важливості кожної з функцій, що входить до згортки і повинен відповідати наступним умовам:

$$a_i \geq 0 \quad (2.4)$$

$$\sum_{i=1}^k a_i = 1$$

У виборі вагових коефіцієнтів присутня велика доля суб'єктивізму, але за допомогою деяких підходів їм можна придати об'єктивний характер: експертні оцінки, функція бажаності та ін.

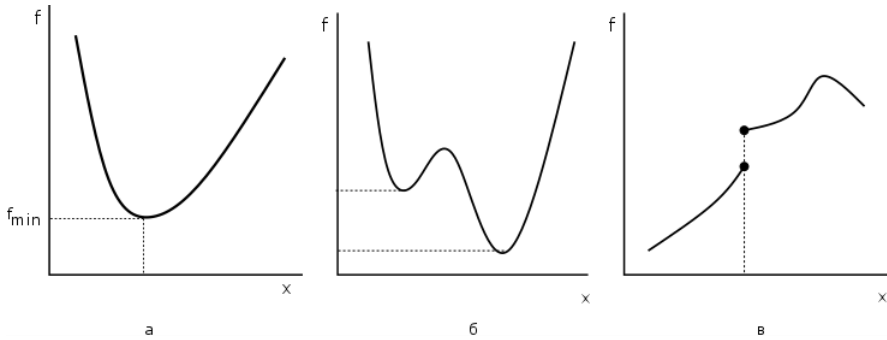
Згортка за рівнянням (2.3) може бути використана коли всі функції f_i мають однакову фізичну природу, тобто розмірність. А якщо розмірності різні, то замість функцій в натуральному вигляді слід використовувати їх нормалізований – безрозмірний варіант $\tilde{f}_i$. Наприклад для нормалізації в межах $[0;1]$ може бути використаний наступний спосіб природньої нормалізації:

$$\tilde{f}_i = \frac{f_i - f_{i \max}}{f_{i \max} - f_{i \min}} \quad (2.5)$$

2. Принцип відповідності. Бажано, щоб цільова функція мала один екстремум – така функція буде **унімодальною** (умову унімодальності буде показано нижче). Функції, що мають розриви та неоднозначності не мають використовуватися в якості цільових. Наявність локальних екстремумів ускладнює пошук оптимального рішення. На Рис. 2.3 функція а) має підходящу форму, а ось функції б) та в) в якості цільових використовувати не бажано.

Особливо вирізняються цільові функції, що зовсім не мають екстремумів, наприклад лінійні цільові функції. При цьому вони повинні бути доповнені обмеженнями. В іншому разі оптимізаційна задача в цьому випадку не має сенсу.

3. Принцип керованості. Функція повинна виражатися через змінні на які можна впливати і якими можна керувати.



а – функція з одним екстремумом; б – функція з локальним екстремумом; в – функція з розривом

Рисунок 2.1 — Деякі види функцій

Питання для самоперевірки

1. В чому полягає моделювання на основі принципу оптимізації?
2. Які елементи присутні в оптимізаційній моделі?
3. Що таке фактори?
4. Дайте визначення терміну «параметри».
5. Що таке критерій оптимальності?
6. Що таке цільова функція?
7. За якими принципами повинна будуватися цільова функція?
8. Яка функція називається унімодальною?
9. Які типи обмежень існують?
10. В чому полягає принцип однозначності?
11. В чому полягає принцип відповідності?
12. В чому полягає принцип керованості?
13. Що таке «жорсткі обмеження»?
14. Що таке «нежорсткі обмеження»?

3.ЗАГАЛЬНІ ВІДОМОСТІ ПРО КРИТЕРІЇ ОПТИМАЛЬНОСТІ ТА ЦІЛЬОВІ ФУНКЦІЇ

Оптимізуватися може будь яка фізична характеристика технічного об'єкту або процесу – температура, сила, стійкість інструменту, довговічність та ін. Але найчастіше оптимізуються нефізичні характеристики - ті що так чи інакше пов'язані з грошима: собівартість, продуктивність, прибуток, економічна ефективність, характеристики якості та ін. В будь якому випадку фактор, що оптимізується повинен оцінюватися якоюсь кількісною мірою.

На підставі обраного критерію оптимальності складається цільова функція, що представляє собою залежність критерію оптимальності від факторів, що впливають на її значення. Вид критерію оптимальності або цільової функції визначається конкретним завданням оптимізації. Таким чином, **задача оптимізації зводиться до знаходження екстремуму цільової функції.**

Як зазначалося вище, найбільш загальною постановкою оптимізаційної задачі є вираз критерію оптимальності у вигляді економічної оцінки (продуктивність, собівартість продукції, прибуток, рентабельність). Однак в окремих задачах оптимізації, коли об'єкт є частиною технологічного процесу, не завжди вдається або не завжди доцільно виділяти прямий економічний показник, який би повністю характеризував ефективність роботи об'єкту що розглядається. У таких випадках критерієм оптимальності може служити технологічна характеристика, яка побічно оцінює економічність роботи агрегату (час контакту, вихід продукту, ступінь перетворення, температура). Наприклад, встановлюється оптимальний температурний профіль, тривалість циклу та ін.

Розглянемо більш докладно вимоги, які повинні пред'являтися до критерію оптимальності.

1. Критерій оптимальності повинен виражатися кількісно.
2. Критерій оптимальності повинен бути єдиним.
3. Критерій оптимальності повинен відображати найбільш істотні сторони процесу.

4. Бажано щоб критерій оптимальності мав ясний зміст і легко розраховувався.

Будь який об'єкт що оптимізується, схематично можна уявити у відповідності до Рис. 3.1.

При постановці конкретних задач оптимізації бажано критерій оптимальності записати у вигляді аналітичного виразу.

У тому випадку, коли вплив випадкових величин, що не контролюються невеликий і його можна не враховувати, критерій оптимальності може бути представлений як функція вихідних, керуючих факторів і параметрів:

$$R = R(X_1, X_2, \dots, X_N, Y_1, Y_2, \dots, Y_N, U_1, U_2, \dots, U_N) \quad (3.1)$$

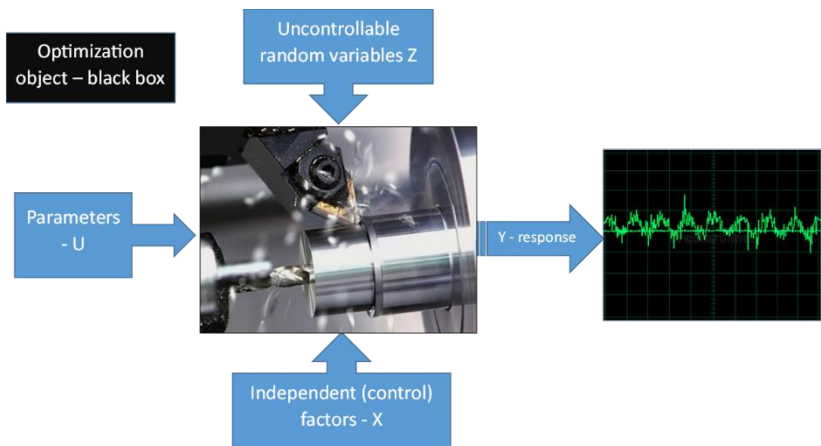


Рисунок 3.1 – Загальна схема об'єкта оптимізації

Так як $\bar{Y} = F(\bar{X})$, то при фіксованих U можна записати:

$$R = R(\bar{X}) \quad (3.2)$$

Якщо ж випадкові впливи досить великі і їх необхідно враховувати, то слід застосовувати експериментально-статистичні

методи, що дозволяють отримати модель об'єкта у вигляді функції, яка справедлива тільки для вивченої локальної області.

В принципі, для оптимізації замість математичної моделі можна використовувати і сам об'єкт, однак оптимізація дослідним шляхом має ряд істотних недоліків:

а) необхідний реальний об'єкт;

б) необхідно змінювати технологічний режим в значних межах, що не завжди можливо;

в) тривалість випробувань і складність обробки даних. Наявність математичної моделі (за умови, що вона досить надійно описує процес) дозволяє значно простіше вирішити завдання оптимізації аналітичним або чисельним методами.

У задачах оптимізації розрізняють прості і складні критерії оптимізації.

Критерій оптимальності називається **простим**, а сама оптимізація **безумовною** якщо потрібно визначити екстремум цільової функції без завдання умов для будь-яких інших величини. Такі критерії зазвичай використовуються при вирішенні окремих задач оптимізації (наприклад, визначення максимальної концентрації цільового продукту, оптимального часу перебування реакційної суміші в апараті і т.п.).

Критерій оптимальності називається **складним**, а сама оптимізація **умовною**, якщо необхідно встановити екстремум цільової функції при деяких умовах, які накладаються на ряд інших величин (наприклад, визначення максимальної продуктивності при заданій собівартості, визначення оптимальної температури при обмеженнях по термостійкості каталізатора і ін.).

Процедура вирішення задачі оптимізації обов'язково включає, крім вибору керуючих факторів, ще й встановлення обмежень на них (термостійкість, вибухонебезпечність, потужність верстата та ін.). Обмеження можуть накладатися як за технологічними, так і з економічних міркувань.

Отже, для вирішення задачі оптимізації необхідно:

а) скласти математичну модель об'єкта оптимізації $\bar{Y} = F(\bar{U}, X)$

б) вибрати критерій оптимальності і скласти цільову функцію

$$R = R(\bar{Y}, X, \bar{U})$$

в) встановити можливі обмеження, які повинні накладатися на змінні;

г) вибрати метод оптимізації, який дозволить знайти значення факторів за яких забезпечується екстремум цільової функції.

Прийнято розрізняти задачі **статичної** оптимізації для процесів, що протікають в сталих режимах, і завдання **динамічної** оптимізації. У першому випадку вирішуються питання створення і реалізації оптимальної моделі процесу, у другому - завдання створення і реалізації системи оптимального управління процесом при невстановлених режимах експлуатації.

Питання для самоперевірки

1. До чого зводиться задача оптимізації цільової функції?
2. Які характеристики зазвичай приймаються як критерії оптимальності?
3. Охарактеризуйте вимоги до критеріїв оптимальності.
4. Чи можна оптимізацію здійснювати дослідним шляхом і якщо да, то які особливості переваги та недоліки цього процесу?
5. Який критерій оптимальності називають простим?
6. Який критерій оптимальності називають складним?
7. Що таке безумовна оптимізація?
8. Що таке умовна оптимізація?
9. Чим відрізняється статична та динамічна оптимізація?

4. ДЕЯКІ ТИПОВІ КРИТЕРІЇ ОПТИМАЛЬНОСТІ ТА ЦІЛЬОВІ ФУНКЦІЇ ПРИ МЕХАНІЧНІЙ ОБРОБЦІ

Найчастіше при визначенні раціональних умов різання при обробці на металорізальних верстатах в якості критерію оптимальності приймається собівартість обробки. При цьому доля собівартості при різанні одним інструментом визначається за формулою [3]:

$$C = \left(t_0 + \frac{t_c}{Q}\right)E + \frac{E_i}{Q} \rightarrow \min \quad (4.1)$$

Де: t_0 – основний технологічний (машинний час); t_c – час на відновлення ріжучих властивостей інструменту та підналадження верстата; E – собівартість однієї хвилини роботи верстата та оператора; E_i – витрати, пов'язані з експлуатацією інструменту за період його стійкості; Q – кількість деталей, що обробляються за період стійкості інструменту.

У свою чергу, витрати, пов'язані з експлуатацією інструменту за період його стійкості можна визначити:

- Для інструменту, що переточується:

$$E_i = t_c E_n + C_{зат} + \frac{(C_{ін} - C_{відх})k_3}{i + 1} \quad (4.2)$$

Де: E_n – номінальна заробітна плата наладчика; $C_{зат}$ – вартість заточування; $C_{ін}$ – вартість інструменту; $C_{відх}$ – вартість відходів інструменту; i – кількість переточувань інструменту до його повного зносу; k_3 – коефіцієнт, що враховує випадкове вибуття інструменту.

- Для інструменту з механічним кріпленням змінних багатогранних пластинок, зважаючи на довгий термін експлуатації державок, затрати на які радше відносяться до капітальних:

$$E_i = t_c E_n + \frac{(C_{\Sigma пл} - C_{\Sigma відх})k_3}{n_{пл}} \quad (4.3)$$

Де: $C_{\Sigma \text{пл}}$ – сумарна вартість пластинок в інструменті; $C_{\Sigma \text{відх}}$ – сумарна вартість відходів пластинок; $n_{\text{пл}}$ – кількість робочих граней пластинки.

Для розрахунків за формулами (4.1), (4.2), (4.3) та взагалі при проектуванні технологічного процесу, велике значення має період стійкості ріжучого інструменту. В залежності від конкретних умов та завдань виробництва можуть використовуватися різні періоди стійкості: T_{max} – максимальний період стійкості; $T_{\text{опт}}$ – оптимальна стійкість, що відповідає оптимальній швидкості різання; $T_{\text{ек}}$ – економічний період стійкості, що відповідає найменшій собівартості обробки; $T_{\text{пр}}$ – період стійкості, що відповідає максимальній продуктивності процесу. На практиці найчастіше використовуються періоди стійкості $T_{\text{ек}}$ та $T_{\text{пр}}$ [3]:

$$T_{\text{ек}} = \frac{1 - m}{m} \left(t_c + \frac{E_i}{E} \right) \quad (4.4)$$

$$T_{\text{пр}} = \frac{1 - m}{m} t_c \quad (4.5)$$

Де m – показник відносної стійкості інструмента.

В подальшому, використовуючи відому залежність для розрахунку швидкості різання, можна оптимізувати режими різання:

$$V = \frac{C_v}{t^{x_v} S^{y_v} T^m} K_v \quad (4.6)$$

Де: t – глибина різання; S – подача; T – стійкість; V – швидкість різання; C_v , K_v – коефіцієнти; x_v , y_v – показники ступеню.

Питання для самоперевірки

1. Який показник найчастіше приймається в якості критерію оптимальності при механічній обробці?
2. Назвіть існуючі різновиди періодів стійкості.
3. Які періоди стійкості частіше використовуються на практиці?

5.ОДНОМІРНА БЕЗУМОВНА ОПТИМІЗАЦІЯ

5.1 Загальні положення

Оптимізація функції однієї змінної за умови відсутності обмежень або одновимірною безумовною оптимізацією вона ж однопараметрична оптимізація - найбільш простий тип оптимізаційних задач. Тим не менш, вона займає важливе місце в теорії оптимізації. Це пов'язано з тим, що завдання однопараметричної оптимізації досить часто зустрічаються в інженерній практиці та, крім того, знаходять своє застосування при реалізації більш складних інтерактивних процедур багатопараметричної оптимізації.

Нехай функція $f(x)$ безперервна на відрізку $[a, b]$. Розглянемо задачу мінімізації цієї функції на цьому відрізку (завдання знаходження найменшого значення на відрізку).

Зауважимо, що якщо функція $f(x)$ має в точці x^ мінімум, то функція $-f(x)$ в x^* має максимум. Тому для відшукування максимуму застосовуються ті ж методи, що і для відшукування мінімуму. Далі ми, як правило, будемо говорити тільки про завдання відшукування мінімуму.*

Для тих, хто вивчав курс математичного аналізу ця задача може здатися елементарною. Так, якщо функція $f(x)$ кусочно-диференційована на відрізку $[a, b]$, то для знаходження її мінімуму досить знайти її похідну, та вирішити рівняння виду:

$$f'(x) = 0 \quad (5.1)$$

Значення x що відповідає (5.1) визначає стаціонарну (критичну) точку, тобто точку в якій вектор градієнту функції дорівнює нулю:

$$\nabla f(x) = 0 \quad (5.2)$$

Будемо вважати, що на відрізку $[a, b]$ існує тільки одна стаціонарна точка яка може бути точкою як локального так і глобального мінімуму на відміну від інших нестаціонарних точок функції, що можуть бути прийняті за мінімум, виходячи з умов задачі (наприклад з урахуванням обмежень про що буде сказано далі).

Зазначимо, що в стаціонарній точці мінімуму повинна виконуватися також умова:

$$f''(x) > 0 \quad (5.3)$$

Однак як показує практика, для широкого класу функцій аналітичне вирішення задачі оптимізації не таке вже просте. З одного боку, навіть якщо функція $f(x)$ безперервна але диференціюється на $[a, b]$ і обчислення її похідної $f'(x)$ не складає труднощів, рішення рівняння (5.1) для знаходження критичних точок виявляється не менш складним завданням, ніж вихідна задача мінімізації. З іншого боку, в практичних завданнях часто буває невідомо, чи є функція $f(x)$ такою що диференціюється, більш того, навіть якщо ми знаємо, що вона диференціюється, ми можемо при цьому не знати аналітичне представлення її похідної $f'(x)$ (Наприклад, якщо функції не задана в аналітичному вигляді). У зв'язку з цим істотне значення мають методи мінімізації, які не потребують обчислення похідної. Тому ми приходимо до застосування чисельних методів оптимізації, які передбачають застосування переважно розрахункових алгоритмів. Такі алгоритми передбачають повторне застосування деяких математичних операцій що має назву **ітерація**, а весь процес називається ітераційним.

Застосування будь якого з чисельних методів починається з визначення **інтервалу або відрізка локалізації** – відрізка в якому знаходиться мінімум функції в межах якого знаходиться **інтервал невизначеності** $[a, b]$. При цьому визначається трійка чисел на відріжку локалізації: x_1, x_2, x_4 . Така трійка чисел називається **вдалою**, якщо виконується вимога: $f(x_2) < \min [f(x_1), f(x_4)]$ див. Рис. 5.1. Тут же запишемо **умову унімодальності функції**: $f(x_2) < f(x_1), f(x_3) < f(x_4)$.

В ході ітераційного процесу розраховане рішення все більше наближається до істинного мінімуму і в цьому контексті йде мова про точність рішення ε , що може визначатися як довжина інтервалу невизначеності на ітерації k , як абсолютна різниця значень знайдених

значень x^* у двох послідовних ітераціях k та $k-1$ (Див. формулу (5.4). Чим вища точність, тим більше ітерацій треба робити.

$$|b_k - a_k| \leq \varepsilon \quad \text{або} \quad |x_k^* - x_{k-1}^*| \leq \varepsilon \quad (5.4)$$

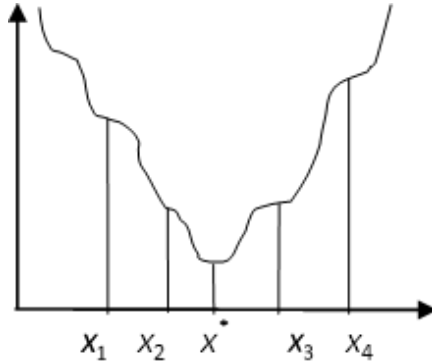


Рисунок 5.1 - Вдала трійка чисел

Зазначений ітераційний процес можна зобразити у вигляді спрощеного алгоритму, зображеному на Рис. 5.2

Розроблено велику кількість методів одновимірної оптимізації і ми розглянемо дві групи таких методів:

- методи звуження інтервалу невизначеності;
- методи з використанням похідних.

5.2. Методи звуження інтервалу невизначеності

Нехай потрібно знайти мінімум функції $f(x)$ на деякому інтервалі $[a, b]$. Задача приблизного відшукування мінімуму в методах звуження інтервалу невизначеності полягає в тому, щоб знайти безліч абсцис $x_1, x_2, \dots, x_k$ в яких обчислюється функція, таке, що мінімальне значення f^* лежить при деякому i в інтервалі $x_{i-1} < x^* < x_i$. Такий інтервал називається інтервалом невизначеності D . Очевидно, що спочатку інтервал невизначеності D збігається з відрізком $[a, b]$.

