Лекція 11. Поняття про багатокритеріальну оптимізацію

11.1. Поняття про задачу багатокритеріальної (векторної) оптимізації.

Широкий спектр задач у промисловості, техніці, економіці та багатьох інших галузях містить необхідність одночасної оптимізації декількох цілей. У багатьох випадках цілі визначаються в непорівняннних одиницях і існує певна ступінь конфлікту між ними, наприклад, одна ціль не може бути покращена без погіршення іншої. Такі задачі називають задачами багатокритеріальної оптимізації (Multiobjective Optimization Problems). Вперше задача багатокритеріальної векторної оптимізації виникла у італійського економіста Вільфредо Парето при математичному дослідженні товарного обміну. В умовах однокритеріальних задач оптимізації зазвичай отримують один оптимальний розв‘язок. Проте, в багатокритеріальній оптимізації не існує прямого метода визначення чи розв‘язок краще, ніж інші. Найчастіше в багатокритеріальній оптимізації для порівняння розв‘язків застосовують метод, який називають відношенням домінантності Парето і який, замість єдиного оптимального розв‘язку, призводить до набору альтернатив з різними компромісами серед цілей. Ці розв‘язки називають Парето оптимальними або недомінованими розв‘язками.

Нехай G - множина можливих рішень (розв‘язків) на якій задано векторний критерій

$$F = \{f_1, ..., f_k\},$$

де $f_1, ..., f_k$ - числові функції (оцінки), які визначені на множині G.

Задача вибору рішень, яка включає множину G та векторний критерій F називається *задачею багатокритеріальної оптимізації*.

Позначимо множину рішень, що обираються через $C(G)$. Ця множина є рішенням задачі вибору і може містити будь-яку підмножину множини можливих рішень.

Особа (група осіб), що здійснює вибір і несе відповідальність за наслідки такого вибору називається *особою, що приймає рішення* (ОПР).

Постановка задачі багатокритеріального вибору включає:

1) множину можливих рішень G;

2) векторний критерій F;

3) відношення переваги $\succ_G$.

Задача полягає у знаходження множини $C(G) \subset G$ з урахуванням відношення переваги $\succ_G$ на основі заданого векторного критерію F.

Нехай потрібно обрати одне з двох можливих рішень X^1 або X^2. Для цих рішень має місце один і тільки один з наступних 3 випадків:

1) $X^1 \succ_G X^2$ - ОПР віддає перевагу першому рішенню (X^1);

2) $X^2 \succ_G X^1$ - ОПР віддає перевагу другому рішенню (X^2);

3) не виконується ані $X^1 \succ_G X^2$, ані $X^2 \succ_G X^1$ - ОПР не може надати переваги жодному рішенню.

У випадку $X^1 \succ_G X^2$ говорять, що рішення X^1 *домінує* рішення X^2, а рішення X^2 *доміноване* рішенням X^1.

В термінах множини рішень, що обираються можна записати так:

$$X^1 \succ_G X^2 \Leftrightarrow C\left\{X^1, X^2\right\} = \left\{X^1\right\}.$$

Аксіома 11.1. (*Аксіома виключення рішень, що домінуються*).
Для будь-якої пари допустимих рішень X^1, $X^2 \in G$, для яких має місце відношення $X^1 \succ_G X^2$, виконується $X^2 \notin C(G)$.

Якщо задано декілька критеріїв оптимальності, то для кожного з них необхідно визначити напрямок зацікавленості ОПР (приміром, прибуток потрібно максимізувати, а ризики – мінімізувати). Тому, надалі будемо вважати побажання ОПР узгодженими, тобто вважаємо, що ОПР зацікавлений в отриманні максимальних значень всіх компонентів векторного критерію F.

Аксіома 11.2. (*Аксіома Парето*).
Для всіх пар допустимих рішень X^1, $X^2 \in G$, для яких мають місце нерівності

$$\begin{aligned} &\forall i \in \{1,\dots,k\}: f_i\left(X^1\right) \geq f_i\left(X^2\right) \\ &\exists j \in \{1,\dots,k\}: f_j\left(X^1\right) > f_j\left(X^2\right), \end{aligned} \qquad (11.1)$$

виконується відношення $X^1 \succ_G X^2$.

