Опты, дз 3

Максим Пасько 776

General optimization problems

1.

$$c^T x \to \min$$
$$\text{s.t. } Ax = b$$

$L(x, \lambda) = c^T x + \lambda^T (Ax - b) = (c^T + \lambda^T A)x - \lambda^T b$.
$\nabla_x L = c + A^T \lambda^* = 0 \Rightarrow \lambda^{*T} = -c^T A^T (AA^T)^{-1}$. Подставляя это λ^* в лагранжиан, получим ответ задачи:
$p^* = -\lambda^{*T} b = -c^T A^T (AA^T)^{-1} b$ - достигается на всех $x : \quad Ax = b$

2.

$$c^T x \to \min$$
$$\text{s.t. } 1^T x = 1$$
$$x \succeq 0$$

Без ограничения общности можно считать, что вектор c отсортирован в порядке по возрастанию, т.е. $c_1 \leq c_2 \leq ... \leq c_n$. Тогда: $c^T x = c_1 x_1 + c_2 x_2 + ... + c_n x_n \geq c_1(x_1 + x_2 + ... + x_n) = c_1(1^T x) = c_1$. Эта оценка достигается при $x : x_1 = 1, x_i = 0, i = \overline{2, n}$. Переходя к общему случаю неотсортированного вектора c, получим, что итоговый ответ будет $p^* = \min(c_1, c_2, ..., c_n) = c_{k_{min}}$, который будет достигаться при $x_{k_{min}} = 1, \quad x_i = 0 \quad \forall i \neq k_{min}$

3.

$$c^T x \to \min$$
$$\text{s.t. } 1^T x = \alpha$$
$$0 \preceq x \preceq 1$$

Такая же оценка, как в предыдущем номере, уже не пройдёт (потому что мы не сможем подобрать такой вектор x, на котором она достигается), но сработает другая:
Опять считаем, что вектор c отсортирован по возрастанию, причём

$$c_1 \leq ... \leq c_{i-1} < c_i = ... = c_\alpha = ... < c_j \leq ... \leq c_n$$

Тогда $c^T x = c_1 x_1 + ... + c_{i-1} x_{i-1} + c_i x_i + ... + c_\alpha x_\alpha + ... + c_j x_j + ... + c_n x_n \geq c_1 + ... + c_\alpha$. Тогда, если $x_1 = ...x_{i-1} = 1$, $\sum\limits_{k=i}^{j-1} x_k = \alpha - i + 1$, а $x_j = ... = x_n = 0$, то $p^* = c_1 + ... + c_\alpha$. Переходя к случаю неотсортированного вектора c, p^* будет суммой α наименьших координат вектора c. В случае же неравенства $1^T x \leq \alpha$, это будет суммой всех неположительных координат вектора c, если таковых меньше, чем α, и суммой α наименьших координат, если их больше или хотя бы равно α

4.

$$c^T x \to min$$
$$\text{s.t. } x^T A x \le 1, \quad A \in \mathbb{S}^n_{++}$$

$L(x, \mu) = c^T x + \mu(x^T A x - 1)$

$\nabla_x L = c + 2\mu A x = 0 \quad \Rightarrow \quad x = -\dfrac{A^{-1}c}{\mu}$

Тогда $x^T A x = (\dfrac{A^{-1}c}{\mu})^T A \dfrac{A^{-1}c}{\mu} = \frac{1}{\mu^2}(A^{-1}c)^T A (A^{-1}c) = \frac{1}{\mu^2} c^T A^{-1} A A^{-1} c = \frac{1}{\mu^2} c^T A^{-1} c$. При этом, $\mu(x^T A x - 1) = 0$. Подставляя сюда $x^T A x$, получаем:
$\frac{1}{\mu} c^T A^{-1} c - \mu = 0 \quad \Rightarrow \quad \mu = \sqrt{c^T A^{-1} c}$.
Тогда $x^* = \dfrac{-A^{-1}c}{\sqrt{c^T A^{-1} c}}$ и решение задачи $p^* = c^T x^* = -\sqrt{c^T A^{-1} c}$

