Методы оптимизации в машинном обучении, ВМК+Физтех, осень 2017

Домашнее задание 3: Условная оптимизация.

Срок сдачи: 1 ноября 2017 (среда), 23:59 для ВМК
3 ноября 2017 (пятница), 23:59 для Физтеха

1 Для каждой из следующих задач найдите множество решений.

(a) (Линейное программирование с одним ограничением)

$$\min_{x \in \mathbb{R}^n} \{\langle c, x \rangle : \langle a, x \rangle \le b\},$$

где $a, c \in \mathbb{R}^n \setminus \{0\}$, $b \in \mathbb{R}$.

(b) (Линейная функция на стандартном симплексе)

$$\min_{x \in \mathbb{R}^n} \left\{ \langle c, x \rangle : x \succeq 0, \, \sum_{i=1}^n x_i = 1 \right\},$$

где $c \in \mathbb{R}^n$.

(c) (Линейная функция с энтропийным регуляризатором)

$$\min_{x \in \mathbb{R}^n_{++}} \left\{ \langle c, x \rangle + \sum_{i=1}^n x_i \ln x_i : \sum_{i=1}^n x_i = 1 \right\},$$

где $c \in \mathbb{R}^n$.

2 Для каждой из следующих задач оптимизации постройте двойственную задачу и выпишите явные формулы, позволяющие по решению двойственной задачи восстановить (вычислить) решение прямой.

(a) (Гребневая регрессия)

$$\min_{x \in \mathbb{R}^n, \, s \in \mathbb{R}^m} \left\{ \frac{1}{2}\|s - b\|_2^2 + \frac{\rho}{2}\|x\|_2^2 : s = Ax \right\},$$

где $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, $\rho > 0$.

(b) (SVM)

$$\min_{x \in \mathbb{R}^n, \, t \in \mathbb{R}^m} \left\{ \langle 1_m, t \rangle + \frac{\rho}{2}\|x\|_2^2 : Ax \succeq 1_m - t, \, t \succeq 0 \right\},$$

где $A \in \mathbb{R}^{m \times n}$, $1_m := (1, \dots, 1) \in \mathbb{R}^m$, $\rho > 0$.

3 Сведите эквивалентным образом следующие негладкие безусловные задачи к гладким условным:

(a) (Максимум из конечного числа гладких функций)

$$\min_{x \in \mathbb{R}^n} \max\{f_1(x), \dots, f_m(x)\},$$

где $f_1, \dots, f_m : \mathbb{R}^n \to \mathbb{R}$ — гладкие функции.

(b) (Наилучшее решение линейной системы в l^∞-норме)

$$\min_{x \in \mathbb{R}^n} \|Ax - b\|_\infty,$$

где $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$.

(c) (Наилучшее решение линейной системы в l^1-норме)

$$\min_{x \in \mathbb{R}^n} \|Ax - b\|_1,$$

где $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$.

(d) (Задача LASSO)

$$\min_{x \in \mathbb{R}^n} \left\{ \frac{1}{2}\|Ax - b\|_2^2 + \rho\|x\|_1 \right\},$$

где $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, $\rho \in \mathbb{R}_{++}$.

4 Для каждой из следующих квадратичных задач найдите множество решений:

(a) (Минимизация линейной формы на эллипсоиде)

$$\min_{x \in \mathbb{R}^n} \{\langle c, x \rangle : \langle Ax, x \rangle \le 1\},$$

где $c \in \mathbb{R}^n \setminus \{0\}$ и $A \in \mathbb{S}^n_{++}$.

(b) (Минимизация квадратичной формы на эллипсоиде)

$$\min_{x \in \mathbb{R}^n} \{\langle Bx, x \rangle : \langle Ax, x \rangle \le 1\},$$

где $A \in \mathbb{S}^n_{++}$, $B \in \mathbb{S}^n_{+}$.

5 Рассмотрим задачу

$$\min_{x \in \mathbb{R}^n} \left\{ \frac{1}{2}\langle Ax, x \rangle - \langle b, x \rangle : \|x\|_2 \le 1 \right\},$$

где $A \in \mathbb{S}^n_{++}$ и $b \in \mathbb{R}^n$. Докажите, что решение в этой задаче равно $(A + \lambda I_n)^{-1}b$, где $\lambda := \max\{0, \bar{\lambda}\}$, и $\bar{\lambda}$ — это наибольшее из решений нелинейного уравнения

$$\langle (A + \lambda I_n)^{-2}b, b \rangle = 1.$$

6 Для каждого из следующих множеств Q в пространстве $\mathbb{R}^n$ найдите евклидову проекцию $\pi_Q(v) := \operatorname{argmin}_{x \in Q} \|x - v\|_2$ заданной точки $v \in \mathbb{R}^n$ на множество:

(a) (Короб) $Q = \{x \in \mathbb{R}^n : a \preceq x \preceq b\}$, где $a \in (\mathbb{R} \cup \{-\infty\})^n$, $b \in (\mathbb{R} \cup \{+\infty\})^n$. (*Замечание:* Допускается, что $a_i = -\infty$ и/или $b_i = +\infty$, т. е. короб может быть неограниченным вдоль некоторых направлений.)

