Методы оптимизации, ВМК, осень 2018

Домашняя работа 3: Условия Каруша–Куна–Таккера. Двойственность.

Срок сдачи: 17 ноября 2018 (суббота), 23:59

1 Для каждого из следующих множеств в пространстве $\mathbb{R}^n$ вычислите ортогональную проекцию[1] точки $v \in \mathbb{R}^n$ на множество:

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

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

2 Для каждой из следующих задач найдите множество всевозможных решений:

(a) $\min_{x\in\mathbb{R}^n}\{\langle c, x\rangle : \langle a, x\rangle \le b\}$, где $a, c \in \mathbb{R}^n \setminus \{0\}$, $b \in \mathbb{R}$.

(b) $\min_{x\in\mathbb{R}^n}\left\{\langle c, x\rangle : x \in \mathbb{R}^n_+;\ \sum_{i=1}^n x_i = 1\right\}$, где $c \in \mathbb{R}^n$.

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

(d) $\min_{x\in\mathbb{R}^n}\{\langle c, x\rangle : \langle Ax, x\rangle \le 1\}$, где $c \in \mathbb{R}^n \setminus \{0\}$ и $A \in \mathbb{S}^n_{++}$.

(e) $\min_{x\in\mathbb{R}^n}\{\langle Bx, x\rangle : \langle Ax, x\rangle \le 1\}$, где $A \in \mathbb{S}^n_{++}$, $B \in \mathbb{S}^n_+$.

(*Подсказка:* Некоторые задачи проще решить без использования теоремы Каруша–Куна–Таккера.)

3 Пусть $a, c \in \mathbb{R}^n_{++}$, $b > 0$. Покажите, что следующая задача имеет единственное решение и найдите его:

$$\min_{x\in\mathbb{R}^n_{++}} \left\{ \sum_{i=1}^n \frac{c_i}{x_i} : \langle a, x\rangle \le b \right\}.$$

4 Пусть $A \in \mathbb{S}^n_{++}$, $b > 0$. Для каждой из следующих задач покажите, что решение единственное и найдите его:

(a) $\max_{X\in\mathbb{S}^n_+}\{\mathrm{Det}(X) : \langle A, X\rangle \le b\}$.

(b) $\min_{X\in\mathbb{S}^n_{++}}\{\langle X^{-1}, I_n\rangle : \langle A, X\rangle \le b\}$.

5 Пусть $e_1, \dots, e_n$ — стандартный базис в $\mathbb{R}^n$. Покажите, что задача

$$\max_{X\in\mathbb{S}^n_+}\{\mathrm{Det}(X) :\ \|Xe_i\| \le 1 \text{ для всех } 1 \le i \le n\}$$

имеет единственное решение, равное I_n. Установите отсюда частный случай *неравенства Адамара*

$$\mathrm{Det}(X) \le \|Xe_1\| \dots \|Xe_n\|,$$

справедливого для любой матрицы $X \in \mathbb{S}^n_+$. (*Подсказка*: Преобразуйте целевую функцию, чтобы она стала строго выпуклой. Используйте тот факт, что задача со строго выпуклой целевой функцией имеет не более одного решения.)

6 * Покажите, что следующая задача имеет единственное решение и найдите его:

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

где $C \in \mathbb{S}^n_{++}$, $a \in \mathbb{R}^n$, $a \neq 0$. В ответе не должно быть обратной матрицы C^{-1}. (*Подсказка*: Используйте формулу Шермана–Моррисона.)

[1] Ортогональная проекция определяется следующим образом. Пусть C — множество в евклидовом пространстве V, и пусть $v \in V$. *Ортогональной проекцией точки v на множество C* называется любая такая точка $\pi_C(v)$, которая минимизирует функционал $x \mapsto \|x - v\|$ на множестве C. Известно, что если C — непустое выпуклое замкнутое множество, то ортогональная проекция $\pi_C(v)$ существует и единственна.

7 * Пусть $A \in \mathbb{S}^n_{++}$ и $b \in \mathbb{R}^n$, $b \neq 0$. Покажите, что задача

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

имеет единственное решение, равное $(A + \lambda_0 I_n)^{-1} b$, где $\lambda_0 := \max\{0, \bar{\lambda}\}$, и $\bar{\lambda}$ — наибольшее из решений нелинейного уравнения

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

8 Пусть $f : E \to \mathbb{R}$ — функция, заданная на множестве E в нормированном вещественном векторном пространстве V. Функция f называется *замкнутой*, если ее надграфик $\operatorname{Epi} f := \{(x,t) \in E \times \mathbb{R} : f(x) \leq t\}$ является замкнутым множеством в пространстве $V \oplus \mathbb{R}$. Покажите, что f является замкнутой, если и только если множество подуровней $f^{-1}((-\infty, a]) := \{x \in E : f(x) \leq a\}$ является замкнутым множеством в пространстве V для любого $a \in \mathbb{R}$.

