ЛЕКЦИЯ № 15

1. Метод Ньютона (метод второго порядка)
2. Метод внешних штрафов
3. Метод внутренних штрафов
4. Метод покоординатного спуска

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: $q(x) = f(x^k) + f'(x^k)(x - x^k) + \frac{1}{2}(x - x^k)^\top f''(x^k)(x - x^k)$.

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

МЕТОД НЬЮТОНА

Предположим, что матрица $f''(x^k)$ — положительно определённая $\Longrightarrow$ функция $q(x)$ — сильно выпукла $\Longrightarrow$ у неё единственный минимум $\Longrightarrow$ его можно найти как решение системы уравнений $q'(x^{k+1}) = 0$, которая по определению $q(x)$ эквивалентна следующей системе линейных уравнений

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

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 19 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 $x^k \longrightarrow x^*$ при $k \longrightarrow \infty$ и метод Ньютона имеет квадратичную скорость сходимости

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

ИДЕЯ МЕТОДА ВНЕШНИХ ШТРАФОВ

Свести решение задачи

$$f(x) \longrightarrow \min_{x \in Q} \qquad (1)$$

$$Q = \{x \in R^n \mid \varphi_i(x) \leq 0,\ i = 1, ..., m\} \qquad (2)$$

к решению последовательности задач минимизации

$$F_k(x) = f(x) + P_k(x) \longrightarrow \min_{x \in R^n},\ \ k = 1, 2, \ldots \qquad (26)$$

где $P_k(x)$ — штрафная функция множества Q.

МЕТОД ВНЕШНИХ ШТРАФОВ

Определение. Функция $P_k(x)$ называется штрафной функцией множества Q, если $P_k(x) \geq 0$ для любых $k = 1, 2, \ldots,\ x \in R^n$ и

$$\lim_{k\to\infty} P_k(x) = \begin{cases} 0, & \text{если } x \in Q \\ +\infty, & \text{если } x \notin Q, \end{cases}$$

где $P_k(x) = kH(x)$, k – коэффициент штрафа,

МЕТОД ВНЕШНИХ ШТРАФОВ

$$H(x) = \sum_{i=1}^{m} [\varphi_i(x)]_+^2 \text{ или } H(x) = \sum_{i=1}^{m} [\varphi_i(x)]_+$$

$([a]_+ = \max(0, a))$.

МЕТОД ВНЕШНИХ ШТРАФОВ

Пусть $\boldsymbol{x}^l = \boldsymbol{x}(\boldsymbol{k}_l)$ – оптимальное решение задачи без ограничений (26) на шаге $\boldsymbol{l}$.

Шаг $\boldsymbol{l+1}$. Найти решение задачи (26) с коэффициентом штрафа $\boldsymbol{k}_{l+1} > \boldsymbol{k}_l$. Если $\boldsymbol{H}(\boldsymbol{x}^{l+1}) \leq \boldsymbol{\varepsilon}$, то алгоритм завершает работу, иначе перейти на следующий шаг.

МЕТОД ВНЕШНИХ ШТРАФОВ

Соглашения:

1. функция $\boldsymbol{H : R^n \longrightarrow R}$ непрерывна и $\boldsymbol{H(x) \geq 0, \forall x \in R^n}$,

2. $\boldsymbol{H(x) = 0 \Longleftrightarrow x \in Q = \{x | \varphi_i(x) \leq 0, i = 1, \ldots, m\}}$

3. $\boldsymbol{f}$ — непрерывная функция, а множество $\boldsymbol{Q}$ замкнуто

МЕТОД ВНЕШНИХ ШТРАФОВ

Теорема 20. Пусть выполняется одно из двух условий

(a). $f(x_k) \longrightarrow +\infty$ для любой последовательности $\{x_k\} \in Q$, $\|x_k\| \longrightarrow +\infty$ или

(b). Q ограничено и $H(x_k) \longrightarrow +\infty$ для любой последовательности $\{x_k\}$, $\|x_k\| \longrightarrow +\infty$

