

# Методы оптимизации.

## Семинар 8. Сопряжённые функции

Александр Катруца

Московский физико-технический институт,  
Факультет Управления и Прикладной Математики

23 октября 2019 г.

- Субградиент и субдифференциал
- Условный субдифференциал
- Способы вычисления субдифференциалов

# Определение

## Снова сопряжённое?

- Ранее были рассмотрены сопряжённые (двойственные) множества и, в частности, конусы
- Сейчас будут рассмотрены сопряжённые (двойственные) функции
- Далее будет введена двойственная оптимизационная задача

## Определение

Пусть  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ . Функция  $f^* : \mathbb{R}^n \rightarrow \mathbb{R}$  называется сопряжённой функцией к функции  $f$  и определена как

$$f^*(\mathbf{y}) = \sup_{\mathbf{x} \in \text{dom } f} (\mathbf{y}^T \mathbf{x} - f(\mathbf{x})).$$

Область определения  $f^*$  — это множество таких  $\mathbf{y}$ , что супремум конечен.

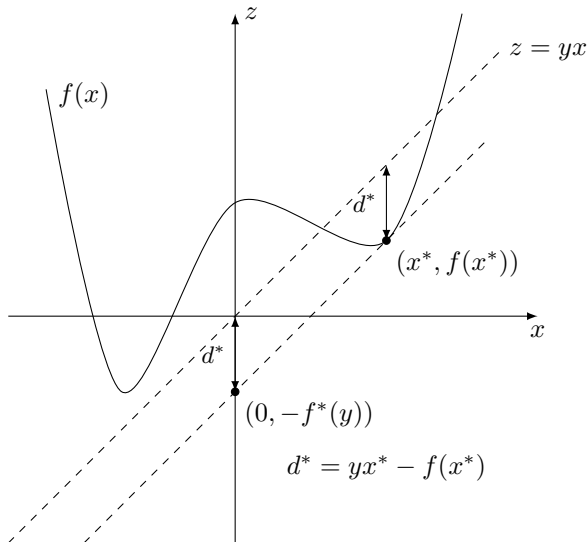
- Сопряжённая функция  $f^*$  всегда выпукла как супремум линейных функций независимо от выпуклости  $f$
- Неравенство Юнга-Фенхеля:  $\mathbf{y}^T \mathbf{x} \leq f(\mathbf{x}) + f^*(\mathbf{y})$   
Обобщение квадратичного случая:  $\mathbf{y}^T \mathbf{x} \leq \frac{1}{2} \mathbf{x}^T \mathbf{x} + \frac{1}{2} \mathbf{y}^T \mathbf{y}$
- Если  $f$  — дифференцируема, то  
 $f^*(\mathbf{y}) = \nabla f^T(\mathbf{x}^*) \mathbf{x}^* - f(\mathbf{x}^*)$ , где  $\mathbf{x}^*$  даёт супремум.
- Если  $f$  выпукла и замкнута, то  $f^{**} = f$

## Определение

Выпуклая функция называется замкнутой, если множество её подуровней замкнутое множество.

Пример:  $f(x) = x \log x$  при  $\text{dom } f = \mathbb{R}_{++}$  — незамкнутая

# Геометрический смысл



# Примеры

1. Линейная функция:  $f(\mathbf{x}) = \mathbf{a}^T \mathbf{x} + b$
2. Отрицательная энтропия:  $f(x) = x \log x$
3. Индикаторная функция множества  $S$ :  $I_S(x) = 0$  iff  $x \in S$
4. Норма:  $f(\mathbf{x}) = \|\mathbf{x}\|$ .
5. Квадрат нормы:  $f(\mathbf{x}) = \frac{1}{2} \|\mathbf{x}\|^2$

# Операции с сопряжёнными функциями

- Разделение переменных:  $f(x_1, x_2) = g(x_1) + h(x_2)$  и  $f^*(y_1, y_2) = g^*(y_1) + h^*(y_2)$
- Сдвиг аргумента:  $f(x) = g(x - a)$  и  $f^*(y) = a^T y + g^*(y)$
- Суперпозиция с обратимым линейным преобразованием:  $f(x) = g(Ax)$  и  $f^*(y) = g^*(A^{-T}y)$
- Инфимальная конволюция (свёртка инфимумом):  
$$f(x) = (h \square g)(x) = \inf_{u+v=x} (h(u) + g(v))$$
  
$$f^*(y) = h^*(y) + g^*(y)$$

# Moreau-Yosida envelope

- $f(\mathbf{x})$  выпуклая, но негладкая
- Moreau-Yosida envelope ( $\lambda > 0$ )

$$M_{\lambda f}(\mathbf{x}) = \inf_{\mathbf{u}} \left( f(\mathbf{u}) + \frac{1}{2\lambda} \|\mathbf{x} - \mathbf{u}\|_2^2 \right) = \left( f \square \frac{1}{2\lambda} \|\cdot\|_2^2 \right) (\mathbf{x})$$

- Функция Хьюбера –  $M_{\lambda f}$  для модуля
  - $f(x) = |x|$
  - $M_{\lambda f}(x) = \begin{cases} \frac{x^2}{2\lambda} & |x| \leq \lambda \\ |x| - \lambda/2 & |x| \geq \lambda \end{cases}$

## Упражнение

- Нарисуйте на одном графике  $f(x)$  и  $M_{\lambda f}(x)$
- Получите выражение  $M_{\lambda f}$  для  $f(\mathbf{x}) = \|\mathbf{x}\|_1$



# Почему получилась гладкая функция?

- $M_{\lambda f}(\mathbf{x})$  – выпукла
- $M_{\lambda f}^*(\mathbf{y}) = f^*(\mathbf{y}) + \frac{\lambda}{2} \|\mathbf{y}\|_2^2$  – сильно выпукла с параметром  $\lambda$
- $M_{\lambda f} = M_{\lambda f}^{**} = (f^* + \frac{\lambda}{2} \|\cdot\|_2^2)^*$
- Сопряжённая функция к сильно выпуклой функции является гладкой  $\Rightarrow M_{\lambda f}$  – гладкая функция и

$$M'_{\lambda f}(\mathbf{x}) = \frac{1}{\lambda}(\mathbf{x} - \mathbf{u}^*), \quad \mathbf{u}^* = \arg \min_{\mathbf{u}} \left( f(\mathbf{u}) + \frac{1}{2\lambda} \|\mathbf{x} - \mathbf{u}\|_2^2 \right)$$

## Важное свойство

Множество точек минимума  $f$  и  $M_{\lambda f}$  совпадает.

- Сопряжённые функции
- Неравенство Юнга-Фенхеля и другие свойства
- Сглаживание негладких функций
- Примеры