Домінування рішення X^1 над X^2 в сенсі умов (11.1) називається *домінуванням за Парето* і позначається $X^1 \succ_P X^2$.

Означення 11.1. Рішення $X^* \in G$ називається *оптимальним за Парето*, (Парето-оптимальним), якщо не існує такого можливого рішення $X \in G$, що $X \succ_P X^*$, тобто не існує X, для якого виконувались нерівності:

$$\begin{aligned} &\forall i \in \{1,\dots,k\}: f_i\left(X\right) \geq f_i\left(X^*\right) \\ &\exists j \in \{1,\dots,k\}: f_j\left(X\right) > f_j\left(X^*\right). \end{aligned}$$

Всі Парето-оптимальні рішення утворюють *множину Парето*, що позначається P^*.

Тобто, Парето-оптимальне рішення, яке не може бути покращене (збільшене) за жодним критерієм без погіршення (зменшення) за будь-яким хоча б одним іншим критерієм. Рішення, що входять до множини Парето також називають Парето-ефективними.

Фронтом Парето $\mathcal{F}P^*$ для множини Парето P^* називається

$$\mathcal{F}P^* = \left\{ F\left(X\right) = \left(f_1\left(X\right),\ldots,f_k\left(X\right)\right) \middle| X \in P^* \right\}.$$

Принцип Еджворта-Парето.

Якщо ОПР веде себе розумно, тобто виконуються умови аксіом 1, 2, то рішення, що їм обираються, повинні бути Парето оптимальними, тобто $C\left(G\right) \subset P^*$.

В багатьох випадках пошук Парето-оптимальних рішень є складною задачею. Тому розглядають поняття слабкого Парето-оптимального рішення, або рішення, оптимального за Слейтером.

Означення 11.2. Рішення $X^* \in G$ називається *оптимальним за Слейтером (слабко-ефективним)*, якщо не існує такого можливого рішення $X \in G$, для якого виконуються нерівності:

$$\forall i \in \left\{1,\ldots,k\right\} : f_i\left(X\right) > f_i\left(X^*\right).$$

Множина рішень, оптимальних за Слейтером позначається S^*. Домінування рішення X^* над X за Слейтером позначається $X^* \succ_S X$.

Рішення, оптимальні за Слейтером, менш цікаві ніж оптимальні за Парето, але в багатьох випадках при розв‘язанні задач багатокритеріальної оптимізації отримують саме такі рішення.

Геометрична інтерпретація принципу Еджворта-Парето (рис.11.1).

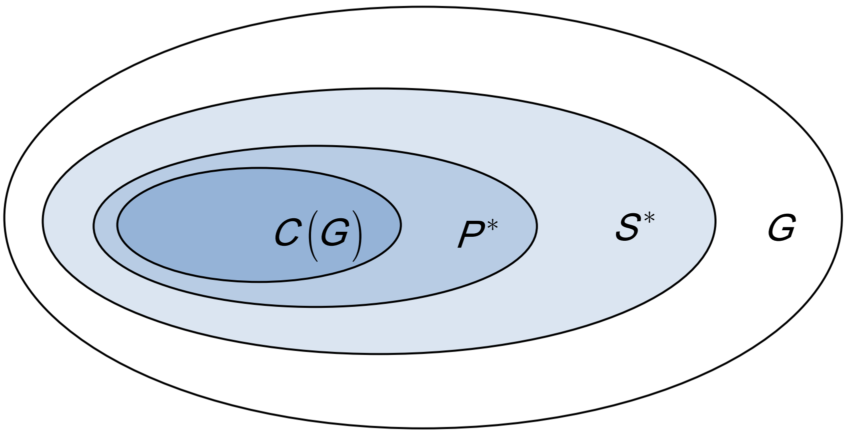

Рис.11.1

11.2. Багатокритеріальна задача математичного програмування.

