25.09.17 момо $\underline{IV}$

$f(x) \to \min\limits_x$, $x \in \mathbb{R}^n$, $f \in C^1$ $\left\{ \begin{matrix} q - \text{ст-ть} \\ f(x) \end{matrix} \right\}$ $\frac{\partial f}{\partial x_i}$, $\frac{\partial^2 f}{\partial x_i \partial x_j}$

Метод	Память	Ст-ть итерации
Градиентный спуск	$O(n)$	$O(qn)$ вычисление градиента
Метод Ньютона	$O(n^2)$	$O(qn^2 + n^3)$ вычисление гессиана; решение СЛАУ

Метод сопряжённых градиентов

$Ax = b$, $A = A^T > 0$ $\iff$ $f(x) = \frac{1}{2} x^T A x - x^T b \to \min\limits_{x \in \mathbb{R}^n}$

$x_{k+1} = x_k + \alpha_k d_k$ $\qquad \nabla f(x) = Ax - b = 0$

$0 = \nabla f(x_k + \alpha d_k)^T d_k = (A(x_k + \alpha d_k) - b)^T d_k = g_k^T d_k + \alpha d_k^T A d_k$

отсюда $\alpha_{opt} = -\dfrac{g_k^T d_k}{d_k^T A d_k}$, d_k ?

$\underline{О}$ $\{d_i\}_{i=0}^{n-1}$ наз. сопряжённым, если $d_i^T A d_j = 0$ $\forall i \neq j$
семейство векторов

$\langle d_i, d_j \rangle_A = d_i^T A d_j$; $\{d_i\}$ сопр. $\iff$ $\{d_i\}$ орт. отн. A

Пример: $\{d_i\}_{i=0}^{n-1}$ ортогональный базис с.в. A

$d_i^T A d_j = d_i^T \lambda_j d_j = \lambda_j d_i^T d_j = \lambda_j \delta_{ij} \Rightarrow \{d_i\}$ сопр. отн. A

$\underline{У}$ Если $\{d_i\}$ — сопр., то $\{d_i\}$ л.н.з.

$\square$ $\sum\limits_{i=0}^{n-1} \alpha_i d_i = 0 \mid d_j^T A \times$, $0 = \sum\limits_{i=0}^{n-1} \alpha_i d_j^T A d_i = \alpha_j \underbrace{d_j^T A d_j}_{>0}$,

тогда $\alpha_j = 0$ $\forall j$ и $\{d_i\}$ л.н.з. $\blacksquare$

$\{d_i\}_{i=0}^{n-1}$ сопр. $\Rightarrow$ $\{d_i\}$ — базис в $\mathbb{R}^n$

$$A x_{opt} = b, \quad x_{opt} - x_0 = \sum_{i=0}^{n-1} \alpha_i d_i \quad \Big| \; d_j^T A \times b$$

$$d_j^T A (x_{opt} - x_0) = \alpha_j d_j^T A d_j, \quad \alpha_j = \frac{d_j^T \overbrace{A (x_{opt} - x_0)}}{d_j^T A d_j},$$

$$\alpha_j = \frac{-g_0^T d_j}{d_j^T A d_j} \quad ! \qquad d_j^T (b - A x_0) = -g_0^T d_j$$

У $g_k^T d_k = g_0^T d_k$

□ $g_k^T d_k = (A x_k - b)^T d_k = (A(x_0 + \sum_{i=0}^{k-1} \alpha_i d_i) - b)^T d_k =$

$= g_0^T d_k + \sum_{i=0}^{k-1} \alpha_i \underbrace{d_i^T A d_k}_{=0} = g_0^T d_k$ ■

Свойство частичной оптимальности

У $g_k^T d_j = 0 \quad \forall j < k$

□ $g_k^T d_j = (A x_k - b)^T d_j = (A(x_0 + \sum_{i=0}^{k-1} \alpha_i d_i) - b)^T d_j =$

$= g_0^T d_j + \sum_{i=0}^{k-1} \alpha_i d_i^T A d_j = g_0^T d_j + \alpha_j d_j^T A d_j = \Big\{$