Тогда 1). последовательность x^l имеет хотя бы одну предельную точку и любая предельная точка этой последовательности есть оптимальное решение задачи; 2). $H(x^l) \longrightarrow 0$

Доказательство. Покажем, что $\exists$ оптимальное решение задачи. (a): Если Q ограничено, то всё очевидно. Иначе $\exists$ последовательность $\{x_k\} \in Q$: $\|x_k\| \longrightarrow +\infty$. Тогда $f(x_k) \longrightarrow +\infty$. Множество Лебега $M(x_0) = \{x | x \in Q, f(x) \leq f(x_0)\}$ – ограничено и замкнуто $\Longrightarrow \exists$ оптимальное решение x^*.

В случае (b) утверждение очевидно.

МЕТОД ВНЕШНИХ ШТРАФОВ

Пусть $-\boldsymbol{k_1 < k_2 < \cdots < k_l < k_{l+1} < \ldots}$ — коэффициенты штрафа (используемые алгоритмом) Имеем

$$\boldsymbol{F_{l+1}(x^{l+1}) = f(x^{l+1}) + k_{l+1}H(x^{l+1}) >}$$

$$\boldsymbol{> f(x^{l+1}) + k_l H(x^{l+1}) \geq f(x^l) + k_l H(x^l) =}$$

$\boldsymbol{= F_l(x^l)}$, т.е.

$$\boldsymbol{F_{l+1}(x^{l+1}) > F_l(x^l)} \text{ для } \boldsymbol{\forall l} \qquad \mathbf{(27)}$$

МЕТОД ВНЕШНИХ ШТРАФОВ

По определению $\boldsymbol{x}^{l}$ и $\boldsymbol{x}^{l+1}$ имеем

$$\boldsymbol{f}(\boldsymbol{x}^{l}) + \boldsymbol{k}_{l}\boldsymbol{H}(\boldsymbol{x}^{l}) \le \boldsymbol{f}(\boldsymbol{x}^{l+1}) + \boldsymbol{k}_{l}\boldsymbol{H}(\boldsymbol{x}^{l+1})$$

$$\boldsymbol{f}(\boldsymbol{x}^{l+1}) + \boldsymbol{k}_{l+1}\boldsymbol{H}(\boldsymbol{x}^{l+1}) \le \boldsymbol{f}(\boldsymbol{x}^{l}) + \boldsymbol{k}_{l+1}\boldsymbol{H}(\boldsymbol{x}^{l})$$

Суммируем, приводим подобные, получим

$$(\boldsymbol{k}_{l+1} - \boldsymbol{k}_{l})\boldsymbol{H}(\boldsymbol{x}^{l+1}) \le (\boldsymbol{k}_{l+1} - \boldsymbol{k}_{l})\boldsymbol{H}(\boldsymbol{x}^{l})$$

$$\Longrightarrow$$

МЕТОД ВНЕШНИХ ШТРАФОВ

$$\boldsymbol{H}(\boldsymbol{x}^{l+1}) \leq \boldsymbol{H}(\boldsymbol{x}^{l}) \text{ для } \forall \boldsymbol{l} \qquad \textbf{(28)}$$

С другой стороны для $\forall \boldsymbol{l}$ имеем

$$\boldsymbol{f}(\boldsymbol{x}^{l}) \leq \boldsymbol{f}(\boldsymbol{x}^{l}) + \boldsymbol{k_l}\boldsymbol{H}(\boldsymbol{x}^{l}) \leq \boldsymbol{f}(\boldsymbol{x}^{*}) + \boldsymbol{k_l}\boldsymbol{H}(\boldsymbol{x}^{*}),$$

следовательно

$$\boldsymbol{f}(\boldsymbol{x}^{l}) \leq \boldsymbol{F_l}(\boldsymbol{x}^{l}) \leq \boldsymbol{f}(\boldsymbol{x}^{*}) \text{ для } \forall \boldsymbol{l} \qquad \textbf{(29)}$$

Условия (28) и (29) гарантируют ограниченность последовательности $\{\boldsymbol{x}^{l}\}$. Пусть $\widehat{\boldsymbol{x}}$ ее предел.