Багатокритеріальна задача математичного програмування полягає в знаходженні максимуму (мінімуму) векторної функції

$$F(X)=\left(f_1(X),\ldots,f_k(X)\right)^T \to \max(\min). \tag{11.2}$$

за умов

$$X \in G, \tag{11.3}$$

де вектор $X \in \mathbb{R}^n$ складається з n змінних розв‘язків, значення яких повинні бути обрані в задачі оптимізації. Допустима область $G \subset \mathbb{R}^n$ визначається рівняннями і нерівностями (для задач лінійного програмування вигляду $AX \leq B$ і $AX = B$), вектор-функція $F:\mathbb{R}^n \to \mathbb{R}^k$ складається з k скалярних функцій $f_i:\mathbb{R}^n \to \mathbb{R}, i=\overline{1,k}, k \geq 2$.

Вибір альтернативи, яка буде розв‘язком задачі багатокритеріальної оптимізації потрібно робити з множини ефективних (Парето оптимальних) або слабко-ефективних (оптимальних за Слейтером) альтернатив.

Геометрична інтерпретація множини ефективних розв‘язків та їх зображення в просторі об‘єктивних функцій-критеріїв - фронт Парето для $n=2, k=2$ показана на рис. 11.2 . Виходячи з означення Парето оптимальних розв‘язків зрозуміло, що фронт Парето в задачі максимізації є північно-східною межею множини розв‘язків в просторі критеріїв.

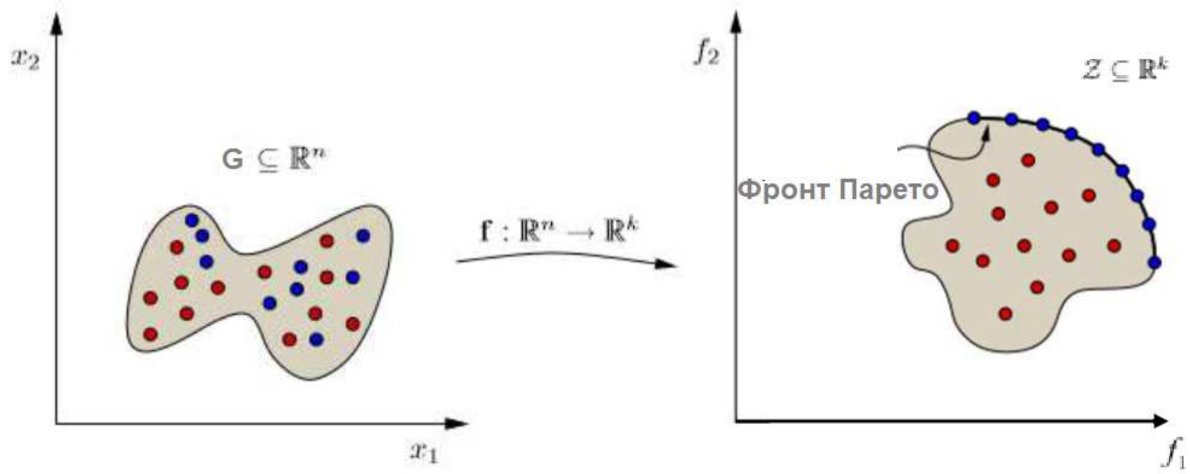

Рис.11.2

Оскільки, зазвичай, оптимальні розв‘язки багатокритеріальної задачі оптимізації за кожним критерієм окремо не співпадають (при співпадінні розв‘язок вважається тривіальним), то розв‘язком може бути тільки деяке компромісне рішення, що задовольняє в тому чи іншому сенсі усі компоненти векторного критерію.

Методи розв‘язку багатокритеріальних задач побудовані таким чином, щоб в результаті одержати одну з точок оптимальних за Парето. Оскільки Парето-оптимальні розв‘язки не порівняні між собою за перевагою, яка задається критеріями задачі, а вибір потрібно здійснити, то для цього потрібна додаткова інформація не на просторі альтернатив, а на просторі критеріїв. Джерелом такої інформації може бути як ОПР, так і специфіка предметної області, у якій розв‘язується задача.

В процесі розв‘язання багатокритеріальних задач доводиться розв‘язувати специфічні проблеми, які пов‘язані з вибором оптимального розв‘язку.

1. Нормалізація критеріїв.