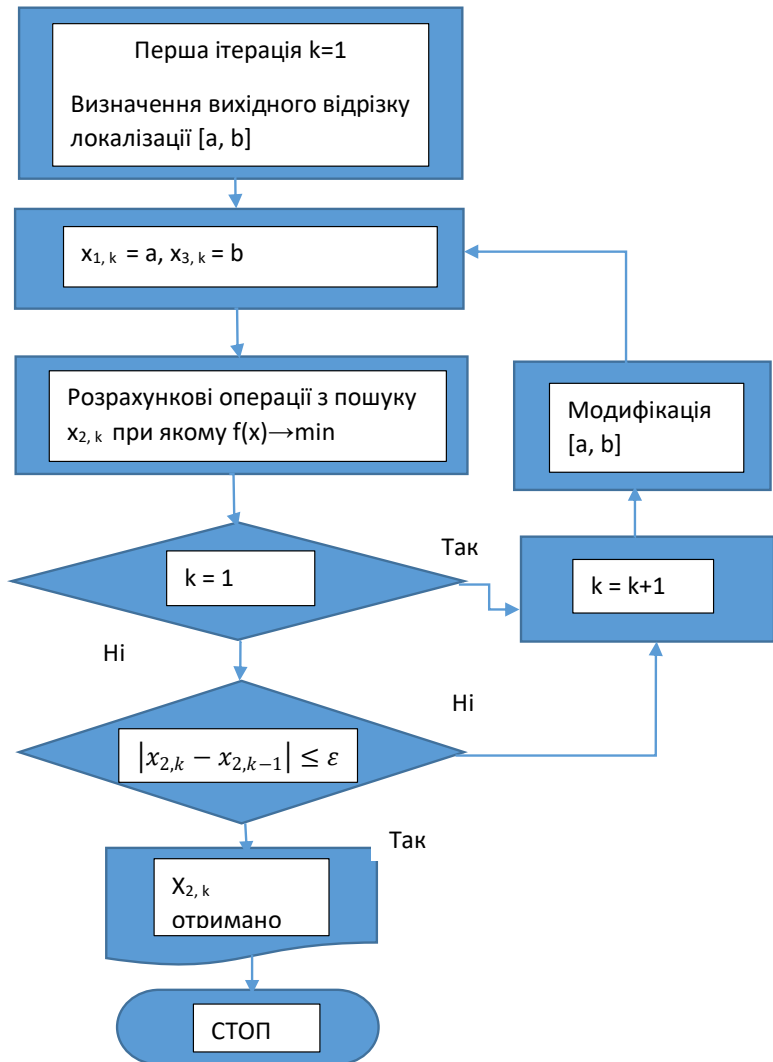


Рисунок 5.2 - Ітераційний процес

Існують кілька способів систематичного звуження інтервалу не- визначеності. Розглянемо три з них.

5.2.1. Загальний пошук

Нехай потрібно знайти мінімум функції $f(x)$ на деякому інтервалі $[a, b]$. Якщо про функцію $f(x)$ на цьому інтервалі ніякої додаткової інформації невідомо, то для пошуку мінімуму на $[a, b]$ можна застосувати найпростіший метод перебору, або, інакше, загального пошуку.

У цьому методі інтервал $[a, b]$ ділиться на кілька рівних частин N з наступним обчисленням значень функції у $n = N+1$ вузлах отриманої сітки. В якості мінімуму приймається абсциса з мінімальним обчисленим значенням функції (рис.5.3).

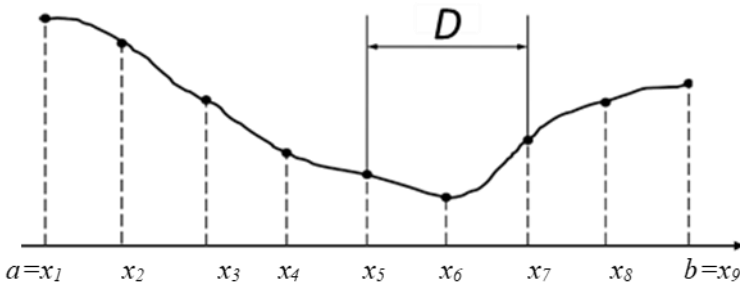


Рисунок 5.3 - Метод загального пошуку

В результаті інтервал невизначеності звужується до двох кроків сітки D . Звужений інтервал знов ділиться на N рівних частин. Процедура повторюється до отримання потрібної точності ε що визначається як різниця знайдених значень x у двох послідовних ітераціях.

Позначимо a_i, b_i – Межі інтервалу невизначеності для ітерації i . Відповідно для першої ітерації: $a_1 = a, b_1 = b$. Далі циклічно виконується ітераційний процес до отримання заданої точності:

➤ Довжина інтервалу невизначеності:

$$D_i = b_i - a_i \quad (5.5).$$

- Шаг розбиття в ітерації i :

$$\delta_i = \frac{D_i}{N} \quad (5.6)$$

- Визначення значення фактору x_i :

$$x_{ij} = a_i + (j - 1) \cdot \delta_i, \quad j = \overline{1, N + 1} \quad (5.7)$$

- Розрахунок значень функції $f(x_{ij})$ та вибір мінімального:

$$x_i^* = x_{ij} \text{ для якого } f(x_i^*) = \min(f(x_{ij})), \quad j = \overline{1, N + 1} \quad (5.8)$$

- Довжина інтервалу невизначеності після першої ітерації:

$$D_{i+1} = 2\delta_i \quad (5.9)$$

- Перевірка точності та повторення циклу у разі невиконання умови (5.10):

$$|b_k - a_k| \leq \varepsilon \text{ або } |x_k^* - x_{k-1}^*| \leq \varepsilon, \quad i = 2, 3, \dots \quad (5.10)$$

- Межі нового інтервалу відносно точки мінімуму в попередній ітерації $i - x_i^*$:

$$a_{i+1} = x_i^* - \delta_i, \quad b_{i+1} = x_i^* + \delta_i \quad (5.11)$$

5.2.2 Метод ділення інтервалу пополам (метод діхотомії)

Призначимо інтервал невизначеності $[a, b]$ (див. Рис. 5.4).

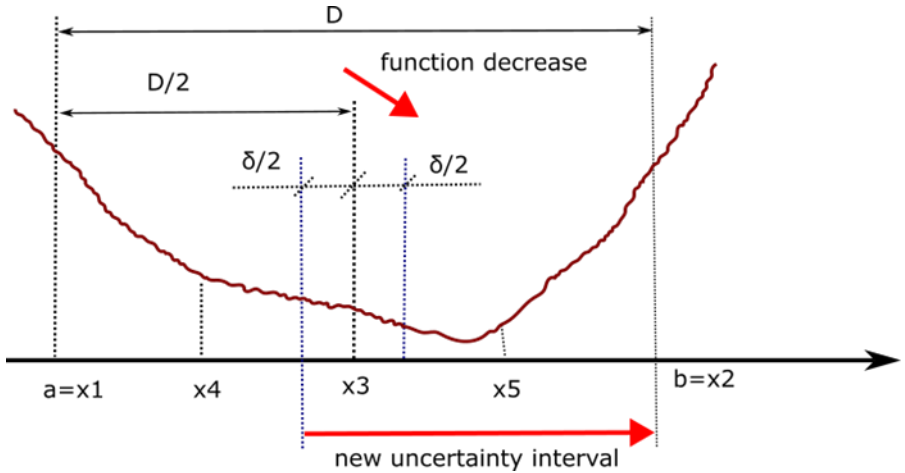


Рисунок 5.4 -. Ілюстрація до методу половинного ділення

На першому етапі (для першої ітерації) позначимо:

$$x_{1(1)} = a, \quad x_{2(1)} = b \quad (5.12)$$

Тут індекс в дужках – номер ітерації.

Визначимо середину інтервалу невизначеності:

$$x_{3(i)} = \frac{x_{1(i)} + x_{2(i)}}{2} \quad (5.13)$$

Та переконуємося, що $x_{1(1)}, x_{2(1)}, x_{3(1)}$ є вдалою трійкою чисел та визначимо малу величину $\delta \leq \varepsilon$, де ε – задана точність розрахунку.

Далі розрахунок відбувається в наступній послідовності:

➤ Розрахувати значень функції:

$$f\left(x_{3(i)} - \frac{\delta}{2}\right), f(x_{3(i)}), f\left(x_{3(i)} + \frac{\delta}{2}\right)$$

- Якщо $f\left(x_{3(i)} + \frac{\delta}{2}\right) < f\left(x_{3(i)} - \frac{\delta}{2}\right)$ то визначити нові межі інтервалу як:

$$x_{1(i+1)} = x_{3(i)} - \frac{\delta}{2}, \quad x_{2(i+1)} = x_{2(i)} \quad (5.14)$$

Такий випадок показано на Рис. 5.4

- Якщо $f\left(x_{3(i)} - \frac{\delta}{2}\right) < f\left(x_{3(i)} + \frac{\delta}{2}\right)$ то визначити нові межі інтервалу як:

$$x_{1(i+1)} = x_{1(i)}, \quad x_{2(i+1)} = x_{3(i)} + \frac{\delta}{2} \quad (5.15)$$

Новий інтервал D буде майже в 2 рази менше за попередній. Після цього визначається середина інтервалу за рівнянням (5.13) і дії повторюються для нового інтервалу. Ітерації виконуються до отримання заданої точності, яка визначається як абсолютна різниця між серединами інтервалів невизначеності в двох послідовних ітераціях:

$$|x_{3(i)} - x_{3(i-1)}| \leq \varepsilon \quad (5.16)$$

5.1.3. Метод золотого перетину

Розподіл інтервалу на нерівні частини дозволяє знайти ще більш ефективний метод. Обчислимо функцію на кінцях відрізка $[a, b]$ і покладемо $a = x_1, b = x_2$, обчислимо також функцію в двох внутрішніх точках x_3, x_4 . Порівняємо всі чотири значення функції і виберемо серед них найменше (рис. 5.5). Нехай, наприклад, найменшим виявилось $f(x_3)$. Очевидно, мінімум буде перебувати в одному з прилеглих до нього відрізків. Тому відрізок $[x_4, b]$ можна відкинути і залишити відрізок $[a, x_4]$.

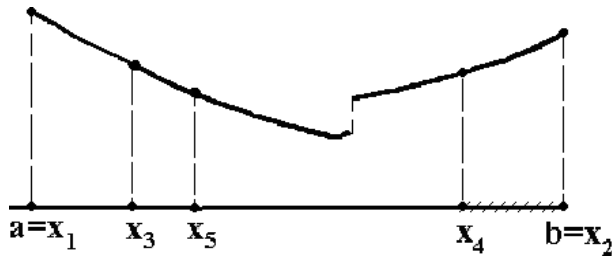


Рисунок 5.5 - Ілюстрація до методу золотого перетину

Перший крок зроблено. На відрізку $[a, x_4]$ знову треба вибрати дві внутрішні точки, обчисливши в них і на кінцях значення функції і зробити наступний крок. Але на попередньому кроці обчислень ми вже знайшли функцію на кінцях нового відрізка $[a, x_4]$ і в одній його внутрішній точці x_3 . Тому досить вибрати всередині $[a, x_4]$ ще одну точку x_5 , визначити в ній значення функції і провести необхідні порівняння. Це вчетверо зменшує обсяг обчислень на одному кроці процесу. Як вигідно розміщувати точки? Кожен раз відрізок що залишається ділиться на три частини і потім відкидається один з крайніх відрізків.

Позначимо початковий інтервал невизначеності через D (рис. 5.6). Так як в загальному випадку може бути відкинтий будь-який з відрізків $[x_1x_3]$ або $[x_4x_2]$ то виберемо точки x_3 і x_4 так, щоб довжини цих відрізків були однаковими: $x_3 - x_1 = x_2 - x_4$. Після відкидання вийде новий інтервал невизначеності довжиною D^1 .

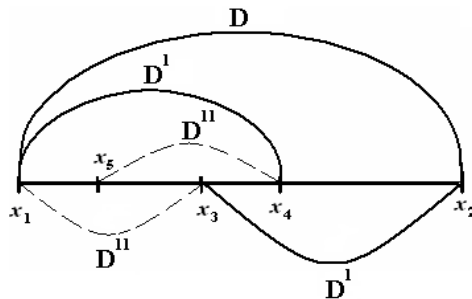


Рисунок 5.6 - Ділення відрізків методом золотого перерізу

Позначимо $D/D' = \varphi$. Далі продовжимо аналогічно – для цього інтервал D^1 розділимо так само як і інтервал D , тобто:

$$D'/D'' = D/D' = \varphi \quad (5.17)$$

Де D'' наступний інтервал невизначеності. При такому діленні весь інтервал відноситься до більшої частини, як більша частина до меншої і $\varphi = 1,618$ – що має назву «золоте відношення». Зазначене ділення відрізків відомо як «золотий перетин». Тому такий метод пошуку має назву «золотого перетину». Відношення φ показує во скільки разів скорочується вихідний розмір інтервалу невизначеності при кожній ітерації. При кількості ітерацій $k \rightarrow \infty$ довжина інтервалу невизначеності буде прагнути до 0 як геометрична прогресія зі знаменником $1/\varphi$, тобто метод золотого перетину завжди сходиться. Вочевидь цей метод більш ефективний ніж метод ділення пополам.

Позначивши координати розрахункових точок $x_{1(i)}, x_{2(i)}, x_{3(i)}, x_{4(i)}$ – де індекс в дужках це номер ітерації, та інтервал невизначеності в ітерації i – D_i , запишемо наступний алгоритм розрахунку: (див. Рис. 5.6):

- $x_{1(1)} = a, x_{2(1)} = b$
- $D_1 = x_{2(1)} - x_{1(1)}$
- $D_{i+1} = D_i/1.618$
- $x_{3(i)} = x_{2(i)} - D_{i+1}$
- $x_{4(i)} = x_{1(i)} + D_{i+1}$
- Розрахувати значення функції $f(x)$ в точках $x_{3(i)}, x_{4(i)}$
- Якщо $f(x_{3(i)}) < f(x_{4(i)})$ – відкинути відрізок x_4x_2 :

$$x_{1(i+1)} = x_{1(i)}, \quad x_{2(i+1)} = x_{4(i)}, \quad x_{(i)}^* = x_{3(i)}$$
- Якщо $f(x_{4(i)}) < f(x_{3(i)})$ – відкинути відрізок x_1x_3 :

$$x_{1(i+1)} = x_{3(i)}, \quad x_{2(i+1)} = x_{2(i)}, \quad x_{(i)}^* = x_{4(i)}$$

- Продовжувати ітерації до отримання заданої точності яка буде оцінюватися як:

$$|x_{(i)}^* - x_{(i-1)}^*| \leq \varepsilon \quad (5.18)$$

або за (5.4)

- Якщо умова (5.18) виконується, то в якості рішення приймається:

$$x^* = x_{3(i)} \text{ при } f(x_{3(i)}) < f(x_{4(i)})$$

$$x^* = x_{4(i)} \text{ при } f(x_{4(i)}) < f(x_{3(i)})$$

5.2.4 Встановлення відрізка локалізації (вихідного інтервалу невизначеності)

Розглянуті вище методи звуження інтервалу невизначеності включають два етапи:

- етап встановлення меж інтервалу;
- етап зменшення інтервалу.

Способи зменшення інтервалу ми вже розглянули. Розглянемо тепер етап встановлення меж інтервалу. Зазвичай використовується евристичний метод, наприклад, метод Свенна.

Отже, нехай потрібно знайти мінімум функції $f(x)$ не на відріжку, а на всій осі x . Припустимо знову, що функція $f(x)$ унімодальна. Виберемо якесь початкове наближення x_0 , і зробимо з нього крок деякої довжини h - $x_1 = x_0 + h$ (Рис. 5.7):

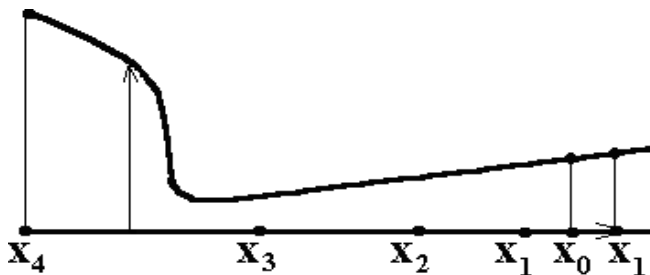


Рисунок 5.7 - Ілюстрація до методу Свенна

Якщо $f(x_1)$ виявиться більше, ніж $f(x_0)$, то змінимо направлення кроку і визначимо $x_1 = x_0 - h$. Нехай тепер $f(x_1) < f(x_0)$. Подвоїмо крок $h' = 2h$ і визначимо $x_2 = x_1 - h'$ і т.д., до тих пір, поки на певному етапі не буде виконана умова $f(x_n) > f(x_{n-1})$.

Тепер зрозуміло, що мінімум унімодальної функції лежить на відрізьку $[x_4, x_3]$ і його можна знайти одним з розглянутих методів.

Головна перевага пошукових методів полягає в тому, що вони засновані на обчисленні тільки значень функції і, отже, не вимагають виконання умови диференційованості і запису цільової функції в аналітичному вигляді. Остання властивість особливо цінна при імітаційному моделюванні. Однак ця перевага пошукових методів є і їх недоліком - швидкість їх збіжності невелика.

Питання для самоперевірки

1. Яка оптимізація називається одномірною?
2. Що таке стаціонарна точка?
3. Що таке одномірна безумовна оптимізація?
4. Чому в розділі розглядається мінімізація функції?
5. Чому є необхідність застосування чисельних методів?
6. Що таке відрізок локалізації?
7. Яка трійка чисел називається вдалою?
8. Що таке ітерація?
9. Як визначається точність чисельного методу?
10. Які групи методів використовуються при одновимірній оптимізації.
11. Який інтервал називається інтервалом невизначеності?
12. В чому полягає метод загального пошуку мінімуму?
13. В чому полягає метод половинного ділення?
14. В чому полягає метод золотого перетину?
15. Чому дорівнює «золоте відношення»?
16. Назвіть спосіб та алгоритм встановлення вихідного інтервалу невизначеності.
17. Як записується умова унімодальності функції.

5.3. Методи з використанням похідних: метод Ньютона-Рафсона

Нехай функція $f(x)$ двічі диференціюється. Як відомо з математичного аналізу, умовою мінімуму функції є рівність:

$$f'(x^*) = 0 \quad (5.19)$$

Це необхідна умова. Однак для того щоб точка x^* була мінімумом, має також виконуватися достатня умова:

$$f''(x^*) > 0 \quad (5.20)$$

Отже, будемо чисельно розв'язувати рівняння:

$$f'(x) = 0 \quad (5.21)$$

Для цього задається деяке початкове наближення x_k а сама функція розкладається в ряд Тейлора який обмежується членами до другого порядку включно, тобто будується квадратична модель функції:

$$\hat{f}(x) \approx f(x_k) + f'(x_k)(x - x_k) + \frac{1}{2}f''(x_k)(x - x_k)^2 \quad (5.22)$$

Рішення цього рівняння відносно x дає чергове $k+1$ наближення до мінімуму:

$$x_{k+1} = x_k - \frac{f'(x_k)}{f''(x_k)} \quad (5.23)$$

Алгоритм, що виражається рівнянням (5.23) з деякими наступними уточненнями має назву «**Метод Ньютона**» або «**Метод Ньютона-Рафсона**». У алгоритму (5.23) є два недоліки:

- По-перше, рівняння (5.19) може визначати не тільки мінімум, але і максимум.

