Методы оптимизации Лекция 2: Сопряжённые конусы. Отделимость. Выпуклые функции

Александр Катруца

Физтех-школа прикладной математики и информатики Московский физико-технический институт



14 сентября 2020 г.

На прошлой лекции

- ▶ О чём этот курс и почему он нужен
- ▶ Примеры постановок задач оптимизации
- Выпуклые множества и их свойства

Напоминание: конусы

Определение

Множество K называется конусом, если для любого $\mathbf{x} \in K$ и произвольного числа $\theta \geq 0$ выполнено $\theta \mathbf{x} \in K$.

Напоминание: конусы

Определение

Множество K называется конусом, если для любого $\mathbf{x} \in K$ и произвольного числа $\theta \geq 0$ выполнено $\theta \mathbf{x} \in K$.

Определение

Множество K называется **выпуклым** конусом, если для любых точек $\mathbf{x}_1, \mathbf{x}_2 \in K$ и любых чисел $\theta_1 \geq 0, \; \theta_2 \geq 0$ выполнено $\theta_1 \mathbf{x}_1 + \theta_2 \mathbf{x}_2 \in K$.

Напоминание: конусы

Определение

Множество K называется конусом, если для любого $\mathbf{x} \in K$ и произвольного числа $\theta \geq 0$ выполнено $\theta \mathbf{x} \in K$.

Определение

Множество K называется **выпуклым** конусом, если для любых точек $\mathbf{x}_1, \mathbf{x}_2 \in K$ и любых чисел $\theta_1 \geq 0, \; \theta_2 \geq 0$ выполнено $\theta_1 \mathbf{x}_1 + \theta_2 \mathbf{x}_2 \in K$.

Важные конусы

- ▶ Неотрицательный октант $\mathbb{R}^n_+ = \{\mathbf{x} \in \mathbb{R}^n \mid x_i \geq 0, \ i=1,\dots,n\} o \mathsf{LP}$
- ▶ Конус второго порядка $\{(\mathbf{x},t) \in \mathbb{R}^{n+1} \mid \|\mathbf{x}\|_2 \leq t\} o \mathsf{SOCP}$
- lacktriangle Конус симметричных положительно полуопределённых матриц $\mathbf{S}^n_+ o \mathsf{SDP}$

Сопряжённый конус (dual cone)

Определение

Пусть K — конус. Тогда множество

$$K^* = \{ \mathbf{y} \mid \langle \mathbf{y}, \mathbf{x} \rangle \ge 0, \ \mathbf{x} \in K \}$$

называется сопряжённым конусом.

Сопряжённый конус (dual cone)

Определение

Пусть K — конус. Тогда множество

$$K^* = \{ \mathbf{y} \mid \langle \mathbf{y}, \mathbf{x} \rangle \ge 0, \ \mathbf{x} \in K \}$$

называется сопряжённым конусом.

Свойства

- ▶ K* конус
- $lacktriangleright K^*$ выпуклый конус для *любого* конуса K
- ▶ Если $K_1 \subseteq K_2$, то $K_2^* \subseteq K_1^*$

Сопряжённый конус (dual cone)

Определение

Пусть K — конус. Тогда множество

$$K^* = \{ \mathbf{y} \mid \langle \mathbf{y}, \mathbf{x} \rangle \ge 0, \ \mathbf{x} \in K \}$$

называется сопряжённым конусом.

Свойства

- ► K* конус
- $lacktriangleright K^*$ выпуклый конус для *любого* конуса K
- ▶ Если $K_1 \subseteq K_2$, то $K_2^* \subseteq K_1^*$

Определение

Если $K=K^{st}$, то конус называется самосопряжённым (self-dual)

Определение

Сопряжённой нормой относительно $\|\cdot\|$ называется

$$\|\mathbf{z}\|_* = \sup_{\|\mathbf{x}\| \le 1} \mathbf{z}^\top \mathbf{x}.$$

Определение

Сопряжённой нормой относительно $\|\cdot\|$ называется

$$\|\mathbf{z}\|_* = \sup_{\|\mathbf{x}\| \le 1} \mathbf{z}^\top \mathbf{x}.$$

Примеры

- $\|\cdot\|_1 \to \|\cdot\|_* = \|\cdot\|_{\infty}$
- $\|\cdot\|_2 \to \|\cdot\|_* = \|\cdot\|_2$
- $\blacksquare \|\cdot\|_{\infty} \to \|\cdot\|_{*} = \|\cdot\|_{1}$

Определение

Сопряжённой нормой относительно $\|\cdot\|$ называется

$$\|\mathbf{z}\|_* = \sup_{\|\mathbf{x}\| \le 1} \mathbf{z}^\top \mathbf{x}.$$

Примеры

- $\|\cdot\|_1 \to \|\cdot\|_* = \|\cdot\|_{\infty}$
- $\|\cdot\|_2 \to \|\cdot\|_* = \|\cdot\|_2$
- $\blacksquare \|\cdot\|_{\infty} \to \|\cdot\|_{*} = \|\cdot\|_{1}$

Самосопряжённые конусы

 $ightharpoonup \mathbb{R}^n_+$

Определение

Сопряжённой нормой относительно $\|\cdot\|$ называется

$$\|\mathbf{z}\|_* = \sup_{\|\mathbf{x}\| \le 1} \mathbf{z}^\top \mathbf{x}.$$

Примеры

- $\|\cdot\|_1 \to \|\cdot\|_* = \|\cdot\|_{\infty}$
- $\|\cdot\|_2 \to \|\cdot\|_* = \|\cdot\|_2$

Самосопряжённые конусы

- $ightharpoonup \mathbb{R}^n_+$
- ▶ Конус второго порядка $\{(\mathbf{x},t) \in \mathbb{R}^{n+1} \mid \|\mathbf{x}\|_2 \leq t\}$

Определение

Сопряжённой нормой относительно $\|\cdot\|$ называется

$$\|\mathbf{z}\|_* = \sup_{\|\mathbf{x}\| \le 1} \mathbf{z}^\top \mathbf{x}.$$

Примеры

- $\|\cdot\|_1 \to \|\cdot\|_* = \|\cdot\|_{\infty}$
- $\|\cdot\|_2 \to \|\cdot\|_* = \|\cdot\|_2$

Самосопряжённые конусы

- $ightharpoonup \mathbb{R}^n_+$
- ▶ Конус второго порядка $\{(\mathbf{x},t) \in \mathbb{R}^{n+1} \mid \|\mathbf{x}\|_2 \leq t\}$
- $ightharpoonup \mathbf{S}_{+}^{n}$

Обобщённое отношение частичного порядка Пусть K — конус. Тогда $\mathbf{x} \leq_K \mathbf{y} \Leftrightarrow \mathbf{y} - \mathbf{x} \in K$.

Обобщённое отношение частичного порядка

Пусть K — конус. Тогда $\mathbf{x} \leq_K \mathbf{y} \Leftrightarrow \mathbf{y} - \mathbf{x} \in K$.

Пример: конус \mathbf{S}^n_+

Пусть $\mathbf{X},\mathbf{Y}\in\mathbf{S}^n$. Тогда $\mathbf{X}\leq_K\mathbf{Y}$ означает, что $\mathbf{Y}-\mathbf{X}\in\mathbf{S}^n_+$

Обобщённое отношение частичного порядка

Пусть K — конус. Тогда $\mathbf{x} \leq_K \mathbf{y} \Leftrightarrow \mathbf{y} - \mathbf{x} \in K$.

Пример: конус \mathbf{S}^n_+

Пусть $\mathbf{X},\mathbf{Y} \in \mathbf{S}^n$. Тогда $\mathbf{X} \leq_K \mathbf{Y}$ означает, что $\mathbf{Y} - \mathbf{X} \in \mathbf{S}^n_+$

Задача линейного программирования

$$egin{aligned} \min_{\mathbf{x} \in \mathbb{R}^n} \mathbf{c}^{\top} \mathbf{x} & \min_{\mathbf{x} \in \mathbb{R}^n} \mathbf{c}^{\top} \mathbf{x} \\ \mathrm{s.t.} \ \mathbf{A} \mathbf{x} = \mathbf{b} & \mathrm{s.t.} \ \mathbf{A} \mathbf{x} = \mathbf{b} \\ x_i \geq 0 & \mathbf{x} \in \mathbb{R}^n_{\perp} \end{aligned}$$

Обобщённое отношение частичного порядка

Пусть K — конус. Тогда $\mathbf{x} \leq_K \mathbf{y} \Leftrightarrow \mathbf{y} - \mathbf{x} \in K$.