9 Пусть $f : E \to \mathbb{R}$ — функция, заданная на множестве E в вещественном нормированном векторном пространстве. Докажите следующий *критерий замкнутости*: f замкнутая тогда и только тогда, когда:

(a) f полунепрерывна снизу;

(b) f обладает *барьерным свойством*: для любого $x_0 \in \partial E \setminus E$ справедливо $f(x) \to +\infty$ при $x \to x_0; x \in E$ (здесь ∂E — граница множества E).

10 Для каждой из следующих функций f, заданных на множестве в пространстве $\mathbb{R}$, вычислите сопряженную функцию f^*:

(a) $f : \mathbb{R} \to \mathbb{R}$ — функция $f(x) := e^x$.

(b) $f : \mathbb{R} \to \mathbb{R}$ — функция $f(x) := |x|^p/p$, где $p > 1$.

(c) $f : (0, +\infty) \to \mathbb{R}$ — функция $f(x) := -\ln x$.

(d) $f : (0, +\infty) \to \mathbb{R}$ — функция $f(x) := \frac{1}{x}$.

(e) $f : \mathbb{R} \to \mathbb{R}$ — функция $f(x) := |x|$.

(f) $f : \mathbb{R} \to \mathbb{R}$ — функция $f(x) := [x]_+ := \max\{x, 0\}$.

11 * Для каждой из следующих функций f, заданных на множестве в пространстве $\mathbb{R}^n$, вычислите сопряженную функцию f^*:

(a) $f : \mathbb{R}^n \to \mathbb{R}$ — функция $f(x) := \ln \sum_{i=1}^n e^{x_i}$.

(b) $f : \mathbb{R}^n_+ \to \mathbb{R}$ — функция $f(x) := -(\prod_{i=1}^n x_i)^{1/n}$.

(c) $f : \mathbb{R}^n \to \mathbb{R}$ — функция $f(x) := \max\{x_1, \ldots, x_n\}$.

(*Подсказка*: Во втором пункте полезно воспользоваться неравенством между средним арифметическим и средним геометрическим.)

12 Для каждой из следующих задач минимизации постройте двойственную задачу Фенхеля и покажите, что выполняется сильная двойственность, и при этом множества решений прямой и двойственной задач непустые:

(a) (Гребневая регрессия) $\min_{x \in \mathbb{R}^n} \frac{1}{2}\|Ax - b\|^2 + \frac{\lambda}{2}\|x\|^2$, где $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, $\lambda > 0$.

(b) (LASSO) $\min_{x \in \mathbb{R}^n} \frac{1}{2}\|Ax - b\|^2 + \lambda\|x\|_1$, где $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, $\lambda > 0$.

(c) (SVM) $\min_{x \in \mathbb{R}^n} \sum_{i=1}^m [1 - \langle a_i, x\rangle]_+ + \frac{\lambda}{2}\|x\|^2$, где $a_1, \ldots, a_m \in \mathbb{R}^n$, $\lambda > 0$.

(d) (l^1-SVM) $\min_{x \in \mathbb{R}^n} \sum_{i=1}^m [1 - \langle a_i, x\rangle]_+ + \lambda\|x\|_1$, где $a_1, \ldots, a_m \in \mathbb{R}^n$, $\lambda > 0$.

13 Покажите, что $\delta^*_{\mathbb{S}^n_+} = \delta_{\mathbb{S}^n_-}$ (сопряженная функция индикатора конуса $\mathbb{S}^n_+$ равна индикатору полярного конуса конуса $\mathbb{S}^n_-$). (*Подсказка*: Воспользуйтесь спектральным разложением и заменой переменных; сперва покажите, что область определения $\delta^*_{\mathbb{S}^n_+}$ вложена в $\mathbb{S}^n_-$; далее докажите и используйте неравенство $\langle A, B\rangle \geq 0$ для $A, B \in \mathbb{S}^n_+$.)

14 Пусть $A_1, \ldots, A_n \in \mathbb{S}^m$, и пусть $A : \mathbb{R}^n \to \mathbb{S}^m$ — линейный оператор

$$Ax := \sum_{i=1}^{n} x_i A_i.$$

Покажите, что сопряженный оператор $A^* : \mathbb{S}^m \to \mathbb{R}^n$ имеет вид

$$A^* U = (\langle A_i, U\rangle)_{1 \leq i \leq n}$$

для всех $U \in \mathbb{S}^m$. (Считаем, что скалярное произведение в $\mathbb{R}^n$ стандартное.)

15 (SDP двойственность) Рассмотрим задачу *полуопределенного программирования* в стандартном виде:

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

где $c \in \mathbb{R}^n$, $A_1, \ldots, A_m, B \in \mathbb{S}^m$.

