МЕТОДЫ ОПТИМИЗАЦИИ
Лекция 7: Численные методы (Задачи безусловной оптимизации)

Панин Артем Александрович
Email: aapanin1988@gmail.com

Новосибирский государственный университет

2025

Рассмотрим задачу безусловной оптимизации:

$$f(x) \longrightarrow \min$$

Численные методы решения являются итерационными и характеризуются двумя параметрами: направлением спуска p^k и длиной шага a_k (k – номер итерации). Решение или приближение на k-й итерации определяется следующим образом:

$$x^k := x^{k-1} + a_k p^k$$

Численные методы. Классификация.

Порядок численного метода – максимальный порядок производной целевой функции, используемой при вычислении направления спуска.

В методах нулевого порядка производные не используются (покоординатный спуск, метод оврагов).

В методах первого порядка используется сама целевая фукнция и ее производная (градиентные методы).

В методах второго порядка используется сама целевая фукнция, ее производная и производная второго порядка (метод Ньютона).

Пусть в точке x^* достигается минимум функции $f(x)$.

Скорость сходимости линейная, если существует q из интервала $(0, 1)$ такое, что для любого k

$$||x^{k+1} - x^*|| \leq q||x^k - x^*||.$$

Например: градиентный спуск.

Скорость сходимости квадратичная, если существует $q \geq 0$ такое, что для любого k

$$||x^{k+1} - x^*|| \leq q||x^k - x^*||^2$$

Например: метод Ньютона.

Покоординатный спуск

Рассмотрим единичные координатные вектора $e^i = (0, ..., 0, 1, 0, ..., 0)$, где единица стоит на i-й позиции.

В методе покоординатного спуска направления спуска совпадают с единичными координатными векторами.

Область применения: минимизируемая функция либо не обладает нужной гладкостью, либо является гладкой, но вычисление производных слишком трудоемко.

Наиболее эффективен, когда функция является гладкой. Метод не требует знания градиента, но сходимость можно гарантировать лишь для гладких функций.

Метод оврагов

Выберем произвольную точку x^0 и спустимся из нее (с малым числом итераций) в точку y^0 . Если рельеф овражный, эта точка окажется вблизи дна оврага.

В окрестности x^0 выберем точку x^1, из которой спустимся в некоторую точку y^1. Эта точка тоже лежит вблизи дна оврага.

Проведем через точки y^0 и y^1 на дне оврага прямую – приблизительную линию дна оврага, передвинемся по этой линии в сторону убывания функции и выберем новую точку x^2 на этой прямой, на расстоянии h от точки y^1 . Величина h называется овражным шагом и для каждой функции подбирается в ходе расчета.

Спустимся из точки x^2 в точку y^2...

Градиентные методы

Все итерационные процессы, в которых направление движения на каждом шаге совпадает с антиградиентом (градиентом) функции, называются градиентными методами:

$$x^{k+1} = x^k - a_k f'(x^k), a_k \geq 0.$$

Методы отличаются способами выбора длины шага a_k.

Метод с постоянным шагом: $a_k = a$.

Метод наискорейшего спуска: при выборе a_k минимизируется по a функция $f(x^k - af'(x^k))$:

$$a_k = \arg\min_{a \geq 0} f(x^k - af'(x^k)).$$

Метод Ньютона

Пусть f дважды непрерывно дифференцируемая функция и есть алгоритм вычисления её вторых производных.

Идея метода: заменить функцию f в окрестности текущего приближения x^k её квадратичной аппроксимацией:

$$q(x) = f(x^k) + f^{'}(x^k)(x - x^k) + 1/2(x - x^k)^T f^{''}(x^k)(x - x^k).$$

Выбрать в качестве нового приближения x^{k+1} точку минимума функции $q(x)$ (если она существует).

Метод Ньютона

Предположим, что матрица $f''(x^k)$ положительно определённая, тогда функция $q(x)$ сильно выпукла, тогда её минимум можно найти как решение системы уравнений

$$q'(x^{k+1}) = 0,$$

которая (по определению $q(x)$) эквивалентна следующей системе линейных уравнений

$$f'(x^k) = -f''(x^k)(x^{k+1} - x^k).$$

Получаем необходимую итерационную формулу:

$$x^{k+1} = x^k - (f''(x^k))^{-1} f'(x^k).$$

Заметим, что здесь как направление, так и длина шага заданы.