Пример: конус \mathbf{S}^n_+

Пусть $\mathbf{X},\mathbf{Y}\in\mathbf{S}^n$. Тогда $\mathbf{X}\leq_K\mathbf{Y}$ означает, что $\mathbf{Y}-\mathbf{X}\in\mathbf{S}^n_+$

Задача линейного программирования

$$\min_{\mathbf{x} \in \mathbb{R}^n} \mathbf{c}^{\top} \mathbf{x} \qquad \min_{\mathbf{x} \in \mathbb{R}^n} \mathbf{c}^{\top} \mathbf{x}$$
s.t. $\mathbf{A} \mathbf{x} = \mathbf{b}$ s.t. $\mathbf{A} \mathbf{x} = \mathbf{b}$ $\mathbf{x} \in \mathbb{R}^n_+$

Введение нелинейности

Использование декартового произведение трёх самосопряжённых конусов позволяет записать многие практически важные выпуклые задачи

Определение

Множества A,B называются отделимыми, если существует вектор $\mathbf{a} \neq 0$ и число b такие что

- $ightharpoonup \mathbf{a}^{ op} \mathbf{x} + b \geq 0$ для всех $\mathbf{x} \in A$
- $ightharpoonup \mathbf{a}^{ op} \mathbf{y} + b \leq 0$ для всех $\mathbf{y} \in B$.

Определение

Множества A,B называются отделимыми, если существует вектор $\mathbf{a} \neq 0$ и число b такие что

- $ightharpoonup \mathbf{a}^{ op} \mathbf{x} + b \geq 0$ для всех $\mathbf{x} \in A$
- $ightharpoonup \mathbf{a}^{ op} \mathbf{y} + b \leq 0$ для всех $\mathbf{y} \in B$.

Теорема

Пусть A и B — выпуклые и непересекающиеся множества. Тогда существует разделяющая их гиперплоскость.

Доказательство

Определение

Множества A,B называются отделимыми, если существует вектор $\mathbf{a} \neq 0$ и число b такие что

- $ightharpoonup \mathbf{a}^{ op} \mathbf{x} + b \geq 0$ для всех $\mathbf{x} \in A$
- $ightharpoonup \mathbf{a}^{ op} \mathbf{y} + b \leq 0$ для всех $\mathbf{y} \in B$.

Теорема

Пусть A и B — выпуклые и непересекающиеся множества. Тогда существует разделяющая их гиперплоскость.

Доказательство

▶ Предположим, что расстояние между A и B положительно: $\inf_{\mathbf{x}\in A,\;\mathbf{y}\in B}\|\mathbf{x}-\mathbf{y}\|_2>0$

Определение

Множества A,B называются отделимыми, если существует вектор $\mathbf{a} \neq 0$ и число b такие что

- $ightharpoonup \mathbf{a}^{ op} \mathbf{x} + b \geq 0$ для всех $\mathbf{x} \in A$
- $ightharpoonup \mathbf{a}^{ op} \mathbf{y} + b \leq 0$ для всех $\mathbf{y} \in B$.

Теорема

Пусть A и B — выпуклые и непересекающиеся множества. Тогда существует разделяющая их гиперплоскость.

Доказательство

- ▶ Предположим, что расстояние между A и B положительно: $\inf_{\mathbf{x} \in A, \ \mathbf{y} \in B} \|\mathbf{x} \mathbf{y}\|_2 > 0$
- ▶ Пусть $\mathbf{c} \in A$ и $\mathbf{d} \in B$ точки, в которых этот минимум достигается

Определение

Множества A, B называются отделимыми, если существует вектор ${\bf a} \neq 0$ и число b такие что

- $ightharpoonup \mathbf{a}^{ op} \mathbf{x} + b \geq 0$ для всех $\mathbf{x} \in A$
- $ightharpoonup \mathbf{a}^{ op} \mathbf{y} + b \leq 0$ для всех $\mathbf{y} \in B$.

Теорема

Пусть A и B — выпуклые и непересекающиеся множества. Тогда существует разделяющая их гиперплоскость.

Доказательство

- ▶ Предположим, что расстояние между A и B положительно: $\inf_{\mathbf{x} \in A, \ \mathbf{y} \in B} \|\mathbf{x} \mathbf{y}\|_2 > 0$
- ▶ Пусть $\mathbf{c} \in A$ и $\mathbf{d} \in B$ точки, в которых этот минимум достигается
- ▶ Рассмотрим $f(\mathbf{x}) = \mathbf{a}^{\top}\mathbf{x} + b$, где $\mathbf{a} = \mathbf{d} \mathbf{c}$ и $b = \frac{\|\mathbf{d}\|_2^2 \|\mathbf{c}\|_2^2}{2}$

lacktriangle Покажем, что $f(\mathbf{y}) \geq 0$ для всех $\mathbf{y} \in B$

- ▶ Покажем, что $f(\mathbf{y}) \geq 0$ для всех $\mathbf{y} \in B$
- lacktriangle Пусть найдётся $\mathbf{u} \in B$ такая что $f(\mathbf{u}) < 0$

- ▶ Покажем, что $f(\mathbf{y}) \ge 0$ для всех $\mathbf{y} \in B$
- lacktriangle Пусть найдётся ${f u} \in B$ такая что $f({f u}) < 0$
- $f(\mathbf{u}) = (\mathbf{d} \mathbf{c})^{\top} (\mathbf{u} \frac{1}{2} (\mathbf{d} + \mathbf{c})) = (\mathbf{d} \mathbf{c})^{\top} (\mathbf{u} \mathbf{d}) + \frac{1}{2} \|\mathbf{d} \mathbf{c}\|_{2}^{2}$

- ▶ Покажем, что $f(\mathbf{y}) \ge 0$ для всех $\mathbf{y} \in B$
- lacktriangle Пусть найдётся ${f u} \in B$ такая что $f({f u}) < 0$
- $\blacktriangleright \ f(\mathbf{u}) = (\mathbf{d} \mathbf{c})^\top (\mathbf{u} \tfrac{1}{2} (\mathbf{d} + \mathbf{c})) = (\mathbf{d} \mathbf{c})^\top (\mathbf{u} \mathbf{d}) + \tfrac{1}{2} \|\mathbf{d} \mathbf{c}\|_2^2$
- $(\mathbf{d} \mathbf{c})^{\top} (\mathbf{u} \mathbf{d}) < 0$

- ▶ Покажем, что $f(\mathbf{y}) \geq 0$ для всех $\mathbf{y} \in B$
- lacktriangle Пусть найдётся ${f u}\in B$ такая что $f({f u})<0$

$$f(\mathbf{u}) = (\mathbf{d} - \mathbf{c})^{\top} (\mathbf{u} - \frac{1}{2}(\mathbf{d} + \mathbf{c})) = (\mathbf{d} - \mathbf{c})^{\top} (\mathbf{u} - \mathbf{d}) + \frac{1}{2} \|\mathbf{d} - \mathbf{c}\|_2^2$$

- $(\mathbf{d} \mathbf{c})^{\top} (\mathbf{u} \mathbf{d}) < 0$
- Заметим, что

$$\left. \frac{d}{dt} \|\mathbf{d} - \mathbf{c} + t(\mathbf{u} - \mathbf{d})\|_{2}^{2} \right|_{t=0} = 2(\mathbf{d} - \mathbf{c})^{\top} (\mathbf{u} - \mathbf{d}) < 0$$

а значит для $t\in(0,1]$

$$\|\mathbf{d} - \mathbf{c} + t(\mathbf{u} - \mathbf{d})\|_2 \le \|\mathbf{d} - \mathbf{c}\|_2.$$

- ▶ Покажем, что $f(\mathbf{y}) \geq 0$ для всех $\mathbf{y} \in B$
- lacktriangle Пусть найдётся ${f u} \in B$ такая что $f({f u}) < 0$

$$f(\mathbf{u}) = (\mathbf{d} - \mathbf{c})^{\top} (\mathbf{u} - \frac{1}{2} (\mathbf{d} + \mathbf{c})) = (\mathbf{d} - \mathbf{c})^{\top} (\mathbf{u} - \mathbf{d}) + \frac{1}{2} \|\mathbf{d} - \mathbf{c}\|_2^2$$

- $(\mathbf{d} \mathbf{c})^{\top} (\mathbf{u} \mathbf{d}) < 0$
- Заметим, что

$$\left. \frac{d}{dt} \|\mathbf{d} - \mathbf{c} + t(\mathbf{u} - \mathbf{d})\|_{2}^{2} \right|_{t=0} = 2(\mathbf{d} - \mathbf{c})^{\top} (\mathbf{u} - \mathbf{d}) < 0$$

а значит для $t\in(0,1]$

$$\|\mathbf{d} - \mathbf{c} + t(\mathbf{u} - \mathbf{d})\|_2 \le \|\mathbf{d} - \mathbf{c}\|_2.$$

lacktriangle Точка $\mathbf{d}+t(\mathbf{u}-\mathbf{d})\in B$ ближе к \mathbf{c} , чем \mathbf{d} , противоречие.

