Градиентным методом называется любой метод спуска, в котором Δx_k является антиградиентом,

т.е. $x_{k+1} = x_k - t_k \nabla f(x_k)$ $\qquad t_k$ - размер шага

Если при этом t_k выбирается по правилу одномерной минимизации, то метод называется методом скорейшего спуска

Метод Ньютона:

применяется к задачам оптимизации для поиска стационарной точки, в которой

$$\nabla f(x) = 0.$$

$$x_{k+1} = x_k - (\nabla^2 f(x_k))^{-1} \nabla f(x_k)$$

Пусть $\{x_k\}$ - последовательность точек в $\mathbb{R}^n$, которая сходится к x^*.

Сходимость

- Линейная, если

$$\exists\, r \in (0,1): \forall \text{ дост. больш. } k \quad \frac{\|x_{k+1} - x^*\|}{\|x_k - x^*\|} \leq r$$

- Сверхлинейная, если $\quad \lim\limits_{k\to\infty} \frac{\|x_{k+1} - x^*\|}{\|x_k - x^*\|} = 0$

- Квадратичная, если $\exists$ константа $M > 0$: $\forall$ дост. больших k

$$\frac{\|x_{k+1} - x^*\|}{\|x_k - x^*\|^2} \leq M$$

Двойственность

Решается задача условной оптимизации.

$$\min_{x \in Q} f_0(x) \qquad Q = \{x \in \mathbb{R}^n \mid f_i(x) \le 0,\ i = \overline{1,m},\ h_j(x) = 0,\ j = \overline{1,q}\}$$

Функция Лагранжа:

$$L(x, \lambda, \nu) = f_0(x) + \sum_{i=1}^{m} \lambda_i \cdot f_i(x) + \sum_{j=1}^{q} \nu_j \cdot h_j(x)$$

Двойственная задача оптимизации:

$$\max_{\lambda \ge 0} g(\lambda, \nu) \quad \text{, где} \quad g(\lambda, \nu) = \inf_{x \in \mathbb{R}^n} L(x, \lambda, \nu)$$

Штрафные функции

Функция $p(x)$ наз-ся (внешним) штрафом для задачи оптимизации $\min_{x \in Q} f_0(x)$, если

$p(x) = 0$, если $x \in Q$

$p(x) > 0$, если $x \notin Q$

Пример: $p(x) = \sum_{i=1}^{m} \left(\max\{f_i(x), 0\}\right)^2 + \sum_{j=1}^{q} \left(h_j(x)\right)^2$

Метод оптимизации с использованием штрафных функций:

$$x_{k+1} = \arg\min_{x \in \mathbb{R}^n} f_0(x) + t_k p(x) \qquad t_k \to \infty$$

Аналитически сначала можно решить задачу на кажд. шаге для произв. t, а затем устремить t к бесконечности.

Функция $f: X \to \mathbb{R}$ наз-ся выпуклой, если $\forall x, y \in X$ и $\alpha \in [0, 1]$:

$$f(\alpha x + (1-\alpha)y) \leq \alpha f(x) + (1-\alpha) f(y)$$

6.09.2016.

$f^* = \min\limits_{x\in Q} f(x)$ $\qquad$ $x^* = \arg\min\limits_{x\in Q} f(x)$

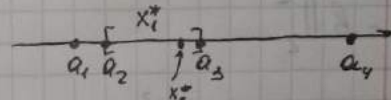
$Q \subset \mathbb{R}^n$

Задачи:

① найти центр набора точек

$a_1, a_2, \dots, a_m \in \mathbb{R}$

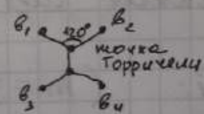
x – центр



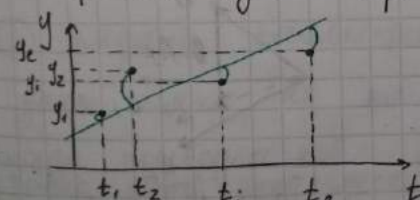
Сумма расстояний от x до всех точек минимальна

$$x_1^* = \arg\min_{x\in\mathbb{R}} \sum_{i=1}^{m} |x-a_i| \quad \text{медиана (центр)}$$

$$x_2^* = \arg\min_{x\in\mathbb{R}} \sum_{i=1}^{m} (x-a_i)^2$$



② даны точки на пл-ти (ℓ штук) провести ч/з них прямую $y(t)$



$f(t) = \alpha t + \beta$

Ошибка:

$(\alpha t_i + \beta - y_i)$

$$\alpha^*, \beta^* = \arg\min_{\alpha,\beta\in\mathbb{R}} \sum_{i=1}^{\ell} (\alpha t_i + \beta - y_i)^2$$

$$\alpha^*, \beta^* = \arg\min_{\alpha,\beta\in\mathbb{R}} \sum_{i=1}^{\ell} |\alpha t_i + \beta - y_i|$$

Линейная регрессия – вся совокупн. стат. данных опис-ся одной прямой.

$$\alpha t_1 + \beta - y_1 \overset{?}{=} 0$$
$$\vdots$$
$$\alpha t_i + \beta - y_i = 0$$
$$\vdots$$
$$\alpha t_m + \beta - y_m = 0$$

$$\begin{cases} \alpha t_1 + \beta = y_1 \\ \quad\vdots \\ \alpha t_i + \beta = y_i \\ \quad\vdots \\ \alpha t_m + \beta = y_m \end{cases}$$

$$A\begin{pmatrix}\alpha \\ \beta\end{pmatrix} = b$$

$b \in \mathbb{R}^m$

$\begin{pmatrix}\alpha \\ \beta\end{pmatrix} = x \in \mathbb{R}^2$

$Ax \overset{?}{=} b$

$$A = \begin{pmatrix} t_1 & 1 \\ \vdots & \vdots \\ t_i & 1 \\ \vdots & \vdots \\ t_m & 1 \end{pmatrix}_{m\times 2}$$

$$b = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix}$$

$$\|Ax - b\|_2^2$$

$$\underset{x \in \mathbb{R}^2}{\operatorname{argmin}} \|Ax - b\|_2^2$$

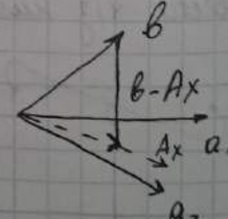
(частн. случай задачи о переопределённой системе ур-й)

Рассм. случай: $(m = 3)$

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \overset{?}{=} \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$$

$\underbrace{}_{\vec{a}_1}$ $\underbrace{}_{\vec{a}_2}$

$$\vec{a}_1 x_1 + \vec{a}_2 x_2 \overset{?}{=} \vec{b}$$



$\|b - Ax\|_2 \to \min$

нормальные уравнения

$$a_1^T(b - Ax) = 0$$
$$a_2^T(b - Ax) = 0 \Rightarrow A^T(b - Ax) = 0 \Rightarrow \boxed{A^T b = A^T A x}$$

$$\Rightarrow \underline{x = (A^T A)^{-1} A^T b} \neq A^{-1} b$$

Для того, чтобы у A^TA $\exists$ обр. дост., чтобы все столбцы м-цы A были лин. нез.

$\underline{(A^TA)^{-1}A^T}$ – „псевдо-обратная" матрица

$$A^+ = (A^TA)^{-1}A^T$$

$$A = \begin{pmatrix} t_1 & 1 \\ \vdots & \vdots \\ t_m & 1 \end{pmatrix} \qquad b = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix}$$

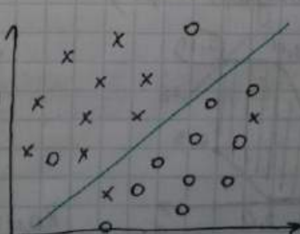
$$\begin{pmatrix} t_1 \dots t_m \\ 1 \dots 1 \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} = \begin{pmatrix} t_1 \dots t_m \\ 1 \dots 1 \end{pmatrix} \begin{pmatrix} t_1 & 1 \\ \vdots & \vdots \\ t_m & 1 \end{pmatrix} (x)$$

$$\begin{pmatrix} \sum_{i=1}^m t_i y_i \\ \sum_{i=1}^m y_i \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^m t_i^2 & \sum_{i=1}^m t_i \\ \sum_{i=1}^m t_i & m \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix}$$

③ задача классификации

даны точки на пл-ти - крестики и нолики

когда приходит нов. точка - дать предсказ-е нолик она или крестик (классифициров.)



„коридор максимальн. ширины"

13.09.2016.

Разрешимо ли ур-е $x^n+y^n=z^n$, $n\in\mathbb{N}$, $n>2$?

Формулир. как задачу оптимизац.:

$$(x^n+y^n-z^n)^2 \to \min_{\substack{x,y,z,n\in\mathbb{N}\\ n>2}}$$ (не реш-ся)

Пример задачи оптимизации, которая решается:

$$\max_{x\in\mathbb{R}^n} z = c^T x$$ стандартная форма задачи ЛП

т.ч. $Ax \le b$ ← покомпонентное сравнение

$x \geqslant 0$

каноническая форма задачи ЛП

$$\max_{x\in\mathbb{R}^n} z = c^T x$$

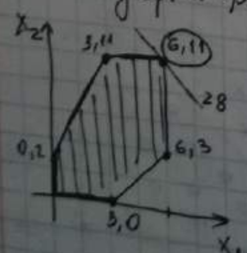
$Ax = b$ $\quad b\in\mathbb{R}^m$

$x\geqslant 0$

Числовой пример: $\max\ x_1+2x_2$

(в стандартн. ф.)

$$\left.\begin{aligned} -3x_1+x_2 &\le 2\\ x_2 &\le 11\\ x_1-x_2 &\le 3\\ x_1 &\le 6\\ x_1,x_2 &\geqslant 0\end{aligned}\right)$$ допустимое множество



Найти max:
а) линии уровня

1947 Данциг - способ реш-я задачи для произвольного количества переменных

Симплекс метод (улучш. способа)

Нет решения, если:

- система $Ax \leq b$ $(x \geq 0)$ несовместна
- ф-ция неогранич. на множ-ве

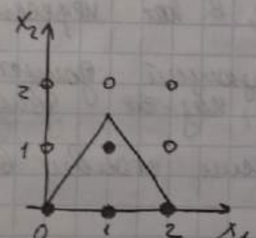
Если рассм. только $x \in \mathbb{Z}^n$, то:

$\max x_2$
$x \in \mathbb{Z}^2$

$3x_1 + 2x_2 \leq 6$

$-3x_1 + 2x_2 \leq 0$

$x_1, x_2 \geq 0$



(задача целочисленного линейного программир-я)

Способ решения задачи в стандартной форме

Словарь задачи ЛП

$$\begin{array}{c|l} x_B = & b - A_N x_N \\ \hdashline z = & c_0 + c^T \underbrace{x_N}_{\text{небазисные переменные}} \end{array} \qquad x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \geq 0$$

x_B - базисные переменные

Переход из стандартной формы в каноническую.
Вводятся фиктивные переменные

$$\langle a_1, x \rangle \leq b_1$$

$$\Downarrow$$

$$\langle a_1, x \rangle + x_{s_1} = b_1$$

Утв. Каждой угловой точке из допустим. мн-ва соответствует какой-н. словарь.

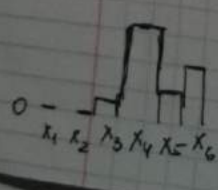
Словарь соотв. точке, в кот. пересек-ся два ограничения.

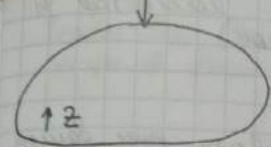
Словарь, соответствующий допустимой точке (попадающей в допустимое мн-во), наз-ся допустимым словарем.

Словарь допустим, если столбец $\vec{b}$ состоит из неотриц чисел.

Можно организовать переход от одного допустим. словаря к другому допустим. словарю.

Пример





до тех пока не дойдём до сост-я, когда увелич. z (целевую ф-ю) уже нельзя

$x_3 = 2 + 3x_1 - x_2$ ok — увелич. x_1 — $x_1 \le +\infty$

$x_4 = 11 + 0 \cdot x_1 - x_2$ ok — $x_1 \le +\infty$

$x_5 = 3 - x_1 + x_2$:c (мешает увелич. x_1) $\Rightarrow x_1 \le 3$

$x_6 = 6 - x_1$ — $x_1 \le 6$

$z = 0 + x_1 + 2x_2$

когда $x_5 = 0$ – в небазисн.

