ЛЕКЦИЯ № 12

1. Метод ветвей и границ

2. Численные методы нелинейного программирования

3. Градиентные методы

Метод ветвей и границ

Рассмотрим произвольный гиперкуб $\boldsymbol{\Gamma[x^c, h]}$ с центром $\boldsymbol{x^c}$ и длиной стороны $\boldsymbol{h}$:

$$\boldsymbol{\Gamma[x^c, h] = \left\{x \in R^n : \|x - x^c\|_\infty \leq \frac{h}{2}\right\}.}$$

Из доказательства теоремы 18 следует, что $\forall\ \boldsymbol{x} \in$ $\boldsymbol{\Gamma[x^c, h]}$:

$$\boldsymbol{f(x) = f(x) - f(x^c) + f(x^c) \geq -L|x - x^c| +}$$
$$\boldsymbol{+ f(x^c) \geq f(x^c) - Lh/2.}$$

Метод ветвей и границ

Естественно, тогда определить нижнюю границу на гиперкубе $\Gamma[x^c, h]$ равенством $H(\Gamma[x^c, h]) = f(x^c) - Lh/2$.

Функцию выбора наилучшего решения определим на г.-кубах со стороной h (играют роль атомарных множеств решений) не превосходящей $\frac{2\varepsilon}{L}$, положив

$$x(\Gamma[x^c, h]) = x^c.$$

Подмножества решений будем задавать в виде набора гиперкубов.

Метод ветвей и границ

На первом шаге имеем $t_1 = \Gamma[x^R, \Delta]$, где первый рекорд x^R – центр г.-куба со стороной $\Delta^R = \Delta$.

Пусть к очередному шагу имеется разбиение

$$t_1 = \Gamma[x^1, h_1], \ldots, t_L = \Gamma[x^L, h_L]$$

и рекорд x^R – центр г.-куба со стороной Δ^R.

Очередной шаг начинается с проверки гиперкуба с номером l. Он считается проверенным и отбрасывается, если выполняется одно из следующих условий:

Метод ветвей и границ

$$1)\ H(\Gamma[x^l, h_l]) \geq f(x^R),$$

2) сторона г.-куба h_l не превосходит величины $\frac{2\varepsilon}{L}$.

При этом, если реализуется второй случай и

$$f(x^l) < f(x^R),$$

то устанавливается новое значение рекорда $x^R = x^l$ и величины $\Delta^R = h_l$.

Дополнительное правило:

Случай 1. (Текущий рекорд хуже)

Если $f(x^l) < f(x^R),$ то пересчитываем рекорд и среди оставшихся гиперкубов разбиения отбрасываем те, которые содержаться в г.-кубе

$$\Gamma[x^R, \frac{2(f(x^R) - f(x^l))}{L}].$$

По определению нижней границы имеем $\forall x \in \Gamma[x^c, h]$ неравенство $f(x) \geq H(\Gamma[x^c, h]) = f(x^c) - Lh/2.$

Поэтому для любой точки $\boldsymbol{x}$ из данного г.-куба имеем

$$\boldsymbol{f(x) \geq H(\Gamma[x^R, \frac{2(f(x^R) - f(x^l))}{L}]) =}$$

$$\boldsymbol{= f(x^R) - \frac{L}{2} \times \frac{2(f(x^R) - f(x^l))}{L} = f(x^l).}$$

Случай 2. (Текущий рекорд лучше)

Если $\boldsymbol{f(x^R) \leq f(x^l),}$ то среди оставшихся гиперкубов разбиения отбрасываем те, которые содержаться в г.-кубе

$$\boldsymbol{\Gamma[x^l, \frac{f(x^l) - f(x^R)}{L}].}$$

Т.к. для любой точки $\boldsymbol{x}$ из данного г.-куба имеем $\boldsymbol{f(x) \geq f(x^R)}$.

Если отброшены все элементы разбиения, то алгоритм заканчивает работу и $\boldsymbol{x^R}$ – требуемое решение.

Метод ветвей и границ

Если есть неотброшеные множества, то выбираем "перспективное"подмножество $\boldsymbol{\Gamma[x^l, h_l]}$. Функция ветвления $\boldsymbol{b(\cdot)}$ разбивает его на $\boldsymbol{2^n}$ одинаковых подкубов со стороной $\boldsymbol{\frac{h_l}{2}}$.