- По-друге, модельна функція (5.22) може сильно відрізнятися від функції що оптимізується $f(x)$ і крок $x_{k+1} - x_k$ може виявитися занадто великим (рис. 5.8):

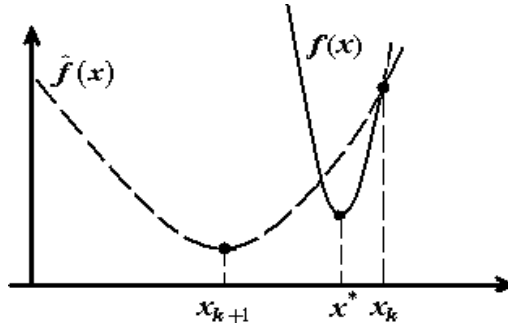


Рисунок 5.8 - Ілюстрація до методу Ньютона

Тому стратегію (5.23) слід уточнити. Щоб бути впевненими, що ми просуваємося до мінімуму, будемо на кожному кроці перевіряти співвідношення $f(x_{k+1}) < f(x_k)$. Якщо воно виконується, то переходимо до наступного кроку і т.д. Якщо ж $f(x_{k+1}) > f(x_k)$, а $f'(x_k)(x_{k+1} - x_k) < 0$, то функція $f(x)$ повинна спочатку зменшуватися в напрямку від x_k до x_{k+1} , тому наступну прийнятну точку можна знайти, дроблячи крок у тому-ж напрямку, наприклад, задавши

$$x'_{k+1} = \frac{x_k + x_{k+1}}{2} \quad (5.24)$$

З формули (5.23) видно, що вираз $f'(x_k)(x_{k+1} - x_k)$ має негативне значення тоді і тільки тоді, коли $f''(x_k) > 0$. Це означає, що якщо локальна модель яка використовується для отримання ньютонівського кроку, має мінімум, а не максимум, то гарантується існування відповідного напрямку кроку. З іншого боку, якщо $f''(x_k) < 0$ та $f'(x_k)(x_{k+1} - x_k) > 0$, то при переході від x_k до x_{k+1} , $f(x)$ збільшується, тому крок потрібно зробити в протилежному напрямку:

$$x'_{k+1} = \frac{x_k + x_{k-1}}{2} \quad (5.25)$$

Критерій припинення ітерацій при оптимізації можна вибрати у вигляді:

$$\left| \frac{f(x_k)}{f'(x_k)} \right| < \varepsilon \quad (5.26)$$

де ε - заздалегідь задана точність.

У деяких задачах похідні функції $f(x)$ недоступні і метод Ньютона можна модифікувати. Виберемо початкове наближення x_k і малий крок h . Тоді:

$$x_{k+1} = x_k - h \frac{f(x_k + h) - f(x_k - h)}{f(x_k + h) - 2f(x_k) + f(x_k - h)} \quad (5.27)$$

Формула (5.27) дає основний крок алгоритму, що має назву **квазіньютонівський метод** або **модифікований метод Ньютона**. Всі міркування щодо кроку $x_{k+1} - x_k$, наведені для методу Ньютона-Рафсона, залишаються в силі.

Питання для самоперевірки

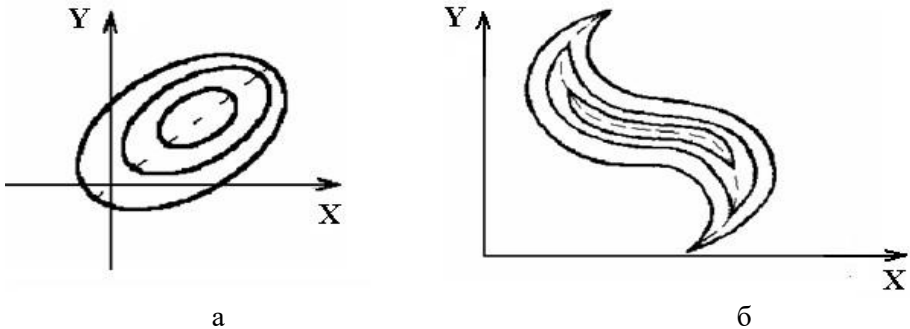
1. Як будується квадратична модель функції?
2. В чому полягає алгоритм методу Ньютона-Рафсона?
3. Назвіть недоліки методу Ньютона та як вони виправляються.
4. Яку перевірку треба робити на кожному кроці методу Ньютона?
5. За яких умов треба дробити крок та в якому напрямку?
6. Як записується критерій припинення ітерацій за методом Ньютона?
7. В яких випадках використовується модифікований метод Ньютона?
8. Як визначається значення аргументу в ітераціях за модифікованим методом Ньютона?

6. МІНІМУМ ФУНКЦІЇ БАГАТЬОХ ЗМІННИХ

6.1 Рельєф функції

Поняття «рельєф функції» зручно розглянути на прикладі функції двох змінних $z = F(x, y)$. Ця функція описує деяку поверхню в тривимірному просторі з координатами x, y, z . Завдання $F(x, y) \rightarrow \min$ означає пошук нижчої точки цієї поверхні.

Як в топографії, зобразимо рельєф цієї поверхні лініями рівня. Проведемо рівновіддалені площини $z = \text{const}$ і знайдемо лінії їх перетину з поверхнею $F(x, y)$. Проекції цих ліній на площину x, y називають **лініями рівня** або **лініями рівного відгуку**. Напрямок спадання функції будемо вказувати штрихами поруч з лініями рівня. По виду ліній рівня умовно виділимо три типи рельєфу: *улоговинний, яружний і невпорядкований* (Рис. 6.1а, б; - 6.5).



а – улоговинний; б - яружний

Рисунок 6.1 - Рельєфи функції

При улоговинному рельєфі лінії рівня схожі на еліпси (рис. 6.1а).

Розглянемо яружний тип рельєфу. Якщо лінії рівня кусочно-гладкі (рис. 6.1б), то виділимо на кожній з них точку зламу. Геометричне місце точок зламу назвемо **справжнім яром**, якщо кут спрямований в бік зростання функції, і **істинним або справжнім гребнем** - якщо в сторону зменшення.

Найчастіше лінії рівня всюди гладкі, але на них є ділянки з великою кривизною. Геометричні місця точок з найбільшою кривизною назвемо **розв'язним яром або гребнем** – зображено пунктирною лінією на Рис. 6.2.

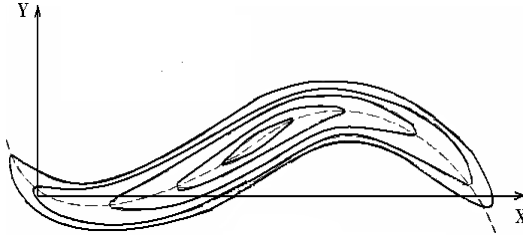


Рисунок 6.2 - Розв'язний яр

Наприклад, рельєф функції $F(x, y) = 10[y - \sin(x)]^2 + 0.1x^2$ (Рис.6.3) має яскраво виражений звивистий розв'язний яр, «дно» якого - синусоїда, а найнижча точка - початок координат.

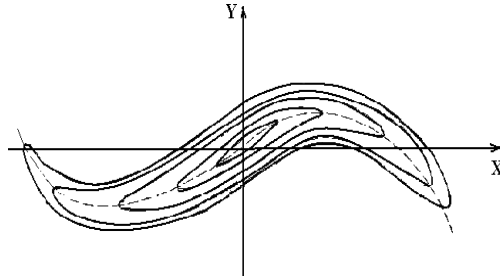


Рисунок 6.3 - Рельєф функції $F(x, y) = 10[y - \sin(x)]^2 + 0.1x^2$

Невпорядкований тип рельєфу характеризується наявністю багатьох максимумів і мінімумів. Так, функція $F(x, y) = (1 + \sin^2 x)(1 + \sin^2 y)$ - Рис. 6.4, має мінімуми в точках $x_k = \pi k, y_k =$

πl і максимуми в точках, зсунутих відносно мінімумів на $\pi/2$ по кожній координаті.

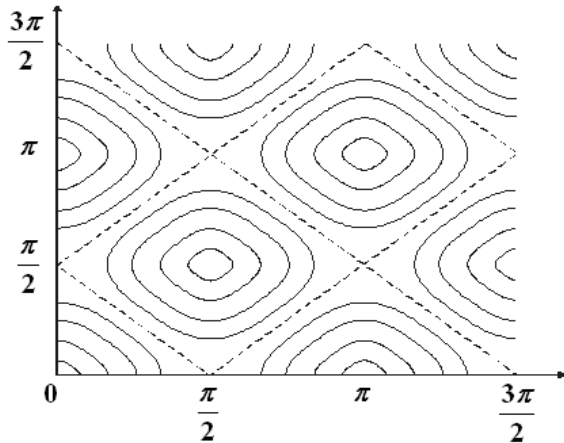


Рисунок 6.4 - Невпорядкований рельєф

Всі ефективні методи пошуку мінімуму зводяться до побудови траєкторій, уздовж яких функція спадає; різні методи відрізняються способами побудови таких траєкторій. Метод, пристосований до одного типу рельєфу, може виявитися поганим на рельєфі іншого типу.

6.2 Метод покоординатного спуску: метод Гауса

Викладемо цей метод на прикладі функції двох змінних $F(x, y)$. Виберемо нульове наближення x_0, y_0 . Фіксуємо значення однієї з координат $y = y_0$. Тоді функція буде залежати тільки від однієї змінної - x ; позначимо її через $f_1(x) = F(x, y_0)$. Використовуючи будь-який спосіб одномірної оптимізації, відшукаємо мінімум функції $f_1(x)$ і позначимо його через x_1 . Ми зробили крок з точки (x_0, y_0) в цілі (x_1, y_0) у напрямку, паралельно осі x ; на цьому кроці значення функції зменшилося.

Тепер з нової точки зробимо спуск у напрямку, паралельному осі y , тобто розглянемо функцію $f_2(y) = F(x_1, y)$, знайдемо її мінімум і

позначимо його через. Другий крок призводить нас в точку (x_1, y_1) що завершує цикл спусків або першу ітерацію.

Далі будемо повторювати цикли. На кожному спуску функція не збільшується, і при цьому значення функції обмежена знизу її значенням в мінімумі $F^* = F(x^*, y^*)$. Отже, ітерації сходяться до деякої межі $F \geq F^*$. Чи буде тут мати місце рівність, тобто чи зійдуться спуски до мінімуму і як швидко? Це залежить від функції і вибору нульового наближення. На прикладі функції двох змінних легко переконатися, що існують випадки збіжності спуску по координатах до мінімуму і випадки, коли цей спуск до мінімуму не сходиться. Справді, розглянемо геометричне трактування цього методу (Рис. 6.5):

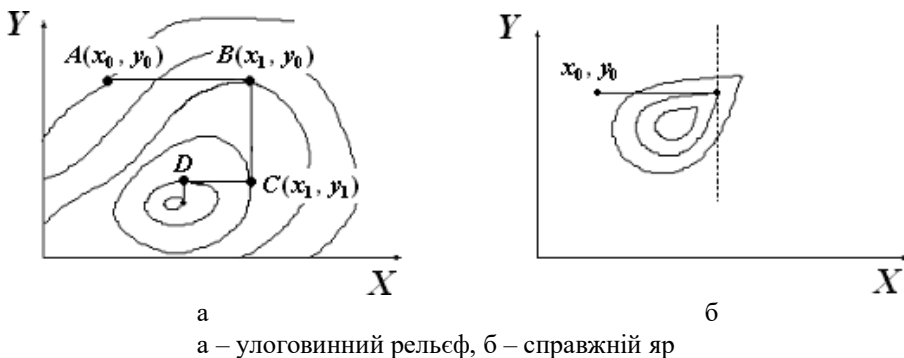


Рисунок 6.5 - Метод Гауса

Будемо рухатися за обраним напрямом, тобто уздовж деякої прямої в площині x, y . У тих ділянках, де пряма перетинає лінії рівня, ми при русі переходимо від однієї лінії рівня до іншої, так що при цьому русі функція змінюється (зростає або убыває, в залежності від напрямку руху). Тільки в тій точці, де дана пряма дотична лінії рівня (Рис. 6.5а), функція має екстремум уздовж цього напрямку. Знайшовши таку точку, ми завершуємо в ній спуск по першому напрямку і повинні почати спуск по другому.

Нехай лінії рівня утворюють справжній яр. Тоді можливий випадок (Рис. 6.5б), коли спуск по одній координаті призводить нас на «дно яру», а будь який рух за такою координаті (пунктирна лінія) веде нас на підйом. Ніякий подальший спуск по координатам неможливий,

хоча мінімум ще не досягнуто. В даному випадку процес спуску по координатам не сходиться до мінімуму.

Навпаки, якщо функція досить гладка, то в деякому околі мінімуму процес спуску по координатах сходиться до цього мінімуму. Однак швидкість збіжності сильно залежить від форми ліній рівня. Так, якщо рельєф функції має тип «розв'язний яр», то при попаданні траєкторії спуску в такий яр збіжність стає настільки повільною, що розрахунок практично вести неможливо.

Зазвичай метод покоординатного спуску використовують в якості першої спроби при знаходженні мінімуму.

Питання для самоперевірки

1. Які утворюються лінії рівня?
2. Які типи рельєфу функції існують?
3. Яким чином зображується рельєф функції на площині?
4. Що таке ісправжній яр?
5. Що таке справжній гребінь?
6. Що таке розв'язний яр або гребінь?
7. Чим характеризується невпорядкований тип рельєфу?
8. В чому полягає метод Гаусса?
9. Чи завжди можна віднайти мінімум за допомогою методу Гаусса?
10. Для яких типів рельєфу найбільш підходить метод Гаусса?
11. В якості чого використовують найчастіше метод Гаусса?

7. УМОВНА ОПТИМІЗАЦІЯ

7.1. Загальні положення

Ряд інженерних задач пов'язаний з оптимізацією при наявності деякої кількості обмежень на керовані змінні. Такі обмеження істотно зменшують розміри області, в якій розшукується оптимум. На перший погляд може здатися, що зменшення розмірів допустимої області має спростити процедуру пошуку оптимуму. Однак, навпаки, процес оптимізації стає більш складним, оскільки при наявності обмежень навіть не можна використовувати застосовувані нами вище умови оптимальності. При цьому може порушуватися навіть основна умова, відповідно до якої оптимум повинен досягатися в **стаціонарній точці**, яка характеризується нульовим градієнтом. Таким чином відрізняємо **безумовний мінімум**, як мінімум без урахування обмежень в стаціонарній точці та **умовний мінімум** – мінімум з урахуванням обмежень навіть якщо він і не знаходиться в стаціонарній точці.

Наприклад, безумовний мінімум функції $f(x) = (x - 2)^2$ має місце в стаціонарній точці $x = 2$. Але якщо завдання мінімізації вирішується з урахуванням обмеження $x > 4$, то буде знайдений умовний мінімум, якому відповідає точка $x = 4$. Ця точка не є стаціонарною точкою функції $f(x)$, так як $f'(x) \neq 0$. Тому потрібно вивчити необхідні і достатні умови оптимуму в задачах з обмеженнями, які інакше називають задачами умовної оптимізації.

В загальному вигляді задачу умовної оптимізації можна записати:

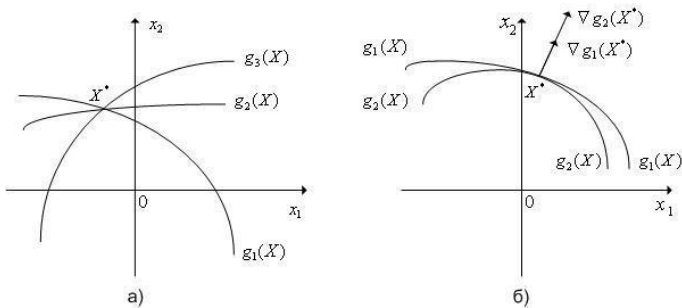
$$f(x) \rightarrow \min, x \in R^n \quad (7.1)$$

$$h_j(x) = 0, j = 1, 2, \dots, K \quad (7.2)$$

$$g_j(x) \leq 0, j = 1, 2, \dots, J \quad (7.3)$$

У точці локального мінімуму x^* задачі (7.1) кожне з обмежень (7.2), (7.3) виконується або у вигляді рівності $h_j(x) = 0$, $g_j(x^*) = 0$, або у вигляді нерівності $g_j(x^*) < 0$. Обмеження у вигляді рівності називаються **активними обмеженнями**. Інші обмеження називаються **неактивними обмеженнями**.

Якщо точка $x^* \in D$ і обмеження активні, то умова лінійної незалежності градієнтів функцій $g_j(x^*)$ (або $h_j(x) = 0$) активних обмежень в точці x^* називається умовою **регулярності обмежуючих функцій в точці x^*** . Ця умова означає, що, наприклад, при кількості незалежних факторів $n = 2$ кількість обмежуючих функцій, що проходять через точку x^* , не повинно перевищувати 2 і в точці x^* вектори $\nabla g_1(x)$, $\nabla g_2(x)$ (або $\nabla h_1(x)$, $\nabla h_2(x)$) не повинні бути колінеарні. Наприклад, на рис. 7.1 в ситуації (а) кількість обмежують функцій, що проходять через точку x^* , перевищує розмірність вектору варійованих факторів, в ситуації (б) в точці x^* градієнти $\nabla g_1(x)$, $\nabla g_2(x)$ обмежуючих функцій колінеарні.



а – перевищення кількості обмежуючих функцій; б – колінеарність векторів

Рисунок 7.1 - Ситуації і яких не виконується умова регулярності обмежуючих функцій

Питання для самоперевірки

1. Що таке умовна оптимізація?
2. Що таке стаціонарна точка?
3. Що таке безумовний оптимум?
4. Що таке умовний оптимум?
5. Які обмеження називаються активними?
6. Які обмеження називаються неактивними?

7. Що означає умова регулярності обмежуючих функцій в точці x^* ?
8. В яких ситуаціях не виконується умова регулярності обмежуючих функцій?

7.2. Задачі з обмеженнями у вигляді рівностей. Множники Лагранжа

За допомогою методу множників Лагранжа по суті встановлюються необхідні умови, що дозволяють ідентифікувати точки оптимуму в задачах оптимізації з обмеженнями-рівностями. При цьому задача з обмеженнями перетвориться в еквівалентну задачу безумовної оптимізації, в якій фігурують деякі невідомі параметри, що називаються множниками Лагранжа.

Розглянемо задачу з одним обмеженням-рівністю:

$$f(x) \rightarrow \min, x \in R^n \quad (7.4)$$

$$h_1(x) = 0 \quad (7.5)$$

Відповідно до методу множників Лагранжа ця задача перетворюється в наступну задачу безумовної мінімізації:

$$L(x, \lambda) = f(x) - \lambda h_1(x) \rightarrow \min, x \in R^n \quad (7.6)$$

Функція $L(x, \lambda)$ називається функцією Лагранжа. Тут λ – множник Лагранжа. Нехай при заданому значенні $\lambda = \lambda^0$ безумовний мінімум функції $L(x, \lambda)$ по змінній x досягається в точці $x = x_0$ і x_0 задовольняє рівнянню (7.5). Тоді, як не важко бачити, x_0 мінімізує (7.4) з урахуванням (7.5), оскільки для всіх значень x , що задовольняють (7.5), $\min L(x, \lambda) = f(x)$.

Зрозуміло, потрібно підібрати значення $\lambda = \lambda^0$ таким чином, щоб координата точки безумовного мінімуму x_0 задовольняла рівності (7.5). Це можна зробити, якщо, розглядаючи λ як змінну, знайти безумовний мінімум функції Лагранжа (7.6) у вигляді функції λ , а потім вибрати значення λ , при якому виконується рівність (7.5).