$x_1 \to$ в базисн.

Критерий остановки:

1. Выбрать входящую в базис переменную. ← та перем., кот. при своём увелич. наиб. увеличит знач. целевой ф-ции
2. Выбрать выходящую из базиса перем-ю. (та перем., чьё ограничение больше всего ограничивает рост входящей)
3. Провести алгебраич. преобраз-я → новая симплекс таблица (соотв. новому базису)

Целевая ф-я выраж-ся с отриц. знаками при всех перем-х → достигнута оптимальн. точка

(Старт из любого допустимого словаря)

1. корректность (если реш-е есть, то оно найдётся и за конечное число шагов)

$B = \{1,2,3,4\}$ ~~базис~~

1) Каждому базису соответствует словарь, если базисная м-це ~~A~~ A_B' невырожд.

те столбцы м-цы A, для кот. $i \in B$.

2) • Если задача ЛП совместна (имеет непустое допустимое множество), то у неё есть по крайней мере одна базисная допустимая точка.

• Если задача ЛП имеет реш-е, то по крайней мере одно из них - это базисная допустимая точка

• Если задача совместна и огранич., то у неё есть оптимальное решение.

(⇒ оптимальн. реш. следует искать в допустим. базисной точке)

<u>Док-во (1):</u> $(B = \{1,2,3,4\})$

$$Ax = A_B x_B + A_N x_N$$

$$A_B x_B + A_N x_N = b \qquad z = c^T x = c_B^T x_B + c_N^T x_N$$

$$x_B = \underbrace{A_B^{-1} b} - A_B^{-1} A_N x_N$$

$$z = c_B^T A_B^{-1} b - c_B^T A_B^{-1} A_N x_N + c_N^T x_N$$

$$z = \underbrace{c_B^T A_B^{-1} b} - (c_B^T A_B^{-1} A_N - c_N^T) x_N$$

A_B вырожд. ⇒ в ~~исх. м-це~~ A некот столбцы линейно завис.

(слишком много информации)

v строки линейно завис

Чтобы A_B была невырожд. дост. и необх., чтобы строки м-цы A были л.н.з.

$$\hat{A}x \le \hat{b} \in \mathbb{R}^m$$
$$\max z = \hat{c}^T x$$
$$x \geq 0$$
$$\mathbb{R}^n$$

20.09.2016.

$$x_B = b + A x_N \qquad x \geq 0$$
$$z = c_0 + c^T x_N$$

$$x_{B_1} = b_1 + a_{11} x_{N_1} + \ldots + a_{1j} x_{N_j} + \ldots \qquad x_j \le \frac{b_1}{-a_{1j}}$$
$$x_{B_2} = b_2 + \ldots$$
$$\vdots$$
$$x_{B_i} = b_i + a_{i1} x_{N_1} + \ldots + a_{ij} x_{N_j} + \ldots \qquad x_j \le \frac{b_i}{-a_{ij}}$$
$$\vdots$$
$$x_{B_m} = b_m + a_{m1} x_{N_1} + \ldots + a_{mj} x_{N_j} + \ldots + a_{mn} x_{N_n}$$
$$z = c_0 + c_1 x_{N_1} + \ldots + c_j x_{N_j} + \ldots + c_n x_{N_n}$$

1. Найти такое $c_j > 0 \Rightarrow x_j$ - входящ. перем.

   Если такого c_j нет $\Rightarrow$ задача решена

2. $a_{ij} > 0$

   если таких нет $\Rightarrow$ задача неогранич.

   $$i = \underset{\substack{i' \in \{1,\ldots,m\} \\ a_{i'j} < 0}}{\operatorname{argmin}} \frac{b_i}{-a_{ij}}$$

x_{B_i} – покидающ. перем.

3. $x_{N_j} = -\frac{b_i}{a_{ij}} - \frac{a_{i1}}{a_{ij}} x_{N_1} - \dots + \frac{x_{B_i}}{a_{ij}}$

x_{B_i} и x_{N_j} мен-ся местами

Утв Если исх. симплекс-таблица допустимая (все $b_i \geq 0$), то новая СТ тоже будет допустимой.

Угловых точек огранич. кол-во ⇒ кол-во СТ огранич. (кажд. СТ соотв. угловой точке)

$T_1 \to T_2 \to T_3 \to \dots \to T_k$

Жадная эвристика: выбирать такой x_{N_j}, кот. на одном шаге приведёт к макс. увелич. целевой функции.

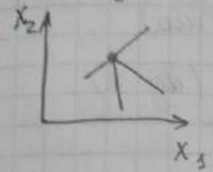
Пример, когда Ж.ал. зацикливается: 1974 Avis, Chvabal

Можно просто выбир. наиб. c_j, но тоже возм. зацикливание.

Зацикливание – если целевая ф-ция не растёт.

Если среди b_i $\exists$ 0, то огр. на рёбр $x = 0$
⇒ x не увелич-ся и z не увел. при перех. к след. симпл. таблице.

Вырожденная СТ (словарь) – то, в кот. $\exists\ b_i = 0$



Можно выбрать различн. огранич-я для одной базисной точки.

Правило Бланда:
- все перем. в задаче нумер-ся индекс
- всегда, когда есть выбор, выбир. ту, у кот. наим. индекс (в лексикографич. порядке)

=> не будет циклов

Пример $\max z = x_1 + 2x_2$

$$\begin{aligned} -2x_1 + x_2 &\le -2 \\ x_2 &\le 4 \\ x_1 - 2x_2 &\le -2 \\ x_1 &\le 4 \\ x_1, x_2 &\ge 0 \end{aligned} \quad \xrightarrow[\text{перем.}]{\text{добавл.}} \quad \begin{aligned} -2x_1 + x_2 + x_3 &= -2 \\ x_2 + x_4 &= 4 \\ x_1 - 2x_2 + x_5 &= -2 \\ x_1 + x_6 &= 4 \\ x_i &\ge 0 \end{aligned}$$

$$x_3 = \underline{-2} + 2x_1 - x_2 + x_0$$

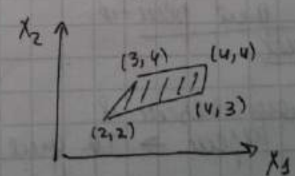
$$x_4 = 4 + 0 \cdot x_2 - x_2 + x_0$$

пример недопустим. симплекс-таблицы

$$x_5 = \underline{-2} - x_1 + 2x_2 + x_0$$

$$x_6 = 4 - x_1 + 0 \cdot x_2 + x_0$$

$$\max z = 0 + x_1 + 2x_2$$



Ограничения избыточны

Нужно найти углов. точку допустим. множ-ва

3 Внешняя (дополнительная) задача ЛП

(Шаг инициализации симплекс-алгоритма)

Берём только ур-я, дающие ограничения (без z).

К кажд. из них справа добав. $x_0 \geq 0$

Получен. область непуста (возьмём $x_0 = -b_1; x_1, x_2 = 0$, бет. выразится)

базисн.

Теперь нужно найти соотв. базисн. угл. точку.

Утв

Если исх. задача совместна, то в доп. области можно найти такой базис, в кот. $x_0 = 0$

Обратно: если найд-ся базисн. допуст. точка, для кот. $x_0 = 0 \Rightarrow$ исх. задача совместна.

Если исх. зад. несовм. $\Rightarrow \forall$ базисе $x_0 > 0$

$\boxed{\max z' = -x_0}$

0

z'

$-z$

$z' = 0$

$\Rightarrow$ точка этого реш-я

$\Rightarrow$ все кроме x_0 - коорд-ты допустим. углов. точки исх. мн-ва

0-й шаг СМ для реш-я доп. задачи:

x_0 - входящ. перем-я

исходящ. перем. $\to$ та для кот b_i - сам. маленький

Если x_0 может покинуть базис, всегда выбир. именно его (для упрощ. работы задачи)

Выраж-е базисных через небазисн. подстав. в целев. ф-ю ⇒ допустим. зад., кот. можно решить симплекс-методом.

Всего шагов м.б. порядка 2^n, где n - кол-во перем. исх. задачи

На практике получ-ся порядка 3m шагов, где m - кол-во ограничен.

Алгоритм имеет полиномиальную оценку сложности в среднем.

СМ позволяет найти точное реш-е в задаче оптимизации.

Проблему чисел с плавающ. точкой можно обойти:
сначала выяснить, какой будет новый базис, а потом восстановить симплекс-таблицу по базису, причём только те знач-я, которые нужны.
по новым данным поймём, какой будет след. базис...

Revised Simplex

27.09.2016.

Время на шаг симплекс-алгоритма:

1. $O(n)$ - выбор входящ. перем.
2. $O(m)$ - выбор исходящ. перем.
3. $O(n\cdot m)$ - пересчитыв. всю м-цу

$\max c^T x$

$\hat{A}x + E_m x_s = \hat{b}$

$x \geq 0$

$x_s \geq 0$

Расход памяти: $O(n\cdot m)$

$$X_B = b + A X_N$$

$$Z = Z_0 + c X_N$$

Если исх. м-ца сильно разрежена - на перв. шаге её размером можно считать $O(m)$.

Ошибки округления накапливаются :с

Шаг:

1. Выбор входящей перем.
2. Выбор покидающей перем.
3. Используя данные исх. задачи и базис (номера перем.) - восстанов. след. симплекс таблицу.

$\max \hat{c}^T x$

$\tilde{A}x = b$

$$c^T x = c_B^T x_B + c_N^T x_N$$

$$A_B x_B + A_N x_N = \hat{b} \qquad A_B \in M_{m\times m}$$

не будет накопления ошибки

$$x_B = A_B^{-1}\hat{b} - A_B^{-1}A_N x_N$$

$$Z = c_B^T A_B^{-1}\hat{b} - \left(c_B^T A_B^{-1} A_N \cancel{x_N} - c_N^T\right) x_N$$

Нужны не все эл-ты симплекс-табл., а только те, кот. прин. участие в выборе.

$$b' = A_B^{-1}\hat{b}$$

$$c' = c_N^T - \overbrace{c_B^T A_B^{-1}}^{\pi} A_N \Rightarrow j\text{- индекс входящ. перем.}$$

$$a'_j = A_B^{-1} a_j \Rightarrow l\text{- индекс покидающей}$$

$\pi = c_B^T A_B^{-1}$ $\quad$ 1. $A_B^T \pi^T = c_B$ $\quad$ Системы ур-й, которые необх. решить

$$\Downarrow$$

$$c' = c_N^T - \pi A_N \Rightarrow j$$

$$\left.\begin{array}{l} 2.\ A_B a_j' = a_j \\ 3.\ A_B b' = b \end{array}\right\} \Rightarrow i$$

При переходе от одного базиса к другому м-ца A_B изм-ся только в одном столбце.

$$A_{B^k} \Rightarrow A_{B^{k+1}} \qquad A_{B^{k+1}} = A_{B^k} + \begin{array}{c} \left(\begin{array}{c|c|c} \begin{matrix}0\dots0\\ \vdots\ \ \vdots\\ 0\dots0\end{matrix} & a_j - a_i & \begin{matrix}0\dots0\\ \vdots\ \ \vdots\\ 0\dots0\end{matrix} \end{array}\right) \\ \text{i-я позиц.} \end{array}$$

$$A_{B^k}^{-1} \Rightarrow A_{B^{k+1}}^{-1}$$

м-ца ранга 1 $= u \cdot v^T = (a_j - a_i) e_i^T$

вектор $(0 \dots 010 \dots 0)$, i

! если у м-цы A есть обратн., то: $(\forall u, v)$

(упр.) $$(A + u \cdot v^T)^{-1} = A^{-1} + \frac{A^{-1} u v^T A^{-1}}{1 + v^T A^{-1} u}$$

