Методы оптимизации. Семинар 7. Двойственность.

Корнилов Никита Максимович

Московский физико-технический институт

15 октября 2024г

Сопряженное множество

Выпуклые множества, содержащие точку 0, допускают двойственное описание: множество может быть описано набором векторов, являющихся нормалями опорных гиперплоскостей.

Сопряженное множество

Выпуклые множества, содержащие точку 0, допускают двойственное описание: множество может быть описано набором векторов, являющихся нормалями опорных гиперплоскостей.

Definition

Пусть $X \subseteq \mathbb{R}^n$ – произвольное непустое множество. Тогда множество

$$X^* = \{y \in \mathbb{R}^n \mid \langle y, x \rangle \geq -1 \quad \forall x \in X\}$$

называется *сопряженным* (*двойственным*) к X.

Самосопряженное множество

Definition

Множества X_1 и X_2 называются взаимосопряженными если

$$X_1^* = X_2 \quad \text{и} \quad X_2^* = X_1$$

Definition

Множество X называется самосопряженным, если

$$X^* = X$$

Definition

Множество

$$X^{**} = \{x \in \mathbb{R}^n \mid \langle y, x \rangle \geq -1 \quad \forall y \in X^*\}$$

называется *вторым сопряженным* к X.

Свойства сопряженных множеств

$$X^* = \{y \in \mathbb{R}^n \mid \langle y, x \rangle \geq -1 \quad \forall x \in X\}$$

- X^* всегда замкнуто, выпукло и содержит точку 0_n

Свойства сопряженных множеств

$$X^* = \{y \in \mathbb{R}^n \mid \langle y, x \rangle \geq -1 \quad \forall x \in X\}$$

- X^* всегда замкнуто, выпукло и содержит точку 0_n
- Для произвольного множества $X \subset \mathbb{R}^n$

$$X^{**} = cl(conv(X \cup \{0_n\}))$$

Свойства сопряженных множеств

$$X^* = \{y \in \mathbb{R}^n \mid \langle y, x \rangle \geq -1 \quad \forall x \in X\}$$

- X^* всегда замкнуто, выпукло и содержит точку 0_n
- Для произвольного множества $X \subset \mathbb{R}^n$

$$X^{**} = cl(conv(X \cup \{0_n\}))$$

- Если множество $X \subset \mathbb{R}^n$ – выпуклое и замкнутое, а также содержит точку 0_n, то $X^{**} = X$

Свойства сопряженных множеств

$$X^* = \{y \in \mathbb{R}^n \mid \langle y, x \rangle \geq -1 \quad \forall x \in X\}$$

- X^* всегда замкнуто, выпукло и содержит точку 0_n
- Для произвольного множества $X \subset \mathbb{R}^n$

$$X^{**} = cl(conv(X \cup \{0_n\}))$$

- Если множество $X \subset \mathbb{R}^n$ – выпуклое и замкнутое, а также содержит точку 0_n, то $X^{**} = X$
- Если $X_1 \subset X_2$, то $X_2^* \subset X_1^*$

Свойства сопряженных множеств

$$X^* = \{y \in \mathbb{R}^n \mid \langle y, x \rangle \geq -1 \quad \forall x \in X\}$$

- X^* всегда замкнуто, выпукло и содержит точку 0_n
- Для произвольного множества $X \subset \mathbb{R}^n$

$$X^{**} = cl(conv(X \cup \{0_n\}))$$

- Если множество $X \subset \mathbb{R}^n$ – выпуклое и замкнутое, а также содержит точку 0_n, то $X^{**} = X$
- Если $X_1 \subset X_2$, то $X_2^* \subset X_1^*$
- $\left(\bigcup_{i=1}^m X_i\right)^* = \bigcap_{i=1}^m X_i^*$
- $X^* = (clX)^* = (conv(X))^*$

Двойственная норма

Definition

Пусть $\|\cdot\|$ – норма в прямом пространстве. Тогда двойственная норма:

$$\|y\|_* = \sup_{\|x\| \leq 1} \langle x, y \rangle.$$

Двойственная норма

Definition

Пусть $\|\cdot\|$ – норма в прямом пространстве. Тогда двойственная норма:

$$\|y\|_* = \sup_{\|x\|\leq 1} \langle x, y \rangle.$$

Прямая и двойственная норма связаны следующим образом: пусть $\|\cdot\|_p$ – это прямая ℓ_p-норма, а двойственная – это ℓ_q-норма, при этом q находится из уравнения $1/p + 1/q = 1$.

