Многомерная градиентная безусловная оптимизация.

Постановка задачи:

$Z = f(X) \to extr$, $X = (x_1, x_2, \dots, x_n)$

Общая схема решения градиентными методами состоит в построении последовательности $X_0, X_1, \dots, X_k, \dots$ решений задачи по следующему принципу: в качестве X_0 выбирается, любая точка области решений и каждая последующая точка получается по формуле: $X_{k+1} = X_k + \lambda l$, где l - некоторое направление (вектор), а λ - число.

Направление l и «длина шага» λ выбираются, чтобы обеспечить сходимость к оптимальному решению X^*.

Так как направление градиента ∇Z целевой функции $Z = f(X)$ является направлением её наискорейшего роста, то при отыскании максимума вогнутой функции (минимума выпуклой функции) в качестве l берётся ∇Z $(-\nabla Z)$ и тогда формула принимает вид:

$$X_{k+1} = X_k + \lambda \nabla Z(X_k), \quad \lambda > 0$$

$$(Z_{\max}) \quad (5)$$

$$X_{k+1} = X_k - \lambda \nabla Z(X_k), \quad \lambda > 0$$

$$(Z_{\min}) \quad (5')$$

Методы спуска, в которых итерационная последовательность находится по формуле (5) или (5'), называются **градиентными**.

Метод наискорейшего спуска

Рассмотрим задачу максимизации функции $f(X) = f(x_1, x_2, \dots, x_n)$, дифференцируемую в окрестности точки $X^{(k)}$. Требуется определить координаты точки **X**, для которой $f(X) \geq f(X^{(k)})$ (1.0)

Значение градиента в точке начального приближения $\nabla f(X^{(k)})$.

Координаты точки X:

$X = X^{(k)} + \lambda \nabla f(X^{(k)}), \lambda > 0$ **(1.1)**

Соответствующее значение целевой функции:

$$f(X) = f(X^{(k)} + \lambda \nabla f(X^{(k)}))$$

Полученное выражение является функцией одной переменной λ:

$$\varphi(\lambda) = f(X^{(k)} + \lambda \nabla f(X^{(k)}))$$

Необходимо найти такое число λ, при котором функция

$$\varphi(\bar{\lambda}) = \max \varphi(\lambda), \lambda > 0 .$$

Для этого используем необходимое условие экстремума функции одной переменной:

$$\frac{d\varphi}{d\lambda} = \frac{\partial f}{\partial x_1} * \frac{dx_1}{d\lambda} + \frac{\partial f}{\partial x_2} * \frac{dx_2}{d\lambda} + \dots + \frac{\partial f}{\partial x_n} * \frac{dx_n}{d\lambda}; \quad \frac{d\varphi}{d\lambda} = 0;$$

Так как координаты функции многих переменных определяются записью **(1.1),** т.е.

$x_j = x_j^{(k)} + \lambda(\partial f(X^{(k)})/\partial x_j), \lambda > 0,$

$j = \overline{1..n}.$

Значения производной определяется как скалярное произведение

$d\varphi / d\lambda = \nabla f(X) * \nabla f(X^{(k)})$ **(*)**

Геометрическая интерпретация метода наискорейшего спуска:

Так как выражение (*) равно нулю, то градиенты целевой функции, вычисленные в данной точке и в точке последующего приближения располагаются взаимно перпендикулярно. Градиент, вычисляемый в некоторой точке линий уровня перпендикулярен касательной в этой точке. Последующее приближение выбирается таким образом, чтобы градиент касался линии уровня.

(Пример)

Условие сходимости метода (критерий окончания)

На практике признаком достижения точки оптимума является малое изменение координат точек и близость к нулю координат вектора градиента.

Проверка критерия окончания итерационного процесса:

$|x^{(k+1)} - x^{(k)}| \leq \varepsilon$ **или** $|\nabla f(X^{(k)})| \leq \varepsilon$

Замечание. МНС - это метод поиска множества локальных экстремумов, среди которых выбирают глобальные. Для выпуклых (вогнутых) функций градиентные методы позволяют найти глобальные экстремумы.