(a) Постройте для этой задачи двойственную задачу Фенхеля.

(b) Покажите, что если исходная задача имеет строго допустимое решение, т. е. существует $x \in \mathbb{R}^n$, такой, что $\sum_{i=1}^n x_i A_i \prec B$, то выполняется сильная двойственность, и супремум в двойственной задаче достигается.

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

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

Постройте для этой задачи двойственную задачу Фенхеля. Покажите, что если ранг системы $a_1, \ldots, a_m$ равен n, то выполняется сильная двойственность, и при этом множества решений прямой и двойственной задач непустые.

17 Пусть $f : E \to \mathbb{R}$ — функция, определенная на множестве E в евклидовом пространстве V, и пусть $f^{**} := (f^*)^* : E_{**} \to \mathbb{R}$ — дважды сопряженная функция с естественной областью определения E_{**}. Покажите, что:

(a) Всегда имеет место *слабая двойственность*: $E \subseteq E_{**}$ и $f^{**}(x) \leq f(x)$ для всех $x \in E$.

(b) Если f *субдифференцируема в точке* $x_0 \in E$, т. е. существует вектор $s \in V$, такой, что $f(x) \geq f(x_0) + \langle s, x - x_0\rangle$ для всех $x \in E$, то имеет место *сильная двойственность*: $f^{**}(x_0) = f(x_0)$.

Бонусные задачи

18 * Для каждой из следующих функций f вычислите сопряженную функцию f^* по определению:

(a) $f : \mathbb{S}^n_{++} \to \mathbb{R}$ — функция $f(X) := -\ln \mathrm{Det}(X)$.

(b) $f : \mathbb{S}^n_{++} \to \mathbb{R}$ — функция $f(X) := \mathrm{Tr}(X^{-1})$.

(c) $f : \mathbb{S}^n_{+} \to \mathbb{R}$ — функция $f(X) := -\mathrm{Det}(X)^{1/n}$.

19 * (Евклидова проекция на симплекс) Пусть $v \in \mathbb{R}^n$, и пусть $\Delta_n := \{x \in \mathbb{R}^n : x \succeq 0;\ \sum_{i=1}^n x_i = 1\}$ — стандартный n-мерный симплекс. Покажите, что $\pi_{\Delta_n}(v) = [v - \nu 1_n]_+$, где $\nu \in \mathbb{R}$ — корень нелинейного уравнения

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

Здесь $1_n := (1, \dots, 1) \in \mathbb{R}^n$, и для $u \in \mathbb{R}^n$ символ $[u]_+$ обозначает поэлементную положительную срезку: $([u]_+)_i := \max\{0, u_i\}$ для всех $1 \leq i \leq n$. Нарисуйте схематичный график левой части вышеприведенного уравнения как функции от ν. (*Подсказка*: Удобно рассмотреть упорядоченные компоненты $v_{[1]} \geq \dots \geq v_{[n]}$.)

20 * (Неравенство Гельдера) Для $p > 1$ и $x \in \mathbb{R}^n$ обозначим $\|x\|_p := \left(\sum_{i=1}^n |x_i|^p\right)^{1/p}$. Пусть $p > 1$, и пусть $s \in \mathbb{R}^n$, $s \neq 0$. Покажите, что задача

$$\max_{x \in \mathbb{R}^n} \{\langle s, x \rangle :\ \|x\|_p \leq 1\}.$$

имеет единственное решение (найдите его), а соответствующее оптимальное значение равно $\|s\|_q$, где $q > 1$ определяется из равенства $1/p + 1/q = 1$. Установите отсюда, что для любых $s, x \in \mathbb{R}^n$ справедливо *неравенство Гельдера*

$$|\langle s, x \rangle| \leq \|s\|_q \|x\|_p.$$

(*Подсказка*: Используйте теорему Каруша–Куна–Таккера. При этом могут оказаться полезными следующие два факта: 1) $(|\,|^p/p)'(u) = u|u|^{p-2}$ для всех $u \in \mathbb{R}$; 2) $u, v \in \mathbb{R}$ удовлетворяют $v = u|u|^{p-2}$, если и только если $u = v|v|^{q-2}$; докажите их. При решении задачи постарайтесь не пользоваться знаниями о том, что $\|\,\|_p$ является нормой.)

21 * (BFGS через дивергенцию Кульбака–Лейблера) Для $\Sigma, \Sigma_0 \in \mathbb{S}^n_{++}$ через $D(\Sigma; \Sigma_0)$ обозначим *дивергенцию Кульбака–Лейблера* между двумя многомерными нормальными распределениями $N(0, \Sigma)$ и $N(0, \Sigma_0)$:

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

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 эта задача имеет единственное решение, равное

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