Двойственная норма

Definition

Пусть $\|\cdot\|$ – норма в прямом пространстве. Тогда двойственная норма:

$$\|y\|_* = \sup_{\|x\|\leq 1} \langle x, y\rangle.$$

Прямая и двойственная норма связаны следующим образом: пусть $\|\cdot\|_p$ – это прямая ℓ_p-норма, а двойственная – это ℓ_q-норма, при этом q находится из уравнения $1/p + 1/q = 1$.
Неравенство Коши-Буняковского-Шварца (КБШ):

$$|\langle x, y\rangle| \leq \|x\| \cdot \|y\|_*.$$

Шар

Example

Пусть $B(0,r) = \{x \in \mathbb{R}^n : \|x\| \le r\}$ - замкнутый шар с центром 0 и радиусом r. Найдите $(B(0,r))^*$.

Решение.

Рассмотрим по определению скалярное произведение и применим неравенство КБШ

$$\langle x, y\rangle \ge -\|x\|\|y\|_* \ge -r\|y\|_*.$$

Если $\|y\|_* \le \frac{1}{r}$, то $\langle x, y\rangle \ge -1$ для любых $x \in B(0,r)$. То есть мы доказали, что

$$\left\{ y \in \mathbb{R}^n : \|y\|_* \le \frac{1}{r} \right\} \subseteq (B(0,r))^*.$$

Докажем включение в обратную сторону от противного.

Шар

Example

Если $\|y\|_* > \frac{1}{r}$, то по определению двойственной нормы существует $e \in \mathbb{R}^n$ с $\|e\| \leq 1$ и

$$\langle y, e \rangle > \frac{1}{r}$$

Взяв $x = \frac{-er}{\|e\|} \in B(0, r)$, получим

$$\langle x, y \rangle = \left\langle \frac{-er}{\|e\|}, y \right\rangle = \frac{r}{\|e\|} \langle -e, y \rangle < -\frac{r}{\|e\|} \frac{1}{r} \leq -\frac{1}{\|e\|} \leq -1.$$

Следовательно, $\langle x, y \rangle < -1$ и $y \notin (B(0, r))^*$.

Сопряженный конус

Proposition

Пусть K – конус в $\mathbb{R}^n$. Тогда

$$K^* = \{y \in \mathbb{R}^n \mid \langle y, x \rangle \geq 0 \quad \forall x \in K\}$$

Сопряженный конус

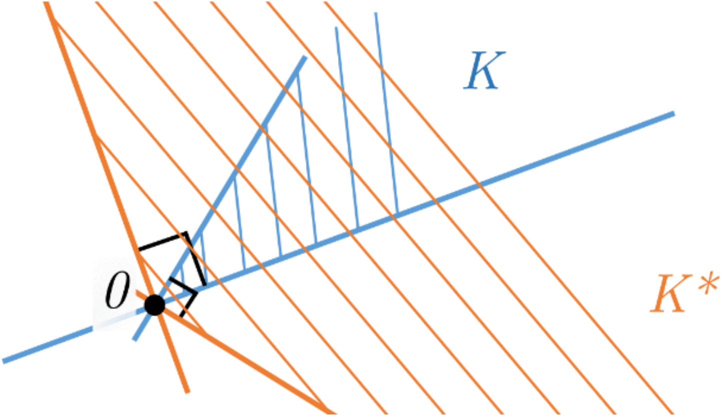
Proposition

Пусть K *– конус в* $\mathbb{R}^n$*. Тогда*

$$K^* = \{y \in \mathbb{R}^n \mid \langle y, x \rangle \geq 0 \quad \forall x \in K\}$$

Часто утверждение 1 понимают как определение сопряженного конуса. Для сопряженных конусов верны все те же свойства, что и для сопряженных множеств (0 всегда принадлежит конусу).