После этого начинается следующий шаг.

Конечность алгоритма следствие следующих фактов:

1. Гиперкуб $\boldsymbol{\Gamma[x^R, \Delta]}$ является компактным множеством.

2. На каждом шаге алгоритма хотя бы один элемент разбиения либо отбрасывается, либо разбивается на подмножества, каждое из которых состоит из меньшего числа атомарных множеств.

3. Атомарные множества всегда отбрасываются.

Замечание. Если в процессе работы алгоритма не происходит смены рекорда по правилу 2, то полученный рекорд – оптимальное решение задачи. В противном случае ε приближенное решение.

Замечание. Наиболее распространены следующие два правила ветвления:

"в ширину", когда ветвляться все или по очереди вершины одного уровня и затем переходят к следующему уровню;

"в глубину", когда ветвится лишь одна вершина уровня (обычно с лучшим значением рекорда, а если его нет, то нижней границы) до конца ветки.

Численные методы НЛП

Задача поиска безусловного минимума:

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

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

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

где $\boldsymbol{p^k}$ — направление спуска, $\boldsymbol{\alpha_k}$ — длина шага вдоль этого направления.

Метод ветвей и границ

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.

$$\|x^k - x^*\| \le q^k \|x^0 - x^*\|.$$

Скорость сходимости сверхлинейна, если

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

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

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

Численные методы НЛП

Методы нулевого порядка $\Longleftrightarrow$ методы, использующие только значения самой целевой функции.

Методы первого порядка $\Longleftrightarrow$ методы, использующие помимо значений целевой функции и её производные.

и так далее.

Численные методы НЛП

Все итерационные процессы, в которых направление движения на каждом шаге совпадает с антиградиентом (градиентом) функции, 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: 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 19 (Первая теорема сходимости)

Пусть функция $f \in C(R^n)$, 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.

Доказательство. Воспользуемся формулой конечных приращений

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

прибавим и вычтем из правой части величину $\boldsymbol{\langle f'(x), y\rangle}$:

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

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

Подставим вместо x и y, соответственно, x^k и $-\alpha f'(x^k)$. Имеем

$$f(x^{k+1}) \le f(x^k) + \langle f'(x^k), -\alpha f'(x^k)\rangle +$$

$$+\int_0^1 |\langle f'(x^k - \tau\alpha f'(x^k)) - f'(x^k),$$

$$-\alpha f'(x^k)\rangle| d\tau \leq \ldots$$

Первое неравенство следствие появления модуля под интегралом. Следующее неравенство вытекает из неравенства Коши-Буняковского $|\langle a, b\rangle| \leq \|a\|\ \|b\|$

$$\dots \le f(x^k) - \alpha \left\|f'(x^k)\right\|^2 + + \int\limits_0^1 \|f'(x^k -$$

$$-\tau\alpha f'(x^k)) - f'(x^k)\|\|\alpha f'(x^k)\| d\tau \le \dots$$

из условия Липшица имеем

$$\le f(x^k) - \alpha \left\|f'(x^k)\right\|^2 +$$

$$+ \int\limits_0^1 L\|\tau\alpha f'(x^k)\| \, \|\alpha f'(x^k)\| d\tau =$$

упрощая, получим

$$= f(x^k) - \alpha \|f'(x^k)\|^2 + L\alpha^2 \|f'(x^k)\|^2 \int_0^1 \tau d\tau$$

$$= f(x^k) - \alpha(1 - L\alpha/2)\|f'(x^k)\|^2 =$$

$$= f(x^k) - \gamma\|f'(x^k)\|^2$$

$(\gamma = \alpha(1 - L\alpha/2)) \Longrightarrow$ и условия $\gamma > 0$ имеем

$$f(x^{k+1}) \le f(x^k).$$

Индукцией легко доказать $\forall s$

$$f(x^{s+1}) \le f(x^0) - \gamma \sum_{k=0}^{s} \|f'(x^k)\|^2.$$

Тогда, учитывая ограниченность функции $f(x) \ge f^* > -\infty$ получим

$$\sum_{k=0}^{s} \|f'(x^k)\|^2 \le (f(x^0) - f(x^{s+1}))/\gamma \le$$

$$\le (f(x^0) - f^*)/\gamma.$$

$$\Longrightarrow f'(x^k) \to 0 \text{ при } k \to \infty. \ \blacksquare$$