$A \in M_{m\times m}$ u-разм. m v-разм. m

$\boxed{\text{ф-ла SMW}}$

$$(A_{B^{k+1}})^{-1} = \underbrace{(E_m - ?)}_{\mathcal{E}_k} A_{B^k}^{-1}$$

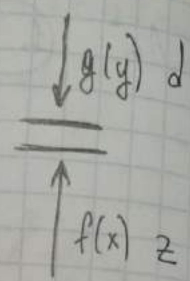
(разреж. м-ца)

(упр.) $\mathcal{E}_k \to \mathcal{E}_{k+1}$ (легко, но появляется ошибка округления)

Двойственная задача ЛП

$\max f(x)$ $\Rightarrow$ $\min g(y)$
$C(x) \leq d$ $\quad$ $P(y) \leq q$

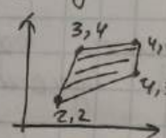
прямая задача (не обязательно линейная) — двойственная задача



Реш-е двойств. задачи даёт оценку исх. задаче

<u>Пример:</u> $\max\ x_1 + 2x_2$

$\cdot 2 \times (-2x_1 + x_2 \leq -2)$
$x_2 \leq 4$
$x_1 - 2x_2 \leq -2$
$\cdot 5 \times (x_1 \leq 4)$
$x_1, x_2 \geq 0$



$z(4,4) = 12$

$\cdot + \cdot \mid x_1 + 2x_2 \leq 16$ — огранич-е на знач. целевой ф-ции

$\mathrm{I} \cdot 2 + \mathrm{V} \cdot 1 \mid x_1 + 2x_2 \leq 12$

правильно подобрав набор коэффиц. можно найти самое жёсткое огранич-е

Так можно сдел. не всегда:

$z = -2x_1 + 3x_2 + x_3$
$(x_1 - x_2 + x_3 \leq 10) \times y_1$
$(2x_1 - 3x_2 - x_3 \leq 10) \times y_2$
$(6x_2 + 2x_3 \leq 10) \times y_3$
$x \geq 0 \rightarrow$ $\left.\begin{array}{l} (-x_1 \leq 0) \times y_4 \\ (-x_2 \leq 0) \times y_5 \\ (-x_3 \leq 0) \times y_6 \end{array}\right\} y_s$

$$\boxed{\begin{array}{l} \min\ d = b^T y \\ A^T y \geq c \\ y \geq 0 \end{array}}$$ двойств. задача

$\Downarrow$

$\max -b^T y$
$-A^T y \leq -c$
$y \geq 0$

$$A^T y - \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} y_s = c$$

$$A = \begin{pmatrix} 1 & -1 & 1 \\ 2 & -3 & -1 \\ 0 & 6 & 2 \end{pmatrix} \quad b = \begin{pmatrix} 10 \\ 10 \\ 10 \end{pmatrix} \quad c = \begin{pmatrix} -2 \\ 3 \\ 1 \end{pmatrix}$$

$$1y_1 + 2y_3 + 0 \cdot y_3 - y_{s_4} = -2$$

...

Прям.: $\max z = c^T x$

$$Ax \leq b$$
$$x \geq 0$$

• Теорема о слабой двойственности

$c^T x \leq b^T y$ ($\forall x, y$ - допустим. точк. исх. и двойств. зад. соотв-но)

• x - допустим. точка прямой задачи, $c^T x = b^T y$ для некот. y - допустим. т. двойств. з.

$\Rightarrow$ в этих точках дости. оптимум и в прям., и в двойств.

• x^* - оптим. реш-е прямой задачи. Тогда

1) у двойств. задачи есть реш-е

2) $\forall$ оптим. реш-е двойств. задачи удовлетворяет $b^T y^* = c^T x^*$

Теорема о сильной двойственности

4.10.2016.

прямая $x \in \mathbb{R}^n$	двойственная
$\max z = c^T x$	$\min q = b^T y \quad y \in \mathbb{R}^m$
$Ax \le b$	$A^T y \ge c$
$x \ge 0$	$y \ge 0$

$\max z = c^T x$	$\min q = b^T y$
$Ax + x_s = b$	$A^T y - y_s = c$
$x, x_s \ge 0$	$y, y_s \ge 0$

Таблица комплементарных переменных

$$\overbrace{\begin{pmatrix} x_1, & x_2, & \dots, & x_n, & \vert & x_{s_1}, & x_{s_2}, & \dots, & x_{s_m} \\ \updownarrow & \updownarrow & & \updownarrow & \vert & \updownarrow & \updownarrow & & \updownarrow \\ y_{s_1} & y_{s_2} & & y_{s_n} & \vert & y_1 & y_2 & & y_m \end{pmatrix}}^{n+m}$$

комплементарное соответствие

<u>Пример</u>

$$\max\ -x_1 - x_2$$
$$-2x_1 - x_2 \le 4$$
$$-2x_1 + 4x_2 \le -8$$
$$-x_1 + 3x_2 \le -7$$
$$x_1, x_2 \ge 0$$

$\iff$

$$\min\ 4y_1 - 8y_2 - 7y_3$$
$$-2y_1 - 2y_2 - y_3 \ge -1$$
$$-y_1 + 4y_2 + 3y_3 \ge -1$$
$$y_1, y_2, y_3 \ge 0$$

$\Downarrow$

$$\max\ -4y_1 + 8y_2 + 7y_3$$
$$2y_1 + 2y_2 + y_3 \le \underline{1}$$
$$y_1 - 4y_2 - 3y_3 \le \underline{1}$$
$$y_1, y_2, y_3 \ge 0$$

c - цены товаров, кот. вы продаёте

a_{ij} - расход i-го ресурса на j-й товар

b_i - общее ограничение на кол-во ресурсов (затраты на ресурс)

x - кол-во товаров

y_i - кол-во ресурса

q - расходы

$\max z$ - прибыль

Теорема 1 (о слабой двойственности)

x – допустим. точка прямой задачи
y – д.т. двойств. задачи
$\Rightarrow \quad z = c^T x \le b^T y = q$

▲ $$c^T x \le y^T A x = x^T A^T y \le b^T y$$
$c \le A^T y$
$c^T \le y^T A$

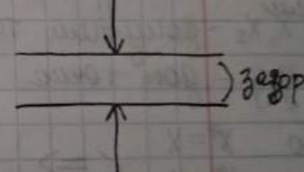
■

Следствие x, y
$$c^T x = b^T y$$
$\Rightarrow$ x – реш-е прямой задачи, y – реш-е двойственной

Теорема 2 (о сильной двойственности)

Если x^* – оптимальная допустим. точка прямой зад. (решение)
Тогда $\exists$ оптим. доп. т. двойств. задачи y^* такая, что $c^T x^* = b^T y^*$

▲



$$x_{s_1} = 4 + 2x_1 + x_2$$
$$x_{s_2} = -8 + 2x_1 - 4x_2$$
$$x_{s_3} = -7 + x_1 - 3x_2$$
$$z = 0 - x_1 - x_2$$

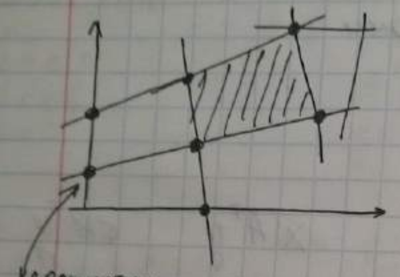
$$y_{s_1} = 1 - 2y_1 - 2y_2 - y_3$$
$$y_{s_2} = 1 - y_1 + 4y_2 + 3y_3$$
$$q = 0 - 4y_1 + 8y_2 + 7y_3$$

шаг на комплементарных переменных

недопустим. с.т.

$B = \{s_1, 1, s_3\}$ $\quad$ $B = \{2, s_1\}$

точка – n из $n+m$ огранич-й



Активное ограничение – достигается равенство.

→ новые с.т. будут точно так же соответствовать друг другу

Утв Такой шаг (на обеих зад.) симплекс алг. всегда сохраняет св-во инверсивной транспонированности

Финальное состояние

q будет. c – ⇒ трансп. инверс. → b c + → допустим внизу нет + ⇒ финальн.

$$-(-q)=z$$

■

Комплементарная эффективность

Если x, x_s – допустим. точка прямой задачи

y, y_s – доп. точка двойств. задачи

То $x^*=x$, $y^*=y$ оптим. реш. $\Longleftrightarrow$ $x_i y_{s_i}=0$ $i=\overline{1,n}$, $x_{s_j} y_j=0$ $j=\overline{1,m}$

▲

$x=0$

$C^T=y^TA \Rightarrow c=A^Ty \qquad c-A^Ty=y_s$

x - нуб. ⇒ x_s - нуб. $\quad x_s=b-Ax$

n нулей
m знач. $x > 0$

$A^T y + y_s = c$ ← система ур, кот. надо решить

$(n+m)$ неизв., n ур-й

x - оптимален ⇒ соотв. y_s - нули $\blacksquare$

11.10.2016.

$$\boxed{\begin{aligned} z &= 0 - x_1 - x_2 \\ x_{s_1} &= 6 \\ x_{s_2} &= -5 - Ax \\ x_{s_3} &= -2 \end{aligned}} \quad \longrightarrow \quad \begin{aligned} -d &= 0 - 6y_1 + 5y_2 + 2y_3 \\ y_{s_1} &= 1 \quad + A^T y \\ y_{s_2} &= 1 \end{aligned}$$

Все коэф. целевой ф-ции неположит. - двойственно допуст.

Прямо допустим. - все в b неотриц.

Прямо допуст. + двойств. допуст. = оптим. точка

(сначала выбир. ~~входящ~~ покидающую. Двойственный симплекс алгоритм

Можно также модифицировать целевую функцию, решить а потом выразить и дорешать

Можно ли воспольз-ся реш-ем задачи ЛП для того чтобы решить похожую задачу?

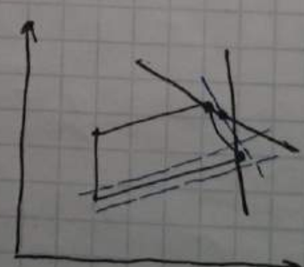
$\max c^T x$
$Ax \leq b$
$x \geq 0$

$$\begin{aligned} z &= 15 - x_1 - x_{s_2} \\ x_{s_1} &= 5 \\ x_2 &= 6 - \hat{A} x_N \\ x_{s_3} &= 3 \end{aligned} \Rightarrow \begin{aligned} -d &= 15 - 5y \\ y_{s_1} &= 1 + \hat{A}^T \\ y_2 &= 1 \end{aligned}$$

$z^* \quad x^* \quad x_s^* \quad y^* \quad y_s^*$

$z^*(c + \Delta c) = z^*(c) + ?$

$z^*(b + \Delta b) = z^*(b) + ?$



неактивные ограничения никак не влияют

$z^*(b+\varepsilon e_i) - z^*(b) = ?$

Рассм. только невырожд. случ.

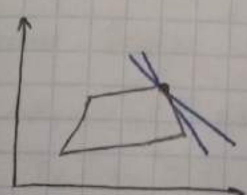
$$z^*(b+\varepsilon e_i) - z^*(b) = \max_{\substack{Ax \le b+\varepsilon e_i \\ x \ge 0}} c^T x - \max_{\substack{Ax \le b \\ x \ge 0}} c^T x =$$

Двойств. з.:

$$\min b^T y, \quad A^T y \ge c, \quad y \ge 0$$

$$= \min_{\substack{A^T y \ge c \\ y \ge 0}} (b+\varepsilon e_i)^T y - \min_{\substack{A^T y \ge c \\ y \ge 0}} b^T y \; \textcircled{=}$$

допустим. мн-ва совпадают



$b_i \ne 0$, y_i

Утв. Если прямая задача невырожд., то в двойств. зад. реш-е одно.