Геометрическая интерпретация



Доп свойства сопряженных конусов

- Для произвольного множества S и конуса K верно

$$(S+K)^* = S^* \cup K^*.$$

- Пусть $K_1, \dots, K_m$ конусы, тогда

$$\left(\sum_{i=1}^{m} K_i\right)^* = \cup_{i=1}^{m} K_i^*.$$

- Пусть $K_1, \dots, K_m$ конусы, а их пересечение имеет внутреннюю точку. Тогда

$$(\cup_{i=1}^{m} K_i)^* = \sum_{i=1}^{m} K_i^*.$$

Примеры на конусы

Example

Дан неотрицательный ортант $K = \mathbb{R}^n_+$. Найдите K^*.

Доказательство.

По определению

$$K^* = \{y \mid \langle x, y\rangle \geq 0 \quad \forall x \succcurlyeq 0\}.$$

Следовательно $K^* = \{y \mid y \succcurlyeq 0\}$. То есть $K^* = \mathbb{R}^n_+ = K$. □

Примеры на конусы

Example

Дано множество симметричных положительно полуопределенных матриц $\mathcal{S}_+^n$. Найдите $(\mathcal{S}_+^n)^*$ в пространстве $\mathcal{S}^n$.

Example

Дано множество симметричных положительно полуопределенных матриц $\mathcal{S}_+^n$. Найдите $(\mathcal{S}_+^n)^*$ в пространстве $\mathcal{S}^n$.

Proof.

На множестве симметричных матриц мы рассматриваем стандартное скалярное произведение $\mathrm{Tr}(XY) = \sum_{i,j=1^n} X_{ij} Y_{ij}$. По определению

$$K^* = \{Y \in \mathcal{S}^n \mid \langle X, Y \rangle \geq 0 \quad \forall X \in \mathcal{S}_+^n\}.$$

Пусть $Y \in \mathcal{S}^n$. Рассмотрим произвольную матрицу $X \in \mathcal{S}_+^n$. Разложим матрицу X в ортонормированном базисе из собственных векторов: $X = \sum_{i=1}^n \lambda_i q_i q_i^\top$, $\lambda_i \geq 0$ для всех $i = 1, \ldots, n$. Следовательно, имеем

$$\mathrm{Tr}(YX) = \mathrm{Tr}\left(Y \sum_{i=1}^n \lambda_i q_i q_i^\top \right) = \sum_{i=1}^n \lambda_i q_i^\top Y q_i.$$

Примеры на конусы

Example

Если все значения $q_i^\top Y q_i$ будут неотрицательными, то неотрицательность останется. Значит, если $Y \in \mathcal{S}_n^n$ то $Y \in (\mathcal{S}_+^n)^*$. Докажем включение и в другую сторону.

Примеры на конусы

Example

Если все значения $q_i^\top Y q_i$ будут неотрицательными, то неотрицательность останется. Значит, если $Y \in \mathcal{S}_n^n$ то $Y \in (\mathcal{S}_+^n)^*$.
Докажем включение и в другую сторону.
Рассмотрим $Y \notin \mathcal{S}_+^n$. Так как матрица Y не является положительно полуопределенной, то существует вектор $q \in \mathbb{R}^n$:

$$q^\top Y q = \mathrm{Tr}(qq^\top Y) < 0$$

Введем $X = qq^\top \in \mathcal{S}^n$. То есть $\mathrm{Tr}(YX) \leq 0$, значит, $Y \notin (\mathcal{S}_+^n)^*$.