Часто локальні критерії мають різний фізичний зміст і вимірюються в різних одиницях. Операція зведення масштабів локальних критеріїв до єдиного, зазвичай безрозмірного, називається нормалізацією.

Нехай $X_i^* = \arg\max_{X \in G} f_i(X)$ - точка утопії (ідеальна точка) для i-го критерію задачі максимізації, $i = \overline{1,k}$, $X_i^0 = \arg\min_{X \in G} f_i(X)$ - точка антиутопії (надір). Тоді нормалізований критерій $\bar{f}_i(X) = \dfrac{f_i(X) - f_i(X_i^0)}{f_i(X_i^*) - f_i(X_i^0)}$ безрозмірний і належить відрізку $[0;1]$, причому $\bar{f}_i(X_i^*) = 1$, а $\bar{f}_i(X_i^0) = 0$.

2. Вибір принципу оптимальності.

3. Урахування пріоритетів критеріїв.

4. Обчислення оптимуму багатокритеріальної задачі.

Усі наведені проблеми, використовуючи різні техніки, зводять багатокритеріальну задачу до однокритеріальної. Розглянемо деякі методи знаходження компромісного розв‘язку.

11.3. Метод ідеальної точки.

Цей метод не використовує допоміжної інформації від ОПР про переваги на множині критеріїв. Правило знаходження компромісу полягає у знаходженні розв‘язку (альтернативи) X^*, значення критерію $F(X)$ для якого є найближчим до ідеальної точки в деякій метриці.

Зазвичай використовують L_p - метрику:

$$\|Y\|_p = \left(\sum_{i=1}^{k} (y_k)^p \right)^{\frac{1}{p}}$$

для $1 \le p \le \infty$. Для $p = 1$ отримаємо Манхетенську метрику, при $p = 2$ - звичайна Евклідова метрика, при $p = \infty$ - метрика Чебишова.

Компромісний розв‘язок знаходять як розв‘язок скалярізованої задачі

$$X^* = \arg\min_{X \in G} \left(\sum_{i=1}^{k} \left(f_i(X) - f_i(X^*) \right)^p \right)^{\frac{1}{p}}.$$

Нехай $p = 2$, тоді розв‘язують задачу

$$\min_{X \in G} \sum_{i=1}^{k} \left(f_i\left(X\right) - f_i\left(X_i^{*}\right) \right)^2 . \qquad (11.4)$$

Якщо $p = 1$, то скаляризована задача набуває вигляду

$$\min_{X \in G} \sum_{i=1}^{k} \left| f_i\left(X\right) - f_i\left(X_i^{*}\right) \right| \qquad (11.5)$$

і розв‘язок X^*:

$$X^* = \arg\min_{X \in G} \sum_{i=1}^{k} \left| f_i\left(X\right) - f_i\left(X_i^{*}\right) \right| = \arg\max_{X \in G} \sum_{i=1}^{k} f_i\left(X\right),$$

якщо $p = \infty$, то

$$\min_{X \in G} \max_{1 \le i \le k} \left| f_i\left(X\right) - f_i\left(X_i^{*}\right) \right| = -\max_{X \in G} \min_{1 \le i \le k} \left(f_i\left(X\right) - f_i\left(X_i^{*}\right) \right). \qquad (11.6)$$

Задача (11.5) обирається, коли ОПР оцінює «відстань» до ідеалу, як сумарну нев‘язку за усіма критеріями. Задача (11.6) оцінює «відстань» до ідеалу, як максимальну нев‘язку за усіма критеріями (тобто за найгіршим за значенням показником).

Якщо критерії векторної задачі мають різні шкали (одиниці виміру, масштаб), то розглядають задачі для нормалізованих критеріїв.