$= \Big\{ \alpha_j = -\frac{g_0^T d_j}{d_j^T A d_j} \Big\} = 0$ ■

$g_k^T (\sum_{i=0}^{k-1} \gamma_i d_i) = 0$, x_k явл-ся минимумом $f(x)$ в ③ $x_0 + L\{d_0 \dots d_{k-1}\}$ $\overset{①}{=} x_k + L\{d_0, \dots, d_{k-1}\}$

② $\hat{x}_0 + L\{d_0, \dots, d_{k-1}\}$

$$\underset{\hat{x}_1}{x_0 + L\{d_0\}} \subset \underset{\hat{x}_2}{x_0 + L\{d_0, d_1\}} \subset \dots \subset \underset{\hat{x}_n}{x_0 + L\{d_0 \dots d_{n-1}\}} = \mathbb{R}^n$$

идём далее: $x = x_k + \sum_{i=k}^{n-1} \gamma_i d_i$

$j < k$, $(Ax - b)^T d_j = \underbrace{(A x_k - b)^T d_j}_{=0} + \sum_{i=k}^{n-1} \gamma_i \underbrace{d_i^T A d_j}_{=0} = 0$

x явл-ся минимумом f в $x + L\{d_0 \dots d_{k-1}\}$

$$f(x) = \frac{1}{2} x^T A x - x^T b \to \min_x$$

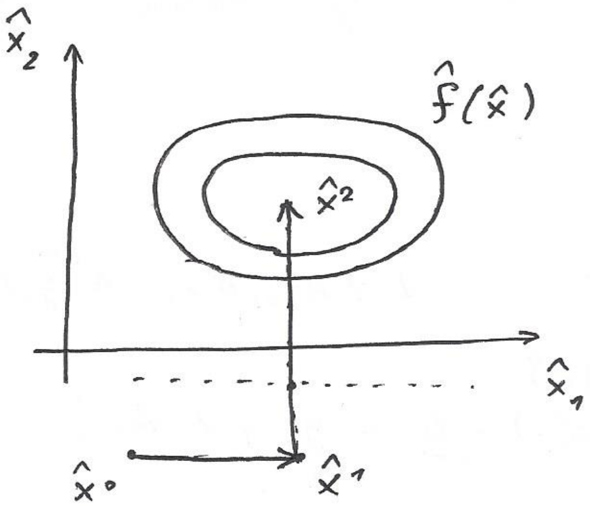
$\{d_i\}_{i=0}^{n-1}$ сопр. A, $C = [d_0 | d_1 | \dots | d_{n-1}] \in \mathbb{R}^{n \times n}$

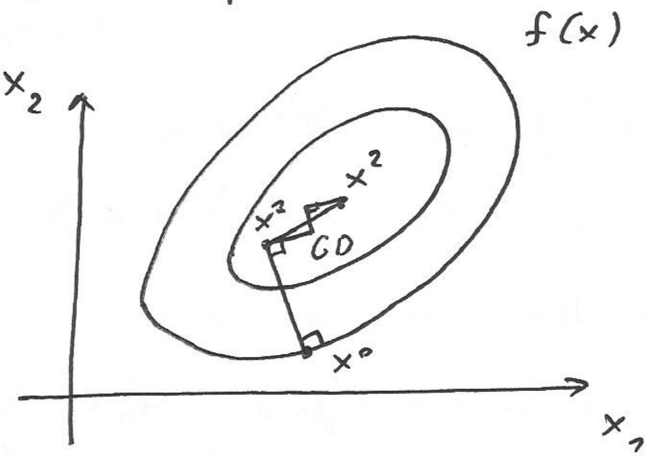
$x = C\hat{x}$, $\hat{x} = C^{-1} x$, $\hat{f}(\hat{x}) = f(C\hat{x}) = \frac{1}{2} \hat{x}^T \underbrace{C^T A C}_{\Lambda} \hat{x} - \hat{x}^T C^T b$

$\hat{x}$ - координаты в $d_0 \dots d_{n-1}$

$\hat f(\hat x) = \frac{1}{2} \hat x^T \Lambda \hat x - \hat x^T C^T b = \frac{1}{2} \sum_{i=0}^{n-1} (\lambda_i \hat x_i^2 - 2 \hat x_i (C^T b)_i) \to \min_{\hat x}$

квадратичная сепарабельная ф-ция

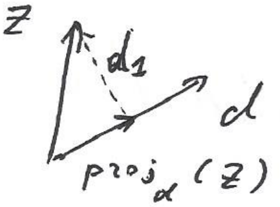




поддержание оптимальности относительно предыдущего подпр-ва

использование свойства частичной оптимальности для экономии памяти

Грама-Шмидта процесс



$proj_d(z) = \frac{z^T d}{d^T d} d$, ~~$\{d_i\}$~~ $\{z_i\}$ л.н.з.