- ▶ Покажем, что $f(\mathbf{y}) \geq 0$ для всех $\mathbf{y} \in B$
- lacktriangle Пусть найдётся ${f u}\in B$ такая что $f({f u})<0$

$$f(\mathbf{u}) = (\mathbf{d} - \mathbf{c})^{\top} (\mathbf{u} - \frac{1}{2} (\mathbf{d} + \mathbf{c})) = (\mathbf{d} - \mathbf{c})^{\top} (\mathbf{u} - \mathbf{d}) + \frac{1}{2} \|\mathbf{d} - \mathbf{c}\|_2^2$$

- $(\mathbf{d} \mathbf{c})^{\top} (\mathbf{u} \mathbf{d}) < 0$
- ▶ Заметим, что

$$\left. \frac{d}{dt} \|\mathbf{d} - \mathbf{c} + t(\mathbf{u} - \mathbf{d})\|_{2}^{2} \right|_{t=0} = 2(\mathbf{d} - \mathbf{c})^{\top} (\mathbf{u} - \mathbf{d}) < 0$$

а значит для $t \in (0,1]$

$$\|\mathbf{d} - \mathbf{c} + t(\mathbf{u} - \mathbf{d})\|_2 \le \|\mathbf{d} - \mathbf{c}\|_2.$$

- lacktriangle Точка $\mathbf{d}+t(\mathbf{u}-\mathbf{d})\in B$ ближе к \mathbf{c} , чем \mathbf{d} , противоречие.
- lacktriangle Случай $f(\mathbf{v}) \leq 0$ для всех $\mathbf{v} \in A$ аналогичен для -f

- ▶ Покажем, что $f(\mathbf{y}) \ge 0$ для всех $\mathbf{y} \in B$
- lacktriangle Пусть найдётся ${f u}\in B$ такая что $f({f u})<0$

$$f(\mathbf{u}) = (\mathbf{d} - \mathbf{c})^{\top} (\mathbf{u} - \frac{1}{2} (\mathbf{d} + \mathbf{c})) = (\mathbf{d} - \mathbf{c})^{\top} (\mathbf{u} - \mathbf{d}) + \frac{1}{2} \|\mathbf{d} - \mathbf{c}\|_2^2$$

- $(\mathbf{d} \mathbf{c})^{\top} (\mathbf{u} \mathbf{d}) < 0$
- ▶ Заметим, что

$$\left. \frac{d}{dt} \|\mathbf{d} - \mathbf{c} + t(\mathbf{u} - \mathbf{d})\|_{2}^{2} \right|_{t=0} = 2(\mathbf{d} - \mathbf{c})^{\top} (\mathbf{u} - \mathbf{d}) < 0$$

а значит для $t \in (0,1]$

$$\|\mathbf{d} - \mathbf{c} + t(\mathbf{u} - \mathbf{d})\|_2 \le \|\mathbf{d} - \mathbf{c}\|_2.$$

- ightharpoonup Точка $\mathbf{d} + t(\mathbf{u} \mathbf{d}) \in B$ ближе к \mathbf{c} , чем \mathbf{d} , противоречие.
- lacktriangle Случай $f(\mathbf{v}) \leq 0$ для всех $\mathbf{v} \in A$ аналогичен для -f

Q: верно ли обратное утверждение?

Теорема

Два выпуклых множества, одно из которых открыто, не пересекаются тогда и только когда, когда они отделимы.

Теорема

Два выпуклых множества, одно из которых открыто, не пересекаются тогда и только когда, когда они отделимы.

Доказательство

 \blacktriangleright Пусть два выпуклых множества A,B отделимы и B является открытым

Теорема

Два выпуклых множества, одно из которых открыто, не пересекаются тогда и только когда, когда они отделимы.

Доказательство

- \blacktriangleright Пусть два выпуклых множества A,B отделимы и B является открытым
- lacktriangle Тогда ${f a}^{ op}{f y}+b<0$ на всех ${f y}\in B$

Теорема

Два выпуклых множества, одно из которых открыто, не пересекаются тогда и только когда, когда они отделимы.

Доказательство

- \blacktriangleright Пусть два выпуклых множества A,B отделимы и B является открытым
- ▶ Тогда $\mathbf{a}^{\top}\mathbf{y} + b < 0$ на всех $\mathbf{y} \in B$
- lacktriangle Если ${f u}\in B$ и ${f a}^{ op}{f u}+b=0$, то в окрестности ${f u}$ нашлась бы точка ${f u}_+$, в которой ${f a}^{ op}{f u}_++b>0$

Критерий отделимости для выпуклых множеств

Теорема

Два выпуклых множества, одно из которых открыто, не пересекаются тогда и только когда, когда они отделимы.

- \blacktriangleright Пусть два выпуклых множества A,B отделимы и B является открытым
- ▶ Тогда $\mathbf{a}^{\top}\mathbf{y} + b < 0$ на всех $\mathbf{y} \in B$
- lacktriangle Если ${f u}\in B$ и ${f a}^{ op}{f u}+b=0$, то в окрестности ${f u}$ нашлась бы точка ${f u}_+$, в которой ${f a}^{ op}{f u}_++b>0$
- ▶ Тогда A и B не пересекаются, так как на элементах A выполнено \geq , а на элементах B выполнено <.

Лемма Фаркаша

Выполнено одно и только одно из следующих условий

- ▶ множество $\{ \mathbf{x} \mid \mathbf{A}\mathbf{x} = \mathbf{b}, \ \mathbf{x} \ge 0 \}$ непусто
- lacktriangle существует вектор f p такой что $f p^{ op} {f A} \ge 0$ и $f p^{ op} {f b} < 0$

Лемма Фаркаша

Выполнено одно и только одно из следующих условий

- ▶ множество $\{ \mathbf{x} \mid \mathbf{A}\mathbf{x} = \mathbf{b}, \ \mathbf{x} \ge 0 \}$ непусто
- ▶ существует вектор $\mathbf p$ такой что $\mathbf p^{\top} \mathbf A \geq 0$ и $\mathbf p^{\top} \mathbf b < 0$

Лемма Фаркаша

Выполнено одно и только одно из следующих условий

- ▶ множество $\{ \mathbf{x} \mid \mathbf{A}\mathbf{x} = \mathbf{b}, \ \mathbf{x} \ge 0 \}$ непусто
- ▶ существует вектор $\mathbf p$ такой что $\mathbf p^{\top} \mathbf A \geq 0$ и $\mathbf p^{\top} \mathbf b < 0$

Доказательство

▶ Первое условие означает, что ${\bf b}$ лежит в конусе C, образованном столбцами матрицы ${\bf A}=[{\bf a}_1,\dots,{\bf a}_m]$

Лемма Фаркаша

Выполнено одно и только одно из следующих условий

- ▶ множество $\{ \mathbf{x} \mid \mathbf{A}\mathbf{x} = \mathbf{b}, \ \mathbf{x} \ge 0 \}$ непусто
- ▶ существует вектор $\mathbf p$ такой что $\mathbf p^{\top} \mathbf A \geq 0$ и $\mathbf p^{\top} \mathbf b < 0$

- Первое условие означает, что b лежит в конусе C,
 образованном столбцами матрицы $\mathbf{A} = [\mathbf{a}_1, \dots, \mathbf{a}_m]$
- ► Если это не так, то существует гиперплоскость, которая *строго* отделяет конус от точки **b**:

$$\mathbf{c}^{\top}\mathbf{y} < d, \ \mathbf{y} \in C \quad \mathbf{c}^{\top}\mathbf{b} > d.$$

Лемма Фаркаша

Выполнено одно и только одно из следующих условий

- ▶ множество $\{ \mathbf{x} \mid \mathbf{A}\mathbf{x} = \mathbf{b}, \ \mathbf{x} \ge 0 \}$ непусто
- ▶ существует вектор \mathbf{p} такой что $\mathbf{p}^{\top}\mathbf{A} \geq 0$ и $\mathbf{p}^{\top}\mathbf{b} < 0$

Доказательство

- ▶ Первое условие означает, что ${f b}$ лежит в конусе C, образованном столбцами матрицы ${f A}=[{f a}_1,\dots,{f a}_m]$
- ► Если это не так, то существует гиперплоскость, которая *строго* отделяет конус от точки **b**:

$$\mathbf{c}^{\top} \mathbf{y} < d, \ \mathbf{y} \in C \quad \mathbf{c}^{\top} \mathbf{b} > d.$$

▶ Поскольку $0 \in C$, то d > 0. Также $\mathbf{a}_i \in C \Rightarrow \alpha \mathbf{a}_i \in C \alpha > 0$

Лемма Фаркаша

Выполнено одно и только одно из следующих условий

- ▶ множество $\{ \mathbf{x} \mid \mathbf{A}\mathbf{x} = \mathbf{b}, \ \mathbf{x} \ge 0 \}$ непусто
- ▶ существует вектор $\mathbf p$ такой что $\mathbf p^{\top} \mathbf A \geq 0$ и $\mathbf p^{\top} \mathbf b < 0$

- ▶ Первое условие означает, что ${\bf b}$ лежит в конусе C, образованном столбцами матрицы ${\bf A}=[{\bf a}_1,\ldots,{\bf a}_m]$
- ► Если это не так, то существует гиперплоскость, которая *строго* отделяет конус от точки **b**:

$$\mathbf{c}^{\top} \mathbf{y} < d, \ \mathbf{y} \in C \quad \mathbf{c}^{\top} \mathbf{b} > d.$$