Приклад.

Вирішити задачу

$$f(x_1, x_2) = x_1^2 + x_2^2$$

За наявності обмеження:

$$h_1(x_1, x_2) = 2x_1 + x_2 - 2 = 0$$

Побудуємо функцію Лагранжа:

$$L(x_1, x_2, \lambda) = x_1^2 + x_2^2 - \lambda(2x_1 + x_2 - 2)$$

Та визначимо її безумовний мінімум. Знайдемо її стаціонарну точку прирівнявши нулю компоненти її градієнта або її частинні похідні:

$$\frac{\partial L}{\partial x_1} = 2x_1 - 2\lambda = 0 \Rightarrow x_1^0 = \lambda$$

$$\frac{\partial L}{\partial x_2} = 2x_2 - \lambda = 0 \Rightarrow x_2^0 = \frac{\lambda}{2}$$

Для того щоб перевірити чи відповідає стаціонарна точка з координатами (x_1^0, x_2^0) глобальному мінімуму, розрахуємо матрицю Гессе функції Лагранжа – ця матриця, що складається з *других частинних похідних* має бути позитивно визначена для того, щоб мати глобальний мінімум в зазначеній точці. В нашому випадку матриця Гессе буде мати вигляд:

$$\begin{pmatrix} \frac{\partial^2 L}{\partial x_1^2} & \frac{\partial^2 L}{\partial x_1 \partial x_2} \\ \frac{\partial^2 L}{\partial x_2 \partial x_1} & \frac{\partial^2 L}{\partial x_2^2} \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$

Матриця є позитивно визначеною, коли є позитивними всі її кутові мінори. В нашому випадку – матриця 2x2 має один кутовий мінор, що дорівнює 2 – тобто матриця визначена позитивно:

$$\begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$

Оптимальне значення λ буде знайдено шляхом підстановки x_1^0, x_2^0 в рівняння для $h_1(x_1, x_2)$ звідки отримаємо $\lambda = \frac{4}{5}$. Таким чином умовний мінімум досягається при: $x_1^0 = \frac{4}{5}, x_2^0 = \frac{2}{5}$ та дорівнює $f(x_1, x_2) = \frac{4}{5}$

Цей метод можна розповсюдити на випадок, коли є декілька обмежень у вигляді рівностей. В даному курсі не розглядається.

Питання для самоперевірки

1. Яким чином враховуються обмеження у вигляді рівностей?
2. Як записується функція Лагранжа?
3. Назвіть умову глобального мінімуму функції Лагранжа?
4. Яка матриця є позитивно визначеною?

7.3. Задачі з обмеженнями у вигляді нерівностей.

Розглянемо задачу з обмеженнями-нерівностями

$$f(x) \rightarrow \min, x \in R^n \quad (7.7)$$

$$g_j(x) \leq 0, j = 1, 2, \dots, J \quad (7.8)$$

Процес оптимізації функції (7.7) за наявності умов (7.8) може виявитися таким, що не дуже відрізняється від розглянутих в розділах 5 та 6. Тобто *почати є сенс з аналітичного вирішення задачі безумовної оптимізації функції (7.7) та знайти її стаціонарні точки, після чого або перевірити знайдене рішення на відповідність умовам (7.8) або при наявності декількох рішень – відібрати ті, що їм відповідають.* Формалізувати цей процес або знайти рішення, що не є стаціонарними точками та відповідають умовам (7.8) можна за допомогою розглянутого вище метод Лагранжа. Але розглянемо спочатку декілька загальних положень.

Велике значення в теорії та обчислювальній практиці має *теорема Куна-Таккера* для задачі умовної оптимізації з обмеженнями типу нерівностей. Не розглядаючи саму теорему, зазначимо, що якщо функція (7.7) при обмеженнях (7.8) має локальний мінімум в точці x^* , що відповідає вимозі регулярності обмежуючих функцій, то замість задачі умовної оптимізації можна вирішувати задачу безумовної оптимізації функції Лагранжа (7.6) - Див. Рис. 7.2.

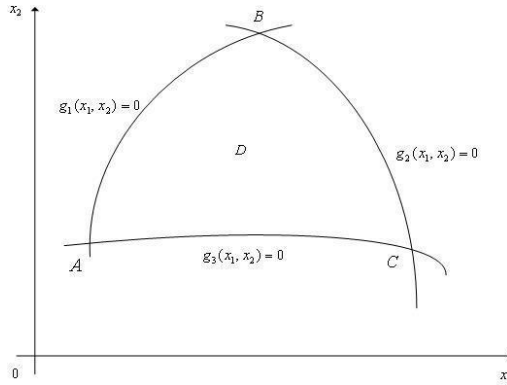


Рисунок 7.2 - До теореми Куна-Таккера

Метод Лагранжа можна розповсюдити на випадок наявності обмежень типу (7.8), перетворивши їх у обмеження в вигляді рівностей шляхом додавання до кожного з них *ослаблюючої змінної* u_j^2 :

$$g_j(x) + u_j^2 = 0 \quad (7.9)$$

Таким чином, задача зводиться до мінімізації функції при наявності обмежень у вигляді рівностей та функція Лагранжа буде мати вигляд:

$$L(x, u, \lambda) = f(x) - \sum_{j=1}^J \lambda_j [g_j(x) + u_j^2] \quad (7.10)$$

Необхідними умовами, що мають виконуватися в стаціонарній точці (якщо це точка мінімуму) є наступні:

$$\left\{ \begin{array}{l} \frac{\partial L}{\partial x} = \frac{\partial f}{\partial x} - \sum \lambda_j \frac{\partial g_j}{\partial x} \\ \frac{\partial L}{\partial \lambda_j} = g_j(x) + u_j^2 \\ \lambda_j \geq 0 \\ \frac{\partial L}{\partial u_j} = 2\lambda_j u_j = 0 \end{array} \right. \quad (7.11)$$

Останнє рівняння перетворимо з урахуванням (7.9) та помножимо його на $u_j/2$. Отримаємо:

$$\lambda_j u_j^2 = 0, \quad -\lambda_j g_j(x) = 0 \text{ або } \lambda_j g_j(x) = 0 \quad (7.12)$$

Отримане рівняння (7.12) є досить принциповим, воно означає:

- Якщо $\lambda_j \neq 0$ то $g_j(x) = 0$ і обмеження є активним
- Якщо обмеження є неактивним, тобто $g_j(x) < 0$ то відповідний множник Лагранжа $\lambda_j = 0$

Приклад:

Знайти умовний мінімум наступної задачі:

$$f(x_1, x_2) = x_1^2 + x_2^2 \rightarrow \min$$

$$g(x_1, x_2) = 2x_1 + x_2 - 2 \leq 0$$

Складемо функцію Лагранжа:

$$L = x_1^2 + x_2^2 - \lambda(2x_1 + x_2 - 2 + u^2)$$

Умови існування умовного мінімуму:

$$\left\{ \begin{array}{l} (x_1 - \lambda) = 0 \\ 2x_2 - \lambda = 0 \\ \lambda(2x_1 + x_2 - 2) = 0 \\ \lambda \geq 0 \end{array} \right.$$

Якщо $\lambda=0$, то $x_1 = 0$ та $x_2 = 0$ що задовольняє умові $2x_1 + x_2 - 2 \leq 0$. Значення функції в точці – $f(0,0)=0$

Якщо $\lambda \neq 0$ то $2x_1 + x_2 - 2 = 0$ і тоді вирішуємо систему з 3 рівнянь:

$$\begin{cases} (x_1 + \lambda) = 0 \\ 2x_2 + \lambda = 0 \\ \lambda(2x_1 + x_2 - 2) = 0 \\ \lambda \geq 0 \end{cases}$$

З якої $x_1 = \frac{4}{5}, x_2 = \frac{2}{5}, \lambda = \frac{4}{5}$ але зазначена точка не є стаціонарною та локальним мінімумом. Значення функції в точці - $f(x)=4/5$.

Питання для самоперевірки

1. З яких дій є сенс почати вирішення задачі умовної оптимізації з обмеженнями у вигляді нерівностей?
2. Яким шляхом обмеження у вигляді нерівностей перетворюються у обмеження – рівності?
3. В чому полягає сутність теореми Куна-Таккера?
4. Які умови для функції Лагранжа повинні виконуватися в стаціонарній точці?

7.4. Методи штрафних функцій. Нежорсткі обмеження

Розглянемо задачу умовної оптимізації:

$$f(x) \rightarrow \min, x \in R^n \quad (7.13)$$

$$h_j(x) = 0, j = 1, 2, \dots, K \quad (7.14)$$

$$g_j(x) \leq 0, j = 1, 2, \dots, J \quad (7.15)$$

$$x_i^{(l)} < x_i < x_i^{(u)} \quad (7.16)$$

Таку задачу також називається задачею **нелінійного програмування**. Кажуть, що точка x відповідає допустимому

вирішення задачі нелінійного програмування, якщо для неї виконуються всі обмеження, тобто співвідношення (7.13-7.16). Передбачається, що для вектору x^* , що є вирішенням задачі нелінійного програмування, відомо деяке початкове наближення x_0 , можливо неприпустиме. У методах штрафних функцій будується послідовність точок x_m , $m = 0, 1, \dots, M$, яка починається з заданої точки x_0 і закінчується точкою x_M , що дає найкраще наближення до x^* серед усіх точок побудованої послідовності. В якості x_m беруться точки рішення допоміжної задачі безумовної мінімізації, отриманої перетворенням вихідної цільової функції за допомогою так званих штрафних функцій. У цих методах вихідна задача умовної оптимізації перетворюється в послідовність задач безумовної оптимізації.

Методи штрафних функцій класифікуються відповідно до способів обліку обмежень-нерівностей. Залежно від того, чи є елементи послідовності x_m допустимими або недопустимими точками, кажуть про методи **внутрішньої** і **зовнішньої** точки відповідно. якщо послідовність x_m містить точки обох типів, метод називають **змішаним**.

Таким чином, для вирішення задачі умовної будують допоміжну функцію виду:

$$Q(x, r, l) = f(x) + \sum_{j=1}^M r_j G_j[g_j(x)] + \sum_{j=1}^K l_j H_j[h_j(x)] \quad (7.17)$$

Або

$$Q(x, r, l) = f(x) + r \sum_{j=1}^M G_j[g_j(x)] + l \sum_{j=1}^K H_j[h_j(x)] \quad (7.18)$$

та вирішують для неї послідовність задач безумовної оптимізації:

$$Q(x, r, l) \rightarrow \min, x \in R^n \quad (7.19)$$

В методі зовнішньої точки який ще називається **метод зовнішніх штрафних функцій** – функції G та H вибирають таким

чином, щоб вони ставали відмінними від нуля (позитивними) при порушенні відповідного обмеження (див. Рис.7.3). А оскільки ми мінімізуємо (7.13) то рух в бік порушення стає не вигідним. Всередині допустимої області функції H та G повинні дорівнювати нулю.

Для обмежень нерівностей:

$$G_j(g_j(x)) > 0 \text{ при } g_j(x) > 0 \quad (7.20)$$

$$G_j(g_j(x)) = 0 \text{ при } g_j(x) \leq 0$$

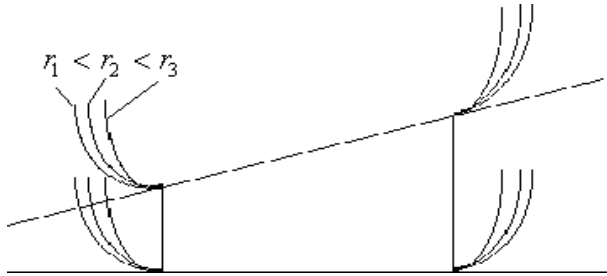


Рисунок 7.3 - Поводження штрафних функцій за методом зовнішньої точки

Одною з форм функції (7.20) є квадратична штрафна функція:

$$G_j(g_j(x)) = \left\{ \frac{1}{2} [g_j(x) + |g_j(x)|] \right\}^2 = \frac{1}{4} [g_j(x) + |g_j(x)|]^2 \quad (7.21)$$

Взагалі, показник ступеню в рівнянні (7.21) може бути будь який парний.

Для обмежень – рівностей:

$$H_j(h_j(x)) \rightarrow 0 \text{ при } h_j(x) \rightarrow 0 \quad (7.22)$$

Вимогі (7.22) відповідають наступні штрафні функції:

$$H_j(h_j(x)) = |h_j(x)| \quad (7.23)$$

$$H_j(h_j(x)) = (h_j(x))^\alpha \quad (7.24)$$

Де α – будь яке парне число. Наприклад $\alpha = 2$

Обмеження-нерівності задачі умовної оптимізації до яких може бути застосовано метод зовнішньої точки можна інтерпретувати як **нежорсткі**, тобто такі які можна подолати. Але треба розуміти, що якість процесу при цьому може істотно погіршитися. В цьому випадку фактори обмежуються як правило верхньою - b_j^U та нижньою - b_j^L межами, а самі обмеження приймають відповідно вигляд (7.25) та (7.26):

$$g_j(x) = x - b_j^U - \varepsilon \leq 0 \quad (7.25)$$

$$g_j(x) = b_j^L - x - \varepsilon \leq 0 \quad (7.26)$$

Де ε – точність оптимізації

В методі внутрішньої точки або **бар'єрних функцій** функції H , G вибираються відмінними від нуля в допустимій області та такими, щоб при наближенні до кордону допустимої області (зсередини) вони зростали, перешкоджаючи виходу при пошуку за кордон області (Рис. 7.4). У цьому випадку ці функції повинні бути малими (позитивними чи негативними) всередині допустимої області та великими позитивними поблизу кордону (всередині області) та називаються **бар'єрними**. Наприклад, для обмежень-нерівностей:

$$G_j(g_j(x)) \rightarrow \infty \text{ при } g_j(x) \rightarrow 0^- \quad (7.27)$$

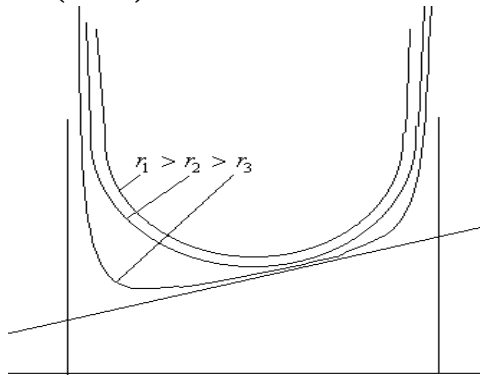


Рисунок 7.4 - Поводження штрафних функцій за методом бар'єрних функцій

В якості бар'єрних функцій для обмежень-нерівностей можуть слугувати функції виду:

$$G_j(g_j(x)) = \frac{1}{-g_j(x)} \quad (7.28)$$

$$G_j(g_j(x)) = -\ln(-g_j(x)) \quad (7.29)$$

Логарифмічний штраф (7.29) - це бар'єрна функція, що не визначена в неприпустимих точках (тобто для таких x , що $g(x) > 0$). Тому в тих випадках, коли доводиться мати справу з неприпустимими точками (наприклад, коли задане початкове наближення x_0 не є допустимим), потрібно спеціальна процедура, що забезпечує потрапляння в допустиму область.

Штраф, заданий функцією (7.28), не має негативних значень в допустимій області. Цей штраф, як і попередній, є бар'єром; в цьому випадку також виникають труднощі, пов'язані з можливою появою недопустимих точок.

Штрафні функції-рівності для методу внутрішньої точки може бути вибрані у вигляді (7.23), (7.24).

В загальному вигляді алгоритм оптимізації методом штрафних функцій виглядає наступним чином:

1. На підставі задачі (7.13 - 7.16) будуємо функцію (7.17) або (7.18), вибираємо початкові наближення x і початкові значення коефіцієнтів штрафу r_j , l_k (r , l), кількість ітерацій, точність безумовної оптимізації, точність дотримання обмежень і т.д.

2. Вирішуємо задачу (7.19) – стаціонарна точка функції $Q(x, r, l)$ повинна відповідати мінімальному значенню функції $f(x)$ що задовольняє обмеженням.

3. Якщо отримане рішення не задовольняє системі обмежень у випадку використання методу штрафних функцій, то:

- Для методу зовнішньої точки **збільшуємо** значення коефіцієнтів штрафу
- Для методу внутрішньої точки **зменшуємо** значення коефіцієнтів штрафу

і знову вирішуємо задачу (7.18).

4. Процес припиняється, якщо знайдене рішення задовольняє системі обмежень з певною точністю.

Приклад:

Методом штрафних функцій знайти умовний мінімум функції:

$$f(x_1, x_2) = 2(x_1 - 2)^2 + x_2^2$$

З обмеженнями:

$$g_1(x_1, x_2) = x_1 - 2x_2 + 1 \leq 0$$

$$g_2(x_1, x_2) = -2x_1 - x_2 + 2 \leq 0$$

Функція $f(x)$ має стаціонарну точку безумовного мінімуму $x_1=2, x_2=0$ яка не задовольняє обмеженням. Штрафну функцію приймаємо у вигляді (7.21) та отримуємо допоміжну функцію (7.18) пошук безумовного мінімуму якої здійснюємо використовуючи розглянуті раніше методи половинного ділення та покоординатного спуску (Гауса), корегуючи при цьому коефіцієнт r в сторону збільшення починаючи з $r = 1$ таким чином, щоб мінімальне значення функції (7.18) дорівнювало значенню функції $f(x)$ що відповідає обмеженням.

Отримуємо рішення: $x_1=2; x_2=1.25; f(x_1, x_2) = 2.5$ $g_1(x_1, x_2) = 0$ $g_2(x_1, x_2) = -3.5; r = 3$. Графічну інтерпретацію рішення та поводження штрафної функції у перетині при $x_1=2$ представлено на Рис. 7.5. Тут зеленим квадратом показана точка безумовного мінімуму функції $f(x)$, а червоними трикутниками точки умовного мінімуму $f(x)$ та безумовного мінімуму $Q(x, r)$ – що є рішенням задачі, а також значення обмежень.

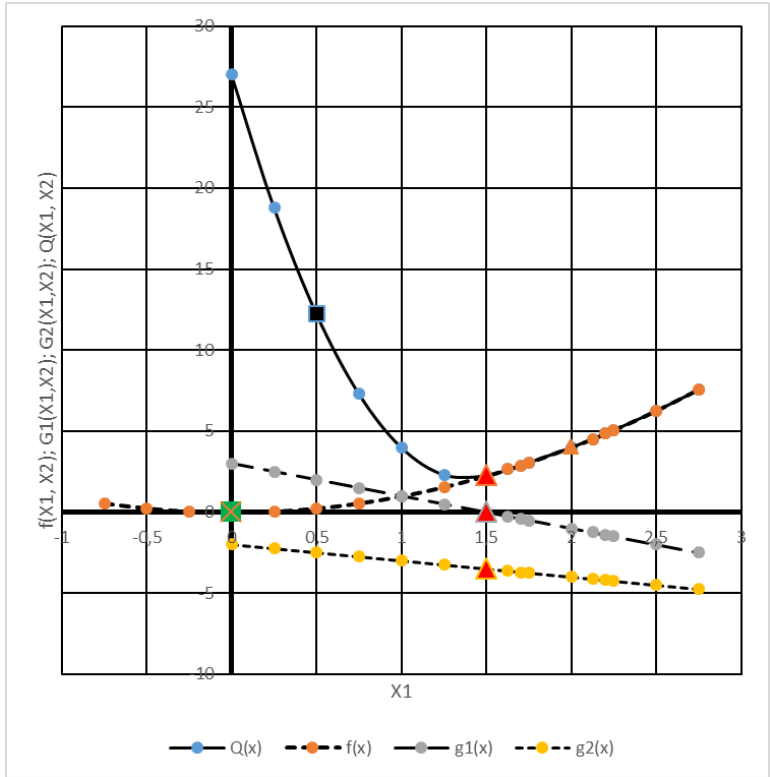


Рисунок 7.5 - Графічна інтерпретація методу штрафних функцій

Питання для самоперевірки

1. Які існують різновиди методу штрафних функцій?
2. В чому полягає метод зовнішньої точки?
3. Які обмеження можна інтерпретувати як нежорсткі?
4. Яку форму мають нежорсткі обмеження факторів?
5. Як записується квадратична штрафна функція?
6. Наведіть приклад штрафної функції для обмеження-рівності за методом внутрішньої точки?
7. В чому полягає метод внутрішньої точки?
8. Які функції називаються бар'єрними?
9. Наведіть приклад бар'єрної функції.
10. Які складнощі можуть виникнути при застосуванні бар'єрних функцій?
11. Який алгоритм пошуку мінімуму методом штрафних функцій?