Теперь рассмотрим случай, когда $A \notin \mathbb{S}^n_{++}$:
$A = Q D(\lambda) Q^T = \sum\limits_{i=1}^{n} \lambda_i q_i q_i^T$. Тогда
$x^T A x = x^T Q D(\lambda) Q^T x = x^T \sum\limits_{i=1}^{n} \lambda_i q_i q_i^T x = \sum\limits_{i=1}^{n} \lambda_i (q_i x)(q_i x)^T = \sum\limits_{i=1}^{n} \lambda_i (q_i x)^2$. Пусть $q_i x = y_i$, т.е. $y = Qx$. Тогда $c^T x c^T (Q^{-1} y) = c^T Q^T y = (Qc)^T y$. Пусть $b = Qc$. Тогда мы получили такую задачу:

$$b^T y \to min$$
$$\text{s.t. } y^T D y \le 1$$

Рассмотрим разные случаи:
1)$\lambda_i > 0 \quad \forall i$ - этот случай это просто наша исходная задача (когда $A \in \mathbb{S}^n_{++}$) и её решение мы уже знаем.
2)$\exists i : \quad \lambda_i < 0$ - в этом случае мы можем взять сколь угодно большой положительный или отрицательный y_i и $p^* = -\infty$
3)$\exists i : \quad \lambda_i = 0$:

3.1)$b_i \ne 0$ - опять же $p^* = -\infty$

3.2)$b_i = 0$ - в этой задаче мы можем просто выкинуть i-ые элементы из y и b, и λ_i, таким образом мы сведём задачу к задаче размера на один меньше. Рекурсивно так спускаясь, мы придём к задаче вида $\lambda_j > 0 \quad \forall j$, решение которой мы уже знаем.

5.

$$c^T x \to \min$$
$$\text{s.t. } (x - x_c)^T A (x - x_c) \le 1, \quad A \in \mathbb{S}^n_{++}$$

$L(x, \mu) = c^T x + \mu((x - x_c)^T A (x - x_c) - 1)$

$\nabla_x L = c + 2\mu A (x - x_c) = 0 \quad \Rightarrow \quad x - x_c = -\dfrac{A^{-1}c}{2\mu}$

При этом, $\mu((x - x_c)^T A (x - x_c) - 1) = 0$. Подставляя сюда найденный $x - x_c$, получаем:
$\mu(\dfrac{(A^{-1}c)^T}{2\mu} A \dfrac{A^{-1}c}{2\mu} - 1) = \mu(\dfrac{c^T A^{-1} A A^{-1} c}{4\mu^2} - 1) = \dfrac{c^T A^{-1} c}{4\mu} - \mu = 0 \quad \Rightarrow \quad \mu = \dfrac{\sqrt{c^T A^{-1} c}}{2}$
Тогда $p^* = c^T x^* = c^T x_c - \dfrac{c^T A^{-1} c}{\sqrt{c^T A^{-1} c}} = c^T x_c - \sqrt{c^T A^{-1} c}$

6.

$$x^T B x \to \min$$
$$\text{s.t. } x^T A x \leq 1, \quad A \in \mathbb{S}^n_{++}, \quad B \in \mathbb{S}^n_+$$

Так как $B \in \mathbb{S}^n_+$, верно, что $x^T B x \geq 0 \quad \forall x \in \mathbb{R}^n$. Тогда на нулевом векторе $x^* = 0$ будет достигаться эта оценка, и условие $x^T A x \leq 1$ тоже будет выполнено, то есть ответом будет $p^* = 0$. Это значение будет достигаться на всех таких x, что $x^T B x = 0$ и $x^T A x \leq 1$

7.

$$\|Ax - b\|_2^2 \to \min$$
$$\text{s.t. } Cx = d$$

$L(x, \lambda) = \|Ax - b\|_2^2 + \lambda^T(Cx - d)$
$\nabla_x L = 2A^T(Ax - b) + C^T \lambda = 0$
$2A^T Ax - 2A^T b + C^T \lambda = 0 = 2Gx - 2A^T b + C^T \lambda \quad \Rightarrow \quad \lambda^T C = 2b^T A - 2x^T G \quad \Rightarrow$
$\Rightarrow \lambda = 2(CC^*)^{-1}(C^*)^T A^T b - 2(CC^*)^{-1}(C^*)^T G x$
При этом, $Cx = d \quad \Rightarrow \quad C^* C x = C^* d \quad \Rightarrow \quad x = ({}^* C)^{-1} C^* d$. Подставим его в выражение для λ:
$\lambda = 2(CC^*)Y - 1\overline{C} A^T b - 2(CC^*)^{-1} \overline{C} G (C^* C)^{-1} C^* d$, где $\overline{C}$ - сопряженная к C матрица (но не транспонированная!)