- ▶ Поскольку $0 \in C$, то d>0. Также $\mathbf{a}_i \in C \Rightarrow \alpha \mathbf{a}_i \in C \ \alpha > 0$
- ▶ Значит $\mathbf{c}^{\top} \alpha \mathbf{a}_i < d \Rightarrow \mathbf{c}^{\top} \mathbf{a}_i < d/\alpha$. При $\alpha \to \infty$, $\mathbf{c}^{\top} \mathbf{a}_i \leq 0$

Лемма Фаркаша

Выполнено одно и только одно из следующих условий

- ▶ множество $\{ \mathbf{x} \mid \mathbf{A}\mathbf{x} = \mathbf{b}, \ \mathbf{x} \ge 0 \}$ непусто
- ▶ существует вектор $\mathbf p$ такой что $\mathbf p^{\top} \mathbf A \geq 0$ и $\mathbf p^{\top} \mathbf b < 0$

- Первое условие означает, что b лежит в конусе C,
 образованном столбцами матрицы $\mathbf{A} = [\mathbf{a}_1, \dots, \mathbf{a}_m]$
- ► Если это не так, то существует гиперплоскость, которая *строго* отделяет конус от точки **b**:

$$\mathbf{c}^{\top} \mathbf{y} < d, \ \mathbf{y} \in C \quad \mathbf{c}^{\top} \mathbf{b} > d.$$

- ▶ Поскольку $0 \in C$, то d>0. Также $\mathbf{a}_i \in C \Rightarrow \alpha \mathbf{a}_i \in C \ \alpha>0$
- ▶ Значит $\mathbf{c}^{\top} \alpha \mathbf{a}_i < d \Rightarrow \mathbf{c}^{\top} \mathbf{a}_i < d/\alpha$. При $\alpha \to \infty$, $\mathbf{c}^{\top} \mathbf{a}_i \leq 0$
- lacktriangle Таким образом, ${f p}=-{f c}$ и выполнено второе условие

Главное в первой части

▶ Сопряжённые конусы и геометрическая интерпретация

Главное в первой части

- ▶ Сопряжённые конусы и геометрическая интерпретация
- Самосопряжённые конусы

Главное в первой части

- ▶ Сопряжённые конусы и геометрическая интерпретация
- ▶ Самосопряжённые конусы
- ▶ Отделимость выпуклых множеств

Выпуклая функция (convex function)

Определение

```
Функция f: X \subset \mathbb{R}^n \to \mathbb{R} называется выпуклой (строго выпуклой), если X — выпуклое множество и для \forall \mathbf{x}_1, \mathbf{x}_2 \in X и \alpha \in [0,1] (\alpha \in (0,1)) выполнено: f(\alpha \mathbf{x}_1 + (1-\alpha)\mathbf{x}_2) \leq (<) \alpha f(\mathbf{x}_1) + (1-\alpha)f(\mathbf{x}_2)
```

Выпуклая функция (convex function)

Определение

```
Функция f: X \subset \mathbb{R}^n \to \mathbb{R} называется выпуклой (строго выпуклой), если X — выпуклое множество и для \forall \mathbf{x}_1, \mathbf{x}_2 \in X и \alpha \in [0,1] (\alpha \in (0,1)) выполнено: f(\alpha \mathbf{x}_1 + (1-\alpha)\mathbf{x}_2) \leq (<) \alpha f(\mathbf{x}_1) + (1-\alpha)f(\mathbf{x}_2)
```

Определение

Функция f вогнута (concave), если функция -f выпукла.

Выпуклая функция (convex function)

Определение

```
Функция f: X \subset \mathbb{R}^n \to \mathbb{R} называется выпуклой (строго выпуклой), если X — выпуклое множество и для \forall \mathbf{x}_1, \mathbf{x}_2 \in X и \alpha \in [0,1] (\alpha \in (0,1)) выполнено: f(\alpha \mathbf{x}_1 + (1-\alpha)\mathbf{x}_2) \leq (<) \alpha f(\mathbf{x}_1) + (1-\alpha)f(\mathbf{x}_2)
```

Определение

Функция f вогнута (concave), если функция -f выпукла.

Примеры выпуклых функций

- ▶ x^p для $x \ge 0$ и $p \ge 1$
- $ightharpoonup x \log x$, где x > 0
- $ightharpoonup \max\{x_1,\ldots,x_n\}$
- ▶ ||x||
- $ightharpoonup \log \left(\sum_{i=1}^n e^{x_i}\right)$
- ightharpoonup $-\log\det\mathbf{X}$ для $\mathbf{X}\in\mathbf{S}^n_{++}$

Определение

Множество $\operatorname{epi} f = \{(\mathbf{x},t) \in \mathbb{R}^{n+1} \mid t \geq f(\mathbf{x})\}$ называется надграфиком (эпиграфом) функции f.

Определение

Множество ${\rm epi}\; f=\{({\bf x},t)\in \mathbb{R}^{n+1}\;|\; t\geq f({\bf x})\}\;$ называется надграфиком (эпиграфом) функции f.

Теорема

Функция f выпукла $\Leftrightarrow \operatorname{epi} f$ выпуклое множество.

Определение

Множество $\operatorname{epi} f = \{(\mathbf{x},t) \in \mathbb{R}^{n+1} \mid t \geq f(\mathbf{x})\}$ называется надграфиком (эпиграфом) функции f.

Теорема

Функция f выпукла $\Leftrightarrow {
m epi} \ f$ выпуклое множество.

Определение

Множество $\operatorname{epi} f = \{(\mathbf{x},t) \in \mathbb{R}^{n+1} \mid t \geq f(\mathbf{x})\}$ называется надграфиком (эпиграфом) функции f.

Теорема

Функция f выпукла $\Leftrightarrow \operatorname{epi} f$ выпуклое множество.

Доказательство

1. Пусть f выпуклая функция

Определение

Множество $\operatorname{epi} f = \{(\mathbf{x},t) \in \mathbb{R}^{n+1} \mid t \geq f(\mathbf{x})\}$ называется надграфиком (эпиграфом) функции f.

Теорема

Функция f выпукла $\Leftrightarrow \operatorname{epi} f$ выпуклое множество.

- 1. Пусть f выпуклая функция
 - Рассмотрим две точки из эпиграфа (\mathbf{x}_1,t_1) и (\mathbf{x}_2,t_2) , где $t_1\geq f(\mathbf{x}_1)$ и $t_2\geq f(\mathbf{x}_2)$

Определение

Множество $\operatorname{epi} f = \{(\mathbf{x},t) \in \mathbb{R}^{n+1} \mid t \geq f(\mathbf{x})\}$ называется надграфиком (эпиграфом) функции f.

Теорема

Функция f выпукла $\Leftrightarrow \operatorname{epi} f$ выпуклое множество.

- 1. Пусть f выпуклая функция
 - Рассмотрим две точки из эпиграфа (\mathbf{x}_1,t_1) и (\mathbf{x}_2,t_2) , где $t_1\geq f(\mathbf{x}_1)$ и $t_2\geq f(\mathbf{x}_2)$
 - ▶ Проверим принадлежность эпиграфу точки $(\alpha \mathbf{x}_1 + (1 \alpha)\mathbf{x}_2, \alpha t_1 + (1 \alpha)t_2)$

Определение

Множество $\operatorname{epi} f = \{(\mathbf{x},t) \in \mathbb{R}^{n+1} \mid t \geq f(\mathbf{x})\}$ называется надграфиком (эпиграфом) функции f.

Теорема

Функция f выпукла $\Leftrightarrow \operatorname{epi} f$ выпуклое множество.

- 1. Пусть f выпуклая функция
 - Рассмотрим две точки из эпиграфа $({f x}_1,t_1)$ и $({f x}_2,t_2)$, где $t_1\geq f({f x}_1)$ и $t_2\geq f({f x}_2)$
 - ▶ Проверим принадлежность эпиграфу точки $(\alpha \mathbf{x}_1 + (1 \alpha)\mathbf{x}_2, \alpha t_1 + (1 \alpha)t_2)$
 - ▶ В силу выпуклости функции $\alpha t_1 + (1 \alpha)t_2 \ge \alpha f(\mathbf{x}_1) + (1 \alpha)f(\mathbf{x}_2) \ge f(\alpha \mathbf{x}_1 + (1 \alpha)\mathbf{x}_2).$

Определение

Множество ${
m epi}\ f=\{({\bf x},t)\in \mathbb{R}^{n+1}\ |\ t\geq f({\bf x})\}$ называется надграфиком (эпиграфом) функции f .

Теорема

Функция f выпукла $\Leftrightarrow \operatorname{epi} f$ выпуклое множество.