$\Rightarrow$ <u>y^* не изм. при дост. малом ε</u>

задача превр. в зад. варьиров.-я наклона линии уровня

$$\textcircled{=} \; (b+\varepsilon e_i)^T y^* - b^T y^* = \varepsilon e_i y^* = \varepsilon y_i^*$$

Что значит „достаточная малость ε"?

$$\boxed{\frac{\partial z^*}{\partial b_i}(b) = y_i^*}$$

$$z^*(b+\varepsilon e_i) = z^* b + y_i^* \varepsilon$$

анализ чувствительности

Можно выявлять критерии, дающие оценки на ε

Теневые цены - перем-е в двойств. задаче в экономической науке

ILP (Целочисленное линейное программирование)

$\max c^T x$
$Ax \leq b$
$x \geq 0$
$x \in \mathbb{Z}^n$

$\longrightarrow$

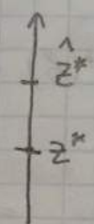
$\max c^T x$
$Ax \leq b$
$x \geq 0$
$x \in \mathbb{R}^n$

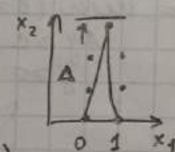
релаксированная задача ЦЛП

$z^* = c^T x^*$
решение ЦЛП

$\hat{z}^* = c^T \hat{x}^*$

$\hat{z}^* \geq z^*$





совместна и огранич. $\longrightarrow$ релаксир. совместна (но м.б. неогранич.)

может и не быть $\longleftarrow$ есть реш-е

бинарные перем-е (0 или 1)

$q_1 \quad q_2 \quad q_3 \quad q_4$

$$\left.\begin{array}{l} q_1 \vee q_3 \vee q_4 \\ q_1 \vee \bar{q}_2 \vee q_3 \\ \vdots \end{array}\right\}$$

$\exists$ ли комбинац. из q_1, q_2, q_3, q_4:
- все усл. истины одноврем.

3-SAT

Задача ЦЛП NP-полн., т.к. с её помощью можн. реш. 3-SAT ~~её можн. реш. с пом. 3-SAT~~

$0 \le x \le 1$

$x \in \mathbb{Z}$

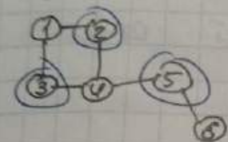
$x_1 + x_3 + x_4 \geq 1$

$x_1 + (1 - x_2) + x_3 \geq 1$

$\vdots$

Найти хотя бы какую-н. допустим. точку → реш-е 3-SAT

Задача о вершинном покрытии



Во всех рёбрах графа встреч. хотя бы одна из выбр. верш.

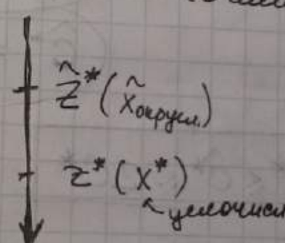
Какое верш. покрытие = ?

$\min x_1 + x_2 + x_3 + x_4 + x_5 + x_6$ $\quad \hat{z}^*$

x_i - участвует ли верш. в покрытии (0 или 1)

$x_1 + x_2 \geq 1$

$\vdots$



$z^*(x^*) \le \hat{z}(\hat{x}_{окр.}) \le 2z^*(x^{\sim})$

Задача о вершинном покрытии

18.10.2016.

$x_1 + x_2 + \ldots + x_n$ n вершин

$0 \le x_i \le 1$ 1 - есть в покрытии

$x \in \mathbb{Z}$ 0 - нет

релаксир. задача

$$\min x_1^I + x_2^I + \ldots + x_n^I \quad \begin{cases} x_i^I + x_j^I \ge 1 \quad \forall (i,j) \in E \\ 0 \le x_i^I \le 1 \\ x^I \in \mathbb{Z}^n \end{cases} \Rightarrow \min \sum_i x_i^R \quad \begin{cases} x_i^R + x_j^R \ge 1 \quad \forall (i,j) \in E \\ \ldots \\ 0 \le x_i^R \le 1 \\ x^R \in \mathbb{R}^n \end{cases}$$

получим дробный ответ

$$x^{R*} \longrightarrow \hat{x}^{R*}$$

$$x_i < \frac{1}{2} \longrightarrow \hat{x}_i := 0$$

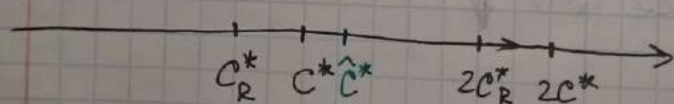
$$x_i \ge \frac{1}{2} \longrightarrow \hat{x}_i := 1$$

$$C^* = \sum_i x_i^{I*}$$

$$C_R^* = \sum_i x_i^R$$

$$\hat{C}^* = \sum_i \hat{x}_i^{R*}$$ ← покрытие, т.к. все $\hat{x}_i$ удовлетв. исходн. неравенствам

Утв. $\hat{C}^*$. $\boxed{C^* \le \hat{C}^* \le 2C^*}$

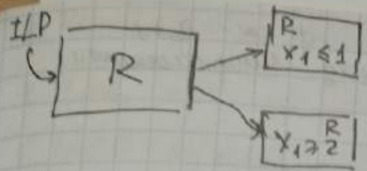


$C_R^* \le C^*$

$2C_R^* \le 2C^*$

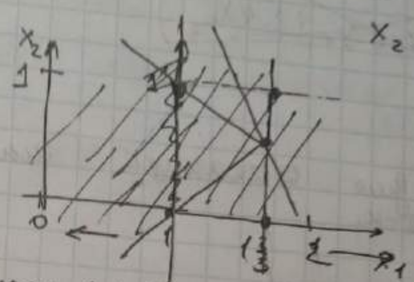
$\hat{C}^* \overset{?}{\le} 2C_R^*$ ← при округлении любой x_i увеличится не более чем в 2 раза

Исх. утв-е доказано.



$x_1 = 1\frac{2}{3}$

$x_2 = 1$



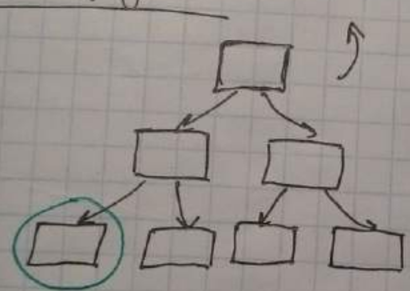
Целочисл. реш-е между 1 и 2 для x_1 искать бессмысл.

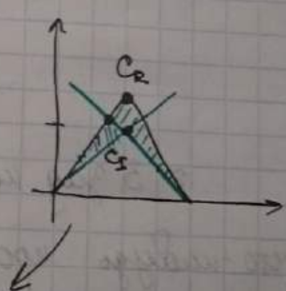
Две подзадачи: $x_1 \le 1$, $x_1 \ge 2$

Пытаемся сдел. переменные целыми на каждом шаге

(Стратегия организации перебора)

Часть подзадач можно отсекать (те, кот. оказ-ся заведомо хуже).





отсекающая гиперплоскость — выкидыв. из задачи область, в кот. точно нет целых чисел

$\cancel{x_1 = 1,2 + 1,6 x_{s_1} - 2,4 x_{s_2}}$

Поиск ф-лы для отсекающ. гиперплоскости

$\cancel{\boxed{1,2 = x_1 - 1,6 x_{s_1} + 2,4 x_{s_2}}}$

$\cancel{1 + 0,2 =}$

Формулу для отсекающей гиперплоскости надо найти где-н.

К/р. ≈ 3 зад. на к/р.

Док-ть что-нибудь про двойств-ть: ≈ 2 зад.

слаб. двойств-ть / комплементарн. эф-ть

Формулировк. ≈ 1 зад.

1. Опр. оптим. точки ← по крайней мере n нерав-в обращ. в рав-ва
мин. кол-во по крайней мере n из этих рав-в
задача – соотв. ли оптим. точка симплекс табл
критерий опт. точки ← сколько нулей в коорд. опт. т. (среди основн. эффект. перем.)

2. Док-во конечн-ти симплекс алг

3. Реконструкц. симпл. табл. по базису
4. Утв. про сильн. и слаб. двойств-ть
5. ЦЛП - оценка реш-я : связь целочисл. реш. с релаксир.

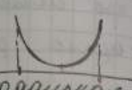
Формулир. в виде прост. задачи.

~~15.10.16~~.

1.11.2016

Задачи безусловной оптимизации

$$\min_{x\in\mathbb{R}^n} f(x)$$



Ф-я от одной переменной унимодальная, если она сначала убывает, а потом возрастает.

Методы тернарного поиска:
- дихотомии
- золотого сечения

Скорость сходимости

$$(r_{k+1}-l_{k+1})=\frac{(r_k-l_k)+\delta}{2}$$

$$\text{if } \left(f\left(\frac{l_k+r_k-\delta}{2}\right) < f\left(\frac{l_k+r_k+\delta}{2}\right)\right)$$

$$\lim_{k\to\infty}(r_{k+1}-l_{k+1})=\delta$$

$$r_{k+1}=\frac{l_k+r_k+\delta}{2}$$

$$l_{k+1}=l_k$$

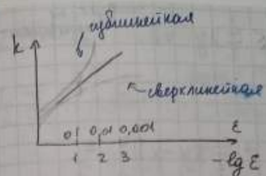
else ...

$$\frac{r_{k+1}-l_{k+1}-\delta}{r_k-l_k-\delta} =$$

- линейная
- сверхлинейная
- сублинейная
- ...

$$= \frac{\frac{r_k-l_k+\delta}{2}-\delta}{r_k-l_k-\delta}=\frac{1}{2} \Rightarrow$$ линейная сходимость

посл-ть типа: $x_k=\delta+\left(\frac{1}{2}\right)^k$



$$|x^*-x_k|\le\varepsilon$$

$$\left(\frac{1}{2}\right)^k\le\varepsilon$$

$$k\ge\log_{1/2}\varepsilon=-\frac{1}{\lg 2}\lg\varepsilon$$

$$k=\frac{1}{\lg 2}(-\lg\varepsilon)$$

точность (ε) увелич-ся на порядок
$\Rightarrow -\lg\varepsilon$ увелич-ся на 1

квадратичн. сх-ть — парабола (сверхлинейн)
...
сублинейная — м.б. логарифм

Алгоритмы градиентного спуска. Метод Ньютона

x_0

$$x_{k+1}=x_k-hf'(x_k)$$

$$x_{k+1}=x_k-[f''(x_k)]^{-1}f'(x_k)$$

Может не сойтись.

Вопрос: когда сходится

Утв $f\in C_L^{1,1}(\mathbb{R}^n) \Rightarrow$ можно подобр. $\exists$ конст. для град. спуска так, чтобы он сходился

непрерывн. перв. произв. с конст. липшицевости L

$$\|f'(y)-f'(x)\|\le L\|y-x\|$$

скаляр. произв.

$$\left| f(y) - f(x) - \langle f'(x), y-x \rangle \right| \leq \frac{L}{2}\|y-x\|^2$$ ← следует из липшицевости

$f(x_{k+1}) < f(x_k)$?

$$f(y) \leq f(x) + \langle f'(x), y-x \rangle + \frac{L}{2}\|y-x\|^2 =$$

$y = x_k - h \cdot f'(x_k)$

$x = x_k$

$y - x = -hf'(x_k)$

$$= f(x_k) - h\|f'(x_k)\|^2 + \frac{Lh^2}{2}\|f'(x_k)\|^2$$

$$f(x_{k+1}) \leq f(x_k) - h\left(1 - \frac{Lh}{2}\right)\|f'(x_k)\|^2 \ \leq$$

$$h\left(1 - \frac{Lh}{2}\right) > 0 \Rightarrow \text{убывание } f(x_k)$$

$$\underset{\substack{(h>0)\\ h \in \mathbb{R}}}{\operatorname{argmax}}\ h\left(1 - \frac{Lh}{2}\right) = \frac{1}{L}$$

$$\leq f(x_k) - \frac{1}{2L}\|f'(x_k)\|^2$$

$$\frac{1}{2L}\|f'(x_k)\|^2 \leq f(x_k) - f(x_{k+1})$$

$\forall N$ верно $\forall N$

$$\frac{1}{2L}\sum_{k=0}^{N}\|f'(x_k)\|^2 \leq f(x_0) - f(x_{N+1}) \leq$$

$$\leq f(x_0) - f^* \quad) \text{ число}$$

↑ минимум ф-ции (предп., что он есть)

$$\Rightarrow f'(x_k) \to 0$$

$$\frac{1}{2L}(N+1)\min_{k=\overline{0,N}}(f'(x_k))^2 \leq f(x_0) - f^*$$

$$\min_{k=\overline{0,N}}(f'(x_k)) \leq \sqrt{\frac{2L}{N+1}(f(x_0) - f^*)} \leq \varepsilon$$

$$\sqrt{\frac{2L}{N+1}(f(x_0) - f^*)} \leq \varepsilon$$ сублинейная скорость сходимости

Скорости сходимости бывают (рассм-ся) двух видов

$x_k \to x^*$

$f_k \to f^*$ ← grad f → 0

Утв. $f \in C^{22}_{M}(\mathbb{R}^n)$

(2 р. дифф-ма, 2-я произв. уд. конст. Липш. M)

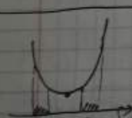
x^* $f'(x^*) > 0$ ($A \geq B$ если $A - B \geq 0$ — неотриц. полож. опр.)