8. ДЕЯКІ ТИПОВІ ОБМЕЖЕННЯ ПРИ МЕХАНІЧНІЙ ОБРОБЦІ

Розглянемо деякі обмеження, що виникають при оптимізації процесів механічної обробки, зокрема токарної.

Обмеження швидкості різання, виходячи з заданої стійкості різального інструменту:

$$V \leq \frac{C_v}{t^{x_v} S^{y_v} T^m} K_v \text{ або } n - \frac{318 C_v}{t^{x_v} S^{y_v} T^m D} K_v \leq 0 \quad (8.1)$$

Де n – частота обертання шпинделя станка, D – діаметр заготовки

При цьому стійкість має бути не менше ніж час обробки однієї деталі t_0 : $t_0 \leq T$, а глибина різання не більше за припуск A : $t \leq A$

Обмеження подачі, виходячи з шорсткості поверхні R_z . В першому наближенні, вибираючи подачу, можна скористатися формулами для розрахунку геометричної складової шорсткості, маючи на увазі, що реальна шорсткість буде вище, оскільки тут не враховується вплив жорсткості технологічної системи, індивідуальних властивостей оброблюваного та інструментального матеріалі, інших режимів та умов різання. Наведені залежності справедливі для повздовжнього точіння. Для гострої вершини різця умова буде мати вигляд:

$$S \leq \frac{R_z (tg\varphi + tg\varphi_1)}{tg\varphi \cdot tg\varphi_1} \quad (8.2)$$

Де φ та φ_1 – відповідно головний та допоміжний кути в плані.

При наявності радіусу заокруглення вершини різця R , що характерно для різців з механічним кріпленням змінних багатограних пластин умова буде мати вигляд:

$$S \leq \sqrt{8R \cdot R_z} \quad (8.3)$$

Обмеження розрахункової частоти обертання заготовки n та обраної подачі S кінематичним можливостям верстата:

$$n \leq n_{max} \text{ або } n - n_{max} \leq 0 \quad (8.4)$$

$$S \leq S_{max} \text{ або } S - S_{max} \leq 0 \quad (8.5)$$

Обмеження обумовлене потужністю верстата:

$$N_e \leq N_{ед} \eta \quad (8.6)$$

Де N_e – ефективна потужність різання; $N_{ед}$ – потужність електродвигуна верстата; η – к.к.д

Оскільки ефективна потужність визначається:

$$N_e = \frac{P_z V}{1000 \cdot 60} \quad (8.7)$$

Де P_z – головна складова сили різання, яка визначається:

$$P_z = C_p t^{x_p} S^{y_p} V^{z_p} K_p \quad (8.8)$$

З (8.6), (8.7), (8.8) отримаємо ще одну залежність, що обмежує швидкість різання:

$$V \leq \left(\frac{1000 \cdot 60 \cdot N_{ед} \eta}{C_p t^{x_p} S^{y_p} K_p} \right)^{\frac{1}{1+z_p}} \quad (8.9)$$

Обмеження пов'язані з точністю обробки:

$$P_y \left(\frac{1}{j_{заг}} + \frac{1}{j_v} + \frac{1}{j_{інстр}} \right) \leq k_3 \frac{\delta_p}{2} \quad (8.10)$$

Де P_y – радіальна складова сили різання; $j_{заг}$, j_v , $j_{інстр}$ – відповідно жорсткість заготовки, верстата та інструмента; k_3 – коефіцієнт, що показує в яку частину допуску повинна укладатися погрішність, викликана деформацією елементів технологічної системи – зазвичай приймається в межах 0,7 ... 0,8; δ_p – допуск на розмір що обробляється. Складові формули (8.10) можна визначити за наступними залежностями:

$$P_y = C_{py} t^{x_{py}} S^{y_{py}} V^{z_{py}} K_{py} \quad (8.11)$$

$$\frac{1}{j_{\text{заг}}} = \frac{L^3}{k_1 J E} \quad (8.12)$$

де L – вільна довжина заготовки, мм; k_1 – коефіцієнт що враховує спосіб кріплення заготовки (в патроні – $k_1=3$, в центрах $k_1=70$, в патроні з піджимом центру задньої бабки $k_1=100$); E – модуль пружності матеріалу заготовки, МПа; J – момент інерції найбільш небезпечного перетину заготовки:

- для суцільного перетину $J = 0,05D^4$
- для кільцевого перетину: $J = 0,05(D^4 - d^4)$

$$\frac{1}{j_{\text{в}}} = 0,000025 k_2 \sqrt[3]{\frac{200}{H_{\text{ц}}}} \quad (8.13)$$

Де $H_{\text{ц}}$ – висота центрів верстата, мм; k_2 – коефіцієнт що враховує тип закріплення:

- при закріпленні в центрах та в патроні з піджимом центром задньої бабки $k_2 = 1$
- При закріпленні в патроні $k_2 = 0.03L$

$$\frac{1}{j_{\text{інстр}}} = \frac{l^3}{3J_{\text{р}} E_{\text{р}}} \quad (8.14)$$

Для випадку токарної обробки тут l – довжина різця, мм; $E_{\text{р}}$ – модуль пружності матеріалу різця, МПа; J – момент інерції перетину різця:

- для круглого перетину $J = 0,05d_{\text{р}}^4$
- для прямокутного перетину: $J = \frac{HB^3}{12}$

Крім того можуть вводитися такі додаткові обмеження:

- Обмеження за температурою різання на основі емпіричних формул або теоретичних розрахунків;
- Обмеження, пов'язані з дослідженням кінетики теплових процесів при різанні тобто зі структурними перетвореннями в оброблюваному матеріалі;

Як бачимо при обробці різанням існують чисельні та досить складні взаємопов'язані обмеження, які треба враховувати при оптимізації режимів різання.

Питання для самоперевірки

1. За якими показниками визначається обмеження швидкост різання?
2. За якими показниками визначається обмеження подачі?
3. Назвіть додаткові обмеження при обробці різанням.

9.3 АДАЧІ ЛІНІЙНОГО ПРОГРАМУВАННЯ

9.1 Загальні положення

Великий клас задач з оптимізації складають задачі лінійного програмування – умовна оптимізація з лінійними цільовими функціями та лінійними ж обмеженнями. На відміну від задач нелінійного програмування де оптимальне рішення може знаходитися як всередині простору окресленому обмеженнями так і на самих обмеженнях, в задачах лінійного програмування рішення завжди знаходиться на лініях обмежень.

Основна задача лінійного програмування може бути записана в наступному вигляді:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2 \\ \dots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n = b_m \end{cases} \quad (9.1)$$

$$F = c_0 + c_1x_1 + c_2x_2 + \dots + c_nx_n \rightarrow \min \quad (9.2)$$

Тут (9.1) – система обмежень задачі; (9.2) – цільова функція.

Окрім обмежень-рівностей до основної задачі входять також обмеження-нерівності:

$$\begin{cases} x_1 \geq 0 \\ x_2 \geq 0 \\ \dots \\ x_m \geq 0 \end{cases} \quad (9.3)$$

Всяке не негативне рішення системи (9.1) - $x_i^{(0)} \geq 0$ називається **допустимим**. Сукупність допустимих рішення називають **планом задачі лінійного програмування**. План, що мінімізує функцію (9.2) називається **оптимальним**. Задача має рішення тільки в тому випадку, коли ранг матриці системи обмежень (складеної з коефіцієнтів при невідомих) r не більше кількості невідомих n . Якщо $r = n$ то система має єдине рішення яке і буде оптимальним якщо воно допустимо. Якщо ж ні – то задача не має рішень. Таким чином інтерес має тільки випадок коли $n < r$, що і буде розглядатися в подальшому.

Кожну задачу лінійного програмування можна звести до форми основної. Для цього:

1. Будь яку задачу максимізації треба звести до задачі мінімізації.
2. Вміти переходити від обмежень, пов'язаних з нерівностями до еквівалентних обмежень-рівностей.

Якщо обмеження задачі задані у вигляді нерівностей:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n + b_1 \geq 0 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n + b_2 \geq 0 \\ \dots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n + b_m \geq 0 \end{cases} \quad (9.4)$$

То (9.4) завжди можна звести до вигляду (9.1) ввівши додаткові **не негативні** змінні: $x_{n+1}, x_{n+2}, \dots, x_{n+m}$:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n + b_1 = x_{n+1} \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n + b_2 = x_{n+2} \\ \dots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n + b_m = x_{n+m} \end{cases} \quad (9.5)$$

Очевидно, що умова не негативності додаткових змінних еквівалентна виконанню нерівностей (9.4). Таким чином обмеження - нерівності (9.4) еквівалентні обмеженням-рівностям (9.5), а отже і (9.1).

Розглянемо декілька прикладів задачі лінійного програмування.

9.2 Задача про використання сировини

Виготовлення продукції двох типів – P_1 та P_2 потребує сировини чотирьох видів – S_1, S_2, S_3, S_4 . Кількість сировини обмежена і складає відповідно – b_1, b_2, b_3, b_4 . Кількість одиниць сировини, що потребує виготовлення одиниці продукції відомо та задається у вигляді таблиці. Треба скласти такий план випуску продукції, щоб дохід від її реалізації був би максимальним. Умови задачі представлені в Табл. 9.1.

Таблиця 9.1 - Умови задачі про використання сировини

Види сировини	Запаси сировини	Види продукції	
		P ₁	P ₂
S ₁	b ₁	a ₁₁	a ₁₂
S ₂	b ₂	a ₂₁	a ₂₂
S ₃	b ₃	a ₃₁	a ₃₂
S ₄	b ₄	a ₄₁	a ₄₂
Дохід		c ₁	c ₂

Якщо позначити кількість продукції, що виготовляється відповідно x_1 та x_2 то цільова функція буде виглядати:

$$F = c_1x_1 + c_2x_2 \rightarrow \max \quad (9.6)$$

Оскільки будь-яку задачу треба звести до задачі мінімізації, то функція (9.6) прийме вигляд:

$$F = -c_1x_1 - c_2x_2 \rightarrow \min \quad (9.7)$$

Умови задачі при цьому можна записати у вигляді системи:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 \leq b_1 \\ a_{21}x_1 + a_{22}x_2 \leq b_2 \\ a_{31}x_1 + a_{32}x_2 \leq b_3 \\ a_{41}x_1 + a_{42}x_2 \leq b_4 \end{cases} \quad (9.8)$$

Систему (9.8) можна звести до умов основної задачі у вигляді:

$$\begin{cases} b_1 - a_{11}x_1 - a_{12}x_2 \geq 0 \\ b_2 - a_{21}x_1 - a_{22}x_2 \geq 0 \\ b_3 - a_{31}x_1 - a_{32}x_2 \geq 0 \\ b_4 - a_{41}x_1 - a_{42}x_2 \geq 0 \end{cases} \quad (9.9)$$

Або шляхом введення додаткових змінних привести до виду:

$$\begin{cases} b_1 - a_{11}x_1 - a_{12}x_2 = x_3 \\ b_2 - a_{21}x_1 - a_{22}x_2 = x_4 \\ b_3 - a_{31}x_1 - a_{32}x_2 = x_5 \\ b_4 - a_{41}x_1 - a_{42}x_2 = x_6 \end{cases} \quad (9.10)$$

Рішення розглянемо на кількісному прикладі:

Таблиця 9.2 - Вихідні дані

Види сировини	Запаси сировини	Види продукції	
		P ₁	P ₂
S ₁	19	2	3
S ₂	13	2	1
S ₃	15	0	3
S ₄	18	3	0
Дохід		7	5

Умови Табл. 9.2 можуть бути записані в наступному вигляді:

$$\begin{cases} 2x_1 + 3x_2 \leq 19 \\ 2x_1 + x_2 \leq 13 \\ 3x_2 \leq 15 \\ 3x_1 \leq 18 \\ x_1 \geq 0 \\ x_2 \geq 0 \end{cases}$$

Або у формах (9.9), (9.10):

$$\begin{cases} x_1 \geq 0 & (I) \\ x_2 \geq 0 & (II) \\ 19 - 2x_1 - 3x_2 \geq 0 & (III) \\ 13 - 2x_1 - x_2 \geq 0 & (IV) \\ 15 - 3x_2 \geq 0 & (V) \\ 18 - 3x_1 \geq 0 & (VI) \end{cases}$$

$$\begin{cases} x_1 \geq 0 \\ x_2 \geq 0 \\ 19 - 2x_1 - 3x_2 = x_3 \\ 13 - 2x_1 - x_2 = x_4 \\ 15 - 3x_2 = x_5 \\ 18 - 3x_1 = x_6 \end{cases}$$

При цьому матриця системи, складена з коефіцієнтів при невідомих буде мати вигляд:

$$I = \begin{pmatrix} 2 & 3 & 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 1 & 0 & 0 \\ 0 & 3 & 0 & 0 & 1 & 0 \\ 3 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Зрозуміло, що система не може бути вирішена однозначно, бо кількість невідомих більша за кількість рівнянь: ранг матриці $r = 4$, кількість невідомих $n = 6$. Кількість ступенів свободи $k = 6 - 4 = 2$

Цільова функція має вигляд:

$$F = -7x_1 - 5x_2 \rightarrow \min$$

Для вирішення задачі скористаємося її геометричною інтерпретацією. Для цього введемо систему координат x_1Ox_2 . Відомо, що геометричне місце точок на площині, координати яких задовольняють системі лінійних нерівностей, утворюють випуклий багатокутник, який назовемо багатокутником рішень системи нерівностей. Сторони цього багатокутника розташовуються на прямих, рівняння яких отримуються, якщо в нерівностях системи знаки нерівності замінити точними рівностями. Сам же цей багатокутник є перетинання полуплощин, на які поділяє площину кожна зі вказаних прямих. В нашому випадку такими прямими є:

$$\begin{cases} x_1 = 0 & (I) \\ x_2 = 0 & (II) \\ 19 - 2x_1 - 3x_2 = 0 & (III) \\ 13 - 2x_1 - x_2 = 0 & (IV) \\ 15 - 3x_2 = 0 & (V) \\ 18 - 3x_1 = 0 & (VI) \end{cases}$$

Прямі показані на Рис. 9.1, стрілки вказують які полуплосчини дають багатокутник рішень.

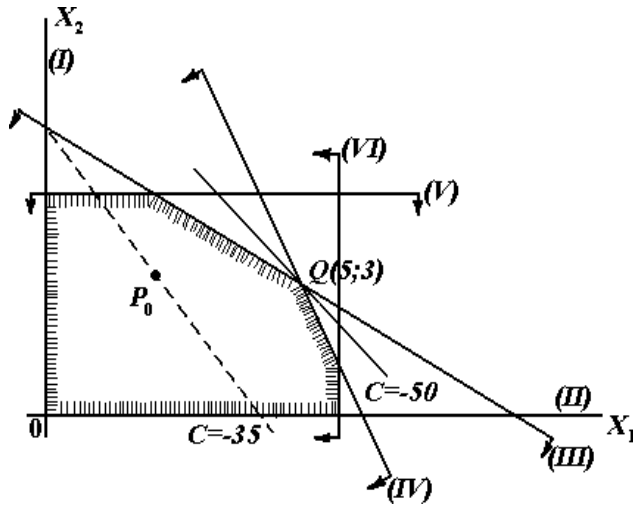


Рисунок 9.1 - Багатокутник рішень

На ряду з цим розглянемо згадану вище цільову функцію F та запишемо її у вигляді:

$$-7x_1 - 5x_2 = C$$

Змінюючи C отримаємо сукупність паралельних прямих. На Рис. 9.2 показані прямі для різних C , а вектор g вказує напрям руху до зменшення значень C . Тому очевидно, що рішення буде лежати на стороні або вершині багатокутника рішень найбільш віддаленої в напрямку вектору g .

Такою вершиною в даному випадку є $Q(5;3)$ (Див. Рис. 9.1) – що може бути визначено сумісним рішенням рівнянь (III) та (IV). При цьому значення цільової функції буде мінімальним, що задовольняє зазначеним вимогам: $F = -50$. Отже найбільш прибутковим рішенням є випуск 5 одиниць продукції P_1 та 3 одиниць P_2 .

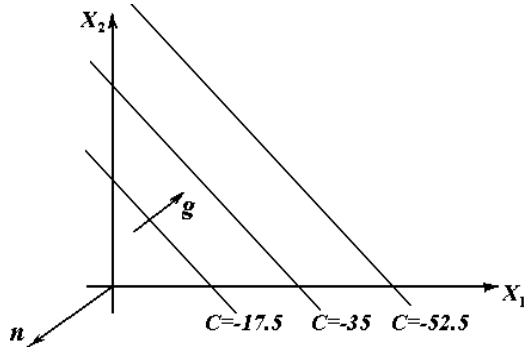


Рисунок 9.2 - Цільова функція

9.3 Задача про використання продуктивності обладнання

Підприємство за час T має виготовити N_1 одиниць продукції P_1 та N_2 одиниць продукції P_2 . Обидва види продукції можна виготовляти на обладнанні A та B , що мають різну продуктивність (кількість одиниць продукції, що виготовляється в одиницю часу) та вартість виготовлення за одиницю часу. Вихідні дані зведені в Табл. 9.3

Таблиця 9.3 Загальний вид вихідних даних до задачі

Обладн.	Продуктивність в одиницю часу		Витрати на виготовл. за одиницю часу		Час на виготовлення програми	
	P_1	P_2	P_1	P_2	P_1	P_2
А	a_1	a_2	α_1	α_2	x_1	x_2
В	b_1	b_2	β_1	β_2	x_3	x_4

За умовами задачі треба спланувати програму виготовлення продукції на відповідному обладнанні з метою мінімізації витрат на виготовлення.

Цільова функція буде мати вигляд:

$$F = \alpha_1 x_1 + \alpha_2 x_2 + \beta_1 x_3 + \beta_2 x_4 \rightarrow \min \quad (9.11)$$

Умови задачі запишуться:

$$\left\{ \begin{array}{l} x_1 \geq 0 \\ x_2 \geq 0 \\ x_3 \geq 0 \\ x_4 \geq 0 \\ x_1 + x_2 \leq T \\ x_3 + x_4 \leq T \\ a_1 x_1 + b_1 x_3 = N_1 \\ a_2 x_2 + b_2 x_4 = N_2 \end{array} \right. \text{ або } \left\{ \begin{array}{l} x_1 \geq 0 \\ x_2 \geq 0 \\ x_3 \geq 0 \\ x_4 \geq 0 \\ x_5 \geq 0 \\ x_6 \geq 0 \\ T - x_1 - x_2 = x_5 \\ T - x_3 - x_4 = x_6 \\ a_1 x_1 + b_1 x_3 = N_1 \\ a_2 x_2 + b_2 x_4 = N_2 \end{array} \right. \quad (9.12)$$

Тобто маємо при 4 рівняннях 6 змінних, що вимагає виразити 4 з них через 2 інші оскільки для графічного рішення треба мати 2 координати. Рішення розглянемо на кількісному прикладі. Вихідні данні наведені в Табл. 9.4. Загальний час на виготовлення всієї програми $T = 6$ год.