- 1. Пусть f выпуклая функция
 - Рассмотрим две точки из эпиграфа (\mathbf{x}_1,t_1) и (\mathbf{x}_2,t_2) , где $t_1\geq f(\mathbf{x}_1)$ и $t_2\geq f(\mathbf{x}_2)$
 - ▶ Проверим принадлежность эпиграфу точки $(\alpha \mathbf{x}_1 + (1 \alpha)\mathbf{x}_2, \alpha t_1 + (1 \alpha)t_2)$
 - ▶ В силу выпуклости функции $\alpha t_1 + (1 \alpha)t_2 \ge \alpha f(\mathbf{x}_1) + (1 \alpha)f(\mathbf{x}_2) \ge f(\alpha \mathbf{x}_1 + (1 \alpha)\mathbf{x}_2).$
- 2. Пусть надграфик выпуклое множество

Определение

Множество $\operatorname{epi} f = \{(\mathbf{x},t) \in \mathbb{R}^{n+1} \mid t \geq f(\mathbf{x})\}$ называется надграфиком (эпиграфом) функции f.

Теорема

Функция f выпукла $\Leftrightarrow \operatorname{epi} f$ выпуклое множество.

- 1. Пусть f выпуклая функция
 - Рассмотрим две точки из эпиграфа (\mathbf{x}_1,t_1) и (\mathbf{x}_2,t_2) , где $t_1\geq f(\mathbf{x}_1)$ и $t_2\geq f(\mathbf{x}_2)$
 - ▶ Проверим принадлежность эпиграфу точки $(\alpha \mathbf{x}_1 + (1 \alpha)\mathbf{x}_2, \alpha t_1 + (1 \alpha)t_2)$
 - ▶ В силу выпуклости функции $\alpha t_1 + (1 \alpha)t_2 \ge \alpha f(\mathbf{x}_1) + (1 \alpha)f(\mathbf{x}_2) \ge f(\alpha \mathbf{x}_1 + (1 \alpha)\mathbf{x}_2).$
- 2. Пусть надграфик выпуклое множество
 - $lack (\mathbf x_1, f(\mathbf x_1))$ и $(\mathbf x_2, f(x_2)) \in \operatorname{epi} f$, то $(\alpha \mathbf x_1 + (1-lpha)\mathbf x_2, lpha f(\mathbf x_1) + (1-lpha)f(\mathbf x_2)) \in \operatorname{epi} f$

Определение

Множество ${
m epi}\ f=\{({\bf x},t)\in \mathbb{R}^{n+1}\ |\ t\geq f({\bf x})\}$ называется надграфиком (эпиграфом) функции f.

Теорема

Функция f выпукла \Leftrightarrow $\mathrm{epi}\ f$ выпуклое множество.

- 1. Пусть f выпуклая функция
 - Рассмотрим две точки из эпиграфа (\mathbf{x}_1,t_1) и (\mathbf{x}_2,t_2) , где $t_1\geq f(\mathbf{x}_1)$ и $t_2\geq f(\mathbf{x}_2)$
 - ▶ Проверим принадлежность эпиграфу точки $(\alpha \mathbf{x}_1 + (1 \alpha)\mathbf{x}_2, \alpha t_1 + (1 \alpha)t_2)$
 - ▶ В силу выпуклости функции $\alpha t_1 + (1 \alpha)t_2 > \alpha f(\mathbf{x}_1) + (1 \alpha)f(\mathbf{x}_2) > f(\alpha \mathbf{x}_1 + (1 \alpha)\mathbf{x}_2).$
- 2. Пусть надграфик выпуклое множество
 - $\mathbf{x}_1, f(\mathbf{x}_1)$) и $(\mathbf{x}_2, f(x_2)) \in \text{ері } f$, то $(\alpha \mathbf{x}_1 + (1 \alpha) \mathbf{x}_2, \alpha f(\mathbf{x}_1) + (1 \alpha) f(\mathbf{x}_2)) \in \text{ері } f$
 - ightharpoonup Из определения надграфика следует выпуклость f

Сильно выпуклая функция (strongly convex function)

Определение

Функция $f: X \subset \mathbb{R}^n \to \mathbb{R}$ называется сильно выпуклой с константой m>0, если X — выпуклое множество и для $\forall \mathbf{x}_1, \mathbf{x}_2 \in X$ и $\alpha \in [0,1]$ выполнено: $f(\alpha \mathbf{x}_1 + (1-\alpha)\mathbf{x}_2) \leq \alpha f(\mathbf{x}_1) + (1-\alpha)f(\mathbf{x}_2) - \frac{m}{2}\alpha(1-\alpha)\|\mathbf{x}_1 - \mathbf{x}_2\|_2^2$

Сильно выпуклая функция (strongly convex function)

Определение

Функция $f:X\subset\mathbb{R}^n\to\mathbb{R}$ называется сильно выпуклой с константой m>0, если X — выпуклое множество и для $\forall \mathbf{x}_1,\mathbf{x}_2\in X$ и $\alpha\in[0,1]$ выполнено: $f(\alpha\mathbf{x}_1+(1-\alpha)\mathbf{x}_2)\leq \alpha f(\mathbf{x}_1)+(1-\alpha)f(\mathbf{x}_2)-\frac{m}{2}\alpha(1-\alpha)\|\mathbf{x}_1-\mathbf{x}_2\|_2^2$

▶ Выпуклость ⊃ строгая выпуклость ⊃ сильная выпуклость

Сильно выпуклая функция (strongly convex function)

Определение

Функция $f:X\subset\mathbb{R}^n o\mathbb{R}$ называется сильно выпуклой с константой m>0, если X — выпуклое множество и для $\forall \mathbf{x}_1, \mathbf{x}_2 \in X$ и $\alpha \in [0, 1]$ выполнено:

$$f(\alpha \mathbf{x}_1 + (1-\alpha)\mathbf{x}_2) \le \alpha f(\mathbf{x}_1) + (1-\alpha)f(\mathbf{x}_2) - \frac{m}{2}\alpha(1-\alpha)\|\mathbf{x}_1 - \mathbf{x}_2\|_2^2$$

- Выпуклость \supset строгая выпуклость \supset сильная выпуклость
- Для сильно выпуклых функций многие утверждения о методах оказываются более сильными, чем просто для выпуклых функций: пример о сходимости градиентного спуска.

Будем считать выпуклую функцию сильно выпуклой с m=0.

Будем считать выпуклую функцию сильно выпуклой с m=0.

Теорема

Пусть функция $f(\mathbf{x})$ дифференцируема и определена на выпуклом множестве $X\subseteq \mathbb{R}^n$. Тогда $f(\mathbf{x})$ сильно выпукла с константой $m\geq 0$ в том и только том случае, если

$$f(\mathbf{x}) - f(\mathbf{x}^*) \ge \langle f'(\mathbf{x}^*), \mathbf{x} - \mathbf{x}^* \rangle + \frac{m}{2} \|\mathbf{x} - \mathbf{x}^*\|_2^2, \quad \forall \mathbf{x}, \mathbf{x}^* \in X.$$

Будем считать выпуклую функцию сильно выпуклой с m=0.

Теорема

Пусть функция $f(\mathbf{x})$ дифференцируема и определена на выпуклом множестве $X\subseteq\mathbb{R}^n$. Тогда $f(\mathbf{x})$ сильно выпукла с константой $m\geq 0$ в том и только том случае, если

$$f(\mathbf{x}) - f(\mathbf{x}^*) \ge \langle f'(\mathbf{x}^*), \mathbf{x} - \mathbf{x}^* \rangle + \frac{m}{2} \|\mathbf{x} - \mathbf{x}^*\|_2^2, \quad \forall \mathbf{x}, \mathbf{x}^* \in X.$$

Доказательство: пусть f выпукла

По определению:

$$f(\alpha \mathbf{x}_1 + (1 - \alpha)\mathbf{x}_2) \le \alpha f(\mathbf{x}_1) + (1 - \alpha)f(\mathbf{x}_2)$$

Будем считать выпуклую функцию сильно выпуклой с m=0.

Теорема

Пусть функция $f(\mathbf{x})$ дифференцируема и определена на выпуклом множестве $X\subseteq\mathbb{R}^n$. Тогда $f(\mathbf{x})$ сильно выпукла с константой $m\geq 0$ в том и только том случае, если

$$f(\mathbf{x}) - f(\mathbf{x}^*) \ge \langle f'(\mathbf{x}^*), \mathbf{x} - \mathbf{x}^* \rangle + \frac{m}{2} \|\mathbf{x} - \mathbf{x}^*\|_2^2, \quad \forall \mathbf{x}, \mathbf{x}^* \in X.$$

Доказательство: пусть f выпукла

По определению:

$$f(\alpha \mathbf{x}_1 + (1 - \alpha)\mathbf{x}_2) \le \alpha f(\mathbf{x}_1) + (1 - \alpha)f(\mathbf{x}_2)$$

Перепишем в виде

$$\begin{split} f(\mathbf{x}_2 + \alpha(\mathbf{x}_1 - \mathbf{x}_2)) &\leq f(\mathbf{x}_2) + \alpha(f(\mathbf{x}_1) - f(\mathbf{x}_2)) \text{ или} \\ \frac{f(\mathbf{x}_2 + \alpha(\mathbf{x}_1 - \mathbf{x}_2)) - f(\mathbf{x}_2)}{\alpha} &\leq f(\mathbf{x}_1) - f(\mathbf{x}_2) \end{split}$$

Будем считать выпуклую функцию сильно выпуклой с m=0.

