ЗАДАЧА МАТЕМАТИЧЕСКОГО ПРОГРАММИРОВАНИЯ

По исторически сложившимся причинам *задачей математического программирования* принято называть задачу поиска в E^n экстремальных элементов функционала $F(x)$, на множестве элементов x удовлетворяющих условиям вида:

$$f_i(x) \geq 0 \text{ , } i = [1, m] \text{ ,}$$

или в развернутой форме

Соответственно в координатной форме эта задача записывается:

Найти максимум $F(\xi_1, \xi_2, \ldots, \xi_n)$ по $\{\xi_1, \xi_2, \ldots, \xi_n\}$,

при условиях:

$$\begin{cases} f_1(x) \geq 0, \\ f_2(x) \geq 0, \\ \quad \ldots \\ f_m(x) \geq 0. \end{cases} \tag{3.5.1}$$

Необходимые условия разрешимости задачи математического программирования

Определение 3.5.1.1	Ограничение $f_i(x) \geq 0$ называется *активным* на элементе x^*, если $f_i(x^*) \leq 0$.

Пусть функционалы $\{ F(x), f_i(x), i = [1, m] \}$ непрерывно дифференцируемы на элементе $x^* \in E^n$, и $J(x^*)$ – множество индексов активных на элементе x^* ограничений, набор элементов $\{\text{grad} f_i(x^*), i \in J(x^*)\}$ линейно независим и, наконец, $\delta x = x - x^*$, тогда оказывается справедливой

Теорема 3.5.1.1. **Пусть функционал $F(x)$ на множестве R принимает максимальное значение на элементе x^*. Тогда необходимое условие оптимальности имеет вид**

$$(\text{grad} F(x^*), \delta x) \leq 0 \quad \text{и} \quad \forall \delta x \in M(x^*) :$$
$$(\text{grad} f_i(x^*), \delta x) \geq 0 , \;\; i \in J(x^*).$$

Утверждение теоремы 3.5.1.1 при замене δx на $x - x^*$ имеет вид

$$(\operatorname{grad} F(x^*), x) \le (\operatorname{grad} F(x^*), x^*) \qquad (3.5.1.1)$$

$$\forall x : (\operatorname{grad} f_i(x^*), x) \ge (\operatorname{grad} f_i(x^*), x^*), i \in J(x^*),$$

что, согласно *теореме Фаркаша*, равносильно существованию набора чисел $\{\lambda_1^*, \lambda_2^*, ..., \lambda_m^*\}$ таких, что

$$\lambda_i^* \ge 0 \quad \text{и} \quad \sum_{i \in J(x^*)} \lambda_i^* \operatorname{grad} f_i(x^*) = \operatorname{grad} F(x^*).$$

Действительно, если систему неравенств (3.5.1.1) привести к виду

$$(\operatorname{grad} F(x^*), \delta x) \le 0$$

$$\forall \delta x : (-\operatorname{grad} f_i(x^*), \delta x) \le 0, \quad i \in J(x^*)$$

или, в координатной форме

$$\sum_{j=1}^{n} \frac{\partial F}{\partial x_j} \delta x_j \le 0$$

$$\forall \delta x : \quad -\sum_{j=1}^{n} \frac{\partial f_i}{\partial x_j} \delta x_j \le 0 \quad \forall i \in J(x^*)$$

и предположить, что каждое решение второй системы удовлетворяет первому неравенству, то по теореме Фаркаша должно существовать решение системы линейных уравнений вида

$$-\sum_{i \in J(x^*)} \frac{\partial f_i}{\partial x_j} \lambda^*{}_i = \frac{\partial F}{\partial x_j} \qquad \forall j = [1, n],$$

такое, что $\lambda^*{}_i \ge 0 \ \forall i \in J(x^*)$.

Иначе говоря, выполняются условия:

$$\lambda_i^* \geq 0 \quad \text{и} \qquad -\sum_{i\in J(x^*)} \lambda_i^* \operatorname{grad} f_i(x^*) = \operatorname{grad} F(x^*)\,.$$

Заметим, что, если использовать обозначения из теоремы 1.1.4.4, то будут справедливы равенства:

$$\| b \| = \left\| \frac{\partial F}{\partial x_j} \right\| \quad \text{и} \quad \| A \| = \left\| -\frac{\partial f_i}{\partial x_j} \right\|.$$

Если доопределить $\lambda_i^* = 0$, $\forall i \notin J(x^*)$ и ввести в рассмотрение элемент $\Lambda \in E^m$, то последнее равенство принимает вид

$$\operatorname{grad} F(x^*) + \sum_{i=1}^{m} \lambda_i^* \operatorname{grad} f_i(x^*) = o;$$

$$\lambda_i^* f_i(x^*) = 0; \ \ \lambda_i^* \geq 0; \ \ \forall i = [1, m].$$

Таким образом оказывается справедливой

Теорема 3.5.1.2. (Куна–Таккера) **Пусть функционал $F(x)$ на множестве R принимает максимальное значение на элементе x^* и элементы $\operatorname{grad} f_i(x^*), i \in J(x^*)$ линейно независимы. Тогда существуют числа λ_i^*, $\forall i = [1, m]$ такие, что**

$$\operatorname{grad} F(x^*) + \sum_{i=1}^{m} \lambda_i^* \operatorname{grad} f_i(x^*) = o;$$

$$\lambda_i^* f_i(x^*) = 0; \;\; \lambda_i^* \geq 0; \;\; \forall i = [1, m].$$

Отметим также, что, равенства

$$\lambda_i^* f_i(x^*) = 0, \;\; \forall i = [1, m]$$

принято называть условиями *дополняющей нежесткости*.