8.

$$trX - \ln \det X \to \min$$
$$\text{s.t. } Xs = y, \quad y^T s = 1$$

$L(X, \lambda) = trX - \ln \det X = \lambda^T(Xs - y)$
$\nabla_x L = E - X^{-1} + \lambda s^T = 0 \quad \Rightarrow \quad X^{-1} = E + \lambda s^T$. При этом, $(X^{-1})^T = X^{-1}, \quad E^T = E \quad \Rightarrow$ $(\lambda s^T)^T = s\lambda^T = \lambda s^T$. Тогда $X^{-1} = E + \frac{1}{2}(\lambda s^T + s\lambda^T)$
$Xs = y \quad \Rightarrow \quad s = X^{-1} y = (E + \frac{1}{2}\lambda s^T + \frac{1}{2} s \lambda^T) y = y + \frac{1}{2}\lambda + \frac{1}{2}s - \frac{1}{2} y^T y \quad \Rightarrow \quad \lambda = -2y + (1 + y^T y)s$. Подставим это в выражение для X^{-1}:
$X^{-1} = E + (-2y + (1 + y^T y)s)s^T = E - 2ys^T + ss^T + y^T y s s^T = E + (1 + y^T y)ss^T - 2ys^T$. При этом, аналогичным образом можно показать, что $ys^T = sy^T$, и тогда $X^{-1} = E + (1 + y^T y)ss^T - ys^T - sy^T$. Теперь покажем, что $(X^*)^{-1} = X^{-1}$:
$X^{-1} X^* = E + yy^T - \frac{1}{s^T s} ss^T + (1 + y^T y)(ss^T + yy^T ss^T - ss^T) - ys^T - ys^T yy^T + \frac{1}{s^T s} ys^T ss^T - sy^T - y^T y sy^T + \frac{1}{s^T s} sy^T ss^T = E + ss^T + ys^T - ss^T + y^T y ss^T + y^T y y s^T - y^T y s s^T - ys^T + ys^T - sy^T - y^T y s y^T = E$

9.

$$f_0(x) \to \min$$
$$\text{s.t. } f_i(x) \leq 0, \quad f_i - \text{convex}$$

Предполагаем, что выполняются условия ККТ.
Так как f_i выпуклые, $f_i(x) - f_i(x^*) \geq \nabla f_i(x^*)^T (x - x^*)$, откуда

$$f_i(x) \geq f_i(x^*) + \nabla f_i(x^*)^T (x - x^*)$$

При этом, $\mu_i^* \geq 0$. Значит, $\sum\limits_{i=1}^n \mu_i^* f_i(x) \geq \sum\limits_{i=1}^n \mu_i^* (f_i(x^*) + \nabla f_i(x^*)^T)(x - x^*) = \sum\limits_{i=1}^n \mu_i^* f_i(x^*) + \mu_i^* \nabla f_i(x^*)^T (x - x^*) = -\nabla f_0(x^*)^T (x - x^*) \leq 0 \quad \Rightarrow \quad \nabla f_0(x^*)^T (x - x^*) \geq 0$

Duality

1.

$$c^T x \to \min$$
$$\text{s.t. } f(x) \le 0$$

$g(\lambda) = \inf\limits_{x\in\mathbb{R}^n} c^T x + \lambda^T f = \inf\limits_{x\in\mathbb{R}^n} \lambda^T((\dfrac{c}{\lambda})^T x + f) = -\sup\limits_{x\in\mathbb{R}^n} \lambda^T((-\dfrac{c}{\lambda})^T x - f) =$
$= -\lambda^T \sup\limits_{x\in\mathbb{R}^n}((-\dfrac{c}{\lambda})^T x - f) = \lambda^T f^*(-\dfrac{c}{\lambda})$
При этом понятно, что если $\lambda_i < 0$, то функция не будет ограниченной снизу, поэтому можно сразу рассматривать $\lambda \succeq 0$. Тогда получаем такую двойственную задачу:

$$-\lambda^T f^*(-\frac{c}{\lambda}) \to \min_{\lambda\succeq 0}$$

Она является выпуклой, т.к. сопряжённая функция всегда выпуклая и функция ограничения $\lambda \succeq 0$ является аффинной.