Теорема

Пусть функция $f(\mathbf{x})$ дифференцируема и определена на выпуклом множестве $X\subseteq\mathbb{R}^n$. Тогда $f(\mathbf{x})$ сильно выпукла с константой $m\geq 0$ в том и только том случае, если

$$f(\mathbf{x}) - f(\mathbf{x}^*) \ge \langle f'(\mathbf{x}^*), \mathbf{x} - \mathbf{x}^* \rangle + \frac{m}{2} \|\mathbf{x} - \mathbf{x}^*\|_2^2, \quad \forall \mathbf{x}, \mathbf{x}^* \in X.$$

Доказательство: пусть f выпукла

По определению:

$$f(\alpha \mathbf{x}_1 + (1 - \alpha)\mathbf{x}_2) \le \alpha f(\mathbf{x}_1) + (1 - \alpha)f(\mathbf{x}_2)$$

Перепишем в виде

$$f(\mathbf{x}_2 + lpha(\mathbf{x}_1 - \mathbf{x}_2)) \le f(\mathbf{x}_2) + lpha(f(\mathbf{x}_1) - f(\mathbf{x}_2))$$
 или $\dfrac{f(\mathbf{x}_2 + lpha(\mathbf{x}_1 - \mathbf{x}_2)) - f(\mathbf{x}_2)}{2} \le f(\mathbf{x}_1) - f(\mathbf{x}_2)$

▶ При $\alpha \to 0$ получим

$$\langle f'(\mathbf{x}_2), \mathbf{x}_1 - \mathbf{x}_2 \rangle < f(\mathbf{x}_1) - f(\mathbf{x}_2)$$

▶ Рассмотрим $\mathbf{z} = \alpha \mathbf{x}_1 + (1 - \alpha) \mathbf{x}_2$

- ▶ Рассмотрим $\mathbf{z} = \alpha \mathbf{x}_1 + (1 \alpha) \mathbf{x}_2$
- ightharpoonup Запишем два неравенства для \mathbf{z}, \mathbf{x}_1 и \mathbf{z}, \mathbf{x}_2

$$f(\mathbf{x}_1) \ge f(\mathbf{z}) + \langle f'(\mathbf{z}), \mathbf{z} - \mathbf{x}_1 \rangle \mid \cdot \alpha$$

$$f(\mathbf{x}_2) \ge f(\mathbf{z}) + \langle f'(\mathbf{z}), \mathbf{z} - \mathbf{x}_2 \rangle \mid \cdot (1 - \alpha)$$

- ▶ Рассмотрим $\mathbf{z} = \alpha \mathbf{x}_1 + (1 \alpha) \mathbf{x}_2$
- ightharpoonup Запишем два неравенства для \mathbf{z}, \mathbf{x}_1 и \mathbf{z}, \mathbf{x}_2

$$f(\mathbf{x}_1) \ge f(\mathbf{z}) + \langle f'(\mathbf{z}), \mathbf{z} - \mathbf{x}_1 \rangle \mid \cdot \alpha$$

$$f(\mathbf{x}_2) \ge f(\mathbf{z}) + \langle f'(\mathbf{z}), \mathbf{z} - \mathbf{x}_2 \rangle \mid \cdot (1 - \alpha)$$

Сложим эти равенства

$$f(\mathbf{z}) = f(\alpha \mathbf{x}_1 + (1 - \alpha)\mathbf{x}_2) \le \alpha f(\mathbf{x}_1) + (1 - \alpha)f(\mathbf{x}_2)$$

- ▶ Рассмотрим $\mathbf{z} = \alpha \mathbf{x}_1 + (1 \alpha) \mathbf{x}_2$
- ightharpoonup Запишем два неравенства для \mathbf{z}, \mathbf{x}_1 и \mathbf{z}, \mathbf{x}_2

$$f(\mathbf{x}_1) \ge f(\mathbf{z}) + \langle f'(\mathbf{z}), \mathbf{z} - \mathbf{x}_1 \rangle \mid \cdot \alpha$$

$$f(\mathbf{x}_2) \ge f(\mathbf{z}) + \langle f'(\mathbf{z}), \mathbf{z} - \mathbf{x}_2 \rangle \mid \cdot (1 - \alpha)$$

Сложим эти равенства

$$f(\mathbf{z}) = f(\alpha \mathbf{x}_1 + (1 - \alpha)\mathbf{x}_2) \le \alpha f(\mathbf{x}_1) + (1 - \alpha)f(\mathbf{x}_2)$$

Доказательство: пусть выполнено неравенство

- ▶ Рассмотрим $\mathbf{z} = \alpha \mathbf{x}_1 + (1 \alpha) \mathbf{x}_2$
- ightharpoonup Запишем два неравенства для \mathbf{z}, \mathbf{x}_1 и \mathbf{z}, \mathbf{x}_2

$$f(\mathbf{x}_1) \ge f(\mathbf{z}) + \langle f'(\mathbf{z}), \mathbf{z} - \mathbf{x}_1 \rangle \mid \cdot \alpha$$

$$f(\mathbf{x}_2) \ge f(\mathbf{z}) + \langle f'(\mathbf{z}), \mathbf{z} - \mathbf{x}_2 \rangle \mid \cdot (1 - \alpha)$$

Сложим эти равенства

$$f(\mathbf{z}) = f(\alpha \mathbf{x}_1 + (1 - \alpha)\mathbf{x}_2) \le \alpha f(\mathbf{x}_1) + (1 - \alpha)f(\mathbf{x}_2)$$

Сильно выпуклый случай

Для случая сильной выпуклости необходимо применить аналогичные выкладки к функции $f(\mathbf{x}) - \frac{m}{2} \|\mathbf{x}\|_2^2.$

Доказательство: пусть выполнено неравенство

- ▶ Рассмотрим $\mathbf{z} = \alpha \mathbf{x}_1 + (1 \alpha) \mathbf{x}_2$
- ightharpoonup Запишем два неравенства для \mathbf{z}, \mathbf{x}_1 и \mathbf{z}, \mathbf{x}_2

$$f(\mathbf{x}_1) \ge f(\mathbf{z}) + \langle f'(\mathbf{z}), \mathbf{z} - \mathbf{x}_1 \rangle \mid \cdot \alpha$$

$$f(\mathbf{x}_2) \ge f(\mathbf{z}) + \langle f'(\mathbf{z}), \mathbf{z} - \mathbf{x}_2 \rangle \mid \cdot (1 - \alpha)$$

Сложим эти равенства

$$f(\mathbf{z}) = f(\alpha \mathbf{x}_1 + (1 - \alpha)\mathbf{x}_2) \le \alpha f(\mathbf{x}_1) + (1 - \alpha)f(\mathbf{x}_2)$$

Сильно выпуклый случай

Для случая сильной выпуклости необходимо применить аналогичные выкладки к функции $f(\mathbf{x}) - \frac{m}{2} \|\mathbf{x}\|_2^2$.

Упражнение

Покажите, что f сильно выпукла $\Leftrightarrow f(\mathbf{x}) - rac{m}{2} \|\mathbf{x}\|_2^2$ выпукла.

Дважды непрерывно дифференцируемая функция f выпукла $\Leftrightarrow f''(\mathbf{x}) \succeq m\mathbf{I}$

Дважды непрерывно дифференцируемая функция f выпукла $\Leftrightarrow f''(\mathbf{x}) \succeq m\mathbf{I}$

Доказательство

$$f(\mathbf{y}) = f(\mathbf{x}) + \langle f'(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle + \frac{1}{2} \langle \mathbf{y} - \mathbf{x}, f''(\mathbf{x}_{\alpha})(\mathbf{y} - \mathbf{x}) \rangle$$

Дважды непрерывно дифференцируемая функция f выпукла $\Leftrightarrow f''(\mathbf{x}) \succeq m\mathbf{I}$

Доказательство

Разложение в ряд Тейлора до второго порядка

$$f(\mathbf{y}) = f(\mathbf{x}) + \langle f'(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle + \frac{1}{2} \langle \mathbf{y} - \mathbf{x}, f''(\mathbf{x}_{\alpha})(\mathbf{y} - \mathbf{x}) \rangle$$

▶ Если $f''(\mathbf{x}) \succeq m\mathbf{I}$, то $\frac{1}{2} \langle \mathbf{y} - \mathbf{x}, f''(\mathbf{x}_{\alpha})(\mathbf{y} - \mathbf{x}) \rangle \geq \frac{m}{2} \|\mathbf{y} - \mathbf{x}\|_2^2$, и по критерию первого порядка f выпукла