$$\max_{X \in G} \sum_{i=1}^{k} \overline{f}_i\left(X\right) = \max_{X \in G} \sum_{i=1}^{k} \frac{f_i\left(X\right) - f_i\left(X_i^{0}\right)}{f_i\left(X_i^{*}\right) - f_i\left(X_i^{0}\right)} = \max_{X \in G} \sum_{i=1}^{k} \frac{f_i\left(X\right)}{f_i\left(X_i^{*}\right) - f_i\left(X_i^{0}\right)},$$

$$\max_{X \in G} \min_{1 \le i \le k} \left(\overline{f}_i\left(X\right) - 1 \right) = \max_{X \in G} \min_{1 \le i \le k} \left(\frac{f_i\left(X\right) - f_i\left(X_i^{0}\right)}{f_i\left(X_i^{*}\right) - f_i\left(X_i^{0}\right)} - 1 \right) =$$

$$= \max_{X \in G} \min_{1 \le i \le k} \frac{f_i\left(X\right) - f_i\left(X_i^{*}\right)}{f_i\left(X_i^{*}\right) - f_i\left(X_i^{0}\right)}.$$

Приклад 11.1. Методом ідеальної точки при $p = 1, 2, \infty$ знайти розв‘язок задачі

$$f_1\left(X\right) = 3x_1 + x_2 \to \max$$

$$f_2\left(X\right) = x_1 + 2x_2 \to \max$$

за умов

$$\begin{cases} x_1 + x_2 \le 5, \\ 0 \le x_1 \le 4, \\ 0 \le x_2 \le 4. \end{cases}$$

Знайдемо розв‘язки за кожним критерієм окремо (рис. 11.3).

$$f_1(X)=3x_1+x_2\to \max \qquad\qquad f_2(X)=x_1+2x_2\to \max$$

$$\begin{cases} x_1+x_2\le 5,\\ 0\le x_1\le 4,\\ 0\le x_2\le 4. \end{cases} \qquad\qquad \begin{cases} x_1+x_2\le 5,\\ 0\le x_1\le 4,\\ 0\le x_2\le 4. \end{cases}$$

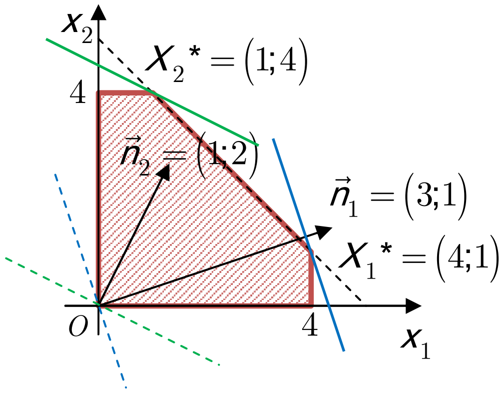

Рис. 11.3

Одержимо значення функцій в ідеальній і найгіршій точках.

$$f_1(X_1^*)=13,\ X_1^*(4,1);f_1(X_1^0)=0,\ X_1^0(0,0)$$

$$f_2(X_2^*)=9,\ X_2^*(1,4);f_2(X_2^0)=0,\ X_2^0(0,0)$$

1) $p=1$ (рис.11.4).

$$\max_{X\in G} f(X)=\max_{X\in G}\sum_{i=1}^{k}\frac{f_i(X)}{f_i(X_i^*)-f_i(X_i^0)}$$

Отже,

$$f(X)=\frac{3x_1+x_2}{13}+\frac{x_1+2x_2}{9}=\frac{40x_1+35x_2}{9\cdot 13}\Rightarrow$$

$$f(X)=8x_1+7x_2\to \max$$

$$\begin{cases} x_1+x_2\le 5,\\ 0\le x_1\le 4,\\ 0\le x_2\le 4. \end{cases}$$

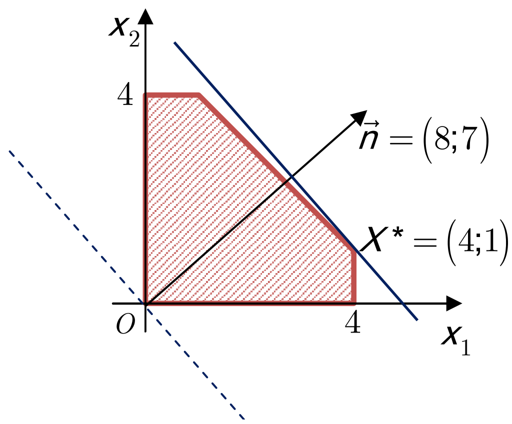

Рис. 11.4

Розв‘язавши задачу графічно, одержимо

$$X^*(4,1),\ f_1(X^*)=13,\ f_2(X^*)=6.$$

2) $p=2$ (рис. 11.5).