Gramm-Schmidt orth process

$d_0 = z_0$, $d_1 = z_2 - proj_{d_0}(z_1)$, $d_2 = z_2 - proj_{d_2}(z_3) - proj_{d_1}(z_3)$

$\ldots$ $d_{n-1} = z_{n-1} - \sum_{j=0}^{n-2} proj_{d_j}(z_{j+1})$, $O(n^2)$ памяти :(

Идея CG: $z_i = -g_i$

$\begin{cases} d_0 = -g_0 \\ d_{k+1} = -g_{k+1} + \beta_k d_k \end{cases}$ $O(n)$ пересчёт и память

$\beta_k = \frac{g_{k+1}^T A d_k}{d_k^T A d_k}$

У 1) $\{g_0, g_1, \dots, g_i\} \subset L\{g_0, Ag_0, A^2 g_0, \dots, A^i g_0\}$

$\{\cancel{g} d_0, d_1, \dots, d_i\} \subset L\{g_0, Ag_0, \dots A^i g_0\}$

подпространство Крылова

2) $g_k^T g_j = 0 \quad \forall j < k$

3) $d_{k+1}^T A d_j = 0 \quad \forall j < k$

□ ① По индукции $\qquad L\{Ag_0, A^2 g_0 \dots A^{i+1} g_0\}$

$i=0$, $g_0 \in L\{g_0\}$, $d_0 = -g_0 \in L\{g_0\}$

$i \to i+1$, $g_{i+1} = A x_{i+1} - b = \underbrace{A(x_i} + \alpha_i d_i) \underbrace{- b} = g_i + \alpha_i A d_i$

тогда $g_{i+1} \in L\{g_0, Ag_0, \dots, A^{i+1} g_0\} \qquad L\{g_0, Ag_0, \dots, A^i g_0\}$

$d_{i+1} = -g_{i+1} + \beta_i d_i \in L\{g_0, Ag_0, \dots, A^{i+1} g_0\}$

② $g_j = \sum_{i=0}^{j} \gamma_i d_i$, $g_k^T g_j = \sum_{i=0}^{j} \gamma_i \underbrace{g_k^T d_i}_{0} = 0 \quad j<k$

③ $g_{j+1} = g_j + \alpha_j A d_j \Rightarrow A d_j = \dfrac{g_{j+1} - g_j}{\alpha_j}$

$d_{k+1}^T A d_j = (-g_{k+1} + \beta_k d_k)^T A d_j = \{j < k\} = -g_{k+1}^T A d_j =$

$= -\dfrac{\overset{0}{g_{k+1}^T g_{j+1}} - \overset{0}{g_{k+1}^T g_j}}{\alpha_j} = 0$ ■

$\alpha_k = -\dfrac{g_k^T d_k}{d_k^T A d_k} = -\dfrac{g_k^T(-g_k + \beta_{k-1} d_{k-1})}{d_k^T A d_k} = \dfrac{g_k^T g_k}{d_k^T A d_k}$

$\beta_k = \dfrac{g_{k+1}^T A d_k}{d_k^T A d_k} = \dfrac{g_{k+1}^T (g_{k+1} - g_k)}{d_k^T A \cancel{d_k}\, \cancel{\alpha_k}} = \dfrac{g_{k+1}^T g_{k+1}}{g_k^T g_k}$

$\leftarrow g_3 \quad \leftarrow$ алгоритм

Схема CG: A, b, x_0, ε $\qquad g_0 = A x_0 - b$, $d_0 = -g_0$

для $k = 0, 1, 2, \dots$

$\alpha_k = \dfrac{g_k^T g_k}{d_k^T A d_k}$, $x_{k+1} = x_k + \alpha_k d_k$, $g_{k+1} = A x_{k+1} - b$,

если $\|g_{k+1}\|^2 < \varepsilon$, выход; иначе $\beta_k = \dfrac{g_{k+1}^T g_{k+1}}{g_k^T g_k}$, $d_{k+1} = -g_{k+1} + \beta_k d_k$