Дважды непрерывно дифференцируемая функция f выпукла $\Leftrightarrow f''(\mathbf{x}) \succeq m\mathbf{I}$

Доказательство

$$f(\mathbf{y}) = f(\mathbf{x}) + \langle f'(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle + \frac{1}{2} \langle \mathbf{y} - \mathbf{x}, f''(\mathbf{x}_{\alpha})(\mathbf{y} - \mathbf{x}) \rangle$$

- ▶ Если $f''(\mathbf{x}) \succeq m\mathbf{I}$, то $\frac{1}{2} \langle \mathbf{y} \mathbf{x}, f''(\mathbf{x}_{\alpha})(\mathbf{y} \mathbf{x}) \rangle \geq \frac{m}{2} \|\mathbf{y} \mathbf{x}\|_2^2$, и по критерию первого порядка f выпукла
- ightharpoonup Если найдётся точка ${f z}$ такая, что $f''({f z}) \not\succeq m{f I}$, тогда найдётся направление ${f d}$ такое, что ${f d}^{ op}f''({f z}){f d} \le m$

Дважды непрерывно дифференцируемая функция f выпукла $\Leftrightarrow f''(\mathbf{x}) \succeq m\mathbf{I}$

Доказательство

$$f(\mathbf{y}) = f(\mathbf{x}) + \langle f'(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle + \frac{1}{2} \langle \mathbf{y} - \mathbf{x}, f''(\mathbf{x}_{\alpha})(\mathbf{y} - \mathbf{x}) \rangle$$

- ▶ Если $f''(\mathbf{x}) \succeq m\mathbf{I}$, то $\frac{1}{2} \langle \mathbf{y} \mathbf{x}, f''(\mathbf{x}_{\alpha})(\mathbf{y} \mathbf{x}) \rangle \geq \frac{m}{2} \|\mathbf{y} \mathbf{x}\|_2^2$, и по критерию первого порядка f выпукла
- ▶ Если найдётся точка \mathbf{z} такая, что $f''(\mathbf{z}) \not\succeq m\mathbf{I}$, тогда найдётся направление \mathbf{d} такое, что $\mathbf{d}^{\top}f''(\mathbf{z})\mathbf{d} \leq m$
- ▶ В таком случае в силу критерия первого порядка f невыпукла противоречие

Дважды непрерывно дифференцируемая функция f выпукла $\Leftrightarrow f''(\mathbf{x}) \succeq m\mathbf{I}$

Доказательство

$$f(\mathbf{y}) = f(\mathbf{x}) + \langle f'(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle + \frac{1}{2} \langle \mathbf{y} - \mathbf{x}, f''(\mathbf{x}_{\alpha})(\mathbf{y} - \mathbf{x}) \rangle$$

- ▶ Если $f''(\mathbf{x}) \succeq m\mathbf{I}$, то $\frac{1}{2} \langle \mathbf{y} \mathbf{x}, f''(\mathbf{x}_{\alpha})(\mathbf{y} \mathbf{x}) \rangle \geq \frac{m}{2} \|\mathbf{y} \mathbf{x}\|_2^2$, и по критерию первого порядка f выпукла
- ▶ Если найдётся точка \mathbf{z} такая, что $f''(\mathbf{z}) \not\succeq m\mathbf{I}$, тогда найдётся направление \mathbf{d} такое, что $\mathbf{d}^{\top}f''(\mathbf{z})\mathbf{d} \leq m$
- ▶ В таком случае в силу критерия первого порядка f невыпукла противоречие

Дважды непрерывно дифференцируемая функция f выпукла $\Leftrightarrow f''(\mathbf{x}) \succeq m\mathbf{I}$

Доказательство

Разложение в ряд Тейлора до второго порядка

$$f(\mathbf{y}) = f(\mathbf{x}) + \langle f'(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle + \frac{1}{2} \langle \mathbf{y} - \mathbf{x}, f''(\mathbf{x}_{\alpha})(\mathbf{y} - \mathbf{x}) \rangle$$

- ▶ Если $f''(\mathbf{x}) \succeq m\mathbf{I}$, то $\frac{1}{2} \langle \mathbf{y} \mathbf{x}, f''(\mathbf{x}_{\alpha})(\mathbf{y} \mathbf{x}) \rangle \geq \frac{m}{2} \|\mathbf{y} \mathbf{x}\|_2^2$, и по критерию первого порядка f выпукла
- ▶ Если найдётся точка \mathbf{z} такая, что $f''(\mathbf{z}) \not\succeq m\mathbf{I}$, тогда найдётся направление \mathbf{d} такое, что $\mathbf{d}^{\top}f''(\mathbf{z})\mathbf{d} \leq m$
- ▶ В таком случае в силу критерия первого порядка f невыпукла противоречие

Напоминание

Для проверки определённости матрицы нужно использовать определение или критерий Сильвестра

lacktriangle Если $f(\mathbf{x})$ — выпукла, то $g(\mathbf{x}) = f(\mathbf{A}\mathbf{x} + \mathbf{b})$ также выпукла

- lacktriangle Если $f(\mathbf{x})$ выпукла, то $g(\mathbf{x}) = f(\mathbf{A}\mathbf{x} + \mathbf{b})$ также выпукла
- lacktriangle Если $f(\mathbf{x})$ выпукла, то $g(t)=f(\mathbf{x}+t\mathbf{y})$ также выпукла

- lacktriangle Если $f(\mathbf{x})$ выпукла, то $g(\mathbf{x}) = f(\mathbf{A}\mathbf{x} + \mathbf{b})$ также выпукла
- lacktriangle Если $f(\mathbf{x})$ выпукла, то $g(t)=f(\mathbf{x}+t\mathbf{y})$ также выпукла
- lacktriangle Если f_i выпуклы, то $\max_{i=1,\dots,m} f_i$ также выпукла

- lacktriangle Если $f(\mathbf{x})$ выпукла, то $g(\mathbf{x}) = f(\mathbf{A}\mathbf{x} + \mathbf{b})$ также выпукла
- lacktriangle Если $f({f x})$ выпукла, то $g(t)=f({f x}+t{f y})$ также выпукла
- lacktriangle Если f_i выпуклы, то $\max_{i=1,\dots,m} f_i$ также выпукла
- Сумма выпуклых функций с неотрицательными коэффициентами — выпуклая функция

- lacktriangle Если $f(\mathbf{x})$ выпукла, то $g(\mathbf{x}) = f(\mathbf{A}\mathbf{x} + \mathbf{b})$ также выпукла
- lacktriangle Если $f(\mathbf{x})$ выпукла, то $g(t)=f(\mathbf{x}+t\mathbf{y})$ также выпукла
- lacktriangle Если f_i выпуклы, то $\max_{i=1,\dots,m} f_i$ также выпукла
- Сумма выпуклых функций с неотрицательными коэффициентами — выпуклая функция
- lacktriangle Скалярная композиция $h(f(\mathbf{x}))$

Теорема

Если f выпуклая функция и \mathbf{x}^* локальный минимум, то \mathbf{x}^* — глобальный минимум.

Теорема

Если f выпуклая функция и \mathbf{x}^* локальный минимум, то \mathbf{x}^* — глобальный минимум.

Доказательство от противного

lacktriangle Пусть $\mathbf{y}^*
eq \mathbf{x}^*$ — глобальный минимум: $f(\mathbf{y}^*) < f(\mathbf{x}^*)$

Теорема

Если f выпуклая функция и \mathbf{x}^* локальный минимум, то \mathbf{x}^* — глобальный минимум.

Доказательство от противного

- lacktriangle Пусть $\mathbf{y}^*
 eq \mathbf{x}^*$ глобальный минимум: $f(\mathbf{y}^*) < f(\mathbf{x}^*)$
- ▶ По определению локального минимума: $f(\mathbf{x}^*) \leq f(\mathbf{x})$, где $\|\mathbf{x}^* \mathbf{x}\|_2 \leq \delta$

Теорема

Если f выпуклая функция и \mathbf{x}^* локальный минимум, то \mathbf{x}^* — глобальный минимум.

Доказательство от противного

- lacktriangle Пусть $\mathbf{y}^*
 eq \mathbf{x}^*$ глобальный минимум: $f(\mathbf{y}^*) < f(\mathbf{x}^*)$
- ▶ По определению локального минимума: $f(\mathbf{x}^*) \leq f(\mathbf{x})$, где $\|\mathbf{x}^* \mathbf{x}\|_2 \leq \delta$
- $f extbf{ }$ Выберем достаточно малое $lpha\in(0,1)$ и рассмотрим точку ${f z}=(1-lpha){f x}^*+lpha{f y}^*$ такую что $\|{f z}-{f x}^*\|_2\leq\delta$

Теорема

Если f выпуклая функция и \mathbf{x}^* локальный минимум, то \mathbf{x}^* — глобальный минимум.

