Общая постановка задачи нелинейного программирования (ЗНП).

$Z = f(X) \to extr,$

$g_i(X) = b_i \ (1 \le i \le k),$

$g_i(X) \le b_i \ (k+1 \le i \le m),$

где $X = (x_1, x_2, \dots, x_n)$

Определить max(min) целевой функции $Z = f(X)$ при заданной системе ограничений, где хотя бы одна из указанных функций нелинейная.

Постановка классической задачи

Найти решение X, удовлетворяющее системе уравнений связи $g_i(X) = 0 \ (1 \le i \le m)$ и обращающие в оптимум целевую функцию

Метод множителей Лагранжа

Метод позволяет свести задачу условной оптимизации к безусловной оптимизации функции Лагранжа:

$$L(X) = f(X) + \lambda_1 g_1(X) + \dots + \lambda_m g_m(X)$$

($\lambda_1, \dots, \lambda_m$ – множители Лагранжа).

Необходимые условия экстремума функции нескольких переменных:

$$\begin{cases} \dfrac{\partial L}{\partial x_1} = 0, \dfrac{\partial L}{\partial x_2} = 0, \dots \dfrac{\partial L}{\partial x_n} = 0, \\ \dfrac{\partial L}{\partial \lambda_1} = 0, \dots \dfrac{\partial L}{\partial \lambda_m} = 0, \end{cases}$$

Если функция Лагранжа и функции в ограничениях связи дифференцируемы, то справедливо следующее утверждение.

Теорема (Необходимое условие экстремума функции Лагранжа) /Принцип Лагранжа/

Точка X* является точкой условного экстремума целевой функции $Z = f(X)$ при заданных уравнениях связи $g_i(X) = 0 \ (1 \le i \le m)$, если ранг матрицы, составленной из частных производных 1-го порядка функций ограничений $g_i(X)$ по переменным $x_1, x_2, \dots, x_n$, равен m, и существуют числа $\lambda_1, \dots, \lambda_m$, не равные одновременно нулю, то выполняется условие:

$$\nabla L(X^*, \lambda^*) = \nabla f(X^*) + \sum_{i=1}^{m} \lambda_i \nabla g_i(X^*) = 0$$

Где $\nabla L(X^*, \lambda^*)$ – градиент функции Лагранжа в точке (X^*, λ^*), $(1 \le j \le n)$.

Геометрическая интерпретация метода множителей Лагранжа

Если целевая функция и уравнения связи зависят от двух переменных, то задаче условной оптимизации можно дать геометрическую интерпретацию.

Согласно принципу Лагранжа при одном уравнении связи:

$\nabla L(X^*, \lambda^*) = \nabla f(X^*) + \lambda^* \nabla g(X^*) = 0$

или $\nabla f(X^*) = -\lambda^* \nabla g(X^*)$

Это эквивалентно тому, что **в точке $X^* = (x_1^*, x_2^*)$ линии уровня функций $f(X^*)$ и $g(X^*)$ касаются, а векторы $\nabla f(X^*)$ и $\nabla g(X^*)$ коллинеарны.**

Пример

$Z = x^2 + y^2 \to extr,$

если $x + y = 1.$

Функция Лагранжа:

$L(x, y, \lambda) = x^2 + y^2 + \lambda(x + y - 1),$

Из решения системы:

$2x + \lambda = 0, \ 2y + \lambda = 0, \ x + y = 1$

Имеем $x = 0.5, y = 0.5, Z = 0.5.$

$\nabla Z(0.5, 0.5) = \{1; 1\}, \nabla g = \{1; 1\}$