Примеры на конусы

Example (Конус Лоренца)

Дан конус $K = \{(x, t) \in \mathbb{R}^{n+1} : \quad \|x\| \leq t\}$. Найдите K^*.

Примеры на конусы

Example (Конус Лоренца)

Дан конус $K = \{(x, t) \in \mathbb{R}^{n+1} : \quad \|x\| \leq t\}$. Найдите K^*.

Proof.

По определению двойственного конуса имеем

$$K^* = \{(u, v) \in \mathbb{R}^{n+1} \mid \langle x, u \rangle + tv \geq 0 \quad \forall x : \|x\| \leq t\}.$$

Докажем, что это множество совпадает с $\{(u, v) \in \mathbb{R}^{n+1} \mid \|u\|_* \leq v\}$ то есть нужно показать

$$\langle x, u \rangle + tv \Longleftrightarrow \|u\|_* \leq v \tag{1}$$

Example

Покажем, что из правой части (1) следует левая. Предположим, что $\|u\|_* \leq v$ и $\|x\| \leq t$ для некоторого $t > 0$. Применяя определение двойственной нормы, а также $\|-x/t\| \leq 1$, получаем

$$\langle u, -x/t\rangle \leq \|u\|_* \leq v,$$

и поэтому $\langle u, x\rangle + tv \geq 0$.

Example

Покажем, что из правой части (1) следует левая. Предположим, что $\|u\|_* \leq v$ и $\|x\| \leq t$ для некоторого $t > 0$. Применяя определение двойственной нормы, а также $\|-x/t\| \leq 1$, получаем

$$\langle u, -x/t \rangle \leq \|u\|_* \leq v,$$

и поэтому $\langle u, x \rangle + tv \geq 0$.
Теперь покажем, что из левой части (1) следует правая. Предположим, что $\|u\|_* > v$, то есть что правая часть (1) не выполняется. Тогда по определению двойственной нормы существует x с $\|x\| \leq 1$ и $\langle x, u \rangle > v$. Беря $t = 1$, получаем

$$\langle u, -x \rangle + v < 0,$$

что противоречит левой части (1).

Сопряженная функция

Позволяет описывать выпуклые замкнутые функции с помощью расстояния до прямой вдоль y через 0.

Сопряженная функция

Позволяет описывать выпуклые замкнутые функции с помощью расстояния до прямой вдоль y через 0.

Definition

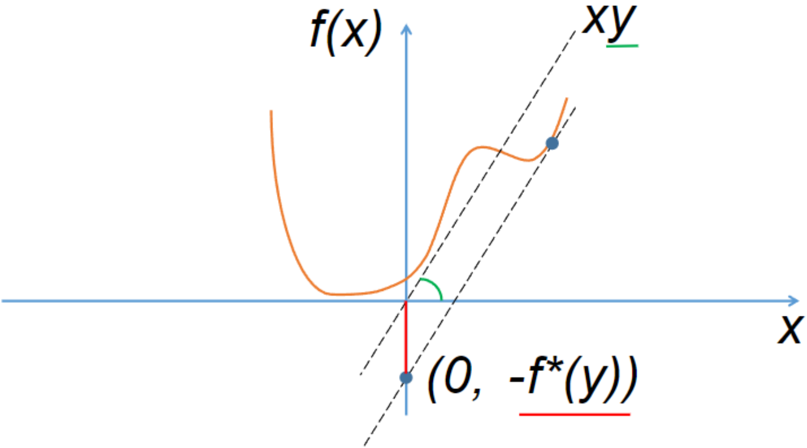
Пусть $f : \mathbb{R}^n \to \mathbb{R}$. Функция $f^* : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$

$$f^*(y) = \sup_{x \in \mathbb{R}^n} \{\langle x, y \rangle - f(x)\}$$

называется сопряженной функцией к f.

Эффективной областью определения $f^*(y)$ является множество, на котором $\sup_{x \in \mathbb{R}^n} \{\langle x, y \rangle - f(x)\} < +\infty$.