Доказательство от противного

- lacktriangle Пусть $\mathbf{y}^*
 eq \mathbf{x}^*$ глобальный минимум: $f(\mathbf{y}^*) < f(\mathbf{x}^*)$
- ▶ По определению локального минимума: $f(\mathbf{x}^*) \leq f(\mathbf{x})$, где $\|\mathbf{x}^* \mathbf{x}\|_2 \leq \delta$
- $f extbf{ }$ Выберем достаточно малое $lpha\in(0,1)$ и рассмотрим точку ${f z}=(1-lpha){f x}^*+lpha{f y}^*$ такую что $\|{f z}-{f x}^*\|_2\leq\delta$
- $f(\mathbf{x}^*) \le f(\mathbf{z}) \le \alpha f(\mathbf{y}^*) + (1 \alpha) f(\mathbf{x}^*) < f(\mathbf{x}^*)$

Определение

Определение

Множество $C^n = \{ \mathbf{A} \in \mathbf{S}^n \mid \mathbf{x}^\top \mathbf{A} \mathbf{x} \ge 0, \ \mathbf{x} \ge 0 \}$ называется copositive cone.

 $ightharpoonup \mathcal{C}^n$ выпукло

Определение

- $ightharpoonup \mathcal{C}^n$ выпукло
- $ightharpoonup \mathbf{S}_+^n \subset \mathcal{C}^n$

Определение

- $ightharpoonup \mathcal{C}^n$ выпукло
- $ightharpoonup \mathbf{S}_+^n \subset \mathcal{C}^n$
- lacktriangle Задача проверки $\mathbf{X}
 ot\in \mathcal{C}^n$ является со-NP полной!

Определение

- $ightharpoonup \mathcal{C}^n$ выпукло
- $ightharpoonup \mathbf{S}_+^n \subset \mathcal{C}^n$
- ▶ Задача проверки $\mathbf{X} \not\in \mathcal{C}^n$ является со-NP полной!
- ightharpoonup Задача конической оптимизации с конусом \mathcal{C}^n является NP-трудной

Определение

- $ightharpoonup \mathcal{C}^n$ выпукло
- $ightharpoonup \mathbf{S}_+^n \subset \mathcal{C}^n$
- ▶ Задача проверки $\mathbf{X} \not\in \mathcal{C}^n$ является со-NP полной!
- ightharpoonup Задача конической оптимизации с конусом \mathcal{C}^n является NP-трудной

Определение

Множество $C^n = \{ \mathbf{A} \in \mathbf{S}^n \mid \mathbf{x}^\top \mathbf{A} \mathbf{x} \ge 0, \ \mathbf{x} \ge 0 \}$ называется copositive cone.

- $ightharpoonup \mathcal{C}^n$ выпукло
- $ightharpoonup \mathbf{S}_+^n \subset \mathcal{C}^n$
- ▶ Задача проверки $\mathbf{X} \notin \mathcal{C}^n$ является со-NP полной!
- ightharpoonup Задача конической оптимизации с конусом \mathcal{C}^n является NP-трудной

Пример

Задача определения максимального независимого множества вершин графа сводится к задаче оптимизации на множестве \mathcal{C}^n . Подробности тут

Дана матрица $\mathbf{Q} \in \mathbf{S}^n_{++}$.

$$\min_{\mathbf{x} \in \mathbb{R}^n} \mathbf{x}^{\top} \mathbf{Q} \mathbf{x}$$

Дана матрица $\mathbf{Q} \in \mathbf{S}^n_{++}.$

$$\min_{\mathbf{x} \in \mathbb{R}^n} \mathbf{x}^\top \mathbf{Q} \mathbf{x}$$
 s.t. $\|\mathbf{x}\|_2 = 1$

▶ Допустимое множество невыпукло

Дана матрица $\mathbf{Q} \in \mathbf{S}^n_{++}.$

$$\min_{\mathbf{x} \in \mathbb{R}^n} \mathbf{x}^{\top} \mathbf{Q} \mathbf{x}$$
 s.t. $\|\mathbf{x}\|_2 = 1$

- ▶ Допустимое множество невыпукло
- ▶ Целевая функция выпукла

Дана матрица $\mathbf{Q} \in \mathbf{S}^n_{++}$.

$$\min_{\mathbf{x} \in \mathbb{R}^n} \mathbf{x}^\top \mathbf{Q} \mathbf{x}$$
 s.t. $\|\mathbf{x}\|_2 = 1$

- ▶ Допустимое множество невыпукло
- Целевая функция выпукла

 ${f Q}$: какая интерпретация у ${f x}^*$ и $f({f x}^*)$?

Теорема

Если функция
$$f$$
 выпукла, то $f\left(\sum\limits_{i=1}^k \alpha \mathbf{x}_i\right) \leq \sum\limits_{i=1}^k \alpha_i f(\mathbf{x}_i)$, где

$$\sum_{i=1}^{k} \alpha_i = 1, \ \alpha_i \ge 0.$$

Теорема

Если функция
$$f$$
 выпукла, то $f\left(\sum\limits_{i=1}^k \alpha \mathbf{x}_i\right) \leq \sum\limits_{i=1}^k \alpha_i f(\mathbf{x}_i)$, где $\sum\limits_{i=1}^k \alpha_i = 1, \; \alpha_i \geq 0.$

Доказательство по индукции

lacktriangle База k=2 выполнена в силу определения

Теорема

Если функция
$$f$$
 выпукла, то $f\left(\sum\limits_{i=1}^k \alpha \mathbf{x}_i\right) \leq \sum\limits_{i=1}^k \alpha_i f(\mathbf{x}_i)$, где $\sum\limits_{i=1}^k \alpha_i = 1, \; \alpha_i \geq 0.$

Доказательство по индукции

- ▶ База k=2 выполнена в силу определения
- ▶ Пусть неравенство выполнено для k=m-1: $f\left(\sum_{i=1}^{m-1}\alpha\mathbf{x}_i\right)\leq\sum_{i=1}^{m-1}\alpha_if(\mathbf{x}_i)\text{ и }\sum_{i=1}^{m-1}\alpha_i=1\text{, }\alpha_i\geq0$

Теорема

Если функция
$$f$$
 выпукла, то $f\left(\sum\limits_{i=1}^k \alpha \mathbf{x}_i\right) \leq \sum\limits_{i=1}^k \alpha_i f(\mathbf{x}_i)$, где $\sum\limits_{i=1}^k \alpha_i = 1, \; \alpha_i \geq 0.$

Доказательство по индукции

- ▶ База k=2 выполнена в силу определения

Рассмотрим
$$k=m$$
: $f\left(\sum\limits_{i=1}^{m}\hat{\alpha}_{i}\mathbf{x}_{i}\right)=f\left(\sum\limits_{i=1}^{m-1}\hat{\alpha}\mathbf{x}_{i}+\hat{\alpha}_{m}\mathbf{x}_{m}\right)=f\left((1-\hat{\alpha}_{m})\sum\limits_{i=1}^{m-1}\frac{\hat{\alpha}_{i}}{1-\hat{\alpha}_{m}}\mathbf{x}_{i}+\hat{\alpha}_{m}\mathbf{x}_{m}\right)\leq$
$$(1-\hat{\alpha}_{m})f\left(\sum\limits_{i=1}^{m-1}\frac{\hat{\alpha}_{i}}{1-\hat{\alpha}_{m}}\mathbf{x}_{i}\right)+\hat{\alpha}_{m}f(\mathbf{x}_{m})\leq\sum_{i=1}^{k}\alpha_{i}f(\mathbf{x}_{i})$$

Следствия и обобщения

lacktriangle Запись неравенства Йенсена для функции $-\log x$

$$\frac{1}{m} \sum_{i=1}^{m} x_i \ge \sqrt[m]{x_1 \cdot \ldots \cdot x_m}$$

Следствия и обобщения

lacktriangle Запись неравенства Йенсена для функции $-\log x$

$$\frac{1}{m} \sum_{i=1}^{m} x_i \ge \sqrt[m]{x_1 \cdot \ldots \cdot x_m}$$

Неравенство Гёльдера

$$\sum_{i=1}^{n} x_i y_i \le \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p} \left(\sum_{i=1}^{n} |y_i|^q\right)^{1/q}$$

Следствия и обобщения

lacktriangle Запись неравенства Йенсена для функции $-\log x$

$$\frac{1}{m} \sum_{i=1}^{m} x_i \ge \sqrt[m]{x_1 \cdot \ldots \cdot x_m}$$

Неравенство Гёльдера

$$\sum_{i=1}^{n} x_i y_i \le \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p} \left(\sum_{i=1}^{n} |y_i|^q\right)^{1/q}$$

 ▶ Обобщение на непрерывный случай даёт неравенство для выпуклой функции от матожидания

$$f(\mathbb{E}(\mathbf{x})) \le \mathbb{E}(f(\mathbf{x}))$$

Главное во второй части

Выпуклые функции и их свойства

Главное во второй части

- Выпуклые функции и их свойства
- ▶ Операции, сохраняющие выпуклость

Главное во второй части

- Выпуклые функции и их свойства
- ▶ Операции, сохраняющие выпуклость
- Неравенство Йенсена