ЛЕКЦИЯ № 15

Методы нулевого порядка

1. Метод покоординатного спуска

2. Метод покрытия (МВГ для липшицевых функций на г.-кубе)

Прямые и двойственные методы

1. Метод Келли

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

Рассмотрим задачу

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

Пусть $e_i = (0, \ldots, 0, 1, 0, \ldots, 0)$ – i-ый единичный координатный вектор, x_0 – начальное приближение, $\alpha_0 > 0$ – начальная длина шага. Пусть $x_k \in R^n$ – текущее приближение, $\alpha_k > 0$ – текущая длина шага, вектор p_k – текущее направление движения, $\lambda, 0 < \lambda < 1$ – фиксированное число.

Метод покоординатного спуска

Выбор направления движения:

$$\boldsymbol{p_k = e_{i_k}, i_k = k - n[k/n] + 1,} \qquad \textbf{(2)}$$

где $[\boldsymbol{k/n}]$ – целая часть, числа $\boldsymbol{k/n}$. Условие (2) гарантирует циклический перебор векторов $\boldsymbol{e_1, e_2, \ldots, e_n}$:

$$\boldsymbol{p_0 = e_1, \ldots, p_{n-1} = e_n,}$$

$$\boldsymbol{p_n = e_1, \ldots, p_{2n-1} = e_n, p_{2n} = e_1, \ldots}$$

Метод покоординатного спуска

Итерация $(k+1)$. Вычислить $f(x)$ в точке $x = x_k + \alpha_k\, p_k$. Если

$$f(x_k + \alpha_k\, p_k) < f(x_k), \tag{3}$$

то положим

$$x_{k+1} = x_k + \alpha_k\, p_k, \alpha_{k+1} = \alpha_k. \tag{4}$$

Если (3) не выполняется, то вычислить $f(x)$ в точке $x = x_k - \alpha_k\, p_k$.

Метод покоординатного спуска

Если

$$f(x_k - \alpha_k\, p_k) < f(x_k), \tag{5}$$

то положим

$$x_{k+1} = x_k - \alpha_k\, p_k, \alpha_{k+1} = \alpha_k. \tag{6}$$

Итерацию $k + 1$ назовем удачной, если выполняется хотя бы одно из неравенств (3) или (5).

Метод покоординатного спуска

Если $k+1$ итерация неудачная, то положим: $x_{k+1} = x_k$,

$$\alpha_{k+1} = \begin{cases} \lambda\alpha_k, & \text{если } i_k = n, x_k = x_{k-n+1}, \\ \alpha_k, & \text{если } i_k \neq n \text{ или } x_k \neq x_{k-n+1} \\ & \text{или } 0 \leq k \leq n-1. \end{cases} \quad (7)$$

Метод покоординатного спуска

В (7) условие $i_k = n, x_k = x_{k-n+1}$ означает, что при последовательном переборе направлений $e_1, \ldots, e_n$ (n последних итераций) не оказалось ни одной удачной $\Longrightarrow$ шаг α_k дробится. В этом случае выполняются неравенства

$$f(x_k+\alpha_k e_i) \geq f(x_k), f(x_k-\alpha_k e_i) \geq f(x_k) \tag{8}$$

для всех $i = 1, \ldots, n$.

Метод покоординатного спуска

Если в данном цикле из $\boldsymbol{n}$ итераций реализовалась хотя бы одна удачная итерация, то тогда на последней итерации цикла $\boldsymbol{i_k = n}$ но $\boldsymbol{x_k \neq x_{k-n+1}}$ $\Longrightarrow$ длина шага $\boldsymbol{\alpha_k}$ не дробится и сохраняется еще на протяжении $\boldsymbol{n}$ итераций следующего цикла (т.к. дробление возможно только на последней итерации цикла)

Теорема 24. Пусть функция $f(x)$ выпукла на R^n и $f \in C^1(R^n)$, а начальное приближение таково, что множество $M(x_0) = \{x \in R^n : f(x) \le f(x_0)\}$ ограничено.

Тогда последовательность x_k имеет хотя бы одну предельную точку и любая предельная точка этой последовательности есть оптимальное решение задачи.