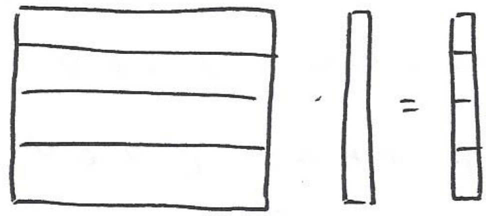
$\{g_{k+1} = A x_{k+1} - b = g_k + \alpha_k \underline{A d_k}\}$

$O(n^2)$ "Ad" 1 шт

$O(n)$ "$y^T y$" 2 шт

$O(n)$ "$x+y$" 3 шт

A:



фиксированная матрица A

скорость сх-ти CG

① Если у A $\exists$ r различных с. зн., тогда CG сх-ся не более чем за r шагов

② Если у A $\exists$ r кластеров с. зн., тогда CG сх-ся „примерно" за r шагов

③ $\mu = \lambda_{min}(A)$, $L = \lambda_{max}(A)$, $f \in C_L^{2,1}$ и μ сил. вып.

тогда $f(x_{k+1}) - f_{opt} \leq 4 \left(\frac{\sqrt{L/\mu} - 1}{\sqrt{L/\mu} + 1} \right)^{2k} \left(f(x_{\cancel{opt}}) - f_{opt} \right)$

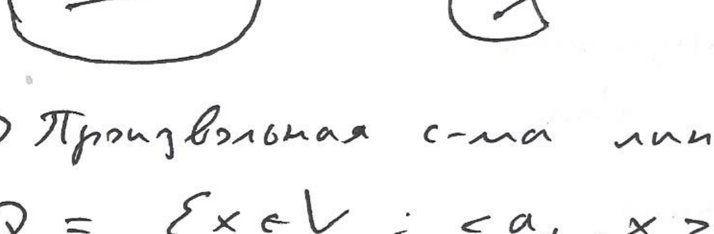
$$\frac{\sqrt{\frac{L}{\mu}} - 1}{\sqrt{\frac{L}{\mu}} + 1} = 1 - \frac{2}{\sqrt{\frac{L}{\mu}} + 1} = 1 - \frac{2\sqrt{\frac{\mu}{L}}}{1 + \sqrt{\frac{\mu}{L}}} = \{(1+x)^{-1} \approx 1 - x\} \approx$$

$$\approx 1 - 2\sqrt{\frac{\mu}{L}}\left(1 - \sqrt{\frac{\mu}{L}}\right) = 1 - 2\sqrt{\frac{\mu}{L}} + 2\frac{\mu}{L} \approx 1 - 2\sqrt{\frac{\mu}{L}}$$

Выпуклые множества и функции, семинар

О V - вект. пр-во, $Q \subseteq V$ $\forall \alpha \in [0,1]$

Q наз-ся выпуклым, если $x, y \in Q \Rightarrow \alpha x + (1-\alpha) y \in Q$

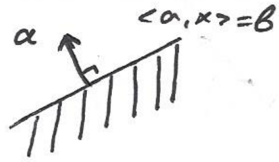


① Произвольная с-ма лин. огр. $(a_\alpha)_{\alpha \in \mathcal{A}} \in V$ $(b_\alpha)_{\alpha \in \mathcal{A}} \in \mathbb{R}$

$Q = \{ x \in V : \langle a_\alpha, x \rangle \leq b_\alpha \ \forall \alpha \in \mathcal{A} \}$

яв-ся выпуклым

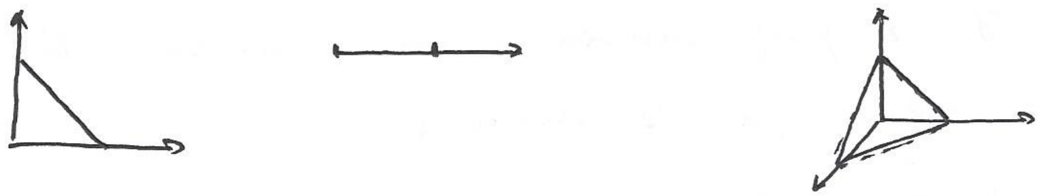
гиперплоскость $\{x \in V : \langle a, x\rangle = b\}$