$l I_n \leq f'(x^*) \leq L I_n$

$\| x_0 - x^* \| \leq \bar{r}$ ← в $\bar{r}$ участв. константы из усл.

то тогда $\| x_{k+1} - x^* \| = r_{k+1} \leq C(L, l, M)\left(1 - q(L, l, M)\right)^k$

линейная сходим-ть



$f'(x^*) = 0$

$f''(x) \geq 0$

$Q \quad Q$

$$Q = \left\{ \begin{array}{ll} f_1(x) \leq 0 & h_1(x) = 0 \\ f_2(x) \leq 0 & \vdots \\ f_m(x) \leq 0 & h_q(x) = 0 \end{array} \right.$$

1. $x^* = \arg\min_{x \in Q} f_0(x)$

1 $\Downarrow$ 2 $\Uparrow$? 3 $\Updownarrow$

усл-я ККТ (Каруш Кун Таккер)

Лагранжиан $L(x, \lambda, \nu) = f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) + \sum_{i=1}^{q} \nu_i h_i(x) =$

$= f_0(x) + \lambda^T f(x) + \nu^T h(x)$

Усл-я ККТ

1. $x^* \in Q$
2. $\exists\, \lambda^* \geq 0 \quad \lambda_i^* f_i(x^*) = 0, \quad i = \overline{1, n}$
3. $\exists\, \nu^*$ $\nabla_{(x)} L(x^*, \lambda^*, \nu^*) = 0$

Boid „Convex optimization"

8.11.2016

Q:

$f_0(x) \to \min$

$f_1(x) \leq 0$

$\vdots$

$f_m(x) \leq 0$

$h_1(x) = 0$

$\vdots$

$h_p(x) = 0$

$$L(x, \lambda, \nu) \triangleq f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) + \sum_{i=1}^{p} \nu_i h_i(x)$$

функция Лагранжа

Двойственная целевая функция:

$$g(\lambda, \nu) = \inf_{x \in \mathbb{R}^n} \left(L(x, \lambda, \nu) \right)$$

Двойственная задача оптимизации:

$$g(\lambda, \nu) \to \max \qquad d^* = g(\lambda^*, \nu^*)$$
$$\lambda \geq 0$$

$$\underset{\lambda \geq 0}{g(\lambda, \nu)} \leq \overset{f_0(x^*)}{f^*} \leq \underset{x \in Q}{f_0(x)}$$

факт аналогичный теореме о слабой двойственности

если $d^* = f^*$

$d^* - f^* = 0$ $\qquad \lambda^* \geq 0$

$L(x, \lambda^*, \nu^*)$ достигает минимума в x^* и равна в этом минимуме f^*

$$\min_{x \in \mathbb{R}^n} L(x, \lambda^*, \nu^*) \overset{?}{=} f^*$$

Предп.: $g(\lambda^*, \nu^*) = f_0(x^*)$ $(d^* = f^*)$

$x \in Q$ $\quad g(\lambda^*, \nu^*) \leq L(x, \lambda^*, \nu^*) = f_0(x) + \sum_{i=1}^{m} \lambda_i^* f_i(x) + \sum_{i=1}^{p} \nu_i^* h_i(x) \leq f_0(x)$

$x^* \in Q$

$$g(\lambda^*, \nu^*) \leq L(x^*, \lambda^*, \nu^*) = f_0(x^*) + \sum_{i=1}^{m} \lambda_i^* f_i(x^*) + \sum_{i=1}^{p} \nu_i^* h_i(x^*)$$

$\shortparallel$

d^*

$\shortparallel$

f^*, т.е. $g(\lambda^*, \nu^*) = L(x^*, \lambda^*, \nu^*)$ $\Rightarrow f_0(x^*) = \inf_{x \in \mathbb{R}^n} \left[f_0(x) + \sum_{i=1}^{m} \lambda_i^* f_i(x) + \sum_{i=1}^{p} \nu_i^* h_i(x) \right]$

$$\Downarrow$$

$$\sum_{i=1}^{m} \lambda_i^* f_i(x^*) = 0$$

$$\Downarrow$$

$$\lambda_i^* f_i(x^*) = 0 \quad \forall i$$

(часть теоремы о комплементарн. эффективн-ти)

$$\Downarrow$$

$$f^* = \min_{x \in \mathbb{R}^n} L(x, \lambda^*, \nu^*)$$

в предп.: $d^* - f^* = 0$

$$f^* = \min_{x \in \mathbb{R}^n} L(x, \lambda^*, \nu^*)$$

$$L(x^*, \lambda^*, \nu^*) \leq L(x, \lambda^*, \nu^*)$$

$\shortparallel$

f^*

$$L(x^*, \lambda^*, \nu^*) \leq L(x, \lambda^*, \nu^*)$$

$\shortparallel$

f^*

~~[illegible]~~

Более аккуратно: $\quad d^* = f^*$

Предп.: $\exists\, \lambda^*, \nu^*, \underline{x^* \in Q}$ $\quad \lambda^* \geq 0 \quad g(\lambda^*, \nu^*) = f_0(x^*)$

Тогда $f_0(x^*) = g(\lambda^*, \nu^*) = \inf\limits_{x \in \mathbb{R}^n} \left(f_0(x) + \lambda^{*T} f(x) + \nu^{*T} h(x)\right)$

$\leq f_0(x^*) + \lambda^{*T} f(x^*) + \nu^{*T} h(x^*) \leq f_0(x^*)$

$$\Downarrow$$

$$\boxed{L(x^*, \lambda^*, \nu^*) \leq L(x, \lambda^*, \nu^*)}$$

Симметричное нерав-во:

$$L(x^*, \lambda, \nu) \leq L(x^*, \lambda^*, \nu^*)$$

Кроме того: $\quad \lambda_i^* f_i(x^*) = 0$

Если $\exists\, \lambda^*, \nu^* : \lambda^* \geq 0$, $x^* \in Q$, такие что

$g(\lambda^*, \nu^*) = f_0(x^*)$, то:

$\Rightarrow f_i(x^*) \leq 0 \;\; \forall i = \overline{1, m}$

$h_j(x^*) = 0 \;\; \forall j = \overline{1, q}$

$\lambda^* \geq 0$

$\lambda_i^* f_i(x^*) = 0, \; i = \overline{1, m}$

$\nabla_x L(x^*, \lambda^*, \nu^*) = 0$

усл. ККТ

Вопрос: когда верно обратное?

Предположим: есть 2 точки

x^*, λ^*, ν^* $\qquad \tilde{x}, \tilde{\lambda}, \tilde{\nu}$

удовлет в. тому, что после „Если" $\qquad$ удовлет в. тому, что после „То"

Можно тогда считать, что $\tilde{x}, \tilde{\lambda}, \tilde{\nu}$ - какая-то точка оптимума?

$$g(\tilde{\lambda}, \tilde{\nu}) = \inf_{x \in \mathbb{R}^n} \left(L(x, \tilde{\lambda}, \tilde{\nu})\right) \leq L(\tilde{x}, \tilde{\lambda}, \tilde{\nu}) \overset{=}{\leq} f_0(\tilde{x})$$

$\nabla_x L(\tilde{x}, \tilde{\lambda}, \tilde{\nu}) = 0$ - min, если ф-ция $L(x)$ выпуклая

f_0, f_i, h_i - выпуклые

$\Downarrow$

<u>верно обратное</u>

$\Leftarrow$

(необх. и дост. усл. минимума)

Вопрос: как понять, что двойств. зазор нулевой?

Квалификации ограничений: ← усл. Слейтера

1) Если $\exists x_s \in Q$: $f_i(x_s)<0$ для всех i, то для выпуклой задачи двойств. зазор равен 0.

f_0, f_i - выпуклые, диф-мые

h_i - аффинные

т.е. $\forall$ точка $\tilde{x}$ удовлетворяющ. усл. ККТ явл-ся точкой условн. минимума исходн. задачи

$$\frac{1}{2}x^T Bx + c^T x + c_0 \to \min \quad \text{- выпуклая}$$

$$Ax = b \qquad B>0$$

Усл. Слейтера выполн., усл. типа равенств аффинны

Тогда ~~мы~~ поиск надо вести среди точек, удовлетвор.:

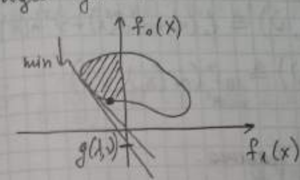
$$Ax^* = b$$
$$Bx^* + c + A^T\nu^* = 0$$

$$h_j(x) = (a_{j_1}, \dots a_{j_n})\begin{pmatrix} x_1^* \\ x_2^* \\ \vdots \\ x_n^* \end{pmatrix} - b_j$$

$$\begin{pmatrix} B & A^T \\ A^T & 0 \end{pmatrix}\begin{pmatrix} x^* \\ \nu^* \end{pmatrix} = \begin{pmatrix} -c \\ b \end{pmatrix}$$

$$f_0(x_k) - g(\lambda_k, \nu_k) \le \varepsilon$$
$$\Downarrow$$
$$f_0(x_k) - f^* \le \varepsilon$$

← Идея док-ва:



15.11.2016.

$$\begin{cases} f_0(x) \to \min \\ f_1(x) \leq 0 \\ \dots \\ f_n(x) \leq 0 \\ h_1(x) = 0 \\ \dots \\ h_q(x) = 0 \end{cases}$$

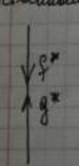
$$L(x, \lambda, \nu) \triangleq f_0(x) + \lambda^T f(x) + \nu^T h(x)$$

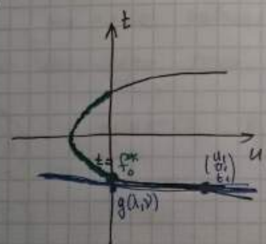
$$g(\lambda, \nu) \triangleq \inf_{x \in \mathbb{R}^n} (L(x, \lambda, \nu))$$

Двойств. задача:

$$g(\lambda, \nu) \to \max$$
$$\lambda \geq 0$$

Двойственный зазор:





$$G = \left\{ \begin{pmatrix} u \\ t \end{pmatrix} \in \mathbb{R}^2 \mid \exists x : f_0(x) = t, f_1(x) = u \right\}$$

$$f_0(x) = x$$
$$f_1(x) = x^2 - 1$$

$f_0(x) \to \min$	$t \to \min$
$f_1(x) \leq 0$	$\begin{pmatrix} u \\ t \end{pmatrix}$
	$u \leq 0$

$$g(\lambda, \nu) = \inf_{x \in \mathbb{R}^n} (\underbrace{f_0(x)}_{t} + \lambda^T \underbrace{f(x)}_{u} + \nu^T \underbrace{h(x)}_{v})$$

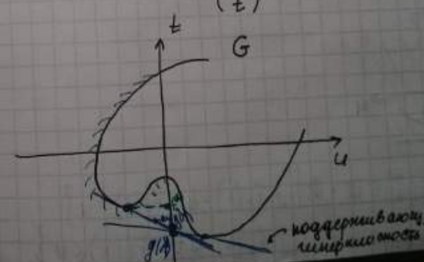
$$G = \left\{ \begin{pmatrix} u \\ v \\ t \end{pmatrix} \in \mathbb{R}^m \times \mathbb{R}^q \times \mathbb{R} \mid \exists x : f_0(x) = t; \ f_i(x) = u_i, \ i = \overline{1, m}, \ h_j(x) = v_j, \ j = \overline{1, q} \right\}$$

