Методы оптимизации. Семинар 8. Сопряженные функции.

Корнилов Никита Максимович

Московский физико-технический институт

23 октября 2025г

Сопряженная функция

Позволяет описывать функции с помощью максимального расстояния до прямой с углом наклона y.

Definition

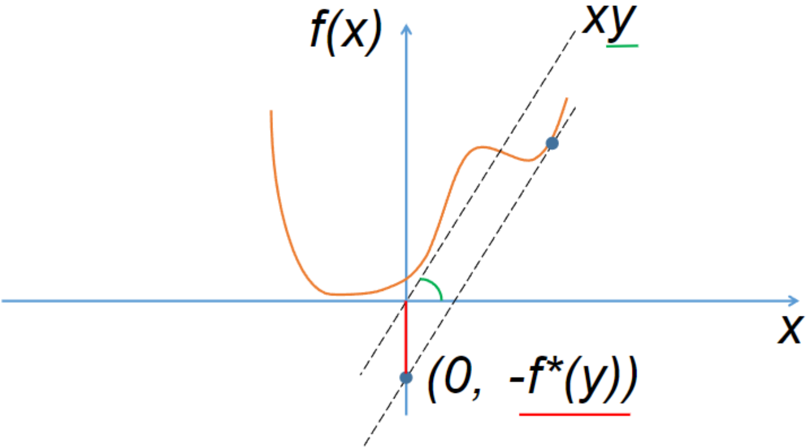
Пусть $f : \mathbb{R}^n \to \overline{\mathbb{R}}$. Функция $f^* : \mathbb{R}^n \to \overline{\mathbb{R}}$

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

называется сопряженной по Фенхелю функцией к f.

Сопряженная функция всегда выпуклая независимо от выпуклости f!

Геометрия сопряженной функции



Примеры сопряженных функций

Попробуем посчитать по определению $f^*(y)$. Если $f(x)$ - выпукла, то $\langle x, y\rangle - f(x)$ - вогнута по x, так что можно пользоваться критериями глобального максимума.

Example

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

Примеры сопряженных функций

Попробуем посчитать по определению $f^*(y)$. Если $f(x)$ - выпукла, то $\langle x, y\rangle - f(x)$ - вогнута по x, так что можно пользоваться критериями глобального максимума.

Example

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

Proof.

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

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 $domf^* = \{a\}$.

1D сопряженные функции

Example

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

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

1D сопряженные функции

Example

Итого сопряженная функция

$$f^*(y) = \begin{cases} y \log y - y, & y \geq 0, \\ +\infty, & y < 0, \end{cases}$$

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

Общий алгоритм подсчета f^*

Для выпуклых и дифференцируемых функций:

1. Для каждого y посчитать градиент $\langle x, y\rangle - f(x)$ по x.
2. Посмотреть, для каких y градиент можно приравнять к 0 и найти глоб максимум по x.
3. Для остальных y надо смотреть, расходится ли супремум или сходится к конечному значения.

*Для субдифференцируемых функций можно считать субдифференциал для $\sup_x\{\langle x, y\rangle - f(x)\} = -\inf_x\{f(x) - \langle x, y\rangle\}$.
**Для сложно дифференцируемых функций, можно исследовать супремум по определению или свойствам сопряженных функций.

1D сопряженные функции

Example

Найти сопряженную функцию для $f(x) = \max\{1 - x, 0\}$, $x \in \mathbb{R}$.

1D сопряженные функции

Example (Логистическая функция)

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

Example (Логистическая функция)

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

Proof.

По определению сопряженной функции

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

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 $f^*(y) = y \log y + (1-y)\log(1-y)$. Остальные случаи рассмотрим отдельно.

1D сопряженные функции

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 $x \to -\infty$. То есть супремум равен $+\infty$.

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 (1) равен $+\infty$.

1D сопряженные функции

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 (1) равен 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 $\log(1 + e^x) = x + \log(1 + e^{-x}), \forall x \in \mathbb{R}$ и $\ln(1 + e^{-x}) \to 0$ при $x \to \infty$. Итого имеем

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

Example (Квадратичная функция)

Найти сопряженную функцию для $f(x) = \frac{1}{2}\langle x, Ax\rangle + \langle b, x\rangle$ для $A \in \mathbb{S}^n_{++}$ и $x, b \in \mathbb{R}^n$.

Многомерные сопряженные функции

Example (Лог-детерминант)

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

Многомерные сопряженные функции

Example (Лог-детерминант)

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

Proof. По определению сопряженная функция

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

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

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 $\mathrm{dom}\, f^* = -\mathbb{S}^n_{++}$.

Многомерные сопряженные функции

Example (Норма)

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

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 $\{\|y\|_* \leq 1\}$.