полупр-во $\{x \in V : \langle a, x\rangle \le b\}$

многогранник $\{x \in V : \langle a_k, x\rangle \le b_k \ , \ k \in \{k\}\}$ конечное мн-во

стандартный симплекс $\Delta_n = \{x \in \mathbb{R}^n , \ x_i \ge 0 \ . \ \sum_i x_i = 1\}$



② Шар в произвольной норме

$Q = \{x \in V : \|x - a\| \le r\}$

$\forall \ x, y \in Q \ , \forall \alpha \in \mathbb{R}$

$\|\alpha x + (1-\alpha) y - a\| = \cancel{\|\alpha(x-y) + (y-a)\| \le |\alpha| \|x-y\| + \|y-a\|}$

$= \|\alpha(x-a) + (1-\alpha)(y-a)\| \le |\alpha| \|x-a\| + |1-\alpha| \|y-a\| \le$

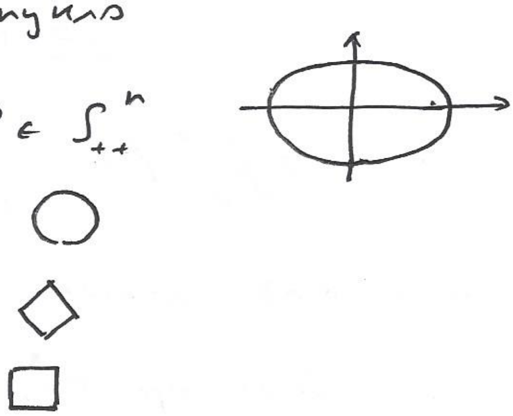
$\le |\alpha| r + |1-\alpha| \cancel{r} = r$, Q - выпукло

эллипсоид : $\|x\|_P = \langle Px, x\rangle^{\frac{1}{2}}$, $P \in S^n_{++}$

l^2-шар : $\{x \in \mathbb{R}^n : \|x-a\|_2 \le r\}$

l^1-шар : $\{x \in \mathbb{R}^n : \|x-a\|_1 \le r\}$

l^∞-шар : $\{x \in \mathbb{R}^n : \|x-a\|_\infty \le r\}$



③ Конус S^n_+ - выпуклый в S^n

$X, Y \in S^n_+ : \ \langle(\lambda X + (1-\lambda) Y)v, v\rangle = \lambda \underbrace{\langle Xv, v\rangle}_{\ge 0} + (1-\lambda) \underbrace{\langle Yv, v\rangle}_{\ge 0} \ge 0$

Операции, сохраняющие выпуклость

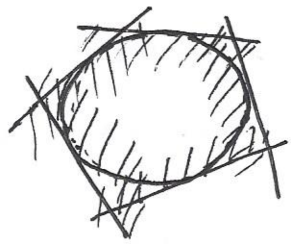
① Пересечение $(Q_\alpha)_{\alpha \in A}$ вып.

$Q = \bigcap_{\alpha \in A} Q_\alpha$ также выпукло

$x, y \in Q \Leftrightarrow x, y \in Q_\alpha \ \forall \alpha$

$\lambda x + (1-\lambda) y \in Q_\alpha \iff \lambda x + (1-\lambda) y \in Q \quad \forall \alpha$

$\{ x \in \mathbb{R}^2 : x_1^2 + x_2^2 \le 2 , \; x_2 \le 0 \}$



выпуклое как пересечение мн-в

$$\bigcap_{\alpha \in [-\pi, \pi]} \{ x \in \mathbb{R}^2 : (\cos \alpha) x_1 + (\sin \alpha) x_2 \le 1 \}$$

② Прямое произведение $A \times B$

$\{ x \in \mathbb{R}^3 , \; x_1^2 + x_2^2 \le 1 , \; -1 \le x_3 \le 1 \}$, цилиндр

③ Проекция. $Q_1 \times \ldots \times Q_n$ выпуклые, $\{ x_i : x \in Q_1 \times \ldots \times Q_n \}$

④ Произвольная коническая комбинация

$\alpha_1 Q_1 + \ldots + \alpha_k Q_k$ сумма Минковского, $\alpha_1 \ldots \alpha_k \ge 0$