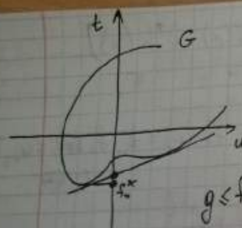
$$g(\lambda, \nu) = \inf_G (t + \lambda^T u + \nu^T v) =$$

$$= \inf_G \left((\lambda^T \ \nu^T \ 1) \begin{pmatrix} u \\ v \\ t \end{pmatrix} \right)$$

$$g(\lambda, \nu) \leq (\lambda^T \ \nu^T \ 1) \begin{pmatrix} u \\ v \\ t \end{pmatrix}$$

$$\begin{pmatrix} u \\ v \\ t \end{pmatrix} \in G$$





$g \le f$

$g(\lambda, \nu) \le t + \lambda u$

tg угла наклона прям. в коорд. (u, t)

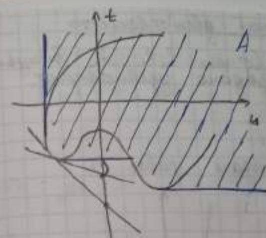
$t = -\lambda u + g(\lambda)$

$-\lambda \le 0$

$A = \left\{ \begin{pmatrix} u \\ v \\ t \end{pmatrix} \in \mathbb{R}^m \times \mathbb{R}^q \times \mathbb{R} \;\middle|\; \exists x: f_0(x) \le t,\ f_i(x) \le u_i,\ h_j(x) = v_j \right\}$, $i = \overline{1, m}$, $j = \overline{1, q}$

$G = \left\{ \begin{pmatrix} u \\ t \end{pmatrix} \in \mathbb{R} \times \mathbb{R} \;\middle|\; \exists\ f_0(x) = t,\ f_1(x) = u \right\}$

$A = \left\{ \begin{pmatrix} u \\ t \end{pmatrix} \in \mathbb{R} \times \mathbb{R} \;\middle|\; \exists\ f_0(x) \le t,\ f_1(x) \le u \right\}$



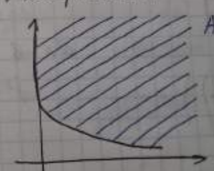
Можно строить поддерживающие гиперплоскости не к множеству G, а к мн-ву A.

← A не яв-ся выпуклым

Если A выпукло, поддерживающ. гиперпл. требуем. вида построить не всегда получится.

Плоскость $u = 0$ построить нельзя (не будет соответствовать $g(\lambda)$)

т.е. поддерживающ. гиперплоскость должна быть невертикальн.



Док-во теоремы о сильной двойственности

Если задача выпуклая (ф-ции $f_0, f_1, \dots, f_m$ - выпуклые, $h_1, \dots, h_q$ - аффинные), и задача удовлетворяет условию Слейтера

$(\exists x_s \in Q: f_i(x_s) < 0,\ i = \overline{1,m})$,

то двойственный зазор в задаче равен 0.

$f_0(x) \to \min$

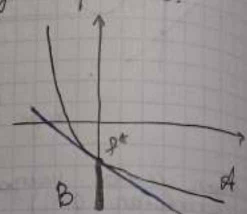
$f_1(x) \le 0$

$\dots$

$f_m(x) \le 0$

$Ax = b$

$\operatorname{rang}(A) = q$



$$\mathcal{A} = \left\{ \begin{pmatrix} u \\ v \\ t \end{pmatrix} \in \mathbb{R}^m \times \mathbb{R}^q \times \mathbb{R} \;\middle|\; \exists x: f_0(x) \le t;\ f_i(x) \le u_i,\ i = \overline{1,m};\ a^T x - b = v_j,\ j = \overline{1,q} \right\}$$

$$g(\lambda, \nu) = \inf_{\mathcal{A}} \left((\lambda^T \nu^T 1) \begin{pmatrix} u \\ v \\ t \end{pmatrix} \right)$$

1) $\mathcal{A}$ - выпуклое множество

2) $B = \left\{ \begin{pmatrix} 0 \\ 0 \\ s \end{pmatrix} \in \mathbb{R}^m \times \mathbb{R}^q \times \mathbb{R} \;\middle|\; s < f^* \right\}$ - выпуклое мн-во (т.к. это луч)

3) $\mathcal{A} \cap B = \varnothing$

4) $\exists$ гиперпл-ть, разделяющ. $\mathcal{A}$ и B:

$\tilde{\lambda}^T u + \tilde{\nu}^T v + \mu t \le \alpha \quad \begin{pmatrix} u \\ v \\ t \end{pmatrix} \in B$

$\tilde{\lambda}^T u + \tilde{\nu}^T v + \mu t \ge \alpha \quad \begin{pmatrix} u \\ v \\ t \end{pmatrix} \in \mathcal{A}$

$\Rightarrow \mu t \le \alpha \quad \forall t < f^*$, если $\begin{pmatrix} u \\ v \\ t \end{pmatrix} \in B$

если $\begin{pmatrix} u \\ v \\ t \end{pmatrix} \in \mathcal{A}$, то $\tilde{\lambda} \ge 0,\ \mu \ge 0$

5.1) Предп., что $\mu > 0$

$$g\left(\frac{\tilde{\lambda}}{\mu}, \frac{\tilde{\nu}}{\mu}\right) = \left(\frac{\tilde{\lambda}}{\mu}\right)^T u + \left(\frac{\tilde{\nu}}{\mu}\right)^T v + t \ge \frac{\alpha}{\mu} \ge f^*$$

при $\frac{\tilde{\lambda}}{\mu} \ge 0$ $\qquad$ $\le f^*$ (по слаб. двойств.)

$\Rightarrow g(\lambda, \nu) = f^*$ (нет двойств. зазора)

5.2) Предп., что $\mu = 0$

Тогда $\tilde{\lambda}^T u + \tilde{\nu}^T v \ge \alpha \quad \begin{pmatrix} u \\ v \\ t \end{pmatrix} \in \mathcal{A}$

$f_i \le 0$

$$\overbrace{\sum_{i=1}^{m} \lambda_i f_i(x)}^{\lambda^T u} + \underbrace{\nu^T (Ax - b)} \ge \alpha$$

$\nu^T(Ax-b) \geq 0 \quad \forall x$

за исключ. случ. $\nu^T A = 0$

$A^T \nu = 0$

но это невозм., т.к. $\ker A = \varnothing$

$\Rightarrow$ противоречие $\Rightarrow \mu \neq 0$.

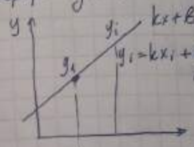
<u>Практика</u> 22.11.2016.

1. $\|Ax - b\|_2^2 \to \min\limits_x$ сумма квадратов ошибок

$\|Ax - b\|_1 \to \min\limits_x$ сумма модулей ошибок

$\|Ax - b\|_\infty \to \min\limits_x$ абс. значение

Генерир. знач.: $Ax + \mathcal{N}(0,6) = b$



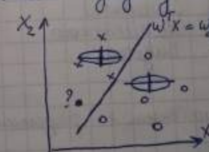
$y_i = kx_i + b + \varepsilon$

$\varepsilon \sim \mathcal{N}(0,10)$

$$\left.\begin{array}{l} y_0 = y_0 + 50 \\ y_{m-1} = y_{m-1} - 50 \end{array}\right\} \text{выбросы в распределении}$$

Посмотреть, как изменятся k и b

2. Решить задачу классификации



$\times \to y = -1$

$\circ \to y = +1$

$x \in \mathbb{R}^n,\ y \in \{-1; +1\}$

$\begin{pmatrix} x_1 \\ y_1 \end{pmatrix}, \dots, \begin{pmatrix} x_m \\ y_m \end{pmatrix}$ обучающая выборка

Алгоритм $a(x) = \mathrm{sign}(w^T x - w_0)$ Линейный дискриминантный анализ

1 Ошибки классификации

$y_i \cdot a(x_i)$

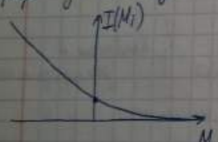
$-\frac{1}{2}\, sign \cdot \underbrace{y_i (w^T x_i - w_0)}_{M_i} + \frac{1}{2}$ индикатор ошибки

0, если не ошибке

1, если ошибке

$\min\limits_{w, w_0} \underbrace{\sum\limits_{i=1}^{m} I(M_i)}_{\text{количество ошибок}}$

Штраф за ошибку



$\min\limits_{w,w_0} \underbrace{\sum\limits_{i=1}^{m} -\ln\left(\frac{1}{1+e^{-M_i}}\right)}_{\text{выпуклая ф-ция}}$

$w \in R^n$

алгоритм градиентного спуска

Логистическая регрессия

$6(x_i) = \frac{1}{1+e^{-\frac{y_i(w^T x_i - w_0)}{M}}}$

Метод наибольш. правдоподобия

$\min\limits_{w} \sum\limits_{i=1}^{m} \ln(1+e^{-y_i \langle w, x \rangle})$

$w \in \mathbb{R}^{n+1}$

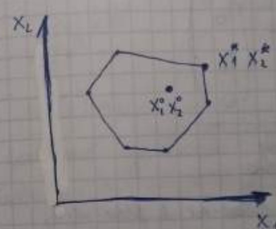
$x \in \mathbb{R}^{n+1}$, коорд. соотв. w_0 (нулевая) – всегда $= -1$

Скользящий контроль

CV – cross validation

machinelearning.ru

3. Решение задачи ЛП методом внутренней точки (метод центрального пути)



$\max\ c_1 x_1 + c_2 x_2$

$Ax \leq b$

$x \geq 0$

$\mu > 0$

$\mu_0 \to x_1^0\ x_2^0$

Док-ть, что если $\mu \to 0$, то $x^k \to x^*$

$solve(F(x, x_s, y, y_s, \mu) = 0)$

комплементарн. перем-е

$Ax + x_s = b$

$x_1\ x_s \geq 0$

Ищем реш-е, т.ч. $x, x_s, y, y_s > 0$ (внутри допустимой обл.)

Вернёт: x^k, x_s^k, y^k, y_s^k

отметить на рисунке

$$F(x, x_s, y, y_s) = \begin{pmatrix} Ax + x_s - b \\ A^T y + y_s - c \\ x_i y_{s_i} - \mu \\ \vdots \\ y_j x_{s_j} - \mu \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

$$X = \begin{pmatrix} x_1 & & 0 \\ & \ddots & \\ 0 & & x_n \end{pmatrix} \qquad X_s = \begin{pmatrix} x_{s_1} & & 0 \\ & \ddots & \\ 0 & & x_{s_m} \end{pmatrix}$$

$$F(x, x_s, y, y_s) = \underbrace{\begin{pmatrix} Ax + x_s - b \\ A^T y + y_s - c \\ X \cdot Y_s \cdot \mathbb{1} - \mu \mathbb{1} \\ Y \cdot Y_s \cdot \mathbb{1} - \mu \mathbb{1} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}}_{\text{решить систему методом Ньютона}}$$

$\mathbb{1}$ - единичн. в-р нужн. размера

Запретить методу Ньютона отправл-ся к реш-ям, которые меньше нуля:

$$\begin{pmatrix} x^{l+1} \\ x_s^{l+1} \\ y^{l+1} \\ y_s^{l+1} \end{pmatrix} = \begin{pmatrix} x^l \\ x_s^l \\ y^l \\ y_s^l \end{pmatrix} - \lambda_l \left(F'(x^l \ldots y_s^l, \mu_k)\right)^{-1} \cdot F(x^l, \ldots, y_s^l, \mu_k)$$

λ_l — препятствует появл-ю отриц. чисел

λ_l выбрать так, чтобы все коорд-ты на $l+1$ шаге были >0.