Геометрия сопряженной функции



Дополнительные определения

Definition

Функция называется замкнутой, если её надграфик является замкнутым множеством

Замкнутость функции равносильна её полунепрерывности снизу:

$$\underline{\lim}_{k\to\infty} f(x_k) \geq f(x_0)$$

для любых $x_k \to x_0$. Непрерывные функции, очевидно, являются полунепрерывными снизу.

Definition

Функция называется собственной, если она не принимает значение $-\infty$ ни в какой точке

Свойства сопряженных функций

$$f^*(y) = \sup_{x \in \mathbb{R}^n} \{\langle x, y \rangle - f(x)\}$$

- f^* замкнутая выпуклая функция.
- Функция $f^{**} = f$ если и только если f — выпуклая, замкнутая, собственная функция
- Пусть f — замкнутая, собственная функция. Тогда следующие два утверждения равносильны при $\mu > 0$:
  1. f является μ-сильно выпуклой
  2. f^* имеет $1/\mu$-липшицев градиент или f^* — $1/\mu$-гладкая

Fenchel–Young inequality

- Пусть f — произвольная функция:

$$f(x) + f^*(y) \geq \langle x, y \rangle, \quad \forall x, y \in \mathbb{R}^n$$

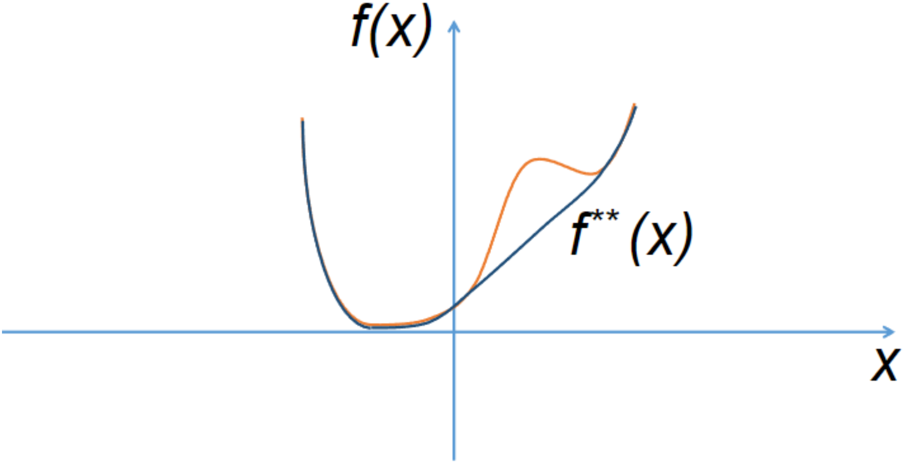
- Равенство достигается только и только если

$$f(x) + f^*(p) = \langle x, p \rangle \longleftrightarrow p \in \partial f(x)$$

- Следствие из Fenchel–Young

$$f(x) \geq f^{**}(x) \quad \forall x \in \mathbb{R}^n$$

Двойное сопряжение



Свойства сопряженных функций 2

- Если $f(x, y) = f_1(x) + f_2(y)$, где f_1, f_2 - выпуклые, то

$$f^*(p, q) = f_1^*(p) + f_2^*(q).$$

Примеры на сопряженные функции

По определению $f^*(y)$ найти y, где функция принимает конечные значения. Для них найти max по известным правилам.
Если $f(x)$ - выпукла, то $\langle x, y\rangle - f(x)$ - вогнута по x.

Example

Найти сопряженную функцию к линейной функции $f(x) = \langle a, x\rangle + b$, где $x \in \mathbb{R}^n$.

Примеры на сопряженные функции

По определению $f^*(y)$ найти y, где функция принимает конечные значения. Для них найти max по известным правилам.
Если $f(x)$ - выпукла, то $\langle x, y\rangle - f(x)$ - вогнута по x.