Таблиця 9.4 - Вихідні данні

Обладн.	Продуктивність в одиницю часу		Витрати на виготовл. за одиницю часу		Час на виготовлення програми		Програма випуску	
Прод.	P_1	P_2	P_1	P_2	P_1	P_2	P_1	P_2
А	6	24	4	47	x_1	x_2	30	96
В	13	13	13	26	x_3	x_4		

Приймаючи за замовчуванням що всі змінні не негативні, запишемо систему (9.12) та цільову функцію (9.11):

$$\begin{cases} x_1 + x_2 \leq 6 & (I') \\ x_3 + x_4 \leq 6 & (II') \\ 6x_1 + 13x_3 = 30 & (III') \\ 24x_2 + 13x_4 = 96 & (IV') \end{cases}$$

$$F = 4x_1 + 13x_3 + 47x_2 + 26x_4$$

Матриця системи обмежень буде мати вигляд:

$$I = \begin{pmatrix} 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 \\ 6 & 13 & 0 & 0 & 0 & 0 \\ 0 & 0 & 24 & 13 & 0 & 0 \end{pmatrix}$$

Виберемо в якості вільних змінних x_1 та x_2 та виразимо через них x_3 та x_4 з рівнянь (III') та (IV')

$$x_3 = \frac{1}{13}(30 - 6x_1) = \frac{6}{13}(5 - x_1) \geq 0 \Rightarrow x_1 \leq 5$$

$$x_4 = \frac{1}{13}(96 - 24x_2) = \frac{24}{13}(4 - x_2) \geq 0 \Rightarrow x_2 \leq 4$$

Підставивши отримані вирази в рівняння (II') отримаємо систему та цільову функцію:

$$\begin{cases} x_1 \geq 0 & (I) \\ x_2 \geq 0 & (II) \\ x_1 \leq 5 & (III) \\ x_2 \leq 4 & (IV) \\ x_1 + x_2 \leq 6 & (V) \\ x_1 + 4x_2 \geq 8 & (VI) \end{cases}$$

$$F = 222 - 2x_1 - x_2 \rightarrow \min$$

Геометрична інтерпретація рішення буде мати вигляд (Рис. 9.3):

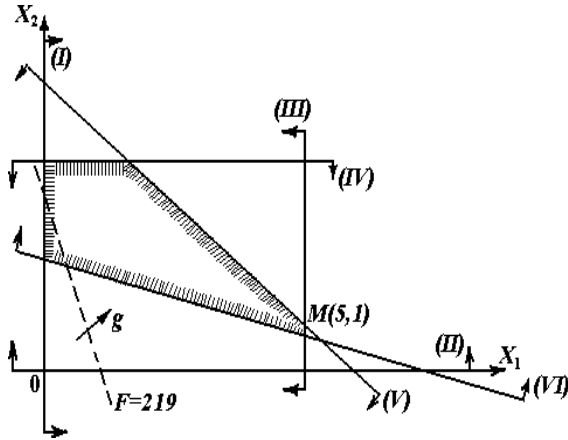


Рисунок 9.3 - Геометрична інтерпретація рішення

З рисунка видно, що рішенням буде точка $M(5;1)$ координати якої визначаються сумісним рішенням рівнянь (III) та (V):

$$x_1 = 5; x_2 = 1; F = 211$$

Тепер можна визначити: $x_3 = 0$; $x_4 = \frac{72}{13}$

Отже задача вирішена повністю. Таким чином на обладнанні А за означений час буде виготовлено $x_1 a_1 = 5 \cdot 6 = 30$ одиниць продукції $P1$ та $x_2 a_2 = 1 \cdot 24 = 24$ одиниць продукції $P2$. На обладнанні В буде виготовлено $x_3 b_1 = 0 \cdot 13 = 0$ одиниць продукції $P1$ та $x_4 b_2 = \frac{72}{13} \cdot 13 = 72$ одиниць продукції $P2$

9.4 Транспортна задача

У двох постачальників А1 та А2 є деяка однорідна продукція у кількості відповідно a_1 та a_2 . Цю продукцію треба доставити до трьох споживачів В1, В2 та В3 в кількості b_1 , b_2 , b_3 . Вартість перевезення з пункту Аі до пункту Вj позначимо відповідно C_{ij} , та кількість продукції до перевезення – x_{ij} . Треба скласти план перевезень з тим щоб, мінімізувати сумарну вартість перевезень. Часто задача доповнюється даними про вартість продукції у кожного з постачальників і тоді мінімізуються сумарні витрати на продукцію та транспортування. Таким чином математичне формулювання транспортної задачі буде мати вигляд:

$$\begin{cases} x_{11} + x_{21} = b_1 \\ x_{12} + x_{22} = b_2 \\ x_{13} + x_{23} = b_3 \\ x_{11} + x_{12} + x_{13} \leq a_1 \\ x_{21} + x_{22} + x_{23} \leq a_2 \end{cases} \quad (9.13)$$

$$F = \sum_{i=1}^2 \sum_{j=1}^3 x_{ij} C_{ij} \rightarrow \min \quad (9.14)$$

Спрощений варіант транспортної задачі полягає в рівності сумарних потреб і запасів:

$$\sum_{i=1}^2 a_i = \sum_{j=1}^3 b_j \quad (9.15)$$

В цьому випадку задача спрощується, два останніх рівняння в системі (9.12) перетворюються на рівності а сама задача повністю приймає форму основної. Безумовно кількість постачальників і споживачів може відрізнитися від наведених в більшу сторону – що безумовно ускладнить рішення задачі.

Питання для самоперевірки

1. Що таке задачі лінійного програмування?
2. Де знаходяться рішення задач лінійного програмування по відношенню до обмежень?
3. Яку форму має основна задача лінійного програмування?
4. Яке рішення називається допустимим?
5. Що таке план задачі лінійного програмування?
6. Яке рішення задачі лінійного програмування називається оптимальним?
7. Коли існує єдине рішення задачі лінійного програмування – воно ж оптимальне?
8. В якому випадку задача лінійного програмування не має рішення?
9. Яким чином кожна задача лінійного програмування може бути представлена як основна?
10. Що таке багатокутник рішень?
11. В чому полягає сутність транспортної задачі?
12. Що передбачає спрощений варіант транспортної задачі?

10. ОПТИМІЗАЦІЯ ЗА ЕКСПЕРЕМЕНТАЛЬНИМИ ДОСЛІДЖЕННЯМИ

10.1. Загальні положення. Повний факторний експеримент

Головна вимога до математичної моделі це передбачати поведінку об'єкту з певною точністю. А якщо цю вимогу конкретизувати для оптимізаційних моделей – то це передбачати напрям руху до тих умов де об'єкт буде функціонувати найкращим чином. У будь якому випадку ми приймаємо наступні допущення: *поверхня відгуку безперервна, вона гладка та має єдиний екстремум (по меншій мірі в межах області визначення).*

Зазвичай при пошуку оптимального рішення в факторному просторі вибирається якась точка C та розглядається декілька симетричних відносно неї точок, що разом формують деяку підобласть (див. Рис.10.1). Тут проводиться експеримент за результатами якого будується перша модель. Ця модель використовується для передбачення поведінки об'єкта в точках, що не входили до експерименту. Якщо точки знаходяться всередині підобласті, то таке передбачення носить назву *інтерполяція*, а якщо за її межами – *екстраполяція*. Чим далі від області експерименту лежить точка тим з меншою впевненістю це можна робити. Тому ми вимушені екстраполювати недалеко, а отримані результати використовувати для проведення подальших експериментів. Якщо змінювати тільки один з факторів - x_1 або x_2 то в якості оптимального рішення можуть бути прийняті точки A або B (Рис. 10.1) тоді як насправді воно лежить в точці M і для її пошуку потрібно змінювати одночасно обидва фактори – це і є один з сенсів багатофакторного експерименту.

називається «*майже стаціонарною*». В цій зоні використання лінійної функції у вигляді полінома першого ступеню стає не ефективним і не призводить до покращення результатів. В цьому випадку, потрапляння в майже стаціонарну область треба вважати завершенням процесу оптимізації або є сенс переходити до поліному другого ступеню, бо він є функцією, що має екстремум і може адекватно описати ділянку поверхні відгуку. Таким чином при плануванні багатofакторного експерименту обмежимося поліномами першого та другого ступенів.

План експерименту характеризується кількістю ступенів свободи:

$$f = N - l - 1 \quad (10.3)$$

де N – кількість незалежних дослідів в плані; l – кількість коефіцієнтів регресії в рівнянні (без вільного коефіцієнту).

В залежності від кількості ступенів свободи плани бувають: *ненасиченими*, коли $f > 0$, тобто кількість дослідів більша ніж потрібно для визначення коефіцієнтів; *насиченими*, коли $f = 0$ та *зверхнасиченими* при $f < 0$.

Перед безпосереднім плануванням експерименту повинно бути виконано ряд:

Аналіз апіорної інформації. Дослідження технічного об'єкту зазвичай починається в умовах, коли такий або аналогічний об'єкти деяким чином досліджувалися раніше. Інформацію яка міститься в результатах попередніх досліджень називають *апіорною*.

Оцінка границь областей визначення. Для цього повинні бути враховані обмеження що накладаються на фактори.

Вибір основного рівня. Найкращім умовам, що отримані з апіорної інформації відповідає комбінація або декілька комбінацій рівнів факторів. Кожна така комбінація є багатомірною точкою в багатомірному факторному просторі. Її можна розглядати як початкову точку для побудови плану експерименту. Визначимо її як основний або нульовий рівень. Побудова плану експерименту зводиться до вибору експериментальних точок симетричних відносно основного

(нульового) рівня які повинні знаходитись в межах області визначення. Позначимо дійсне значення основного рівня j -го фактору як $\tilde{x}_{0j}$.

Вибір інтервалів варіювання. Наступне завдання – вибрати експериментальні точки симетричні відносно основного рівня. Положення цих точок визначається інтервалом варіювання - I_j . Якщо інтервал варіювання додати до нульового рівня то отримаємо верхній рівень (ВР), а якщо його відняти від нульового – нижній (НР), тобто відповідно:

$$\tilde{x}_{Uj} = \tilde{x}_{0j} + I_j \quad (10.4)$$

$$\tilde{x}_{Lj} = \tilde{x}_{0j} - I_j$$

Графічно це можна зобразити як показано на Рис.10.2

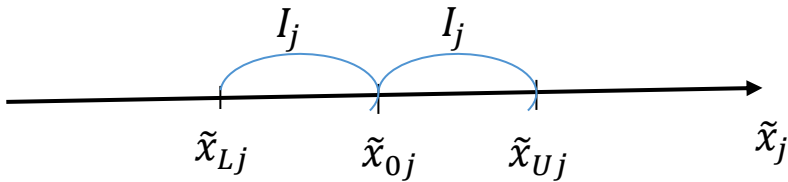


Рисунок 10.2. — Рівні факторів та інтервал варіювання

На вибір інтервалів варіювання накладають обмеження як знизу так і зверху. З одного боку, інтервал варіювання не може бути менше за ту похибку з якою експериментатор фіксує значення рівня, інакше рівні не можна буде відрізнити. З іншого боку вони не повинні бути настільки великими, щоб значення рівнів опинилися за межами області визначення. В цих межах безумовно зберігається значна невизначеність, вибір інтервалу варіювання є складною задачею бо вона зв'язана з неформалізованим етапом планування експерименту тому вона повинна вирішуватися виключно за рахунок досвіду та інтуїції дослідника. Вибрана величина в ході дослідів часто має

корегуватися. Орієнтовно, якщо інтервал варіювання складає не більше 10% від області визначення – то він вважається *вузьким*; від 10% до 30% – *середнім*; більше 30% – *широким*.

При плануванні експерименту використовують не дійсні значення факторів, а кодовані, які визначаються наступним чином:

$$x_j = \frac{\tilde{x}_j - \tilde{x}_{0j}}{I_j} \quad (10.5)$$

З рівнянь (10.4) та (10.5) видно, що кодоване значення основного рівня дорівнює 0 – тому його і називають ще нульовим рівнем. Кодоване значення нижнього рівня – -1 (може позначатися просто « - »), а верхнього – +1 (може позначатися « + »).

При повному факторному експерименті фактори варіюються на двох рівнях – верхньому та нижньому. При цьому повинні бути реалізовані *всі можливі комбінації значень факторів*. Кількість таких комбінацій, а отже і кількість незалежних дослідів визначається:

$$N = 2^k \quad (10.6)$$

тут 2 – це кількість рівнів, k – кількість факторів.

На формулі (10.6) базується і умовне позначення ПФЕ – 2^k . Так для двох факторів ПФЕ позначається 2^2 , а кількість незалежних дослідів які мають бути проведені дорівнює чотирьом. За результатами ПФЕ будується модель (рівняння регресії) у вигляді *полінома першого ступеню*. Для двох факторів ця модель має вигляд (10.1):

План 2^2 для побудови рівняння регресії (10.1) є ненасиченим, бо має $f = 1$ ступінь свободи. Умови проведення експерименту записуються в таблицю, що називається матрицею планування експерименту (Див. Табл. 10.1). В цю матрицю додатково вводиться стовпчик, що відповідає фіктивній змінній x_0 при вільному коефіцієнті b_0 , що у всіх експериментах має значення +1 – він буде використовуватися виключно для розрахунків, а не як умова проведення експерименту.

Таблиця 10.1 — Матриця планування експерименту 2^2

N	x_0	Умови експерименту		Результати експерименту
		x_1	x_2	
1	+ 1	+ 1	+ 1	y_1
2	+ 1	- 1	+ 1	y_2
3	+ 1	+ 1	- 1	y_3
4	+ 1	- 1	- 1	y_4

Кожен стовбець матриці планування експерименту називається *вектор-стовбець*, а рядок – *вектор-рядок*.

Повний факторний експеримент в поєднанні з рівнянням регресії у вигляді алгебраїчного полінома дозволяє досить легко визначати коефіцієнти регресії. **Треба пам'ятати, що фактори в такому рівнянні мають кодовані значення, а не дійсні.** У наведеній нижче формулі (10.7) j - це індекс коефіцієнта регресії:

$$b_j = \frac{1}{N} \sum_{i=1}^N (x_{ji} \cdot y_i) \quad (10.7)$$

Виходячи з рівняння (10.6), коефіцієнти регресії з рівняння (10.1) на основі матриці Табл. 10.1 будуть визначатися як сума y зі знаками відповідного x :

$$b_0 = \frac{1}{4}(y_1 + y_2 + y_3 + y_4)$$

$$b_1 = \frac{1}{4}(y_1 - y_2 + y_3 - y_4)$$

$$b_2 = \frac{1}{4}(y_1 + y_2 - y_3 - y_4)$$

Наявність однієї ступені свободи дозволяє до рівняння (10.1) ввести ще один елемент – і цей елемент буде нелінійним. Рівняння (10.1) прийме вигляд:

$$y = b_0 + b_1 \cdot x_1 + b_2 \cdot x_2 + b_{12} \cdot x_1 \cdot x_2 \quad (10.8)$$

Тут елемент $b_{12} \cdot x_1 \cdot x_2$ характеризує взаємодію факторів і називається ефект взаємодії. Він відображає зміну ефекту одного з факторів при зміні рівня іншого фактору. Для розрахунку коефіцієнту b_{12} до матриці планування додатково вводиться відповідний стовбець (Табл. 10.2)

Таблиця 10.2 — Матриця планування експерименту 2^2 з ефектом взаємодії

N	x_0	Умови експерименту		$x_1 x_2$	Результати експерименту
		x_1	x_2		y
1	+ 1	+ 1	+ 1	+ 1	y_1
2	+ 1	- 1	+ 1	- 1	y_2
3	+ 1	+ 1	- 1	- 1	y_3
4	+ 1	- 1	- 1	+ 1	y_4

Коефіцієнт при ефекті взаємодії відповідно буде визначатися:

$$b_{12} = \frac{1}{4}(y_1 - y_2 - y_3 + y_4)$$

При наявності двох факторів перебрати усі можливі комбінації значень рівнів не важко, але при їх збільшенні потрібні деякі правила, щоб формалізувати цю процедуру. Розглянемо два:

- 1) Чергування знаків: перший стовпчик – знаки через один; другий через два; третій через чотири і т.д. І взагалі, знаки змінюються через 2^{j-1} де j – номер стовпчика.
- 2) Домовимось, що коли $k > 2$, для плану 2^k план 2^{k-1} є вихідним. При появі нового фактору, вихідний план повторюється двічі: в поєднанні з нижнім та верхнім рівнем вихідного плану.

Зі збільшенням кількості факторів, кількість можливих взаємодій стрімко зростає. При наявності трьох факторів елемент $x_1 x_2 x_3$ має назву *ефект взаємодії другого порядку*. Матриця

планування експерименту з урахуванням ефектів взаємодії звісно буде розширюватися. Принципово важливим моментом при плануванні експерименту та аналізі отриманих результатів є те, що *як правило чим більше порядок взаємодії, тим менше імовірність його виникнення*.

Отримана модель має перевіряється на адекватність (в цьому курсі не розглядається).

10.2. Ортогональні центральні композиційні плани

Потрапивши в майже стаціонарну зону в ході вирішення оптимізаційної задачі, поліном першого ступеню вже не дає поліпшення результату. Тепер треба або зупинитися і вважати отриманні результати найліпшими або спробувати описати відгук більш складною моделлю – поліномом другого ступеню, який на відміну від поліному першого ступеню має екстремум. Цей підхід має сенс, якщо на основі апріорної інформації, включаючи теоретичні уявлення про процес, є підстави вважати наявність такого екстремуму у функції відгуку.

Спрощено цю ситуацію можна проілюструвати, як показано на Рис. 10.3. На рисунку зображено перетин поверхні відгуку в напрямку градієнту. Дійсно, градієнт у вигляді поліному першого ступеню вірно вказує напрям до оптимального рішення, зображеного точкою *M*. Але рухаючись в цьому напрямку ми потрапляємо в т. *P* та очікуємо відповідний результат. Тоді як насправді отримуємо результат що відповідає точці *Q* і який є гіршим (або не кращим) за результат в початковій точці *C*. Тобто, для пошуку оптимального рішення в межах нижнього та верхнього рівнів факторів (зображених відповідно точками *L* та *U*) поліному першого ступеню вже не досить і треба мати більш точний опис поверхні відгуку з відображенням наявного екстремуму. І цей опис може дати вже поліном другого ступеню, який з урахуванням можливих взаємодій факторів для двох факторів має вигляд (10.9):

$$y = \beta_0 + b_1x_1 + b_2x_2 + b_{11}x_1^2 + b_{22}x_2^2 + b_{12}x_1x_2 \quad (10.9)$$



Рисунок 10.3 – Майже стаціонарна область

А в загальному вигляді для k факторів:

$$y = \beta_0 + \sum_{j=1}^k b_j x_j + \sum_{j=1}^k b_{jj} x_j^2 + \sum_{j < u} b_{ju} x_j x_u \quad (10.10)$$

Для отримання моделі (10.10) змінні повинні варіюватися не менше ніж на 3 рівнях, а кількість серій незалежних дослідів без урахування повторів N повинно відповідати вимозі:

$$N \geq \frac{(k+2)(k+1)}{2} \quad (10.11)$$

Ортогональність будь якого плану це властивість що забезпечує незалежність оцінок коефіцієнтів регресії та в даному випадку

досягається використанням *ортогональних центральних композиційних планів* (ОЦКП) другого порядку – які з великої кількості можливих є одними з простіших. В план експерименту входять досліді на основному рівні та матриця ПФЕ (або його дробова репліка при $k > 5$) що складають *ядро плану*, та декілька додаткових точок що звуться *зіркові*. Отже для k факторів план буде складатися з наступних елементів:

1. Ядро 2^k або 2^{k-p}
2. Центр з координатами $(0,0,0, \dots, 0)$
3. Зіркові точки з координатами $(+d, 0, 0, \dots, 0)$; $(-d, 0, 0, \dots, 0)$; $(0, +d, 0, \dots, 0)$; $(0, -d, 0, \dots, 0)$ і т.д. Тут d – так зване плече (кодоване значення), причому $d \geq 1$.