МЕТОД ВНЕШНИХ ШТРАФОВ

Тогда $\lim_{l\to+\infty} f(x^l) = f(\widehat{x})$ и $f(\widehat{x}) \le f(x^*)$. (27), (29) $\Longrightarrow$ последовательность $\{F_l(x^l)\}$ имеет предел $F^* \le f(x^*) \Longrightarrow$

$$f(x^l) + k_l H(x^l) \longrightarrow F^* \text{ при } l \to +\infty$$

$$\Longrightarrow \lim_{l\to+\infty} k_l H(x^l) = F^* - f(\widehat{x}) \Longrightarrow$$

$H(\widehat{x}) = 0$, т.е. $\widehat{x} \in Q$, но тогда $f(x^*) \le f(\widehat{x})$ и $f(\widehat{x}) = f(x^*) \Longrightarrow \widehat{x}$ оптимальное решение. $\blacksquare$

МЕТОД ВНУТРЕННИХ ШТРАФОВ ИЛИ МЕТОД БАРЬЕРНЫХ ФУНКЦИЙ

Проблема метода внешних штрафов: приближения

$$x^1 = x^1(k_1), x^2 = x^2(k_2), \dots, x^l = x^l(k_l), \dots$$

не являются допустимыми решениями задачи $\Longrightarrow$ попробуем аппроксимировать оптимум изнутри.

Идея метода, как и ранее, заключается в сведении исходной задачи (1)–(2) к последовательности задач минимизации следующего вида:

$$F_k(x) = f(x) + a_k B(x) \to \min_{x \in R^n}, \quad k = 1, 2, \ldots \tag{30}$$

где $B(x)$ — барьерная функция, $a_k > 0$ — барьерный коэффициент, $a_k \to 0$ при $k \to \infty$.

Определение. Функция $B(x)$ называется барьерной функцией для множества Q, если $B(x)$ определена, конечна и неотрицательна во всех точках из Int$Q$ и

$$\lim_{x \to \partial Q} B(x) = +\infty.$$

Соглашение: $B(x) = +\infty$, для $x \in \partial Q$(граница множества).

Примеры барьерных функций:

$$-\sum_{i=1}^{m}\varphi_i(x)^{-1},\ \sum_{i=1}^{m}|\varphi_i(x)|^{-1},\ \sum_{i=1}^{m}|\varphi_i(x)|^{-2}.$$

Соглашение: Задачу (30) решаем методом градиентов: $x^{k+1} = x^k - \alpha_k f'(x_k), \alpha_k \geq 0$.

Если $x^k \in \mathrm{Int}Q$, то при достаточно малом α_k $x^{k+1} \in \mathrm{Int}Q$.

Таким образом, если $x^0 \in \mathrm{Int}Q$, то из определения следует, что все приближения x^k – допустимые решения (1)-(2)

ИДЕЯ МЕТОДА

Пусть $\boldsymbol{x}^{\boldsymbol{k}}$ — решение задачи (30) на шаге $\boldsymbol{k}$ и $\boldsymbol{x}^{\boldsymbol{k}} \in \mathrm{Int}\boldsymbol{Q}$.

Итерация $(\boldsymbol{k}+\mathbf{1})$. Находим решение задачи (30) со значением барьерного коэффициента $\boldsymbol{a}_{\boldsymbol{k}+\mathbf{1}} < \boldsymbol{a}_{\boldsymbol{k}}$. Если $\boldsymbol{a}_{\boldsymbol{k}+\mathbf{1}}\boldsymbol{B}(\boldsymbol{x}^{\boldsymbol{k}+\mathbf{1}}) \leq \boldsymbol{\varepsilon}$, то алгоритм завершает работу. Иначе начинаем новую итерацию.