$$\min_{X\in G} f(X)=\min_{X\in G}\sum_{i=1}^{k}\left(f_i(X)-f_i(X_i{}^*)\right)^2$$

$$f(X)=(3x_1+x_2-13)^2+(x_1+2x_2-9)^2\to \min$$

Лінії рівня функції $f(X)$ – еліпси з центром в точці $O\left(\frac{17}{5},\frac{14}{5}\right)$ - точка безумовного мінімуму функції, знаходимо з системи

$$\begin{cases}3x_1+x_2=13\\ x_1+2x_2=9\end{cases}.$$

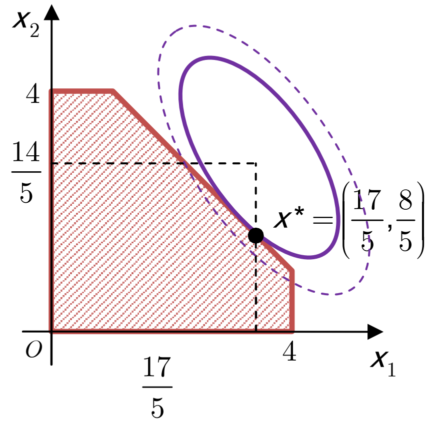

Рис. 11.5

Одержуємо задачу знаходження умовного екстремуму функції

$$\min f(X)=(3x_1+x_2-13)^2+(x_1+2x_2-9)^2$$

за умов $x_1+x_2=5$.

Склавши функцію Лагранжа

$$L(x_1,x_2,\lambda)=(3x_1+x_2-13)^2+(x_1+2x_2-9)^2+\lambda(x_1+x_2-5)$$

знаходимо оптимальну точку $X^*\left(\frac{17}{5},\frac{8}{5}\right)$ і значення функцій в цій точці

$$f_1(X^*)=\frac{59}{5}=11{,}8;\ f_2(X^*)=\frac{33}{5}=6{,}6.$$

3) $p=\infty$ (рис. 11.6)

$$\min_{X\in G}\max_{1\le i\le k}\left|\overline{f}_i(X)-1\right|=\min_{X\in G}\max_{1\le i\le k}\left|1-\frac{f_i(X)-f_i(X_i^0)}{f_i(X_i^*)-f_i(X_i^0)}\right|=$$

$$=-\max_{X\in G}\min_{1\le i\le k}\frac{f_i(X)-f_i(X_i^*)}{f_i(X_i^*)-f_i(X_i^0)}$$

Розв'язуємо задачу:

$$\min\left\{\frac{3x_1+x_2-13}{13},\frac{x_1+2x_2-9}{9}\right\}\to\max$$

Лінії рівня цільової функції скаляризованої задачі мають вигляд кутів, вершини яких знаходяться на прямій:

$14x_1=17x_2$.

Ця пряма задається умовою рівності функцій в $\{\ldots\}$, а бокові сторони кутів паралельні лініям рівня відповідних критеріїв початкової задачі.

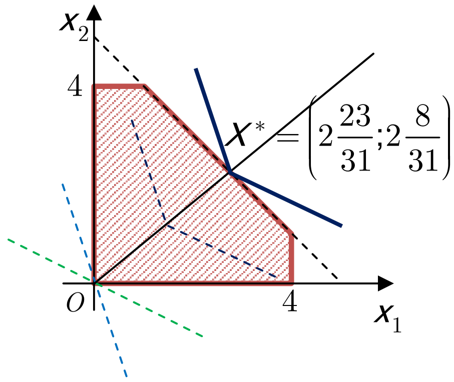

Рис. 11.6

Максимум досягається в точці $X^*\left(2\frac{23}{31};2\frac{8}{31}\right)$, яка є розв'язком системи

$$\begin{cases} 14x_1 - 17x_2 = 0, \\ x_1 + x_2 = 5. \end{cases}$$

$$f_1\left(X^*\right) = \frac{325}{31} = 10,484;\ f_2\left(X^*\right) = \frac{225}{31} = 7,2581.$$

11.4. Метод скалярної згортки.