Загальна кількість незалежних дослідів без урахування повторів при цьому буде визначатися:

$$N = 2^k + 1 + 2k \quad (10.12)$$

або

$$N = 2^{k-p} + 1 + 2k \quad (10.13)$$

Наприклад, для двох факторів маємо план показаний в табл. 10.3, де $N = 2^2 + 1 + 2 \cdot 2 = 9$.

Таблиця 10.3 — Ортогональний центральний композиційний план
 $2^2 + 1 + 2 \cdot 2$

N	Елементи плану	x_1	x_2	y
1	Ядро	+ 1	+ 1	y_1
2		- 1	+ 1	y_2
3		+ 1	- 1	y_3
4		- 1	- 1	y_4
5	Центр	0	0	y_5
6	Зіркові точки	+ d	0	y_6
7		- d	0	y_7
8		0	+ d	y_8
9		0	- d	y_9

Для забезпечення ортогональності плану (сума добутків відповідних елементів будь-яких двох стовбців повинна дорівнювати нулю) в рівняння (10.10) вводяться наступні заміни $b_0 = (\beta_0 + \lambda \sum_{j=1}^k b_{jj})$ та $x'_j = x_j^2 - \lambda$, і воно приймає вигляд:

$$y = b_0 + \sum_{j=1}^k b_j x_j + \sum_{j=1}^k b_{jj} x'_j + \sum_{j < u}^k b_{ju} x_j x_u \quad (10.14)$$

Матриця планування експерименту з таблиці 10.3 з урахуванням вектор-стовбців для розрахунку прийме вигляд показаний в таблиці 10.4.

Невідомі λ та d визначаються з умов забезпечення ортогональності стовбців:

$$d = \sqrt{\frac{1}{2} [\sqrt{N \cdot 2^{k-p}} - 2^{k-p}]} \quad (10.15)$$

$$\lambda = \frac{2^{k-p} + 2d^2}{N}$$

Деякі значення λ та d в залежності від кількості факторів та ядра експерименту наведені в Табл. 10.5

Таблиця 10.4 — Ортогональний центральний композиційний план $2^2 + 1 + 2 \cdot 2$ в загальному вигляді з вектор-стовбцями для розрахунків

N	x_0	x_1	x_2	x'_1 $= x_1^2$ $-\lambda$	x'_2 $= x_2^2$ $-\lambda$	$x_1 x_2$	y
1	+ 1	+ 1	+ 1	$1-\lambda$	$1-\lambda$	+ 1	y_1
2	+ 1	- 1	+ 1	$1-\lambda$	$1-\lambda$	- 1	y_2
3	+ 1	+ 1	- 1	$1-\lambda$	$1-\lambda$	- 1	y_3
4	+ 1	- 1	- 1	$1-\lambda$	$1-\lambda$	+ 1	y_4
5	+ 1	0	0	$-\lambda$	$-\lambda$	0	y_5
6	+ 1	+ d	0	$d^2-\lambda$	$-\lambda$	0	y_6
7	+ 1	- d	0	$d^2-\lambda$	$-\lambda$	0	y_7
8	+ 1	0	+ d	$-\lambda$	$d^2-\lambda$	0	y_8
9	+ 1	0	- d	$-\lambda$	$d^2-\lambda$	0	y_9

Коефіцієнти регресії перетвореного рівняння (10.14) визначаються наступним чином [36]:

$$b_j = \frac{1}{C_j} \sum_{i=1}^N (x_{ji} y_i) \quad (10.16)$$

Де C_j - сума квадратів елементів кожного вектор-стовбця:

$$C_j = \sum_{i=1}^N x_{ji}^2 \quad (10.17)$$

Виходячи з Табл. 10.5 план наведений в Табл.10.4 у загальному вигляді представлено в Табл.10.6 в нижньому рядку якої наведена сума квадратів елементів кожного вектор-стовбця – C_j .

Таблиця 10.5 — Характеристики деяких ОЦКП

Кількість факторів k	p	Ядро плану	Кількість дослідів N	Плече експ. d	λ	$1-\lambda$	$d^2 - \lambda$
2	0	2^2	9	1	0,667	0,333	0,333
3	0	2^3	15	1,215	0,730	0,270	0,746
4	0	2^4	25	1,414	0,800	0,200	1,199
5	0	2^5	43	1,596	0,863	0,137	1,684
5	1	2^{5-1}	27	1,547	0,770	0,230	1,623
6	0	2^6	77	1,761	0,912	0,088	2,189
6	1	2^{6-1}	45	1,724	0,843	0,157	2,129
6	2	2^{6-2}	29	1,664	0,743	0,257	2,026
7	0	2^7	143	1,909	0,946	0,054	2,698
7	1	2^{7-1}	79	1,885	0,900	0,100	2,653
7	2	2^{7-2}	47	1,841	0,825	0,175	2,564
7	3	2^{7-3}	31	1,771	0,718	0,282	2,418

Таблиця 10.6 — Ортогональний центральний композиційний план $2^2 + 1 + 2 \cdot 2$

N	Елементи плану	x_0	x_1	x_2	x'_1 $= x_1^2$ $-\lambda$	x'_2 $= x_2^2$ $-\lambda$	x_1x_2	y
1	Ядро	+ 1	+ 1	+ 1	0,333	0,333	+ 1	y_1
2		+ 1	- 1	+ 1	0,333	0,333	- 1	y_2
3		+ 1	+ 1	- 1	0,333	0,333	- 1	y_3
4		+ 1	- 1	- 1	0,333	0,333	+ 1	y_4
5	Центр	+ 1	0	0	-0,667	-0,667	0	y_5
6	Зоряні точки	+ 1	+ 1	0	0,333	-0,667	0	y_6
7		+ 1	- 1	0	0,333	-0,667	0	y_7
8		+ 1	0	+ 1	-0,667	0,333	0	y_8
9		+ 1	0	- 1	-0,667	0,333	0	y_9
$C_j = \sum_{i=1}^N x_{ji}^2$		9	6	6	2	2	4	

За відомих коефіцієнтів рівняння (10.14), можна знайти вільний член вихідного рівняння (10.10):

$$\beta_0 = b_0 - \lambda \sum_{j=1}^k b_{jj} \quad (10.18)$$

10.3. Планування експерименту та оптимізація за концепцією Тагучі

Концепція була висунута д-ром Генічі Тагучі в той час коли він працював в Лабораторії Електричного Зв'язку (ECL), заснованій в Японії після другої світової війни. Займаючись оптимізацією планування експериментів він запропонував плани при яких **кількість дослідів значно скорочується** по відношенню до традиційних планів особливо при великій кількості факторів. Порівняння кількості необхідних дослідів при проведенні повного факторного експерименту і за концепцією Тагучі наведено в Табл. 10.7.

Таблиця 10.7 - Порівняння кількості необхідних дослідів для повного факторного експерименту та за планами Тагучі

Факторів	Рівнів	Кількість дослідів	
		Повний факторний експеримент	План Тагучі
2	2	4	4
3	2	8	4
4	2	16	8
7	2	128	8
15	2	32768	16
4	3	81	9

Початковим призначенням планів було забезпечення показників якості продукції та керування ними. Методика Тагучі почала широко застосовуватися в Сполучених Штатах на початку 80-х років XX сторіччя та принесла суттєві зміни в інженерію якості та

забезпечила вражаючі результати, що дало декому привід вважати запропонований підхід найбільш значущим вкладом в статистичні методи управління якістю за декілька останніх десятирічь.

Принципово підхід до планування експерименту (Design of experiment - DOE) практично не відрізняється від багатофакторних експериментів, розглянутих вище – послідовність дослідів є набір різних комбінацій рівнів факторів. Планування здійснюється на основі планів у вигляді стандартизованих ортогональних масивів (Orthogonal Arrays - OA). Результатом застосування методики є пошук оптимального співвідношення факторів, що забезпечує **мінімальне або максимальне значення відгуку або результуючого показника**. Але є деякі відмінності:

- Кількість рівнів факторів може бути будь-яким;
- Кількість рівнів для різних факторів може бути різним;

• Фактори поділяються на ті якими можна керувати або *зовнішній масив* і ті якими керувати не можна – *внутрішній масив*. **Фактори, внутрішнього масиву мають назву шумових**. Наявність шумів призводить до відхилення роботи системи від ідеальної. В свою чергу фактори внутрішнього масиву можна поділити на ті що піддаються контролю та вимірюванню і випадкові варіації. Існує три категорії факторів шуму:

- а) Зовнішня середа (вологість, температура та ін.);
- б) Старіння та зміна властивостей об'єкта в часі;
- в) Виробничі варіації.

• Деякі фактори з внутрішнього масиву - ті, значення яких піддаються контролю та вимірюванню можуть вводитися до плану експерименту. Кількість рівнів для таких факторів як правило дорівнює 2 (наприклад вище чи нижче деякого порогового значення);

• В результаті експерименту модель об'єкту або процесу **не будується**, а тільки визначається оптимальне поєднання значень факторів. Оптимальним поєднанням вважається таке, що забезпечує мінімальне або максимальне значення результуючого показника. Виходячи з цього, збільшення кількості рівнів підвищує точність

визначення оптимального рішення, але збільшує необхідну кількість дослідів.

•Оптимальне поєднання значень факторів, визначене в ході експерименту та шляхом нескладного розрахунку, **як правило не співпадає з тими, що використовувалися для проведення дослідів.** Це цілком зрозуміло, бо плани Тагучі є лише маленьким фрагментом повного факторного експерименту. Такий результат є передбаченим, а не таким що був отриманий в ході експерименту. Якщо так, обов'язковим елементом концепції Тагучі є **необхідність проведення перевірного досліді** на визначеній оптимальній комбінації рівнів факторів.

Різниця у використанні підходу Тагучі до управління якістю та до оптимізації будь-яких фізичних показників буде полягати тільки в тому, **що використовується в якості результуючого показника.** Якщо мова йде про оптимізацію як таку, то результуючим показником буде будь-який фізичний показник процесу або виробу, наприклад величина зносу ріжучого інструменту або його стійкість, сила різання та ін. Такий відгук може як мінімізуватись так і максимізуватись. При оцінці якості результуючими показниками є співвідношення сигнал-шум – SN, що характеризують розсіювання (дисперсію) чи розподілення результуючого показника відносно номіналу (тут не розглядаються).

Ми також не будемо тут розглядати як саме генеруються ортогональні масиви Тагучі. З цією метою можна користуватися спеціальним програмним забезпеченням, таблицями та ін. Позначають масиви наступним чином: $OA_N(s^m)$ де N – кількість дослідів в масиві, s – кількість рівнів, m – кількість факторів. Якщо для різних факторів встановлена різна кількість рівнів, то позначення буде мати вигляд: $OA_N(s_1^{m_1} \times s_2^{m_2} \dots)$. Рівні частіше позначаються послідовністю 1, 2, 3, ..., s або іноді 0, 1, 2, 3, ..., $s-1$.

Деякі з ОА з однаковою кількістю рівнів для всіх факторів наведені в Табл.10.8. Приклади ОА з різною кількістю рівнів наведені в Табл. 10.9, 10.10. Треба зазначити, що будь який ОА який влаштовує

за кількістю рівнів можна використовувати для меншої кількості факторів, відкинувши зайві стовбці.