Example

Найти сопряженную функцию к линейной функции $f(x) = \langle a, x\rangle + b$, где $x \in \mathbb{R}^n$.

Доказательство.

$$f^*(y) = \sup_{x\in\mathbb{R}^n}\{\langle y, x\rangle - \langle a, x\rangle - b\} = \sup_{x\in\mathbb{R}^n}\{\langle y - a, x\rangle - b\}.$$

Величина $\langle y - a, x\rangle - b$ как функция по x ограничена в том и только в том случае, когда $y = a$, в этом случае она является константой, равной $-b$. Тогда получаем, что сопряженная функция $f^*(y) = -b$ с областью определения $domf^* = \{a\}$. $\square$

Примеры на сопряженные функции

Example

Найти сопряженную функцию к экспоненте $f(x) = e^x$, где $x \in \mathbb{R}$.

Примеры на сопряженные функции

Example

Найти сопряженную функцию к экспоненте $f(x) = e^x$, где $x \in \mathbb{R}$.

Proof.

По определению сопряженной функции

$$f^*(y) = \sup_{x \in \mathbb{R}} \{xy - e^x\}.$$

Дифференцируем $xy - e^x$ по x и приравниваем к нулю:

$$y - e^x = 0.$$

Такое возможно только при $y > 0$, а именно $xy - e^x$ достигает своего максимуму в точке $x = \log y$. Поэтому $f^* = y \log y - y$. Остальные случаи рассматриваем отдельно:

При $y < 0$ функция $xy - e^x$ не ограничена.

При $y = 0$, $f^*(y) = \sup_{x \in \mathbb{R}} -e^x = 0$.

Example

Итого сопряженная функция

$$f^*(y) = \begin{cases} y \log y - y & , \quad y \in \mathbb{R}_+ \\ +\infty & , \quad \text{иначе} \end{cases}$$

с областью определения $dom f^*(y) = \mathbb{R}_+$ (мы доопределили $0 \log 0 = 0$)

Примеры на сопряженные функции

Example

Логистическая функция. Найти сопряженную функцию для $f(x) = \log(1 + e^x)$, $x \in \mathbb{R}$

Примеры на сопряженные функции

Example

Логистическая функция. Найти сопряженную функцию для
$f(x) = \log(1 + e^x)$, $x \in \mathbb{R}$

Proof.

По определению сопряженной функции

$$f^*(y) = \sup_{x \in \mathbb{R}} \{xy - \log(1 + e^x)\}. \tag{2}$$

Беря производную от $xy - \log(1 + e^x)$ по x и приравнивая градиент к 0, получаем

$$x = \log y - \log(1 - y).$$

Эта формула корректно определена только при $0 < y < 1$. Поскольку функция $xy - \log(1 + e^x)$ вогнутая по x, то найденное значение — это и есть супремум. Тогда $f^*(y) = y \log y + (1 - y) \log(1 - y)$. Отстальные случаи рассмотрим отдельно.

Example

Рассмотрим случай, когда $y < 0$. Покажем, что в этом случае выражение $xy - \log(1 + e^x)$ как функция по x будет не ограничено при $x \to -\infty$. Действительно, из монотонности логарифма и того, что $e^x < 1$ при $x < 0$ следует, $\log(1 + e^x) < \log 2$ для всех $x < 0$. Поэтому $xy - \log(1 + e^x) > xy - \log 2$. Поскольку $yx \to \infty$ при $x \to -\infty$, то $xy - \log(1 + e^x) \to \infty$ при $x \to -\infty$. Таким образом, супремум равен ∞.

Пусть теперь $y > 1$. Аналогичные рассуждения дают неравенство $\log(1 + e^x) < \log(e^x + e^x) = \log 2 + x$ при $x > 0$. Отсюда $xy - \log(1 + e^x) > (y - 1)x - \log 2$ для всех $x > 0$. Устремляя $x \to \infty$, получаем, что супремум 2 равен ∞.