Доказательство. По теореме Вейерштрасса в задаче (1) $\exists$ оптимальное решение $\boldsymbol{x^* \in R^n}$. С другой стороны из (2)-(7) $\Longrightarrow$

$$\boldsymbol{f(x_{k+1}) \leq f(x_k), k = 0, 1, \ldots,} \Longrightarrow$$

$$\boldsymbol{\{x_k\} \in M(x_0)} \text{ и } \exists \lim_{\boldsymbol{k \to \infty}} \boldsymbol{f(x_k) \geq f^* = f(x^*).}$$

Покажем

$$\lim_{\boldsymbol{k \to \infty}} \boldsymbol{\alpha_k = 0.}$$

Метод покоординатного спуска

Допустим противное: $\boldsymbol{\alpha_k = \alpha > 0, \forall k \geq k_0}$ (т.е. процесс дробления конечен).

Пусть $\boldsymbol{M_\alpha = \{u | u = (x_{k_0} + \alpha\ r\ e_i) \in M(x_0), i = 1, \ldots, n, r = 0, \pm 1, \pm 2, \ldots\}}$ – сетка с шагом $\boldsymbol{\alpha}$.

Понятно, что начиная с номера $\boldsymbol{k_0}$ любой цикл из $\boldsymbol{n}$ итераций содержит хотя бы одну удачную итерацию. На каждой удачной итерации переходим от текущей точки сетки $\boldsymbol{x_k}$ к соседней $\boldsymbol{x_{k+1}}$.

Метод покоординатного спуска

Т.к. при этом $f(x_k) > f(x_{k+1})$, 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_k бесконечен $\Longrightarrow$

$\lim_{k\to\infty} \alpha_k = 0$.

Метод покоординатного спуска

Пусть $k_1 < k_2 < \dots < k_m < \dots$ – 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

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

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

Это означает, что градиент функции $\boldsymbol{f}$ в точке $\widehat{\boldsymbol{x}}$ равен 0, т.е. $\boldsymbol{f}'(\widehat{\boldsymbol{x}}) = \mathbf{0}$. Для выпуклой функции $\boldsymbol{f}$ это необходимое условие локальной оптимальности является достаточным. Следовательно $\widehat{\boldsymbol{x}}$ является оптимальным решением задачи. $\blacksquare$

Прямые методы: метод Келли

Данный метод используется для решения задач выпуклого программирования вида:

$$\min f(x)$$

при условии, что

$$\varphi_i(x) \leq 0, i = \overline{1, m}.$$

Здесь $f, \varphi_i : R^n \longrightarrow R$ – выпуклые функции и $f, \varphi_i \in C^1$.

Прямые методы: метод Келли или метод секущих плоскостей

Вводя дополнительные переменную и ограничение, сделаем функционал задачи линейным:

$$\min y$$

$$f(x) \leq y$$

$$\varphi_i(x) \leq 0, i = 1, ..., m,$$

$\Longrightarrow$ без ограничения общности считаем, что $f(x) = \langle c, x\rangle$ и ограничимся изучением выпуклых задач вида:

Метод Келли

$$(c, x) = \sum_{j=1}^{n} c_j x_j \longrightarrow \min$$

при условии, что

$$\varphi_i(x) \le 0, i = \overline{1, m}.$$

Также считаем, что в задаче существует оптимальное решение x^*, которое содержится в многогранном множестве $Q_0 = \{x \in R^n | Ax = b, x \ge 0\}$, а также $Q \subseteq Q_0$.

Итерация k.

1. Решаем задачу ЛП

$$(c, x) \longrightarrow \min$$

$$x \in Q^k,$$

где Q^k – текущее многогранное приближение множества Q, причём $Q \subseteq Q^k$.

Пусть x^k – оптимальное решение этой задачи.

Метод Келли

Если $\boldsymbol{x^k}$ – допустимое решение исходной задачи, то оно является его оптимальным решением (т.к. $\boldsymbol{Q \subseteq Q^k}$). Алгоритм заканчивает работу. В противном случае переходим к следующему шагу.