Згідно з методом скалярної згортки задається вектор вагових коефіцієнтів критеріїв

$$\alpha = \left\{\alpha_i, i = \overline{1,k}\right\},$$

компоненти якого характеризують відносну важливість відповідного критерію. Вагові коефіцієнти задовольняють умовам

$$\sum_{i=1}^{k} \alpha_i = 1, \alpha_i \geq 0, i = \overline{1,k}.$$

Далі розглядають однокритеріальну (скалярну задачу лінійного програмування

$$\sum_{i=1}^{k} \alpha_i f_i\left(X\right) \to \max$$

$$X \in G.$$

Зазвичай, замість функцій $f_i\left(X\right)$ розглядають задачу для нормалізованих функцій

$$\sum_{i=1}^{k} \alpha_i \overline{f}_i\left(X\right) \to \max.$$

11.5. Метод послідовних поступок.

Проводять попередній аналіз відносної, за думкою ОПР, важливості критеріїв і впорядковують їх. Будемо вважати, що $f_1 \succ f_2 \succ ... \succ f_k$.

Розв‘язують задачу за першим критерієм і визначають його найбільше значення $f_1^* = \max\limits_{X \in G} f_1\left(X\right) = f_1\left(X_1^*\right)$ на множині $G_1 = G$. Аналізують значення інших критеріїв в одержаній точці X_1^*: $f_i\left(X_1^*\right), i = \overline{2,k}$. Якщо вони не влаштовують ОПР, то призначається величина «допустимого» зниження (поступки) значення першого критерію Δf_1 і розв‘язують однокритеріальну задачу за другим критерієм

$$f_2\left(X\right) \to \max$$

на уточненій множині

$$G_2 = \left\{X \in G_1 \middle| f_1\left(X\right) \geq f_1^* - \Delta f_1\right\}.$$

Розв‘язок задачі максимізації за критерієм $f_k\left(X\right)$ на множині $G_k=\left\{X\in G_{k-1}\middle|f_{k-1}\left(X\right)\geq f_{k-1}^{*}-\Delta f_{k-1}\right\}$, одержаний на останньому кроці вважають оптимальним розв‘язком векторної задачі

$$X^{*}=\arg\max_{X\in G_k} f_k\left(X\right).$$

11.6. Метод максиміну (гарантованого результату).

Позначимо нормалізований критерій через $\lambda_i\left(X\right)$:

$$\lambda_i\left(X\right)=\frac{f_i\left(X\right)-f_i\left(X_i^0\right)}{f_i\left(X_i^*\right)-f_i\left(X_i^0\right)},i=\overline{1,k}$$

і назвемо його відносною оцінкою, яка призначена для порівняння критеріїв між собою за числовою величиною, $0\leq\lambda_i\left(X\right)\leq 1$.

Відносний рівень λ - найнижча оцінка серед усіх відносних оцінок $\lambda_i\left(X\right)$:

$$\forall X\in G,\quad \lambda\leq\lambda_i\left(X\right), i=\overline{1,k},$$

тобто,

$$\forall X\in G,\quad \lambda=\min_{1\leq i\leq k}\lambda_i\left(X\right).$$

Якщо критерії $f_i\left(X\right)$ рівнозначні, то розв‘язок задачі (11.2), (11.3) полягає у знаходженні максимального рівня λ^* серед усіх відносних оцінок, тобто

$$\lambda^*=\max_{X\in G}\min_{1\leq i\leq k}\lambda_i\left(X\right).$$

Максимінну задачу перетворюють в однокритеріальну екстремальну задачу (будують λ-задачу):

$$\lambda^*=\max_{X\in G}\lambda$$

додаючи до основних обмежень задачі (11.3) нерівності

$$\lambda\leq\lambda_i\left(X\right), i=\overline{1,k}.$$

Тобто, розв‘язують задачу

$$\lambda\to\max$$

за умов

$$\lambda-\frac{f_i\left(X\right)-f_i\left(X_i^0\right)}{f_i\left(X_i^*\right)-f_i\left(X_i^0\right)}\leq 0, i=\overline{1,k},$$

$$X\in G.$$