ЛЕКЦИЯ № 16

1. Метод покоординатного спуска

2. Метод Ньютона (продолжение)

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 единичны 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 $(\boldsymbol{k}-\mathbf{1})\,\boldsymbol{n}+\mathbf{1}$. Последняя итерация этой группы имеет номер $\boldsymbol{kn}$.

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

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

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 t неудачная, то положим: $x^t = x^{t-1}$,

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

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 22. 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.

МЕТОД ПОКООРДИНАТНОГО СПУСКА

Доказательство. В задаче (1) $\exists$ оптимальное решение $x^* \in R^n$ (теор. Вейерштрасса): из (3)-(7) $\Longrightarrow$

$$f(x^t) \leq f(x^{t-1}), t = 1, \ldots, \Longrightarrow$$

$$\{x^t\} \in M(x^0) \text{ и } \exists \lim_{t\to\infty} f(x^t) \geq f^* = f(x^*).$$

Покажем

$$\lim_{t\to\infty} \alpha_t = 0.$$

МЕТОД ПОКООРДИНАТНОГО СПУСКА

Допустим противное: $\alpha_t = \alpha > 0, \forall t \geq t_0$ (т.е. процесс дробления конечен).

Пусть $M_\alpha = \{u \in M(x^0) | u = x^{t_0} + \alpha \sum_i r_i e_i, r_i = 0, \pm 1, \pm 2, \ldots, i = 1, \ldots, n\}$ – сетка с шагом α.

Понятно, что начиная с номера t_0 любой цикл из n итераций содержит хотя бы одну удачную итерацию. На каждой удачной итерации переходим от текущей точки сетки x^t к соседней x^{t+1}.

МЕТОД ПОКООРДИНАТНОГО СПУСКА

Т.к. при этом $f(x^t) > f(x^{t+1})$, то посетим каждую точку сетки не более одного раза $\Longrightarrow$

сетка M_α содержит бесконечное множество разных точек $\Longrightarrow$

противоречие с ограниченностью множества $M(x^0)$ $\Longrightarrow$

процесс дробления длины шага α_t бесконечен $\Longrightarrow$

$\lim_{t\to\infty} \alpha_t = 0$.

Метод покоординатного спуска

Пусть $k_1 < k_2 < \cdots < k_m < \ldots$ – номера итераций, на которых дробится длина шага и выполняются неравенства (8).

Т.к. $\{x^k\} \in M(x^0)$ и множество $M(x^0)$ ограничено, то без ограничения общности можно считать, что $\exists \lim x^{k_m}$ при $m \longrightarrow +\infty$ и равен $\widehat{x}$. Из формулы конечных приращений и (8) имеем

$$\forall m \ \langle f'(x^{k_m} + \theta_{k_m}\alpha_{k_m} e_i), \alpha_{k_m} e_i\rangle =$$
$$= f(x^{k_m} + \alpha_{k_m} e_i) - f(x^{k_m}) \geq 0 \Longrightarrow$$

Метод покоординатного спуска

$f'_{x_i}(x^{k_m}+\theta_{k_m}\alpha_{k_m} e_i) \geq 0,$ где $i = 1, \ldots, n,$
$0 \leq \theta_{k_m} \leq 1.$

Аналогично получим, что

$$\forall m \ \langle f'(x^{k_m} - \overline{\theta}_{k_m}\alpha_{k_m} e_i), -\alpha_{k_m} e_i\rangle =$$

$$= f(x^{k_m} - \alpha_{k_m} e_i) - f(x^{k_m}) \geq 0$$

и, следовательно, для $\forall m$ и $\forall i = 1, \ldots, n$

$$f'_{x_i}(x^{k_m} - \overline{\theta}_{k_m}\alpha_{k_m} e_i) \leq 0,$$

где $0 \leq \overline{\theta}_{k_m} \leq 1.$

Метод покоординатного спуска

По условию, частные производные $f'_{x_i}(x)$ – непрерывные функции на R^n. Поэтому из условий:

$$\lim_{m\to\infty} \alpha_{k_m} = 0, \lim_{m\to\infty} x^{k_m} = \widehat{x}$$

$$\forall m \, f'_{x_i}(x^{k_m}+\theta_{k_m}\alpha_{k_m} e_i) \geq 0, i = 1, \ldots, n$$

$$\forall m \, f'_{x_i}(x^{k_m}-\overline{\theta}_{k_m}\alpha_{k_m} e_i) \leq 0, i = 1, \ldots, n$$

имеем $f'_{x_i}(\widehat{x}) = 0, i = 1, \ldots, n.$

Метод покоординатного спуска

$\Longrightarrow \boldsymbol{f'(\widehat{x}) = 0} \Longrightarrow \boldsymbol{\widehat{x}}$ является оптимальным решением задачи, т.к. $\boldsymbol{f}$ – выпуклая функция. $\blacksquare$