(b) (Аффинное подпространство) $Q = \{x \in \mathbb{R}^n : Ax = b\}$, где $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, $\operatorname{Rank}(A) = m$.

(c) (Полупространство) $Q = \{x \in \mathbb{R}^n : \langle a, x \rangle \le b\}$, где $a \in \mathbb{R}^n$, $b \in \mathbb{R}$, $a \ne 0$.

Воспользуйтесь полученными выше результатами и выпишите ответ для следующих случаев:

- (Неотрицательный ортант) $Q = \mathbb{R}^n_+ = \{x \in \mathbb{R}^n : x \succeq 0\}$.
- (Единичный l^∞-шар) $Q = B_{l^\infty}(0, 1) = \{x \in \mathbb{R}^n : \|x\|_\infty \le 1\}$.
- (Гиперплоскость) $Q = \{x \in \mathbb{R}^n : \langle a, x \rangle = b\}$, где $a \in \mathbb{R}^n$, $b \in \mathbb{R}$, $a \ne 0$.

Бонусная часть

7 Найдите в явном виде решение следующей матричной задачи:

$$\min_{X \in \mathbb{S}^n_{++}} \{\langle C, X \rangle - \ln \operatorname{Det}(X) : \langle Xa, a \rangle \le 1\},$$

где $C \in \mathbb{S}^n_{++}$, $a \in \mathbb{R}^n \setminus \{0\}$. Ответ выпишите в терминах матрицы C^{-1}.

8 Пусть $a_1, \dots, a_m \in \mathbb{R}^n$ — заданные ненулевые точки в пространстве $\mathbb{R}^n$. Рассмотрим задачу поиска эллипсоида минимального объема, накрывающего эти точки:

$$\min_{X \in \mathbb{S}^n_{++}} \{-\ln \mathrm{Det}(X) : \langle X a_i, a_i \rangle \le 1 \text{ для всех } 1 \le i \le m\}.$$

Постройте для этой задачи двойственную.

9 Обозначим через Δ_n стандартный n-мерный симплекс:

$$\Delta_n := \left\{ x \in \mathbb{R}^n : x \succeq 0, \sum_{i=1}^n x_i = 1 \right\}.$$

Рассмотрим задачу поиска евклидовой проекции заданной точки $v \in \mathbb{R}^n$ на Δ_n:

$$\pi_{\Delta_n}(v) := \operatorname*{argmin}_{x \in \Delta_n} \|x - v\|_2.$$

Докажите, что $\pi_{\Delta_n}(v) = [v - \nu 1_n]_+$, где $\nu \in \mathbb{R}$ — корень нелинейного уравнения

$$\langle 1_n, [v - \nu 1_n]_+ \rangle = 1. \tag{0.1}$$

Здесь $1_n := (1, \dots, 1) \in \mathbb{R}^n$, а $[u]_+$ обозначает поэлементную положительную срезку: $([u]_+)_i := \max\{0, u_i\}$ для всех $1 \le i \le n$. Нарисуйте схематичный график левой части уравнения (0.1) как функции от ν.

Подсказка. Удобно рассмотреть упорядоченные компоненты $v_{[1]} \ge \cdots \ge v_{[n]}$.

10 Пусть для $\Sigma, \Sigma_0 \in \mathbb{S}^n_{++}$ символ $D(\Sigma; \Sigma_0)$ обозначает *дивергенцию Кульбака–Лейблера* между двумя многомерными нормальными распределениями $N(0, \Sigma)$ и $N(0, \Sigma_0)$, т. е.

$$D(\Sigma; \Sigma_0) := \frac{1}{2}(\langle \Sigma_0^{-1}, \Sigma \rangle - \ln \mathrm{Det}(\Sigma_0^{-1} \Sigma) - n).$$

Пусть $H \in \mathbb{S}^n_{++}$, и пусть $y, s \in \mathbb{R}^n$, причем $\langle y, s \rangle > 0$. Рассмотрим задачу поиска матрицы $H_+ \in \mathbb{S}^n_{++}$, удовлетворяющей условию $H_+ y = s$ и минимизирующей дивергенцию $X \mapsto D(X^{-1}; H^{-1})$:

$$\min_{X \in \mathbb{S}^n_{++}} \{D(X^{-1}; H^{-1}) : Xy = s\}.$$

Покажите, что решение этой задачи существует, единственно и выражается по *формуле обновления BFGS*

$$H_+ = \left(I_n - \frac{s y^T}{\langle y, s \rangle} \right) H \left(I_n - \frac{y s^T}{\langle y, s \rangle} \right) + \frac{s s^T}{\langle y, s \rangle}.$$