Example

Пусть теперь $y = 0$. Поскольку $\log(1 + e^x) \geq 0$ для всех $x \in \mathbb{R}$ и $\log(1 + e^x) \to 0$ при $x \to -\infty$, то супремум 2 равен 0. Значит, $f^*(y) = 0$.

Пусть $y = 1$, покажем, что в этом случае супремум так же равен нулю. Из неравенства $\log(1 + e^x) \geq x$ для всех $x \in \mathbb{R}$ следует, что супремум не может быть больше нуля. Он равен нулю, поскольку $\log(1 + e^x) = x + \log(1 + e^{-x})$ для всех $x \in \mathbb{R}$ и $\ln(1 + e^{-x}) \to 0$ при $x \to \infty$.
Итого имеем

$$f^*(y) = \begin{cases} y \log y + (1 - y)\log(1 - y) & , y \in [0, 1] \\ +\infty & , \text{иначе} \end{cases}.$$

Примеры на сопряженные функции 2

Example (Лог-детерминант)

Найти сопряженную функцию для $f(X) = -\log\det X$ на $X \in \mathcal{S}^n_{++}$.

Примеры на сопряженные функции 2

Example (Лог-детерминант)

Найти сопряженную функцию для $f(X) = -\log\det X$ на $X \in \mathcal{S}^n_{++}$.

Proof. По определению сопряженная функция

$$f^*(Y) = \sup_{X \succ 0} \{\mathrm{Tr}(XY) + \log\det X^{-1}\},$$

где $\mathrm{Tr}(XY)$ – стандартное скалярное произведелние на $\mathcal{S}^n$.

Вычисляя градиент под супремумом по X и приравнивая его к нулю, получаем

$$\nabla_X(\mathrm{Tr}(XY) + \log\det X) = Y + X^{-1} = 0,$$

значит, $X = -Y^{-1}$, а поскольку X является положительной определенной матрицей, то $Y \prec 0$. В этом случае

$$f^*(Y) = \log\det(-Y)^{-1} - n.$$

Example

Покажем, что если $Y \not\preceq 0$, то супремум равен бесконечности. Если $Y \not\preceq 0$, то Y имеет собственный вектор v с $\|v\|_2 = 1$ и собственным значением $\lambda \geq 0$. Возьмем $X = I + tvv^\top$, тогда

$$\begin{aligned} \mathrm{Tr}(XY) + \log\det X^{-1} &= \mathrm{Tr}(Y) + t\lambda + \log\det(I + tvv^\top) \\ &= \mathrm{Tr}(Y) + t\lambda + \log(1+t). \end{aligned}$$

То есть супремум равен бесконечности при $t \to \infty$.
Область определения $domf^* = -\mathcal{S}^n_{++}$.

Примеры на сопряженные функции 2

Example (Норма)

Найти сопряженную функцию для произвольной нормы $f(x) = \|x\|$ на $x \in \mathbb{R}^n$.

Примеры на сопряженные функции 2

Example (Норма)

Найти сопряженную функцию для произвольной нормы $f(x) = \|x\|$ на $x \in \mathbb{R}^n$.

Proof.

Если $\|y\|_* > 1$, тогда по определению двойственной нормы существует $z \in \mathbb{R}^n$ с $\|z\| \leq 1$ и $y^\top z > 1$. Беря $x = tz$ и устремляя $t \to \infty$, получаем

$$y^\top x - \|x\| = t(y^\top z - \|z\|) \to \infty.$$

То есть $y^\top x - \|x\|$ не ограничено.

Пусть теперь $\|y\|_* \leq 1$, тогда $\langle y, x \rangle \leq \|x\| \|y\|_*$ для всех $x \in \mathbb{R}^n$. Тогда

$$y^\top x - \|x\| \leq 0.$$

При $x = 0$, выражение $y^\top x - \|x\| = 0$, то есть $f^*(y) = 0$.

Итого $f^*(y)$ – это индикатор множества $\{\|y\|_* \leq 1\}$.