Лемма 15 Если f — дважды непрерывно дифференцируемая сильно выпуклая функция с константой l, то выполняется неравенство:

$$\|[f''(x)]^{-1}\| \le l^{-1}.$$

Доказательство. Воспользуемся следующей формулой для конечных приращений функции f:

$$f(x+y) - f(x) = \langle f'(x), y \rangle +$$

Метод Ньютона

$$+\frac{\langle f''(x+\tau_2 y)y, y\rangle}{2},$$

где $0 \le \tau_2 \le 1.$ Из неравенства (23) (определения сильной выпуклости) $\Longrightarrow$

$$\langle f''(x+\tau_2 y)y, y\rangle/2 = f(x+y) - f(x) -$$
$$-\langle f'(x), y\rangle \ge l\|y\|^2/2.$$

Заменяя y на ty, получим:

$$\langle f''(x+\tau_2 ty)ty, ty\rangle \ge l\|ty\|^2.$$

Следовательно,

$$t^2\langle f''(x+\tau_2 ty)y, y\rangle \geq t^2 l\|y\|^2.$$

Поделив на t^2 и устремляя t к нулю, будем иметь

$$\langle f''(x)y, y\rangle \geq l\|y\|^2.$$

Пусть $y = (f''(x))^{-1}z \Longrightarrow$

$$\langle z, (f''(x))^{-1}z\rangle \geq l\|(f''(x))^{-1}z\|^2$$

$\Longrightarrow$ (используя неравенство Коши-Буняковского)

Метод Ньютона

$$\|z\| \|(f''(x))^{-1}z\| \geq l\|(f''(x))^{-1}z\|^2 \Longrightarrow$$

$$l\|(f''(x))^{-1}z\| \leq \|z\| \ \forall z.$$

Это означает, что $\|[f''(x)]^{-1}\| \leq l^{-1}$. ■

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

Метод Ньютона

Теорема 19 Пусть f – дважды непрерывно дифферен-цируемая сильно выпуклая (с константой $l > 0$) функция, 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 $x^k \longrightarrow x^*$ при $k \longrightarrow \infty$ и метод Ньютона имеет квадратичную скорость сходимости

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

Доказательство. Рассмотрим следующую формулу конечных приращений:

$$\boldsymbol{g(x+y) = g(x) + \langle g'(x), y\rangle + \int\limits_0^1 \langle g'(x+\tau y) - g'(x), y\rangle d\tau.}$$

Заменим $\boldsymbol{g}$ на $\boldsymbol{f'}$:

Метод Ньютона

$$f'(x+y) = f'(x)+\langle f''(x), y\rangle+\int\limits_0^1 \langle f''(x+\tau y)-$$

$$-f''(x), y\rangle d\tau.$$

Из неравенства Коши-Буняковского и условия Липшица следует, что

$$\|f'(x+y)-f'(x)-\langle f''(x), y\rangle\| \leq L\|y\|^2/2.$$

Метод Ньютона

Подставим $x = x^k$ и $y = -[f''(x^k)]^{-1} f'(x^k)$

$$\|f'(x^{k+1})\| \le (L/2)\|[f''(x^k)]^{-1}\|^2 \|f'(x^k)\|^2.$$

Применяя лемму 15, получим

$$\|f'(x^{k+1})\| \le (L/2l^2)\|f'(x^k)\|^2.$$

Итерируем это неравенство по k

$$\|f'(x^{k+1})\| \le (2l^2/L)\,(\underbrace{L\|f'(x^0)\|/2l^2}_{q})^{2^{k+1}}.$$

Метод Ньютона

f — сильно выпуклая функция $\Longrightarrow$

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

y заменим на $y - x \Longrightarrow$

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

меняем x на y, а y на $x \Longrightarrow$

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

Сложим оба неравенства:

$$\langle f'(x) - f'(y), x-y\rangle \geq l\|x-y\|^2.$$

Метод Ньютона

Подставим $\boldsymbol{y} = \boldsymbol{x}^*, \boldsymbol{x} = \boldsymbol{x}^{k+1}$, и учитывая равенство $\boldsymbol{f}'(\boldsymbol{x}^*) = \mathbf{0}$, получим

$$l\|x^{k+1} - x^*\|^2 \leq \langle f'(x^{k+1}), x^* - x^{k+1}\rangle \leq$$

$$\leq \|f'(x^{k+1})\| \, \|x^* - x^{k+1}\| \leq$$

$$\leq (2l^2/L)q^{2^{k+1}}\|x^* - x^{k+1}\| \Longrightarrow$$

$$\|x^{k+1} - x^*\| \leq (2l/L)q^{2^{k+1}} \blacksquare$$