2.

$$\ln\det X^{-1} \to \min$$
$$\text{s.t. } a_i^T X a_i \le 1$$

$L(X,\mu) = \ln\det X^{-1} + \sum\limits_{i=1}^m \mu_i(a_i^T X a_i - 1)$
$g(\mu) = \inf\limits_{X\in\mathbb{S}^n_{++}}(\ln\det X^{-1} + \sum\limits_{i=1}^m \mu_i(a_i^T X a_i - 1))$
$\nabla_x L = -X^{-1} + \sum\limits_{i=1}^m \mu_i a_i a_i^T = 0 \quad \Rightarrow \quad X^{-1} = \sum\limits_{i=1}^m \mu_i a_i a_i^T$
Тогда двойственной задачей будет

$$g(\mu) = \ln\det(\sum_{i=1}^m \mu_i a_i a_i^T) + \sum_{j=1}^m \mu_j(a_j^T(\sum_{i=1}^m \mu_i a_i a_i^T)^{-1} a_j - 1) \to \min_{\mu\succeq 0}$$

Проверим наличие сильной двойственности:
Возьмём $X = \dfrac{E}{\max\limits_i \|a_i\|^2 + 1}$. Тогда для этой матрицы будет так:

$$a_i^T X a_i = a_i^T E a_i \frac{1}{\max\limits_j \|a_j\|^2 + 1} = \frac{\|a_i\|^2}{\max\limits_j \|a_j\|^2 + 1} < 1$$

То есть сильная двойственность есть в данной задаче.
(далее надо решить двойственную задачу, но что-то не очень ясно как её решить)

3.

Так как $\tilde{x}$ - точка минимума, то $\nabla_x \phi(\tilde{x}) = 0$:
$(*) \nabla f(\tilde{x})^T + 2\alpha A^T(A\tilde{x} - b) = 0$.
Пусть $\lambda = 2\alpha(A\tilde{x} - b)$. Тогда $(*) \to \nabla f(\tilde{x})^T + \lambda A^T = \nabla_x(f(x) + \lambda^T(Ax - b))(\tilde{x}) = 0$, то есть $\tilde{x}$ также является точкой минимума функции $f(x) + \lambda^T(Ax - b)$. Тогда λ будет точкой минимума у двойтсвенной функции $g(\lambda) = \inf\limits_{x\in\mathbb{R}^n} f(x) + \lambda^T(Ax - b) = f(\tilde{x}) + 2\alpha\|A\tilde{x} - b\|_2^2$, откуда следует, что $\min f(x) \ge f(\tilde{x}) + 2\alpha\|A\tilde{x} - b\|_2^2$.

4.

$$-\sum_{i=1}^{m} \ln(b_i - a_i^T x) \to \min$$

Пусть $y_i = b_i - a_i^T x$, то есть $y = b - Ax$(A - матрица, составленная из столбцов a_i). Тогда $-\sum\limits_{i=1}^{m} \ln(b_i - a_i^T x) \to -\sum\limits_{i=1}^{m} \ln y_i$, то есть наша задача теперь выглядит так:

$$-\sum_{i=1}^{m} \ln y_i \to \min_{y=b-Ax}$$

$L = -\sum\limits_{i=1}^{m} \ln y_i + \lambda^T(y - b + Ax)$

$g(\lambda) = \inf\limits_{x \in \mathbb{R}^n, y \in \mathbb{R}} (-\sum\limits_{i=1}^{m} \ln y_i + \lambda^T(y - b + Ax))$. Функция под инфимумом ограничена снизу, если $\lambda^T A = 0$ и $\lambda \succ 0$.

$\nabla_{y_i}(-\sum\limits_{i=1}^{m} \ln y_i + \lambda_i y_i) = -\frac{1}{y_i} + \lambda_i = 0 \quad \Rightarrow \quad y_i = \frac{1}{\lambda_i}$. Тогда, с ограничениями $\lambda^T A = 0$ и $\lambda \succ 0$, двойственная функция $g(\lambda) = -\sum\limits_{i=1}^{m} \ln \lambda^{-1} + \lambda^T y - \lambda^T b = \sum\limits_{i=1}^{m} \ln \lambda_i + m - \lambda^T b$. Тогда двойственная задача будет выглядит так:

$$\sum_{i=1}^{m} \ln \lambda_i + m - \lambda^T b \to \max$$

$$\text{s.t. } \lambda^T a = 0$$

$$\lambda \succ 0$$