Example

Найти сопряженную функцию для $f(x) = \frac{\|x\|^2}{2}$ на $x \in \mathbb{R}^n$.

Свойства сопряженных функций

- Пусть дан набор собственных функций $f_i : \mathbb{R}^n \to \overline{\mathbb{R}}, i \in \overline{1, m}$ с сопряженными функциями $f_i^*, i \in \overline{1, m}$. Тогда для функции $g : \mathbb{R}^n \times \cdots \times \mathbb{R}^n \to \overline{\mathbb{R}}$, заданной по правилу

$$g(x_1, \ldots, x_m) = \sum_{i=1}^{m} f_i(x_i),$$

сопряженная функция g^* считается как

$$g^*(y_1, \ldots, y_m) = \sum_{i=1}^{n} f_i^*(y_i).$$

Example

Найдите сопряженную функцию к $f : \mathbb{R}^n \to \mathbb{R}$

$$f(x) = \sum_{i=1}^{n} x_i \log x_i, \quad \mathrm{dom\ f} = \{x \in \mathbb{R}^n | x \geq 0\}.$$

Свойства сопряженных функций

- Пусть $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ и $\alpha > 0$. Тогда для функций $g(x) = \alpha f(x)$ и $h(x) = \alpha f(\frac{x}{\alpha})$ сопряженные функции считаются как

$$g^*(y) = \alpha f^*(\frac{y}{\alpha}), \quad h^*(y) = \alpha f^*(y).$$

Example

Найти сопряженную функцию для $f(x) = \frac{\alpha\|x\|^2}{2}$ на $x \in \mathbb{R}^n$.

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.

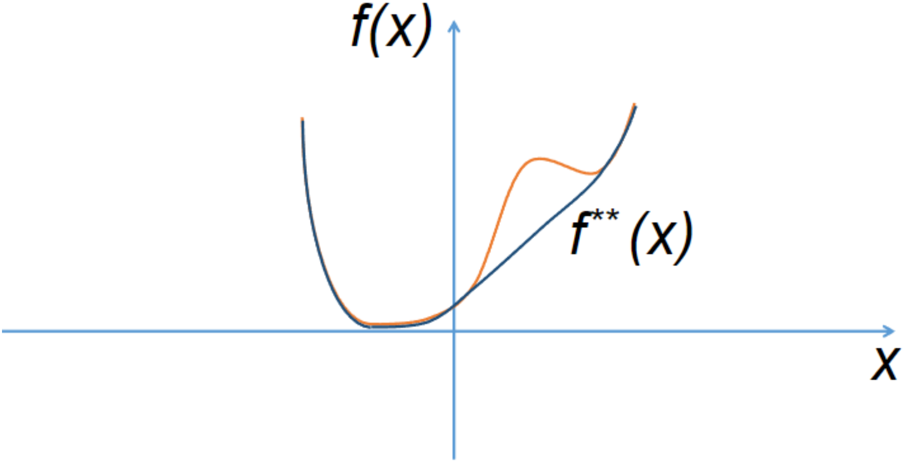
Свойства сопряженных функций

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

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

- 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.

Двойное сопряжение



Fenchel–Young inequality

- Пусть f — произвольная функция, тогда:

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

- Равенство достигается только и только если

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

- Следствие из Fenchel–Young

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

Example

Можно показать, что для $p > 1$, $\frac{1}{p} + \frac{1}{q} = 1$ и $\forall x, y \in \mathbb{R}$ верно

$$\frac{|x|^p}{p} + \frac{|y|^q}{q} \geq xy.$$

Infimal convolution

Для двух собственных функций f_1 и f_2 определим их свертку как

$$(f_1 \Box f_2)(x) := \inf_z \{f_1(z) + f_2(x - z)\}.$$

- Эта операция порождает функцию, эпиграф которой равен сумме Минковского эпиграфов изначальных функций.
- Для выпуклых функций свертка сохраняет выпуклость.
- Для произвольных функций $f_1, f_2, \ldots, f_k$ верно

$$(f_1 \Box f_2 \Box \ldots \Box f_k)^* = f_1^* + f_2^* + \cdots + f_k^*.$$

- Если функции выпуклые и замкнутые, то верно ещё и

$$(f_1 + f_2 + \cdots + f_k)^* = f_1^* \Box f_2^* \Box \ldots \Box f_k^*.$$

Пример на свертку

Example

Найдите сопряженную

$$f(x) = \|Ax\|_1 + \frac{1}{2}\|x\|_2^2, \; A \in \mathbb{R}^{d\times d}, \; \det A \neq 0.$$

Для выпуклых замкнутых функций субдифференциал имеет вид

$$\partial f(x) = \arg\max_{y}\{\langle x, y\rangle - f^*(y)\}.$$