$\boxed{\text{7.12. – дедлайн}}$

22.11.2016

Наука про двойственность дает массу интересных результатов! :)

Чувствительность решения

$p^* = \min f_0(x)$ $\quad g(\lambda^*, \nu^*) = d^* = p^* = f_0(x^*)$

$f_1(x) \le 0$ $\quad$ реш-е двойств. задачи

$\vdots$

$f_m(x) \le 0$

$h_1(x) = 0$

$\vdots$

$h_q(x) = 0$

Немного изменим условие задачи

$p^*(u, v) = \min f_0(x)$ $\quad g(\lambda^*, \nu^*) = d^* = p^*(0,0) = f_0(x^*)$

$f_1(x) \le u_1$

$\vdots$

$f_m(x) \le u_m$ $\quad p^*(u, v) \ge p^*(0,0) - \lambda^{*T} u - \nu^{*T} v$

$h_1(x) = v_1$

$\vdots$

$h_q(x) = v_q$

Примеры:

- $u_i > 0$ $\quad \lambda_i^*$ маленькое $\Rightarrow$ такого рода поведение бессмысл.
- $u_i < 0$ $\quad \lambda_i^*$ большое $\Rightarrow$ существ. ухудшение ситуации

Локальная чувствительность:

$$\frac{\partial p^*(0,0)}{\partial u_i} = -\lambda_i^*$$

$$\frac{\partial p^*(0,0)}{\partial v_i} = -\nu_i^*$$

Соотв. ограничение не явл-ся активным в точке $(0,0)$ - неравенство не становится равенством.

Огранич-е явл-ся ~~простым~~ активным $\Rightarrow$ пропорциональное реш-е

(к/р через пару) → двойственность и ее приложения

к/р на семинаре в след. четверг

- Метод Ньютона и градиентн. спуска
  - → скорость сх-ти
  - → барьерные и штрафные ф-ции
- проекция на простые множества !!!
- задачи на чувствительность решения

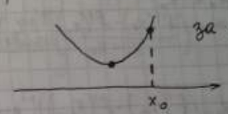
29.11.2016.

В методе Ньютона:

$$x_{k+1} = x_k - \alpha_k [f''(x_k)]^{-1} f'(x_k)$$

в классическом методе Ньютона $\alpha_k = 1$

строит квадратичное приближение ф-ции в точке



$[f''(x_k)]^{-1}$ счит-ся за $O(n^3)$, где n - размерность простр-ва

Другой метод (там 3 или 4 какие-то фамилии) Бройдена Флетчера Гольдфарба Шанно

вместо $[f''(x_k)]^{-1}$ предложено считать м-цу H_k

она должна давать направление убывания

$$s_k = x_{k+1} - x_k = \alpha_k H_k^{-1} f'(x_k)$$

$$s_{k+1} = x_{k+2} - x_{k+1} = \alpha_{k+1} \cdot H_{k+1}^{-1} f'(x_{k+1})$$

$$H_k s_k = \alpha_k f'(x_k)$$

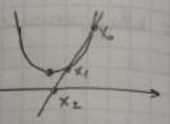
$$H_{k+1} s_{k+1} = \alpha_{k+1} f'(x_{k+1})$$

$$f'(x_k) - f'(x_{k+1}) \geq 0$$ ← можно гарантировать, когда выполнено:

$$\left(f'(x_{k+1}) - f'(x_k)\right)^T (x_{k+1} - x_k) > 0$$

и

$H_{k+1} > 0$ и $H_k > 0$ и ещё $H_k y = s$



x_0
$x_1 = x_0 - \alpha_0 f'(x_0)$
$$x_2 = x_1 - \alpha_1 \cdot \left(\frac{f'(x_1) - f'(x_0)}{x_1 - x_0}\right)^{-1} f'(x_1)$$

$$x_{k+1} = x_k - \alpha_k \cdot \left(\frac{f'(x_k) - f'(x_{k-1})}{x_k - x_{k-1}}\right)^{-1} f'(x_k)$$

$y = f'(x_k) - f'(x_{k-1})$ $s = x_k - x_{k-1}$

$H_k y = s$

Не по касательной, а по секущей

Теперь про двойственность

Теорема (о слабой альтернативе в задачах о совместности системы нестрогих неравенств)

$$(1) \begin{cases} f_i(x) \leq 0, & i = \overline{1,m} \\ h_j(x) = 0, & j = \overline{1,q} \end{cases}$$

(Интересует наличие x, удовлетв. системе)

Если для двух систем

$$(2) \begin{cases} g(\lambda, \nu) > 0 \\ \lambda \geq 0 \end{cases}$$

хотя бы одна из них несовместна,

то, если $\exists x$ удовлетвор. первой системе, то $\nexists \lambda, \nu$ - удовлетворяющ. втор. сист.,

если $\exists \lambda, \nu$ удовлетв. втор. сист., то $\nexists x$ удовлетв. первой.

▲ $\min 0$
$f_i(x) \leq 0$
$h_j(x) = 0$

$$g(\lambda, \nu) = \inf_{x \in \mathbb{R}^n} \left(\lambda^T f(x) + \nu^T h(x)\right) \to \max_{\lambda \geq 0}$$

$$\forall \alpha > 0 \quad g(\alpha\lambda, \alpha\nu) = \alpha g(\lambda, \nu)$$

$$\max_{\lambda \geq 0} g(\lambda, \nu) = \begin{cases} \infty, & \exists \lambda, \nu \; g(\lambda, \nu) > 0, \lambda \geq 0 \\ 0, & \forall \lambda, \nu \; g(\lambda, \nu) \leq 0, \lambda \geq 0 \end{cases}$$

есть реш-е (2) => двойств. задача [illegible] огранич.

=> прямая задача не совместна

$\left(\underset{\lambda \geq 0}{g(\lambda, \nu)} \leq \underset{x \in Q}{f_0(x)} \underset{(\text{по усл.})}{=0} \right)$

В обратную сторону: $\left.\begin{array}{c} A \to \neg B \\ \Downarrow \\ B \to \neg A \end{array}\right\}$ Т. о слабой альтернативе

$A = (2)$ совместна

$B = (1)$ совместна ∎

Для системы строгих неравенств:

(1) $f_i(x) < 0, \ i = \overline{1,m}$ $h_j(x) = 0, \ j = \overline{1,q}$	$g(\lambda, \nu) \geq 0$ $\lambda \neq 0, \lambda \geq 0$ (2)

аналогичное утв-е ((1) совместна)

▲ Пусть $\exists \tilde{x}: f_i(\tilde{x}) < 0, h_j(\tilde{x}) = 0$

$\lambda_1 f_1(\tilde{x}) + \lambda_2 f_2(\tilde{x}) + \ldots + \lambda_k f_k(\tilde{x}) + \nu_1 h_1(\tilde{x}) + \ldots + \nu_q h_q(\tilde{x})$

Тогда для $\underset{\text{положительно}}{\lambda \geq 0}, \lambda \neq 0 \ \forall \nu$ < 0

=> Inf ещё меньше => $g(\lambda, \nu) < 0$

$\Rightarrow (B \Rightarrow \neg A) \Rightarrow (A \Rightarrow \neg B)$ ∎

$\begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_m \end{pmatrix} \geq \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$, при этом хотя бы одно из $\lambda_i > 0$

это и означает запись $\lambda \geq 0, \lambda \neq 0$

$\exists$ ситуации в которых можно гарантировать совместность за счёт несовместности другой системы.

Теорема (о сильной альтернативе)

В точности одна из этих систем (1), (2) (со строгими нерав.) не совместна

$f(x) < 0$ - выпуклые

$Ax = b$ - аффинные

$\exists \tilde{x}: A\tilde{x} = b$ (провер-ся методом Гаусса)

Задача оптимизации:

$f_i(x) - s \leq 0$

$Ax = b$

$p^* = \min s$

Двойственная:

$\inf\limits_{x \in \mathbb{R}^n, s \in \mathbb{R}} (s + \lambda^T (f(x) - s) + \nu^T h(x)) = g_s(\lambda, \nu) = \begin{cases} g(\lambda, \nu), & \sum\limits_{i=1}^{m} \lambda_i = 1 \\ -\infty, & \text{иначе} \end{cases}$

$\max\limits_{\lambda \geq 0} g_s(\lambda, \nu) =$

Эта задача совместна. Если получится, что $s < 0$, то система (1) совместна.

Есть реш-е (2) => двойств. задача не огранич.

=> прямая задача не совместна

$\left(\underset{\lambda \geq 0}{g(\lambda,\nu)} \leq \underset{x \in Q}{f_0(x)} \underset{(\text{по усл.})}{=0} \right)$

В обратную сторону: $A \to \neg B \Downarrow B \to \neg A$ — Т. о слабой альтернативе

A = (2) совместна

B = (1) совместна ■

Для системы строгих неравенств:

$$(1)\ \begin{cases} f_i(x) < 0, & i = \overline{1,m} \\ h_j(x) = 0, & j = \overline{1,q} \end{cases} \qquad \begin{cases} g(\lambda,\nu) \geq 0 \\ \lambda \neq 0, \lambda \geq 0 \end{cases}\ (2)$$

аналогичное утв-е

▲ Пусть $\exists \tilde{x}$: $f_i(\tilde{x}) < 0$, $h_j(\tilde{x}) = 0$ ((1) совместна)

$$\lambda_1 f_1(\tilde{x}) + \lambda_2 f_2(\tilde{x}) + \ldots + \lambda_k f_k(\tilde{x}) + \nu_1 h_1(\tilde{x}) + \ldots + \nu_q h_q(\tilde{x})$$

Тогда для $\lambda \geq 0, \lambda \neq 0$ < 0

неположительно $\forall \nu$

=> Inf ещё меньше => $g(\lambda,\nu) < 0$

$$\Rightarrow (B \Rightarrow \neg A) \Rightarrow (A \Rightarrow \neg B) \ ■$$

$$\begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_m \end{pmatrix} \geq \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$$, при этом хотя бы одно из них $\lambda_i > 0$

это и означает запись $\lambda \geq 0, \lambda \neq 0$

$\exists$ ситуации в которых можно гарантировать совместность за счёт несовместности другой системы.

Теорема (о сильной альтернативе)

(1), (2) (со строгими нерав.)

В точности одна из этих систем не совместна

$f(x) < 0$ - выпуклые

$Ax = b$ - аффинные

$\exists \tilde{x}: A\tilde{x} = b$ (провер-ся методом Гаусса)

Задача оптимизации:

$f_i(x) - S \leq 0$

$Ax = b$

$p^* = \min S$

Двойственная:

$$\inf_{\substack{x \in \mathbb{R}^n, \\ s \in \mathbb{R}}} (s + \lambda^T f(x) + \nu^T h(x)) = g_s(\lambda,\nu) = \begin{cases} g(\lambda,\nu), & \sum_{i=1}^m \lambda_i = 1 \\ -\infty, & \text{иначе} \end{cases}$$

$\max_{\lambda \geq 0} g_s(\lambda,\nu)$

Эта задача совместна. Если получится, что $S < 0$, то система (1) совместна.

$p^* < 0 \Rightarrow \exists x: f_i(x) < 0,\ Ax = b$
$\Leftarrow$
Обратное тоже верно.

В двух последних задачах имеет место сильная двойственность. (т.к. ф-ции выпукл. и выполн. усл. Слейтера)

Это верно, т.к.
ф-ции $f_i - s$ выпуклые.

Если $p^* < 0$, то для задачи выполнено усл. Слейтера.

$\Rightarrow (p^* < 0 \Rightarrow$ сильная двойств-ть$)$

$$\Rightarrow p^* = d^* = \max_{\lambda \geq 0} g_s(\lambda, \nu)$$

$$\Rightarrow d^* < 0,\ \text{т.е. } \max_{\lambda \geq 0} g_s(\lambda, \nu) < 0, \text{ и } \neq -\infty$$
(т.к. p^* - число)

$\Rightarrow g(\lambda, \nu) \geq 0,$
$\nexists \lambda\ \sum \lambda_i = 1,\ \lambda \geq 0$, что

И тогда что $g(\lambda, \nu) < 0 \quad \forall \lambda \geq 0, \lambda \neq 0$
$\Rightarrow d^* < 0 \Rightarrow p^* < 0.$ ∎

<u>Пример</u>
$Ax \leq b$
система линейных нерав-в

$$g(\lambda) = \inf_x \lambda^T (Ax - b) = \begin{cases} -b^T\lambda, & A^T\lambda = 0 \\ -\infty, & \text{иначе} \end{cases}$$

$$A^T\lambda = 0 \qquad \lambda \geq 0$$
$$-b^T\lambda > 0 \quad (b^T\lambda < 0)$$

Слабая альтернатива тоже,
сильная альтернатива, т.к. имеет место сильная двойств.