Таблиця 10.8 – Приклади ортогональних масивів з однаковою кількістю рівнів

<table><tr><th colspan="3">$OA_4(2^2)$</th></tr><tr><th>N</th><th>X_1</th><th>X_2</th></tr><tr><td>1</td><td>1</td><td>1</td></tr><tr><td>2</td><td>1</td><td>2</td></tr><tr><td>3</td><td>2</td><td>1</td></tr><tr><td>4</td><td>2</td><td>2</td></tr></table>	$OA_4(2^2)$			N	X_1	X_2	1	1	1	2	1	2	3	2	1	4	2	2	<table><tr><th colspan="4">$OA_4(2^3)$</th></tr><tr><th>N</th><th>X_1</th><th>X_2</th><th>X_3</th></tr><tr><td>1</td><td>1</td><td>1</td><td>1</td></tr><tr><td>2</td><td>1</td><td>2</td><td>2</td></tr><tr><td>3</td><td>2</td><td>1</td><td>2</td></tr><tr><td>4</td><td>2</td><td>2</td><td>1</td></tr></table>	$OA_4(2^3)$				N	X_1	X_2	X_3	1	1	1	1	2	1	2	2	3	2	1	2	4	2	2	1	<table><tr><th colspan="5">$OA_8(2^4)$</th></tr><tr><th>N</th><th>X_1</th><th>X_2</th><th>X_3</th><th>X_4</th></tr><tr><td>1</td><td>1</td><td>1</td><td>1</td><td>1</td></tr><tr><td>2</td><td>1</td><td>1</td><td>1</td><td>2</td></tr><tr><td>3</td><td>1</td><td>2</td><td>2</td><td>1</td></tr><tr><td>4</td><td>1</td><td>2</td><td>2</td><td>2</td></tr><tr><td>5</td><td>2</td><td>1</td><td>2</td><td>1</td></tr><tr><td>6</td><td>2</td><td>1</td><td>2</td><td>2</td></tr><tr><td>7</td><td>2</td><td>2</td><td>1</td><td>1</td></tr><tr><td>8</td><td>2</td><td>2</td><td>1</td><td>2</td></tr></table>	$OA_8(2^4)$					N	X_1	X_2	X_3	X_4	1	1	1	1	1	2	1	1	1	2	3	1	2	2	1	4	1	2	2	2	5	2	1	2	1	6	2	1	2	2	7	2	2	1	1	8	2	2	1	2																																																																																																																												
	$OA_4(2^2)$																																																																																																																																																																																																																									
	N	X_1	X_2																																																																																																																																																																																																																							
	1	1	1																																																																																																																																																																																																																							
	2	1	2																																																																																																																																																																																																																							
3	2	1																																																																																																																																																																																																																								
4	2	2																																																																																																																																																																																																																								
$OA_4(2^3)$																																																																																																																																																																																																																										
N	X_1	X_2	X_3																																																																																																																																																																																																																							
1	1	1	1																																																																																																																																																																																																																							
2	1	2	2																																																																																																																																																																																																																							
3	2	1	2																																																																																																																																																																																																																							
4	2	2	1																																																																																																																																																																																																																							
$OA_8(2^4)$																																																																																																																																																																																																																										
N	X_1	X_2	X_3	X_4																																																																																																																																																																																																																						
1	1	1	1	1																																																																																																																																																																																																																						
2	1	1	1	2																																																																																																																																																																																																																						
3	1	2	2	1																																																																																																																																																																																																																						
4	1	2	2	2																																																																																																																																																																																																																						
5	2	1	2	1																																																																																																																																																																																																																						
6	2	1	2	2																																																																																																																																																																																																																						
7	2	2	1	1																																																																																																																																																																																																																						
8	2	2	1	2																																																																																																																																																																																																																						
<table><tr><th colspan="3">$OA_9(3^2)$</th></tr><tr><th>N</th><th>X_1</th><th>X_2</th></tr><tr><td>1</td><td>1</td><td>1</td></tr><tr><td>2</td><td>1</td><td>2</td></tr><tr><td>3</td><td>1</td><td>3</td></tr><tr><td>4</td><td>2</td><td>1</td></tr><tr><td>5</td><td>2</td><td>2</td></tr><tr><td>6</td><td>2</td><td>3</td></tr><tr><td>7</td><td>3</td><td>1</td></tr><tr><td>8</td><td>3</td><td>2</td></tr><tr><td>9</td><td>3</td><td>3</td></tr></table>	$OA_9(3^2)$			N	X_1	X_2	1	1	1	2	1	2	3	1	3	4	2	1	5	2	2	6	2	3	7	3	1	8	3	2	9	3	3	<table><tr><th colspan="4">$OA_9(3^3)$</th></tr><tr><th>N</th><th>X_1</th><th>X_2</th><th>X_3</th></tr><tr><td>1</td><td>1</td><td>1</td><td>1</td></tr><tr><td>2</td><td>1</td><td>2</td><td>2</td></tr><tr><td>3</td><td>1</td><td>3</td><td>3</td></tr><tr><td>4</td><td>2</td><td>1</td><td>2</td></tr><tr><td>5</td><td>2</td><td>2</td><td>3</td></tr><tr><td>6</td><td>2</td><td>3</td><td>1</td></tr><tr><td>7</td><td>3</td><td>1</td><td>3</td></tr><tr><td>8</td><td>3</td><td>2</td><td>1</td></tr><tr><td>9</td><td>3</td><td>3</td><td>2</td></tr></table>	$OA_9(3^3)$				N	X_1	X_2	X_3	1	1	1	1	2	1	2	2	3	1	3	3	4	2	1	2	5	2	2	3	6	2	3	1	7	3	1	3	8	3	2	1	9	3	3	2	<table><tr><th colspan="5">$OA_9(2^4)$</th></tr><tr><th>N</th><th>X_1</th><th>X_2</th><th>X_3</th><th>X_4</th></tr><tr><td>1</td><td>1</td><td>1</td><td>1</td><td>1</td></tr><tr><td>2</td><td>1</td><td>2</td><td>2</td><td>2</td></tr><tr><td>3</td><td>1</td><td>3</td><td>3</td><td>3</td></tr><tr><td>4</td><td>2</td><td>1</td><td>2</td><td>3</td></tr><tr><td>5</td><td>2</td><td>2</td><td>3</td><td>1</td></tr><tr><td>6</td><td>2</td><td>3</td><td>1</td><td>2</td></tr><tr><td>7</td><td>3</td><td>1</td><td>3</td><td>2</td></tr><tr><td>8</td><td>3</td><td>2</td><td>1</td><td>3</td></tr><tr><td>9</td><td>3</td><td>3</td><td>2</td><td>1</td></tr></table>	$OA_9(2^4)$					N	X_1	X_2	X_3	X_4	1	1	1	1	1	2	1	2	2	2	3	1	3	3	3	4	2	1	2	3	5	2	2	3	1	6	2	3	1	2	7	3	1	3	2	8	3	2	1	3	9	3	3	2	1																																																																																				
	$OA_9(3^2)$																																																																																																																																																																																																																									
	N	X_1	X_2																																																																																																																																																																																																																							
	1	1	1																																																																																																																																																																																																																							
	2	1	2																																																																																																																																																																																																																							
3	1	3																																																																																																																																																																																																																								
4	2	1																																																																																																																																																																																																																								
5	2	2																																																																																																																																																																																																																								
6	2	3																																																																																																																																																																																																																								
7	3	1																																																																																																																																																																																																																								
8	3	2																																																																																																																																																																																																																								
9	3	3																																																																																																																																																																																																																								
$OA_9(3^3)$																																																																																																																																																																																																																										
N	X_1	X_2	X_3																																																																																																																																																																																																																							
1	1	1	1																																																																																																																																																																																																																							
2	1	2	2																																																																																																																																																																																																																							
3	1	3	3																																																																																																																																																																																																																							
4	2	1	2																																																																																																																																																																																																																							
5	2	2	3																																																																																																																																																																																																																							
6	2	3	1																																																																																																																																																																																																																							
7	3	1	3																																																																																																																																																																																																																							
8	3	2	1																																																																																																																																																																																																																							
9	3	3	2																																																																																																																																																																																																																							
$OA_9(2^4)$																																																																																																																																																																																																																										
N	X_1	X_2	X_3	X_4																																																																																																																																																																																																																						
1	1	1	1	1																																																																																																																																																																																																																						
2	1	2	2	2																																																																																																																																																																																																																						
3	1	3	3	3																																																																																																																																																																																																																						
4	2	1	2	3																																																																																																																																																																																																																						
5	2	2	3	1																																																																																																																																																																																																																						
6	2	3	1	2																																																																																																																																																																																																																						
7	3	1	3	2																																																																																																																																																																																																																						
8	3	2	1	3																																																																																																																																																																																																																						
9	3	3	2	1																																																																																																																																																																																																																						
<table><tr><th colspan="3">$OA_{16}(4^2)$</th></tr><tr><th>N</th><th>X_1</th><th>X_2</th></tr><tr><td>1</td><td>1</td><td>1</td></tr><tr><td>2</td><td>1</td><td>2</td></tr><tr><td>3</td><td>1</td><td>3</td></tr><tr><td>4</td><td>1</td><td>4</td></tr><tr><td>5</td><td>2</td><td>1</td></tr><tr><td>6</td><td>2</td><td>2</td></tr><tr><td>7</td><td>2</td><td>3</td></tr><tr><td>8</td><td>2</td><td>4</td></tr><tr><td>9</td><td>3</td><td>1</td></tr><tr><td>10</td><td>3</td><td>2</td></tr><tr><td>11</td><td>3</td><td>3</td></tr><tr><td>12</td><td>3</td><td>4</td></tr><tr><td>13</td><td>4</td><td>1</td></tr><tr><td>14</td><td>4</td><td>2</td></tr><tr><td>15</td><td>4</td><td>3</td></tr><tr><td>16</td><td>4</td><td>4</td></tr></table>	$OA_{16}(4^2)$			N	X_1	X_2	1	1	1	2	1	2	3	1	3	4	1	4	5	2	1	6	2	2	7	2	3	8	2	4	9	3	1	10	3	2	11	3	3	12	3	4	13	4	1	14	4	2	15	4	3	16	4	4	<table><tr><th colspan="4">$OA_{16}(4^3)$</th></tr><tr><th>N</th><th>X_1</th><th>X_2</th><th>X_3</th></tr><tr><td>1</td><td>1</td><td>1</td><td>1</td></tr><tr><td>2</td><td>1</td><td>2</td><td>2</td></tr><tr><td>3</td><td>1</td><td>3</td><td>3</td></tr><tr><td>4</td><td>1</td><td>4</td><td>4</td></tr><tr><td>5</td><td>2</td><td>1</td><td>2</td></tr><tr><td>6</td><td>2</td><td>2</td><td>1</td></tr><tr><td>7</td><td>2</td><td>3</td><td>4</td></tr><tr><td>8</td><td>2</td><td>4</td><td>3</td></tr><tr><td>9</td><td>3</td><td>1</td><td>3</td></tr><tr><td>10</td><td>3</td><td>2</td><td>4</td></tr><tr><td>11</td><td>3</td><td>3</td><td>1</td></tr><tr><td>12</td><td>3</td><td>4</td><td>2</td></tr><tr><td>13</td><td>4</td><td>1</td><td>4</td></tr><tr><td>14</td><td>4</td><td>2</td><td>3</td></tr><tr><td>15</td><td>4</td><td>3</td><td>2</td></tr><tr><td>16</td><td>4</td><td>4</td><td>1</td></tr></table>	$OA_{16}(4^3)$				N	X_1	X_2	X_3	1	1	1	1	2	1	2	2	3	1	3	3	4	1	4	4	5	2	1	2	6	2	2	1	7	2	3	4	8	2	4	3	9	3	1	3	10	3	2	4	11	3	3	1	12	3	4	2	13	4	1	4	14	4	2	3	15	4	3	2	16	4	4	1	<table><tr><th colspan="5">$OA_{16}(4^4)$</th></tr><tr><th>N</th><th>X_1</th><th>X_2</th><th>X_3</th><th>X_4</th></tr><tr><td>1</td><td>1</td><td>1</td><td>1</td><td>1</td></tr><tr><td>2</td><td>1</td><td>2</td><td>2</td><td>2</td></tr><tr><td>3</td><td>1</td><td>3</td><td>3</td><td>3</td></tr><tr><td>4</td><td>1</td><td>4</td><td>4</td><td>4</td></tr><tr><td>5</td><td>2</td><td>1</td><td>2</td><td>3</td></tr><tr><td>6</td><td>2</td><td>2</td><td>1</td><td>4</td></tr><tr><td>7</td><td>2</td><td>3</td><td>4</td><td>1</td></tr><tr><td>8</td><td>2</td><td>4</td><td>3</td><td>2</td></tr><tr><td>9</td><td>3</td><td>1</td><td>3</td><td>4</td></tr><tr><td>10</td><td>3</td><td>2</td><td>4</td><td>3</td></tr><tr><td>11</td><td>3</td><td>3</td><td>1</td><td>2</td></tr><tr><td>12</td><td>3</td><td>4</td><td>2</td><td>1</td></tr><tr><td>13</td><td>4</td><td>1</td><td>4</td><td>2</td></tr><tr><td>14</td><td>4</td><td>2</td><td>3</td><td>1</td></tr><tr><td>15</td><td>4</td><td>3</td><td>2</td><td>4</td></tr><tr><td>16</td><td>4</td><td>4</td><td>1</td><td>3</td></tr></table>	$OA_{16}(4^4)$					N	X_1	X_2	X_3	X_4	1	1	1	1	1	2	1	2	2	2	3	1	3	3	3	4	1	4	4	4	5	2	1	2	3	6	2	2	1	4	7	2	3	4	1	8	2	4	3	2	9	3	1	3	4	10	3	2	4	3	11	3	3	1	2	12	3	4	2	1	13	4	1	4	2	14	4	2	3	1	15	4	3	2	4	16	4	4	1	3
	$OA_{16}(4^2)$																																																																																																																																																																																																																									
	N	X_1	X_2																																																																																																																																																																																																																							
	1	1	1																																																																																																																																																																																																																							
	2	1	2																																																																																																																																																																																																																							
3	1	3																																																																																																																																																																																																																								
4	1	4																																																																																																																																																																																																																								
5	2	1																																																																																																																																																																																																																								
6	2	2																																																																																																																																																																																																																								
7	2	3																																																																																																																																																																																																																								
8	2	4																																																																																																																																																																																																																								
9	3	1																																																																																																																																																																																																																								
10	3	2																																																																																																																																																																																																																								
11	3	3																																																																																																																																																																																																																								
12	3	4																																																																																																																																																																																																																								
13	4	1																																																																																																																																																																																																																								
14	4	2																																																																																																																																																																																																																								
15	4	3																																																																																																																																																																																																																								
16	4	4																																																																																																																																																																																																																								
$OA_{16}(4^3)$																																																																																																																																																																																																																										
N	X_1	X_2	X_3																																																																																																																																																																																																																							
1	1	1	1																																																																																																																																																																																																																							
2	1	2	2																																																																																																																																																																																																																							
3	1	3	3																																																																																																																																																																																																																							
4	1	4	4																																																																																																																																																																																																																							
5	2	1	2																																																																																																																																																																																																																							
6	2	2	1																																																																																																																																																																																																																							
7	2	3	4																																																																																																																																																																																																																							
8	2	4	3																																																																																																																																																																																																																							
9	3	1	3																																																																																																																																																																																																																							
10	3	2	4																																																																																																																																																																																																																							
11	3	3	1																																																																																																																																																																																																																							
12	3	4	2																																																																																																																																																																																																																							
13	4	1	4																																																																																																																																																																																																																							
14	4	2	3																																																																																																																																																																																																																							
15	4	3	2																																																																																																																																																																																																																							
16	4	4	1																																																																																																																																																																																																																							
$OA_{16}(4^4)$																																																																																																																																																																																																																										
N	X_1	X_2	X_3	X_4																																																																																																																																																																																																																						
1	1	1	1	1																																																																																																																																																																																																																						
2	1	2	2	2																																																																																																																																																																																																																						
3	1	3	3	3																																																																																																																																																																																																																						
4	1	4	4	4																																																																																																																																																																																																																						
5	2	1	2	3																																																																																																																																																																																																																						
6	2	2	1	4																																																																																																																																																																																																																						
7	2	3	4	1																																																																																																																																																																																																																						
8	2	4	3	2																																																																																																																																																																																																																						
9	3	1	3	4																																																																																																																																																																																																																						
10	3	2	4	3																																																																																																																																																																																																																						
11	3	3	1	2																																																																																																																																																																																																																						
12	3	4	2	1																																																																																																																																																																																																																						
13	4	1	4	2																																																																																																																																																																																																																						
14	4	2	3	1																																																																																																																																																																																																																						
15	4	3	2	4																																																																																																																																																																																																																						
16	4	4	1	3																																																																																																																																																																																																																						

Таблица 10.9 – Ортогональный массив $OA_{18}(6^1 \times 3^6)$

N	X_1	X_2	X_3	X_4	X_5	X_6	X_7
1	1	1	1	1	1	1	1
2	1	2	2	2	2	2	2
3	1	3	3	3	3	3	3
4	2	1	1	2	2	3	3
5	2	2	2	3	3	1	1
6	2	3	3	1	1	2	2
7	3	1	2	1	3	2	3
8	3	2	3	2	1	3	1
9	3	3	1	3	2	1	2
10	4	1	3	3	2	2	1
11	4	2	1	1	3	3	2
12	4	3	2	2	1	1	3
13	5	1	2	3	1	3	2
14	5	2	3	1	2	1	3
15	5	3	1	2	3	2	1
16	6	1	3	2	3	1	2
17	6	2	1	3	1	2	3
18	6	3	2	1	2	3	1

Таблиця 10.10 – Ортогональний масив $OA_{18}(2^1 \times 3^7)$

N	X ₁	X ₂	X ₃	X ₄	X ₅	X ₆	X ₇	X ₈
1	1	1	1	1	1	1	1	1
2	1	1	2	2	2	2	2	2
3	1	1	3	3	3	3	3	3
4	1	2	1	1	2	2	3	3
5	1	2	2	2	3	3	1	1
6	1	2	3	3	1	1	2	2
7	1	3	1	2	1	3	2	3
8	1	3	2	3	2	1	3	1
9	1	3	3	1	3	2	1	2
10	2	1	1	3	3	2	2	1
11	2	1	2	1	1	3	3	2
12	2	1	3	2	2	1	1	3
13	2	2	1	2	3	1	3	2
14	2	2	2	3	1	2	1	3
15	2	2	3	1	2	3	2	1
16	2	3	1	3	2	3	1	2
17	2	3	2	1	3	1	2	3
18	2	3	3	2	1	2	3	1

Метою аналізу результатів експерименту за Тагучі є перш за все пошук відповіді на три ключових питання:

1. Які умови є оптимальними?
2. Які фактори впливають на варіабельність результатів і як сильно?
3. Який буде очікуваний результат при оптимальних умовах і наскільки сильно кожен фактор впливає на поліпшення результатів

Для відповіді на ці питання на основі отриманих за результатами експерименту результатів та самого ортогонального масиву розраховується середній ефект кожного фактору для кожного з

рівнів. Розрахунок середнього ефекту фактору X_i на рівні l - $\bar{\bar{X}}_i^l$ виконується наступним чином:

$$\bar{\bar{X}}_i^l = \frac{1}{r\{l \mid X_i = l\}} \sum_{\substack{j=1 \\ \{j \mid X_{ij}=l\}}}^N Y_j \quad (10.19)$$

Тут $r\{l \mid X_i = l\}$ – кількість рівнів l у відповідному стовбці X_i
 $\sum_{j=1}^N Y_j$ - сума значень результируючих показників Y для яких $\{j \mid X_{ij}=l\}$

рівень у відповідному рядку j стовбця X_i дорівнює l

Оптимальним рішенням буде комбінація рівнів факторів, що забезпечують мінімальні або максимальні ефекти (в залежності від потреби). Розглянемо приклад розрахунку середніх ефектів.

Приклад

Розглянемо план експерименту $OA_4(2^2)$ та його результати (Табл. 10.11). Результируючий показник Y отримано за результатами трьох паралельних дослідів ($n = 3$) та його потрібно мінімізувати.

Таблиця 10.11 - План Тагучі $OA_4(2^3)$ та його результати

N	Фактори			Результат
	X_1	X_2	X_3	Y_j
1	1	1	1	30
2	1	2	2	25
3	2	1	2	34
4	2	2	1	27

Розрахуємо ефекти факторів для кожного з рівнів за формулою (10.18):

$$\bar{\bar{X}}_1^1 = \frac{1}{2}(Y_1 + Y_2) = \frac{1}{2}(30 + 25) = 27,5$$

$$\bar{\bar{X}}_1^2 = \frac{1}{2}(Y_3 + Y_4) = \frac{1}{2}(34 + 27) = 30,5$$

$$\bar{\bar{X}}_2^1 = \frac{1}{2}(Y_1 + Y_3) = \frac{1}{2}(30 + 34) = 32,0$$

$$\bar{\bar{X}}_2^2 = \frac{1}{2}(Y_2 + Y_4) = \frac{1}{2}(25 + 27) = 26,0$$

$$\bar{\bar{X}}_3^1 = \frac{1}{2}(Y_1 + Y_4) = \frac{1}{2}(30 + 27) = 28,5$$

$$\bar{\bar{X}}_3^2 = \frac{1}{2}(Y_2 + Y_3) = \frac{1}{2}(25 + 34) = 29,5$$

На основі виконаного розрахунку будують діаграми ефектів (Рис. 10.4). Виходячи з умовою мінімізації вибираємо значення рівнів факторів, які забезпечують мінімальний середній ефект – Табл. 10.12

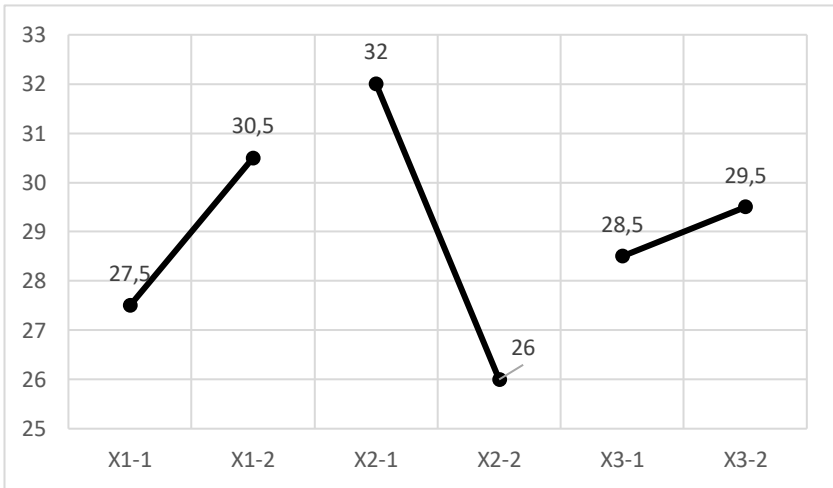


Рисунок 10.4 – Діаграма ефектів

Таблиця 10.12 – Рівні факторів з мінімальним середнім ефектом

Фактор	X ₁	X ₂	X ₃
Рівень	1	2	1

Як видно з Табл. 10.12 рівні факторів не співпадають ні з однією комбінацією умов дослідів, визначених ОА в Табл. 10.11. В цьому випадку за концепцією Тагучі обов'язково необхідно провести перевірочний дослід за умов зазначених в Табл. 10.12 для підтвердження або спростування отриманого результату.

Додатково за методом Тагучі, використовуючи отримані результати, для оцінки ступеню впливу факторів на відгук проводиться специфічний дисперсійний аналіз – ANOVA (Analysis of variance) – тут не розглядається. Окрім того планування за концепцією Тагучі дозволяє визначати взаємодію факторів, а також розмір допуску на фактор що досліджується, виходячи з відомих витрат на виправлення браку (тут не розглядається).

Питання для самоперевірки

1. Яка послідовність пошуку оптимального рішення при експериментальних дослідженнях?
2. Що таке інтерполяція?
3. Що таке екстраполяція?
4. Які види моделі приймаються при плануванні експерименту?
5. Який вид моделі приймається при повному факторному експерименту?
6. Яке призначення моделі при оптимізації за повним факторним експериментом?
7. Що таке «рівень»?
8. Що таке інтервал варіювання?
9. Як отримують «верхній рівень»?
10. Як отримують «нижній рівень»?
11. Що таке «основний» або «нульовий» рівень?
12. Як будується план (матриця) повного факторного експерименту?
13. Як визначаються значення коефіцієнтів регресії при повному факторному експерименті?
14. Що таке апіорна інформація?
15. Яким виразом записується взаємодія факторів?

16. Як пов'язаний порядок взаємодії факторів з імовірністю її виникнення?
17. Як визначається кількість дослідів при повному факторному експерименті?
18. Які дії дослідника передбачаються при потраплянні в майже стаціонарну область?
19. Що приймається в якості моделі при плануванні експерименту за ОЦКП?
20. Як визначається кількість дослідів при плануванні експерименту за ОЦКП?
21. Яка структура ОЦКП?
22. В якому випадку буває необхідним планування експерименту за ОЦКП?
23. Що забезпечує ортогональності плану експерименту?
24. В чому перевага планів Тагучі у порівнянні з повним факторним експериментом?
25. Які особливості планів Тагучі?
26. Що таке фактори внутрішнього масиву?
27. Що таке фактори зовнішнього масиву?
28. Для яких цілей використовуються плани Тагучі?
29. Що є планом експерименту за Тагучі?
30. Як позначаються плани за Тагучі?
31. Що є метою аналізу результатів експерименту за Тагучі?
32. Як визначається середній ефект фактору за Тагучі?
33. Чи будується модель при плануванні експерименту за Тагучі і якщо так, то в якому вигляді?
34. Яке рішення приймається за результатами розрахунку середніх ефектів факторів?
35. Що потрібно робити якщо комбінація рівнів потрібних ефектів не співпадає ні з одним з запланованих дослідів?

ЛІТЕРАТУРА

1. Методи одновимірної оптимізації: практикум з дисципліни «Дослідження операцій» [Електронний ресурс]: навч. посіб. для студ. спеціальності 113 «Прикладна математика», спеціалізації «Наука про дані та математичне моделювання» / Т. С. Ладогубець, О. Д. Фіногенов; КПІ ім. Ігоря Сікорського. – Електронні текстові дані (1 файл: 453 Кбайт). – Київ: КПІ ім. Ігоря Сікорського, 2020. – 47 с.
2. Пляскин, И. И. Оптимизация технических решений в машиностроении // И.И. Пляскин. – М.: Машиностроение, 1982. – 176 с.
3. Roy, Ranjit K. A primer on Taguchi method [Текст] // Ranjit K Roj. – Dearborn, MI: SME, 2010 – 336 p.
4. Kaker, Raghu N. Taguchi's orthogonal arrays are classical designs of experiments [Текст] // Raghu N Kaker, Eric S. Lagergren, James L. Filliben / Journal of Research of the National Institute of Standards and Technology. – 1991. – Vol.96, № 5. – P.577 – 591.