МОМО-16. Домашняя работа 4

Срок сдачи: 21 ноября 2016, 10:30

Напомним определение сопряженной функции. Для произвольной функции $f : \mathrm{Dom}\, f \subseteq \mathbb{R}^n \to \mathbb{R}$ ее сопряженной (по Фенхелю) называется функция $f_* : \mathrm{Dom}\, f_* \subseteq \mathbb{R}^n \to \mathbb{R}$, определенная по правилу

$$f_*(s) = \sup_{x \in \mathrm{Dom}\, f} \{s^\top x - f(x)\}.$$

По определению областью определения функции f_* является множество $\mathrm{Dom}\, f_* = \{s \in \mathbb{R}^n : \sup_{x \in \mathrm{Dom}\, f}\{s^\top x - f(x)\} < +\infty\}$.

1 Для каждой из следующих функций найти ее сопряженную:

(a) (Модуль в степени) $f(x) = \frac{1}{p}|x|^p,\ x \in \mathbb{R}$, где $p \in (1, \infty)$ (*Подсказка:* Запишите ответ через сопряженный параметр q, задаваемый равенством $1/p + 1/q = 1$.)

(b) (Отрицательная энтропия) $f(x) = x \ln x + (1-x)\ln(1-x),\ x \in (0,1)$.

(c) (Постоянная функция) $f(x) = c,\ x \in \mathbb{R}^n$, где $c \in \mathbb{R}$.

(d) (Линейная функция) $f(x) = a^\top x + \beta,\ x \in \mathbb{R}^n$, где $a \in \mathbb{R}^n$, $\beta \in \mathbb{R}$, $a \neq 0$.

(e) (Квадрат нормы) $f(x) = \frac{1}{2}\|x\|^2,\ x \in \mathbb{R}^n$, где $\|\cdot\|$ — произвольная норма на $\mathbb{R}^n$. (*Подсказка:* Используйте двойственную норму.)

2 Докажите следующие свойства сопряженной функции:

(a) (Сепарабельность) Пусть $f_i : \mathrm{Dom}\, f_i \subseteq \mathbb{R} \to \mathbb{R}$. Рассмотрим функцию

$$f(x) = \sum_{i=1}^n f_i(x_i),$$

определенную на множестве $\mathrm{Dom}\, f = \mathrm{Dom}\, f_1 \times \cdots \times \mathrm{Dom}\, f_n \subseteq \mathbb{R}^n$. Тогда

$$f_*(s) = \sum_{i=1}^n f_{i*}(s_i).$$

Областью определения f_* является $\mathrm{Dom}\, f_* = \mathrm{Dom}\, f_{1*} \times \cdots \times \mathrm{Dom}\, f_{n*}$.

(b) (Масштабирование) Пусть $f : \mathrm{Dom}\, f \subseteq \mathbb{R}^n \to \mathbb{R}$, $\alpha \in (0, +\infty)$, $\beta \in \mathbb{R}$. Рассмотрим функцию

$$\phi(x) = \alpha f(x) + \beta,$$

определенную на множестве $\mathrm{Dom}\, \phi = \mathrm{Dom}\, f$. Тогда

$$\phi_*(s) = \alpha f_*((1/\alpha)s) - \beta.$$

Областью определения ϕ_* является $\mathrm{Dom}\, \phi_* = \alpha\, \mathrm{Dom}\, f_*$.

(c) (Замена переменной) Пусть $f : \mathrm{Dom}\, f \subseteq \mathbb{R}^n \to \mathbb{R}$, $A \in \mathbb{R}^{n \times n}$, $b \in \mathbb{R}^n$ Рассмотрим функцию

$$\phi(x) = f(Ax + b)$$

определенную на множестве $\mathrm{Dom}\,\phi = \{x \in \mathbb{R}^n : Ax + b \in \mathrm{Dom}\, f\}$. Если матрица A является невырожденной, то

$$\phi_*(s) = f_*(A^{-\top}s) - s^\top A^{-1}b.$$

Областью определения ϕ_* является $\mathrm{Dom}\,\phi_* = \{s \in \mathbb{R}^n : A^{-\top}s \in \mathrm{Dom}\, f_*\}$.

3 Рассмотрим задачу минимизации эмпирического риска

$$(P) \qquad \min_{w \in \mathbb{R}^d} \left\{ \frac{1}{n} \sum_{i=1}^n f_i(a_i^\top w) + \tau R(w) \right\},$$

где $a_i \in \mathbb{R}^d$, $f_i : \mathbb{R} \to \mathbb{R}$, $R : \mathbb{R}^d \to \mathbb{R}$, $\tau > 0$. На семинаре мы получили общий результат о том, что двойственной задачей к (P) является

$$(D) \qquad \max_{\mu \in \mathbb{R}^n} \left\{ -\frac{1}{n} \sum_{i=1}^n f_{i*}(\mu_i) - \tau R_* \left(-\frac{1}{\tau n} A\mu \right) \right\},$$

где $A = [a_1, \ldots, a_n] \in \mathbb{R}^{d \times n}$ — матрица со столбцами a_i, а f_{i*} и R_* — соответствующие сопряженные функции для f_i и R.

Используя этот результат, выпишите двойственную задачу для задачи бинарной логистической регрессии с L_2-регуляризатором:

$$f_i(z_i) = \ln(1 + \exp(-b_i z_i)), \qquad R(w) = \frac{1}{2}\|w\|_2^2,$$

где $b_i \in \{-1, 1\}$.

(*Указание:* Воспользуйтесь полученным на семинаре выражением $f_*(s) = s \ln s + (1-s)\ln(1-s)$ для сопряженной функции для $f(x) = \ln(1 + e^x)$ и результатом о замене переменной.)

4 Для каждого из следующих множеств $Q \subseteq \mathbb{R}^n$ найти евклидову проекцию заданной точки $v \in \mathbb{R}^n$ на множество Q (т. е. найти $\mathrm{argmin}_{x \in Q} \|x - v\|_2^2$):

(a) (Короб) $Q = \{x \in \mathbb{R}^n : x_i \in [l_i, r_i],\ i = 1, \ldots, n\}$, где $-\infty \le l_i \le r_i \le +\infty$. (*Замечание:* Допускается, что $l_i = -\infty$ и/или $r_i = +\infty$, т. е. короб может быть неограниченным вдоль некоторых направлений.)

(b) (Единичный L_2-шар) $Q = B_2(0, 1) = \{x \in \mathbb{R}^n : \|x\|_2 \le 1\}$.

(c) (Афинное многообразие) $Q = \{x \in \mathbb{R}^n : Ax = b\}$, где $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, $\mathrm{Rank}(A) = m$.

(d) (Полупространство) $Q = \{x \in \mathbb{R}^n : a^\top x \le \beta\}$, где $a \in \mathbb{R}^n$, $\beta \in \mathbb{R}$, $a \ne 0$.

Воспользуйтесь полученными выше результатами и выпишите ответ для следующих случаев:

- (Неотрицательный ортант) $Q = \mathbb{R}^n_+ = \{x \in \mathbb{R}^n : x_i \ge 0,\ i = 1, \ldots, n\}$.
- (Единичный L_∞-шар) $Q = B_\infty(0, 1) = \{x \in \mathbb{R}^n : \|x\|_\infty \le 1\}$.
- (Гиперплоскость) $Q = \{x \in \mathbb{R}^n : a^\top x = \beta\}$, где $a \in \mathbb{R}^n$, $\beta \in \mathbb{R}$, $a \ne 0$.