Если есть такой x, то нет таких λ.
Если есть такой λ, то нет таких x.

$$\min b^T\lambda$$
$$A^T\lambda = 0$$
$$\lambda \geq 0$$

$$Ax < b \quad \Rightarrow \quad \begin{matrix} A^T\lambda = 0 & \lambda \geq 0 \\ b^T\lambda \leq 0 & \lambda \neq 0 \end{matrix}$$

$Ax \le 0$ альтерн. $A^T y + c = 0$
при $\textcircled{>}$ $y \geqslant 0$
$c^T x < 0$

Это удобно :) Где? Играем на бирже.

p_i - цены у нескольких активов (в начале дня)

x - решение о покупках или продажах

$p^T x$ результат от решений в деньгах (получили или потратили)

Хотелось, чтобы к концу дня осталась та же сумма активов, пересчитанная новыми ценами

σ_i - цены активов в конце дня

$p^T x < 0$

$$m \begin{pmatrix} \sigma_1^{(1)} \\ \sigma_1^{(2)} \\ \sigma_1^{(3)} \end{pmatrix} \quad Vx \geqslant 0$$

m различных раскладов

по вложениям ничего не потеряли, а в начале дня получили прибыль - это замечательно, но так не будет, рынок так устроен, что он не позволит этого сделать.

Сертификат того, что так не будет - наличие решения у системы

$$V^T y + p = 0$$
$$y \geqslant 0$$

Допустим, какая-нибудь начальная цена неизвестна: p_n - ?

Теория позвол. получить оценки на знач-е p_n, при усл., что нет реш-й системы

$Vx \geqslant 0$
$p^T x < 0$ (нет арбитража)

$\Rightarrow$ исходя из знаний всех σ и почти всех p, можно дать оценку на max и min p_n:

$$\begin{array}{c|c} \min p_n & \max p_n \\ V^T y = p & V^T y = p \\ y \geqslant 0 & y \geqslant 0 \end{array}$$

6.12.2016.

Задача 2

Метод центрального пути.

Позволяет сгенерировать последовательность точек, которая сходится к решению, если это решение есть.

При этом каждая из точек генерируется отдельно от предыдущих по некот. правилу с уменьшающимся параметром μ_k

$\mu_k \to 0$

$\mu_{k+1} = \mu_k \cdot \frac{1}{2}$

$x^k = solve\_$

x - обычные перем. прямой задачи

x^s - фиктивн. перем. пр. зад.

$$\max c^T x \quad Ax \le b, \quad x \ge 0$$

$$\Downarrow$$

$$Ax + x_s = b$$
$$x, x_s \ge 0$$

Метод пытается решить одновременно и прямую задачу и двойственную.

$$\min b^T y$$
$$A^T y - y_s = b$$
$$y, y_s \ge 0$$

$$Ax + x_s - b = 0$$
$$A^T y - y_s - b = 0$$

+

$$x_i y_{s_i} \overset{?}{=} 0, \quad i = 1, \dots, m$$
$$y_j x_{s_j} \overset{?}{=} 0, \quad j = 1, \dots, m$$

нелинейная система уравнений

В д/з надо находить такие решения системы:

$$\left(\begin{array}{l} Ax + x_s - b = 0 \\ A^T y - y_s - b = 0 \\ x_i y_{s_i} = \mu, \ i = 1, \dots, n \\ y_j x_{s_j} = \mu, \ j = 1, \dots, m \end{array}\right)$$

, где μ - некотор. малое число

$$\max c^T x \quad Ax + x_s = b, \quad x, x_s \ge 0$$

$$L(x, x_s, y, y_s) = c^T x + y_s^T (Ax + x_s - b) + y^T x_s - \overset{?}{y_s^T} x$$

что-то другое тут

Предлагается модифицировать лагранжиан.

$$\nabla L(x, x_s, y, y_s) = 0$$

Забудем про ограничения $x, x_s \ge 0$.

Тогда $L(x, x_s, y, y_s) = c^T x + y_s^T (Ax + x_s - b)$

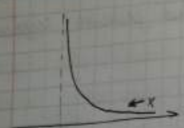
Условие $\min_x L$, $\max_y L$ $\to$ седловая точка

$\nabla L = 0$

$\nabla_y L = 0$

Если потом вспомнить, что $x, x_s \ge 0$.

Можно модифицир. задачу и ~~решать~~ искать седлов. точку только для положительных x методом барьерных функций.



Q

$$L(x, x_s, y, y_s) = c^T x + y_s^T (Ax + x_s - b) - \mu \ln x - \mu \ln x_s$$

$\ln x$ определён покомпонентно для вектора.

$\nabla_x L = 0$

$\nabla_y L = 0$

$\nabla_{x_s} L = 0$

$\nabla_{y_s} L = 0$

Кривая седлов. точка ф-ции лагранжа

min $\quad$ max

$x, x_s \quad y, y_s$

Если задача несовместна, то не будет найдено реш-е $Ax + x_s - b = 0$.

Если неограни. (двойств. несовм. = неогр.), то не найдётся реш-е

$$A^T y - y_s - b = 0.$$

Метод сопряжённых направлений

Придуман для реш-я систем линейных уравнений.

$Ax = b$ $\quad$ эквив. задаче $\quad \underset{x}{\operatorname{argmin}} \left(\frac{1}{2} x^T A x - b x \right)$

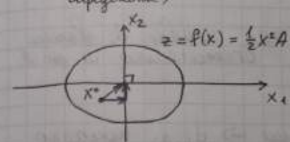
$A^T = A$

$A > 0$

(положительно определённ)

$x \in \mathbb{R}^2$

$z = f(x) = \frac{1}{2} x^T A x$



Линии уровня - эллипсы

A-диаг. $\Rightarrow$ перпендик. оси

$\forall x^0$ метод Ньютона (способ отыскания минимума квадратичн. привлекательн. ф-ции) попадёт при x^0 сразу в решение (при $A > 0$)

если не диаг., могут быть наклон.

Другая идея, не связанная с необходимостью обращения матрицу:

оптимизироваться вдоль координатных направлений

$$x^{k+1} = x^k + \alpha_k d_k \qquad \alpha_k = \underset{\alpha}{\operatorname{argmin}} f(x_k + \alpha d_k)$$

направление на k-м шаге методе

$$\alpha_k = \frac{\nabla f(x^k)^T f(x^k)}{d_k^T A d_k}$$ по любому направлению d_k

Для diag направление (чередовать) & повторять не надо, м-цы A

всё сойдётся не больше чем за размерность пространства.

$d_0, \dots, d_{n-1}$ $\quad x \in \mathbb{R}^n$

Предлагается выбрать их так, чтобы d_i были попарно сопряжёнными относительно м-цы A.

$$\forall i, j \quad i \neq j \quad d_i^T A d_j = 0$$

М-ца A задаёт метрику $\Rightarrow d_i, d_j$ попарно ортогональны в этой метрике

$x^0 \quad d_0 = -\nabla f(x^0)$

Можно гарантир., что в итоге будет получено решение задачи оптимизации

$$\underset{x}{\operatorname{argmin}} \left(\frac{1}{2} x^T A x - bx \right)$$

Сопряжённые вектора d_i генерир-ся по итеративному правилу:

$$d_{k+1} = -\nabla f(x^{k+1}) - \beta_k d_k, \quad \beta_k = \frac{\nabla f(x^{k+1})^T \cdot \nabla f(x^{k+1})}{\nabla f(x^k)^T \cdot \nabla f(x^k)}$$

$$\sigma_0 d_0 + \dots + \sigma_{n-1} d_{n-1} = x^* - x^0$$

$$x^k - x^0 = \alpha_0 d_0 + \dots + \alpha_{k-1} d_{k-1}$$

можно показать, что $\alpha_0 = \sigma_0, \dots, \alpha_k = \sigma_k$:

$$x^* - x^0 = \sigma_0 d_0 + \dots + \sigma_{n-1} d_{n-1}$$

домножс. лев. и прав. часть слева на $d_k^T A$

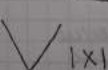
$$\to d_k^T A (x^* - x^0) = \sigma_k d_k^T A d_k$$

$$\Rightarrow \sigma_k = \frac{d_k^T A (x^* - x^0)}{d_k^T A d_k} \qquad \alpha_k = \dots$$

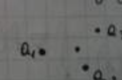
Оптимизация выпуклой, но не гладкой ф-ции

непрерывн.

В одномерн. случ. можно решать методом дихотомии или методом золотого сечения.



Пример двухмерная з. оптимизац.

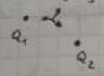


Точка, миним. расст. между ними

$$\min_x \sum_{i=1}^{3} \|x - a_i\|$$

Если min не попадает ни в одну из точек a_1, a_2, a_3, то всё ok

$$(\|p\|)' = \frac{p}{\|p\|}$$

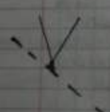


Если есть тупой угол в Δ, т.ч. $\alpha \geq 120°$,

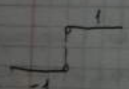
то min может находиться в одной из точек a_1, a_2, a_3.

Надо расширить понятие градиента

$$f(x+y) \geq f(x) + \nabla f(x)^T y$$ — для выпуклой ф-ции всегда справедливо



В тех местах, где ф-я не диф-ма скажем, что произв. $= c$ (от -1 до 1 для модуля).



Тогда выражение выше всё равно будет верным

$$\partial f(x) = \begin{cases} \nabla f(x), & \text{там, где он определён} \\ c, & \text{для котор.} \end{cases}$$

субградиент (подиффер. отобр.)

$$f(x+y) \geq f(x) + \partial f(x)^T y$$

$$f(x+y) \geq f(x) + \partial f(x)^T y$$

$\partial f(x^k) = 0 \iff x^k$ — точка минимума

Строго:

$0 \in \partial f(x^k) \iff x^k$ — точка минимума

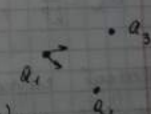
$$\partial(|x|) = \begin{cases} -1, & x < 0 \\ |c| \leq 1, & x = 0 \\ 1, & x > 0 \end{cases}$$

$$\partial(\|x\|) = \begin{cases} \frac{x}{\|x\|}, & x \neq 0 \\ \|c\| < 1, & x = 0 \end{cases}$$

Упр Док-ть, что

$$\partial(\|x - Pr_Q(x)\|) = \begin{cases} \frac{x - Pr_Q(x)}{\|x - Pr_Q(x)\|}, & x \notin Pr_Q(x) \\ ?, & x \in Pr_Q(x) \end{cases}$$

$$\frac{x - a_1}{\|x - a_1\|} + \frac{x - a_2}{\|x - a_2\|} + \frac{x - a_3}{\|x - a_3\|}$$



$|y - a| \geq f(a) + cy$

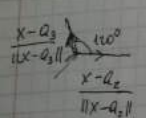
$|y - a| > cy$

$c \leq \frac{|y-a|}{y}$ $c \leq \frac{y-a}{y}$, $y > a$ $y < a$, $c \leq \frac{a-y}{y}$

$|c| \leq 1$

$c \le \frac{a}{y} - 1 \quad y < a \qquad \Rightarrow |c| \le 1$

$c \le 1 - \frac{a}{y} \quad y > a$



$$x^{k+1} = x^k - \alpha_k \partial f(x^k)$$

$\alpha_k \to 0 \qquad \partial f(x^k) \le L$

$\sum_{k=0}^{\infty} \alpha_k \to \infty$ гарантир. $f^k \to f^*$

$\sum_{k=0}^{\infty} \alpha_k^2 \to \infty$ гарантир. $x^k \to x^*$

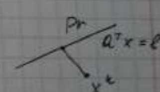
$\alpha_k = \frac{1}{k^p}, \quad \frac{1}{2} \le p \le 1$

Частный случай, когда можно выбрать шаг самым хорошим способом.

$$f^* = \min_x \max_i \left(\|x - Pr_{Q_i}(x)\| \right) = 0$$

$$\alpha_k = \frac{f(x^k) - f^*}{\|\partial f(x^k)\|^2} \qquad x^{k+1} = Pr_{Q_i(x)}(x^k)$$

Метод последовательного проектирования



$$Pr_Q(x^k) = x^k + \lambda a$$

$$\lambda = \frac{b - a^T x^k}{a^T a}$$