Соглашения:

1.$\boldsymbol{Q}$ замкнуто;

2.$\mathrm{Int}\boldsymbol{Q} \neq \emptyset$;

3. любая точка $\boldsymbol{x} \in \boldsymbol{Q}$ есть предел последовательности точек из внутренности $\boldsymbol{Q}$;

4. $\boldsymbol{f}$ — непрерывная функция на всем $\boldsymbol{R}^{\boldsymbol{n}}$;

5. функция $\boldsymbol{B}(\boldsymbol{x})$ непрерывна на множестве $\mathrm{Int}\boldsymbol{Q}$.

О СХОДИМОСТИ МЕТОДА БАРЬЕРНЫХ ФУНКЦИЙ

Теорема 21. Пусть выполняется одно из двух условий

(a). $f(x_k) \longrightarrow +\infty$ для любой последовательности $\{x_k\} \in Q$, $\|x_k\| \longrightarrow +\infty$ или

(b). Q ограничено.

Тогда 1) последовательность x^k имеет хотя бы одну предельную точку и любая предельная точка этой последовательности есть оптимальное решение задачи; 2) $a_k B(x^k) \longrightarrow 0$

Доказательство. В задаче (1)-(2) $\exists$ оптимальное решение $x^* \in Q$ (По теореме Вейерштрасса).

Пусть $x^1, x^2, \ldots, x^k, \ldots$ — приближения, а $a_1, a_2, \ldots, a_k, \ldots$ – соответствующие барьерные коэффициенты.

$\forall\, k$ имеем

$$f(x^*) \le f(x^k) \le f(x^k) + a_k B(x^k) = F_k(x^k) \tag{31}$$

В силу принятых соглашений $\forall\, \varepsilon > 0$ найдется $\widetilde{x}_\varepsilon \in \mathrm{Int}Q$: $f(\widetilde{x}_\varepsilon) \le f(x^*) + \varepsilon$

$$\Longrightarrow f(x^*) + \varepsilon +$$

$$+ a_k B(\widetilde{x}_\varepsilon) \geq f(\widetilde{x}_\varepsilon) + a_k B(\widetilde{x}_\varepsilon) \geq F_k(x^k) \tag{32}$$

Из определения величин x^k, x^{k+1} и условия $a_k > a_{k+1}$ имеем

$$F_k(x^k) = f(x^k) + a_k B(x^k) > f(x^k) +$$

$$+ a_{k+1} B(x^k) \geq f(x^{k+1}) + a_{k+1} B(x^{k+1}) =$$

$$= F_{k+1}(x^{k+1})$$

Т.к. $\{F_k(x^k)\}_{k \in N}$ монотонно убывает и ограничена снизу (31), то существует $\lim\limits_{k \to \infty} F_k(x^k)$.

Т.к. $\lim\limits_{k\to\infty} a_k B(\widetilde{x}_\varepsilon) = 0,$ то из (32) имеем

$$\lim_{k\to\infty} F_k(x^k) \le f(x^*) + \varepsilon.$$

Т.к. это верно $\forall\ \varepsilon > 0$, то из (31) имеем

$$\lim_{k\to\infty} F_k(x^k) = f(x^*),$$

$$\lim_{k\to\infty} f_k(x^k) = f(x^*),$$

$$\lim_{k\to\infty} a_k B(x^k) = 0.$$

Последовательность x^k – ограничена (из условий (a), (b) теоремы).

Без ограничения общности считаем $x^k \longrightarrow \widehat{x}$.

Из непрерывности $f \Longrightarrow f(\widehat{x}) = f(x^*)$.

Но Q – замкнуто и $\forall\ k\ x^k \in Q \Longrightarrow \widehat{x} \in Q$.

Следовательно, любая предельная точка последовательности приближений x^k является оптимальным решением задачи. ■

МЕТОД ПОКООРДИНАТНОГО СПУСКА

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

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

МЕТОД ПОКООРДИНАТНОГО СПУСКА

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

Пусть $e_i = (0, \ldots, 0, 1, 0, \ldots, 0)$ – i-ый единичн координатный вектор,

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

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

$\alpha_t > 0$ – текущая длина шага,

$\lambda, 0 < \lambda < 1$ – фиксированное число.

МЕТОД ПОКООРДИНАТНОГО СПУСКА

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

Опишем итерацию с номером t, где

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

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.