ЛЕКЦИЯ № 7

Методы решения конечномерных задач оптимизации (Задачи безусловной оптимизации)

1. Метод покоординатного спуска

2. Градиентный метод

3. Метод Ньютона

Численные методы НЛП

Задача поиска безусловного минимума:

$$f(x) \longrightarrow \min_{x \in R^n}.$$

Для решения используются численные методы, в которых текущее приближение вычисляется по формуле

$$x^{k+1} = x^k + \alpha_k p^k,$$

где p^k — направление спуска, α_k — длина шага вдоль этого направления.

Численные методы НЛП

Методы, в которых последовательность векторов $x^0, x^1, \ldots, x^k, \ldots,$ удовлетворяет условию

$$f(x^0) \geq f(x^1) \geq \ldots \geq f(x^k) \geq \ldots$$

называются релаксационными.

Пусть x^*– минимум функции $f(x)$. Скорость сходимости линейная: если для $k = 0, 1, \ldots$

$$\|x^{k+1} - x^*\| \leq q\ \|x^k - x^*\|, 0 < q < 1,$$

Численные методы НЛП

или говорят, что метод сходится со скоростью геометрической прогрессии, т.к.

$$\|x^k - x^*\| \leq q^k \, \|x^0 - x^*\|.$$

Скорость сходимости сверхлинейна, если

$$\|x^{k+1} - x^*\| \leq q_k \, \|x^k - x^*\|,$$

где $q_k \to 0$ при $k \to \infty$, и квадратична, если

$$\|x^{k+1} - x^*\| \leq C \, \|x^k - x^*\|^2, C \geq 0.$$

Численные методы НЛП

Метод нулевого порядка, если в процессе вычислений используются только значения целевой функции.

В методах первого порядка, помимо значений целевой функции используются её производные.

и так далее.

МЕТОД ПОКООРДИНАТНОГО СПУСКА

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 ПОКООРДИНАТНОГО СПУСКА

$$\boldsymbol{f(x) \longrightarrow \inf_{x \in R^n}} \qquad (1)$$

Пусть $\boldsymbol{e_i = (0, \ldots, 0, 1, 0, \ldots, 0)}$ – $\boldsymbol{i}$-ый единичный координатный вектор,

$\boldsymbol{x^0}$ – начальное приближение, $\boldsymbol{\alpha_0 > 0}$ – начальная длина шага.

Пусть $\boldsymbol{x^t \in R^n}$ – текущее приближение,

$\boldsymbol{\alpha_t > 0}$ – текущая длина шага,

$\boldsymbol{\lambda, 0 < \lambda < 1}$ – фиксированное число.

МЕТОД ПОКООРДИНАТНОГО СПУСКА

Метод покоординатного спуска – итеративный процесс. Все итерации разбиты на группы. Каждая группа содержит столько итераций сколько координатных векторов. $\boldsymbol{k}$-ая группа начинается с итерации с номером $(\boldsymbol{k}-1)\,\boldsymbol{n}+1$. Последняя итерация этой группы имеет номер $\boldsymbol{kn}$.

Опишем итерацию с номером $\boldsymbol{t}$, где

$$(\boldsymbol{k}-1)\,\boldsymbol{n}+1 \leq \boldsymbol{t} \leq \boldsymbol{kn}. \tag{2}$$

ИТЕРАЦИЯ T

Если

$$f(x^{t-1}+\alpha_{t-1}\, e_{t-(k-1)n}) < f(x^{t-1}), \quad (3)$$

то положим

$$x^{t} = x^{t-1} + \alpha_{t-1}\, e_{t-(k-1)n}, \alpha_{t} = \alpha_{t-1}. \quad (4)$$

ИТЕРАЦИЯ T

Если

$$f(x^{t-1} - \alpha_{t-1}\, e_{t-(k-1)n}) < f(x^{t-1}), \quad (5)$$

то положим

$$x^t = x^{t-1} - \alpha_{t-1}\, e_{t-(k-1)n}, \alpha_t = \alpha_{t-1}. \quad (6)$$

Если выполняется (3) или (5), то итерация t — удачная.

ИТЕРАЦИЯ T

Если итерация $\boldsymbol{t}$ неудачная, то положим: $\boldsymbol{x^t = x^{t-1}}$,

$$\boldsymbol{\alpha_t} = \begin{cases} \boldsymbol{\lambda\alpha_{t-1}}, & \text{если } \boldsymbol{t = kn}, \text{ и все итерации группы неудачны}, \\ \boldsymbol{\alpha_{t-1}}, & \text{если } \boldsymbol{t \neq kn} \text{ или были удачные итерации внутри группы} \end{cases} \qquad \mathbf{(7)}$$

МЕТОД ПОКООРДИНАТНОГО СПУСКА

Пусть в k-ой группе не оказалось ни одной удачной итерации и шаг дробится. В этом случае выполняются неравенства

$$\begin{aligned} f(x^{t-1} + \alpha_{t-1}\, e_i) &\geq f(x^{t-1}), \\ f(x^{t-1} - \alpha_{t-1}\, e_i) &\geq f(x^{t-1}), \end{aligned} \qquad (8)$$

для всех $i = 1, \ldots, n$.

Метод покоординатного спуска

Теорема 15. Пусть функция $\boldsymbol{f(x)}$ выпукла на $\boldsymbol{R^n}$ и $\boldsymbol{f \in C^1(R^n)}$, а начальное приближение таково, что множество $\boldsymbol{M(x^0) = \{x \in R^n : f(x) \le f(x^0)\}}$ ограничено.

