Методы оптимизации, ФКН ВШЭ, зима 2017

Домашнее задание 4: Условная оптимизация

Срок сдачи: 15 марта (среда) 2017, 23:59, после срока не принимается.

Выполненное задание следует отправить письмом на почту своей группы[1] с заголовком:

Домашнее задание 4, Фамилия Имя.

либо сдать 14 марта на семинаре в письменном виде.

1 Для каждой из следующих задач найдите оптимальное значение и множество оптимальных решений:

(a) (Линейное программирование с одним ограничением)

$$\min_{x \in \mathbb{R}^n} \{c^T x : a^T x \leq \beta\},$$

где $a, c \in \mathbb{R}^n$, $\beta \in \mathbb{R}$, $a \neq 0$, $c \neq 0$.

(b) (Линейная функция на стандартном симплексе)

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

где $c \in \mathbb{R}^n$.

(c) (Линейная функция с энтропийным регуляризатором)

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

где $c \in \mathbb{R}^n$.

2 Для каждой из следующих задач оптимизации: 1) Построить двойственную задачу. 2) Выписать 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 $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, $\rho \in \mathbb{R}_{++}$.

(b) (SVM)

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

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

[1] Для 141 группы: opt.homework+141@gmail.com. Для 142 группы: opt.homework+142@gmail.com. Для 145 группы: opt.homework+145@gmail.com.

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

(a) (Максимум из конечного числа гладких функций)

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

где $f_i : \mathbb{R}^n \to \mathbb{R}$ — заданные гладкие функции.

(b) (Наилучшее решение линейной системы в ℓ_∞-норме)

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

где $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$. Для произвольного вектора $y \in \mathbb{R}^m$: $\|y\|_\infty := \max_{i=1}^m |y_i|$.

(c) (Наилучшее решение линейной системы в ℓ_1-норме)

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

где $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$. Для произвольного вектора $y \in R^m$: $\|y\|_1 := \sum_{i=1}^m |y_i|$.

(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 (QCQP) найдите аналитическое решение:

(a) (Минимизация линейной формы на эллипсоиде)

$$\min_{x \in \mathbb{R}^n} \{c^T x : x^T A x \leq 1\},$$

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

(b) (Минимизация квадратичной формы на эллипсоиде)

$$\min_{x \in \mathbb{R}^n} \{x^T B x : x^T A x \leq 1\},$$

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

5 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 $\Pi_Q(v) := \operatorname{argmin}_{x \in Q} \|x - v\|_2^2$):

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

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

(c) (Полупространство) $Q = \{x \in \mathbb{R}^n : a^T x \leq \beta\}$, где $a \in \mathbb{R}^n$, $\beta \in \mathbb{R}$, $a \neq 0$.

Воспользуйтесь полученными выше результатами и выпишите ответ для следующих случаев:

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

Бонусная часть (6 баллов)

6 Рассмотрим QCQP:

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

где $A \in \mathbb{S}^n_{++}$ и $b \in \mathbb{R}^n$. Докажите, что оптимальное 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

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

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

$$\Pi_{\Delta_n}(v) := \operatorname*{argmin}_{x \in \mathbb{R}^n} \left\{ \|x - v\|_2 : x \succeq 0, \sum_{i=1}^n x_i = 1 \right\}.$$

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

$$1_n^T (v - \nu 1_n)_+ = 1. \tag{1}$$

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

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

8 Пусть $\Sigma, \Sigma_0 \in \mathbb{S}^n_{++}$. Обозначим через $D(\Sigma; \Sigma_0)$ *дивергенцию Кульбака–Лейблера* между нормальными распределениями $\mathcal{N}(0, \Sigma)$ и $\mathcal{N}(0, \Sigma_0)$:

$$D(\Sigma; \Sigma_0) = \frac{1}{2} (\operatorname{Tr}(\Sigma_0^{-1} \Sigma) - \ln \operatorname{Det}(\Sigma_0^{-1} \Sigma) - n).$$

Пусть $H \in \mathbb{S}^n_{++}$. Пусть также $y, s \in \mathbb{R}^n$, причем $y^T s > 0$. Рассмотрим задачу поиска матрицы $X \in \mathbb{S}^n_{++}$, минимизирующую дивергенцию $D(X^{-1}; H^{-1})$ при условии $Xy = s$:

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

Решите эту задачу и убедитесь, что ее решение выражается по формуле обновления обратной матрицы в схеме BFGS:

$$X = (I_n - \rho s y^T) H (I_n - \rho y s^T) + \rho s s^T,$$

где $\rho := 1/(y^T s)$.