⑤ Образ при аффинном преобразовании

Q - выпуклое в пр-ве V

линейное

$\mathcal{A} : V \to W$ аффинное пр-е: $\mathcal{A}(x) = L(x) + w$

мн-во $\mathcal{A}(Q) = \{ \mathcal{A}(x), x \in Q \}$ выпуклое

<u>Пример</u> $\{ Lx + a , \; x \in \mathbb{R}^n, L \in \mathbb{R}^{n \times n} , \; \|x\|_2 \le 1 \}$

⑥ Прообраз при аффинном преобразовании

$Q' = \{ x \in V : \mathcal{A}(x) \in Q \}$ ← выпукло в пр-ве W

<u>Пример</u> мн-во реш. LMI

$\{ x \in \mathbb{R}^k : x_1 A_1 + \ldots + x_k A_k \le B \}$ $A_1 \ldots A_k, B \in S^n$

$\underbrace{B - x_1 A_1 - \ldots - x_k A_k}_{\mathcal{A}(x)} \in S^n_+ = Q$ прообраз!

II Выпуклые ф-ии

О V - вещ. векторное пр-во, $X \subseteq V$, $X \neq \varnothing$, $f: X \to \mathbb{R}$

вып. ф-ия, если $f(\lambda x + (1-\lambda) y) \leq \lambda f(x) + (1-\lambda) f(y)$ $\forall x, y \in X$, $\forall \lambda \in [0,1]$

$\lambda x + (1-\lambda) y \in X \Rightarrow X$ выпукло

① Аффинные ф-ии $f(x) = \langle a, x \rangle + b$

выпуклая и вогнутая

$\langle a, \lambda x + (1-\lambda) y \rangle + b = \lambda(\langle a, x \rangle + b) + (1-\lambda)(\langle y, x \rangle + b)$

$\{ b = \lambda b + (1-\lambda) b \}$

тогда $\operatorname{tr} AX = \langle A^T, X \rangle$ выпукла!

② Норма

$f(x) = \|x\|$ выпуклая

У Нер-во Йенсена, f - вып. ф-ия

$$f\left(\sum_{i=1}^{k} \lambda_i x_i\right) \leq \sum_{i=1}^{k} \lambda_i f(x_i)$$

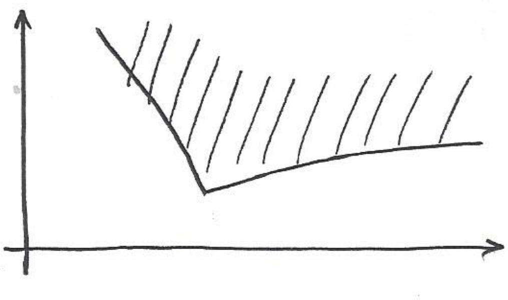
$\lambda_i \geq 0$, $\sum_{i=1}^{k} \lambda_i = 1$, $\sum_{i=1}^{k} \lambda_i x_i$ выпуклая комбинация

$f(\mathbb{E}\xi) \leq \mathbb{E} f(\xi)$, ξ - сл. в.

$\mathbb{E} \langle X^{-1} v, v \rangle \geq \langle \mathbb{E} X^{-1} v, v \rangle$ $\forall v \in \mathbb{R}^n$ $\mathbb{E} X^{-1} \geq (\mathbb{E} X)^{-1}$

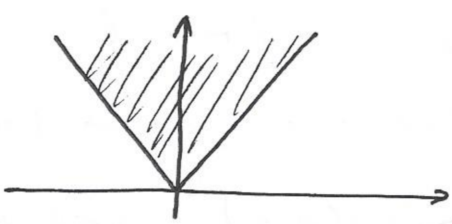
Надграфик ф-ии

$Epi(f) = \{ (x, t) \in X \times \mathbb{R}, f(x) \leq t \}$



У ф-ия f выпукла $\Leftrightarrow$ $Epi\, f$ яв-ся выпуклым мн-вом

Пример $\|\cdot\|$ выпукла



$Q = \{ (x, t) \in V \times \mathbb{R} : \|x\| \le t \}$

конус нормы

У $f: X \to \mathbb{R}$ выпуклая

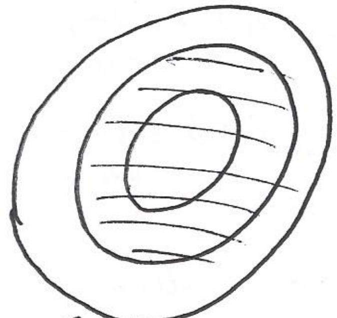
мн-во линий уровня ф-ии f:

$Lev_f(\alpha) = \{ x \in X : f(x) \le \alpha \}$ вып. мн-во

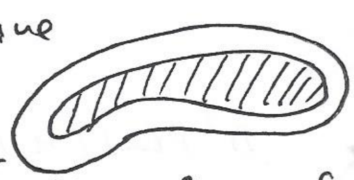
Пример $\{ x \in \mathbb{R}^n : c + \langle b, x \rangle - \frac{1}{2} \langle Ax, x \rangle \ge 0 \}$

$f(x) = \frac{1}{2} \langle Ax, x \rangle - \langle b, x \rangle - c$ выпуклые

$F(x) = \ln (\overbrace{c + \langle b, x \rangle - \tfrac{1}{2} \langle Ax, x \rangle}^{>0})$



Операции, сохраняющие выпуклость ф-ий



1. Коническая комбинация $f_1 \dots f_n$ - выпуклые

$\varphi = \alpha_1 f_1 + \dots + \alpha_n f_n$, $\alpha_1 \dots \alpha_n \ge 0$

2. Аффинная подстановка аргумента

f - вып. $\mathcal{A}: V \to W$ аффинное, $\varphi(x) = f(\mathcal{A}(x))$ φ - вып.

$\varphi: \underset{\text{вып.?}}{X'} \to \mathbb{R}$, $X' = \{ x \in V : \mathcal{A}(x) \in X \}$

3. Поточечный супремум

$(f_\alpha)_{\alpha \in \mathcal{A}}$ - выпуклые ф-ии, $\varphi(x) = \sup\limits_{\alpha \in \mathcal{A}} f_\alpha(x)$

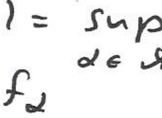
φ - выпуклая, $Epi\, \varphi = \bigcap\limits_{\alpha \in \mathcal{A}} Epi\, f_\alpha$

4. Монотонная композиция

$\varphi(x) = g(f_1(x) \dots f_n(x))$, $f_1 \dots f_n : X \to \mathbb{R}$ вып.

$g : \mathbb{R}^n \to \mathbb{R}$ вып.

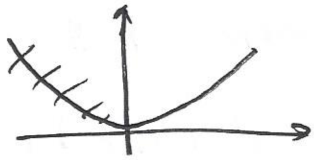
выпукла, если g монотонно возрастающая по каждому из аргументов



$\|x\|^3$, $t \mapsto t^3$ вып. на $[0, \infty)$

$x \mapsto \|x\|$ вып.

$\|x\|^2$, $t \mapsto t^2$ вып. на $[0, \infty)$



~~$\|x\| \to$~~ $x \mapsto \|x\|$

$(-\ln x)^2$ уже не вып.

← аффинное

S^n $\lambda_{max}(X) \neq \cancel{\sup_{\|x\|=1} \|Ax\| = \max\{\lambda_{max}(A), -\lambda_{min}(A)\}}$

$\lambda_{min}(X) \cancel{= \inf_{\|x\|=1} \|Ax\|}$, $\cancel{-\lambda_{min}(X) = -\inf_{\|x\|=1} \|Ax\| = \sup_{\|x\|=1}(-\|Ax\|)}$

$$\lambda_{min}(A) = \inf_{\|x\|=1} \langle Ax, x\rangle = \inf_{\|x\|\neq 0} \frac{\langle Ax, x\rangle}{\|x\|^2}$$

$$\lambda_{max}(A) = \sup_{\|x\|=1} \langle Ax, x\rangle = \sup_{\|x\|\neq 0} \frac{\langle Ax, x\rangle}{\|x\|^2}$$

$$-\lambda_{min}(A) = \sup_{\|x\|=1} \underbrace{(-\langle Ax, x\rangle)}_{\text{вып}}$$

ф-ия выпукла $\Leftrightarrow$ $D^2 f(x)[h, h] \geq 0$

$\langle \nabla^2 f(x)\, dx, dx \rangle \geq 0$