Тогда последовательность $\boldsymbol{x^k}$ имеет хотя бы одну предельную точку и любая предельная точка этой последовательности есть оптимальное решение задачи.

Численные методы НЛП

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

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

Методы отличаются способами выбора длины шага α_k.

Численные методы НЛП

Метод с постоянным шагом: $\alpha_k = \alpha$.

Метод с дроблением шага: на каждом шаге проверяется неравенство

$$f(x^k - \alpha_k f'(x^k)) - f(x^k) \le -\epsilon\, \alpha_k \|f'(x^k)\|^2,$$

где $0 < \epsilon < 1$.

Метод наискорейшего спуска: при выборе α_k минимизируется по α функция $f(x^k - \alpha\, f'(x^k))$:

$$\alpha_k = \arg\min_{\alpha \ge 0} f(x^k - \alpha\, f'(x^k)).$$

Теорема 16 (Первая теорема сходимости)

Пусть функция $f \in C^1(R^n)$, ограничена снизу $f(x) \geq f^* > -\infty$, выполняется условие Липшица для градиента $f'(x)$:

$$\|f'(x) - f'(y)\| \leq L \,\|x - y\|$$

и длина шага α удовлетворяет условию $0 < \alpha < 2/L$. Тогда $f'(x^k) \to 0$ при $k \to \infty$ и $f(x^{k+1}) \leq f(x^k)$, при любом выборе начального приближения x_0.

Определение Дифференцируемая функция $\boldsymbol{f}$ называется сильно выпуклой (с константой $\boldsymbol{l > 0}$), если для любых $\boldsymbol{x}$ и $\boldsymbol{y}$ из $\boldsymbol{R^n}$ справедливо

$$\boldsymbol{f(x+y) \geq f(x) + \langle f'(x), y\rangle + l\|y\|^2/2.} \quad \mathbf{(9)}$$

Лемма 10. Если функция $\boldsymbol{f}$ является сильно выпуклой (с константой $\boldsymbol{l > 0}$), то она имеет глобальный минимум на $\boldsymbol{R^n}$.

Доказательство. Перепишем (9), используя неравенство Коши - Буняковского

$$f(x+y) \geq f(x) - \|f'(x)\|\|y\| + l\|y\|^2/2.$$

Вынесем величину $l\|y\|/2$ за скобки

$$f(x+y) \geq f(x) + l\|y\|/2(\|y\| - 2\|f'(x)\|/l).$$

$\Longrightarrow \forall y : \|y\| > r = 2\|f'(x)\|/l$ имеем неравенство $f(x+y) > f(x)$, из которого следует, что минимум $\exists$ и достигается на шаре $B(x,r)$.

$\blacksquare$

Теорема 17 (Вторая теорема сходимости) Пусть функция $\boldsymbol{f}$ дифференцируема в $\boldsymbol{R^n}$, является сильно выпуклой, выполняется условие Липшица для градиента $\boldsymbol{f'(x)} : \|\boldsymbol{f'(x) - f'(y)}\| \leq \boldsymbol{L}\,\|\boldsymbol{x - y}\|$ и длина шага $\boldsymbol{\alpha}$ удовлетворяет условию $\boldsymbol{0 < \alpha < 2/L}$.

Тогда $\boldsymbol{x^k} \to \boldsymbol{x^*}$ при $\boldsymbol{k} \to \infty$ и $\|\boldsymbol{x^k - x^*}\| \leq \boldsymbol{Cq^k},\ \boldsymbol{0 \leq q < 1}$.

МЕТОД НЬЮТОНА

Пусть $\boldsymbol{f}$ — дважды непрерывно дифференцируемая функция и есть алгоритм вычисления её вторых производных.

Идея метода: заменить функцию $\boldsymbol{f}$ в окрестности текущего приближения $\boldsymbol{x^k}$ её квадратичной аппроксимацией: $\boldsymbol{q(x) = f(x^k) + f'(x^k)(x - x^k) + \frac{1}{2}(x - x^k)^\top f''(x^k)(x - x^k).}$

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

МЕТОД НЬЮТОНА

Предположим, что матрица $\boldsymbol{f''(x^k)}$ — положительно определённая $\Longrightarrow$ функция $\boldsymbol{q(x)}$ — сильно выпукла $\Longrightarrow$ её минимум можно найти как решение системы уравнений

$\boldsymbol{q'(x^{k+1}) = 0}$, которая по определению $\boldsymbol{q(x)}$ эквивалентна следующей системе линейных уравнений

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

МЕТОД НЬЮТОНА

$\Longrightarrow$ 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.

Пусть последовательность $\{x^k\}$ получена с помощью метода Ньютона и точка x^* — глобальный минимум функции f.

МЕТОД НЬЮТОНА

Теорема 18 Пусть дважды непрерывно дифференцируемая функция f сильно выпукла (с константой $l > 0$), вторая производная удовлетворяет условию Липшица: для любых $x, y \in R^n$

$$\|f''(x) - f''(y)\| \leq L \|x - y\|,$$

и $q = L\|f'(x^0)\|/2l^2 < 1$. Тогда $x^k \to x^*$ при $k \to \infty$ и метод Ньютона имеет квадратичную скорость сходимости

$$\|x^k - x^*\| \leq (2l/L)q^{2^k}.$$