Полученные выше необходимые условия экстремальности удобнее формулировать, введя в рассмотрение специальный функционал, зависящий как от x, так и от $\Lambda = \{\lambda_1, \lambda_2, \ldots, \lambda_m\}$ и имеющий вид:

$$L(x, \Lambda) = F(x) + \sum_{i=1}^{m} \lambda_i f_i(x).$$

В этом случае необходимые условия экстремальности формулируются как

Теорема 3.5.1.3. **Пусть функционал $F(x)$ на множестве R принимает максимальное значение на элементе x^* и элементы**

$$\operatorname{grad} f_i(x^*) \;;\; i \in J(x^*)$$

линейно независимы. Тогда существуют числа $\lambda_i^* \geq 0 \; \forall i = [1, m]$ такие, что

$$\operatorname{grad}_x L(x^*, \Lambda^*) = o\,;$$

$$\lambda_i^* f_i(x^*) = 0; \quad \forall i = [1, m].$$

В теории математического программирования числа $\lambda_i^* \geq 0$, $i = [1, m]$ принято называть *множителями Лагранжа*, а функционал $L(x, \Lambda)$ – *функцией Лагранжа*.

Функция Лагранжа и ее свойства

При исследовании свойств функции Лагранжа полезными оказываются следующие теоремы.

Теорема 3.5.2.1. **Справедливо равенство**

$$\max_{x\in R} F(x) = \max_{x} \min_{\Lambda\geq o} L(x,\Lambda) .$$

Доказательство.

В силу неотрицательности чисел λ_i , $i=[1,m]$ и структуры функции Лагранжа справедливо равенство

$$\min_{\Lambda\geq o} L(x,\Lambda) = \begin{cases} F(x) & x\in R, \\ -\infty & x\notin R. \end{cases}$$

Тогда, если x^* – решение исходной задачи, то $x^* \in R$ и

$$F(x^*) = \max_{x\in R} F(x) = \max_{x} \min_{\Lambda\geq o} L(x,\Lambda).$$

Обратно: если x^* – оптимальный элемент функционала

$$\max_{x} \min_{\Lambda\geq o} L(x,\Lambda) ,$$

то $x^* \in R$ и справедливо $\max_{x\in R} F(x) = \max_{x} \min_{\Lambda\geq o} L(x,\Lambda)$.

Теорема доказана.

Следствие 3.5.2.1. **Задачи поиска $\max_{x\in R} F(x)$ и $\max_{x} \min_{\Lambda\geq o} L(x,\Lambda)$ равносильны.**

Задачу поиска $\min_{\Lambda\geq o} \max_{x} L(x,\Lambda)$ логично назвать *двойственной* к *прямой* задаче $\max_{x} \min_{\Lambda\geq o} L(x,\Lambda)$.

Теорема 3.5.2.2. **Справедливо соотношение**

$$\min_{\Lambda \geq o} \max_{x} L(x, \Lambda) \geq \max_{x} \min_{\Lambda \geq o} L(x, \Lambda).$$

Доказательство.

Очевидно, что $L(x, \Lambda) \geq \min_{\Lambda \geq o} L(x, \Lambda)$.

Но тогда, как частный случай, верна оценка

$$\forall \Lambda \geq o: \max_{x} L(x, \Lambda) \geq \max_{x} \min_{\Lambda \geq o} L(x, \Lambda),$$

из которой следует, что

$$\min_{\Lambda \geq o} \max_{x} L(x, \Lambda) \geq \max_{x} \min_{\Lambda \geq o} L(x, \Lambda).$$

Теорема доказана.

Теоремы 3.5.2.1 и 3.5.2.2 справедливы без каких-либо предположений о выпуклости прямой задачи. Их утверждения позволяют получать в общем случае *верхнюю* оценку ее решения. Наложение же дополнительных условий, приводит к более сильным оценкам, таким как

Теорема 3.5.2.3. **Пусть $F(x)$ *выпукла вверх* на *выпуклом* множестве R, имеющем внутренние элементы (условие регулярности Слейтера), тогда верны равенства**

$$\min_{\Lambda \geq o} \max_{x} L(x, \Lambda) = L(x^*, \Lambda^*) = \max_{x} \min_{\Lambda \geq o} L(x, \Lambda).$$