ЛЕКЦИЯ № 10

Методы штрафов

1. Метод внешних штрафов
2. Метод внутренних штрафов

Численные методы НЛП

3. Методы спуска и градиентные методы
4. Метод Ньютона
5. Метод покоординатного спуска

МЕТОД ШТРАФОВ = (ДУАЛИЗАЦИЯ / СПУСК)

$$\min f(x)$$

при условии, что

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

Рассмотрим нелинейную функцию Лагранжа

$$L(x, \lambda) = f(x) + \sum_{i=1}^{m} \lambda_i(\varphi_i(x)),$$

где для каждого i выполняется $\lambda_i(\varphi_i(x)) = 0$, если $\varphi_i(x) \leq 0$ и $\lambda_i(\varphi_i(x)) = +\infty$ иначе. Очевидно, что исходная задача эквивалентна задаче

$$\min_x L(x, \lambda)$$

ИДЕЯ МЕТОДА ВНЕШНИХ ШТРАФОВ

Свести решение задачи

$$f(x) \to \min_{x \in Q} \tag{1}$$

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

к решению последовательности задач минимизации

$$F_k(x) = f(x) + P_k(x) \to \min_{x \in R^n},\ k = 1, 2, \dots \tag{26}$$

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

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

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

МЕТОД ВНЕШНИХ ШТРАФОВ

Пусть $x^l = x(k_l)$ – оптимальное решение задачи без ограничений (26) на шаге l.

Шаг $l+1$. Найти решение задачи (26) с коэффициентом штрафа $k_{l+1} > k_l$. Если $H(x^{l+1}) \leq \varepsilon$, 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

Теорема 17. Пусть выполняется одно из двух условий:

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

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

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

МЕТОД ВНЕШНИХ ШТРАФОВ

Задача. Показать, что в случаях a) и b) существует оптимальное решение задачи.

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.

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

МЕТОД ВНЕШНИХ ШТРАФОВ

Очевидно, что для $\forall l$ имеем

$$f(x^l) \le f(x^l) + k_l H(x^l) \le f(x^*) + k_l H(x^*).$$

Таким образом, аппроксимация координирующей задачи происходит снизу.

МЕТОД ВНУТРЕННИХ ШТРАФОВ ИЛИ МЕТОД БАРЬЕРНЫХ ФУНКЦИЙ

Проблема метода внешних штрафов: приближения

$$\boldsymbol{x^1 = x^1(k_1), x^2 = x^2(k_2), \ldots, x^l = x^l(k_l), \ldots}$$

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

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

Соглашения:

1.Q замкнуто;

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

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

4. f — непрерывная функция на всем R^n;

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

О СХОДИМОСТИ МЕТОДА БАРЬЕРНЫХ ФУНКЦИЙ

Теорема 18. 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) $a_k B(x^k) \longrightarrow 0$

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

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

$\forall$ $\boldsymbol{k}$ имеем

$$\boldsymbol{f(x^*) \leq f(x^k) \leq f(x^k) + a_k B(x^k) = F_k(x^k)}$$

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

$$\boldsymbol{F_k(x^k) = f(x^k) + a_k B(x^k) > f(x^k) +}$$

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

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

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 $k = 0, 1, \ldots$

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

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.

Численные методы НЛП

Все итерационные процессы, в которых направление движения $\boldsymbol{p^k}$ на каждом шаге выбирается из конуса $\boldsymbol{K_d(x^k)}$, называются методами спуска (подъёма). Если $\boldsymbol{p^k = -f'(x_k)}$ ($\boldsymbol{p^k = f'(x_k)}$,) $\forall \boldsymbol{k}$, то такие методы называются градиентными методами:

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

Методы отличаются способами выбора длины шага $\boldsymbol{\alpha_k}$.

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 (подъёма):

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

Теорема 19 (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

$$\boldsymbol{f(x+y) \geq f(x) + \langle f'(x), y \rangle + l\|y\|^2/2.} \tag{23}$$

Лемма 12. Если функция $\boldsymbol{f}$ является сильно выпуклой (с константой $\boldsymbol{l > 0}$), то она имеет глобальный минимум на $\boldsymbol{R^n}$.

Задача. Пусть функция $\boldsymbol{f}$ является сильно выпуклой (с константой $\boldsymbol{l > 0}$). Доказать, что:

1. Она не может быть константой.
2. Она имеет единственный глобальный минимум.

Теорема 20 (Вторая теорема сходимости) Пусть функция f дифференцируема в R^n, является сильно выпуклой, 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 $x^k \to x^*$ при $k \to \infty$ и $\|x^k - x^*\| \leq Cq^k$, $0 \leq q < 1$.

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 21 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

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

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

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

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 $x^t \in R^n$ – 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. $\boldsymbol{k}$-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

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

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: $x^t = x^{t-1}$,

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

МЕТОД ПОКООРДИНАТНОГО СПУСКА

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

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

для всех $\boldsymbol{i = 1, \ldots, n}$.

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.