2. Найдём номер ограничения $\boldsymbol{i_k}$, для которого величина $\boldsymbol{\varphi_{i_k}(x^k) > 0}$ максимальна. Перейдём к выполнению следующей итерации с

$$\boldsymbol{Q^{k+1} = Q^k \cap \{x | \varphi_{i_k}(x^k) + (\varphi'_{i_k}(x^k), x - x^k) \leq \leq 0.\}}$$

Корректность определения множества Q^{k+1}.

1. Т.к. ограничение $\varphi_{i_k}(x^k) + (\varphi'_{i_k}(x^k), x - x^k) \le 0$ линейно, то множество Q^{k+1} является многогранным.

2. В силу выпуклости множества Q^k и функции φ_{i_k} имеем

$$\varphi_{i_k}(x^k) + (\varphi'_{i_k}(x^k), x - x^k) \le \varphi_{i_k}(x)$$

для всех $x \in Q^k$ (Лемма 3, лек. № 2).

Но для $\boldsymbol{x} \in \boldsymbol{Q}$ $\boldsymbol{\varphi_{i_k}(x) \leq 0}$. Следовательно, $\boldsymbol{Q} \subseteq \boldsymbol{Q^{k+1}}$. (Другими словами ограничение

$$\boldsymbol{\varphi_{i_k}(x^k) + (\varphi'_{i_k}(x^k), x - x^k) = 0}$$

является отсечением, которое отсекает точку $\boldsymbol{x^k}$.) Если алгоритм останавливается через конечное число шагов, то текущее приближение – оптимальное решение задачи. Рассмотрим случай, когда последовательность $\boldsymbol{\{x^k\}}$ бесконечна.

Метод Келли

Теорема 24. Любая предельная точка последовательности $\{x^k\}$, порождённая методом секущих плоскостей, есть оптимальное решение задачи.

Доказательство. Последовательность значений $\{(c, x^k)\}$ монотонно неубывающая и ограничена сверху (т.к. существует оптимальное решение). Поэтому без ограничения общности считаем, что последовательность $\{x^k\}$ ограничена. Тогда найдётся сходящаяся подпоследовательность.

Метод Келли

Для упрощения обозначений, считаем, что это и есть наша последовательность $\{x^k\}_{k \in N}$. Пусть $\overline{x}$ – её предел.

Выберем произвольное ограничение с номером i. Рассмотрим подпоследовательность элементов $\{x^k\}_{k \in T}$, где $T \subseteq N$, для которых секущая плоскость порождалась с помощью i-го ограничения. Возможны два случая.

1. Найдётся номер k_0 такой, что $\varphi_i(x^k) \leq 0$ для всех $k \geq k_0$. Тогда

$$\varphi_i(x^k) \longrightarrow \varphi_i(\overline{x}) \leq 0.$$

Либо

2. Подпоследовательность $\{x^k\}_{k \in T}$ бесконечна. Тогда для любого $k' > k$

$$\varphi_i(x^k) + (\varphi_i^{'}(x^k), x^{k'} - x^k) \leq 0.$$

Следовательно,

$$\varphi_i(x^k) \leq \|\varphi_i^{'}(x^k)\| \|x^{k'} - x^k)\|.$$

Т.к. $\|x^{k'} - x^k)\| \longrightarrow 0$ (из-за сходимости последовательности), $\|\varphi_i^{'}(x^k)\| \longrightarrow \|\varphi_i^{'}(\overline{x})\|$ ($\varphi_i \in C^1(R^n)$).

Таким образом

$$\varphi_i(x^k) \longrightarrow \varphi_i(\overline{x}) \leq 0.$$

Метод Келли

Следовательно, $\overline{\boldsymbol{x}}$ – допустимое решение задачи (в силу произвольного выбора $\boldsymbol{i}$). Но для любого $\boldsymbol{k}$ имеем $(\boldsymbol{c}, \boldsymbol{x}^{\boldsymbol{k}}) \leq (\boldsymbol{c}, \boldsymbol{x}^*) \Longrightarrow (\boldsymbol{c}, \overline{\boldsymbol{x}}) \leq (\boldsymbol{c}, \boldsymbol{x}^*)$. Т.е. $\overline{\boldsymbol{x}}$ – оптимальное решение задачи. $\blacksquare$