ДЗ-3 по Методам Оптимизации
Попов Николай, 776 группа

General optimization problems

(№1)
Рассмотрим все возможныве случаи:

а) Если система $Ax = b$ несовместна (вне зависимости от размеров матрицы А), то решения задачи нет.

б)$L(x, \lambda) = c^T x + \lambda^T (Ax - b)$

$$\begin{cases} L_x = c + A^T\lambda = 0; \\ L_\lambda = Ax - b = 0 \end{cases} \Rightarrow \text{ если обе эти системы совместны, то}$$

$c^T x = -\lambda^T A \cdot x = -\lambda^T (Ax) = -\lambda^T b = min(c^T x)$ по ККТ

в)Если система $Ax = b$ совместна, но система $A^T\lambda = -c$ несовместна, то $\exists x_0 : Ax_0 = b$ и $\forall v \neq 0 : Av = 0, x' = x_0 + vt$ $(t \in R) \hookrightarrow Ax' = b$. Тогда значение $c^T x' = c^T (x_0 + tv)$ неограничено снизу (при $t \to \infty$).

(№2)
Изменим порядок компонент вектора с и соотвественных компонент вектора х в возрастающем порядке: $c_1 = c_2 = \cdots = c_k < c_{k+1} \leq \cdots c_n$.
Так как в силу условия задачи $c^T x \geq c_1 \sum_i x_i = \underset{i}{min}\ c_i = c_{min}$, то для достижения равенства в неравенстве $c^T x \geq c_{min}$ и удовлетворения условия получаем: $x_1 + x_2 + \cdots + x_k = 1,\ x_1 \geq 0, \cdots, x_k \geq 0, x_{k+1} = \cdots = x_n = 0$

(№3)
В случае равенства $1^T x = \alpha$ в ограничениях при целом или вещественном α, чтобы получить меньшую величину $\sum_i c_i x_i$, нужно применить "жадный алгоритм"и делать x_i больше, если он соответсвует меньшему c_i. Поэтому начинаем "тратить"α начиная с меньших c_i. Отсортируем координаты вектора с в порядке возрастания.

Из-за ограничения $0 \leq x \leq 1$ получаем при целом α, что $min(c^T x) = c_1 + \cdots + c_\alpha$ (сумма минимальных α координат), что достигается, например, при $x_1 = \cdots = x_\alpha = 1,\ x_{\alpha+1}, \cdots, x_n = 0$ (порядок соответсвует отсортированному вектору с). При этом точка минимума может быть и другой, если среди минимальных компонент вектора есть равные. В таком случае, среди минимальных α компонент с $x_i = 1$ для всех c_i, которые уникальны (пусть их число равно k), а для остальных равных между собой компонент c_i оставшийся вес можно распределить произвольно: $\{\underline{c_i, \cdots, c_{i+k-1}}\}: \ \forall p, q \in \overline{i, i+k-1} c_p = c_q \hookrightarrow x_i + \cdots + x_{i+k-1} = \alpha - k$ и $\forall p \in \overline{i, i+k-1} \hookrightarrow 0 \leq x_p \leq 1$. В любом случае, остальные координаты вектора x равны нулю $\forall p \in \overline{\alpha+1, n} \hookrightarrow x_p = 0$.

При вещественном α аналогично:
$min(c^T x) = c_1 + \cdots + c_{\lfloor\alpha\rfloor} + (\alpha - \lfloor\alpha\rfloor) c_{\lfloor\alpha\rfloor+1}$

В случае нервенства $1^T x \leq \alpha$ и целого α минимальным значением $c^T x$ будет сумма α наименьших отрицательных компонент вектора c, если их число не меньше α, а если меньше, то $min(c^T x$ равен сумме всех отрицательных компонент.

(№4)
Пусть $A \in S^n_{++}, \ \lambda \in R$

$$L = < c, x > + < \lambda x, Ax > - \lambda \begin{cases} L_x = c + 2\lambda Ax = 0 & (1) \\ L_\lambda = x^T Ax - 1 = 0 & (2) \end{cases}$$

$$(1) \Rightarrow x = -\frac{1}{2\lambda} A^{-1} c \Rightarrow (2) \Leftrightarrow \frac{1}{(2\lambda)^2} c^T A^{-T} A A^{-1} c = 1 \Rightarrow 2\lambda = \sqrt{c^T A^{-1} c} \Rightarrow$$

$$x_{\text{опт}} = -\frac{1}{\sqrt{c^T A^{-1} c}} A^{-1} c \Rightarrow c^T x_{\text{опт}} = -\frac{c^T A^{-1} c}{\sqrt{c^T A^{-1} c}} = -\sqrt{c^T A^{-1} c}$$

$L_{xx} = 2\lambda A \succ 0$
При таком λ по ККТ получаем, что $x_{\text{опт}}$ - точка минимума функции.

В случае, если матрица A не является положительно определенной, то

$A = QDQ^T, Q = [u_1, \cdots, u_n], u_i -$ собственный вектор, соответсвующий λ_i

Вводя обозначение $C^{(k)} = u_k u_k^T$, получим строку матрицы $B = QD$:

$$B_{i[1,n]} = (\lambda_1 u_1^i; \lambda_2 u_2^i; \cdots; \lambda_n u_n^i) \Rightarrow$$

$$A_{ij} = \sum_{k=1}^n \lambda_k u_k^i u_k^j = \sum_{k=1}^n \lambda_k C_{ij}^{(k)} \Rightarrow A = \sum_{k=1}^n \lambda_k u_k u_k^T \qquad (*)$$

Сделаем замену переменных в задаче, используя свойство ортого-

нальных матриц $Q^T = Q^{-1}$: $c^T x = c^T Q Q^T x = (Q^T c)^T \cdot (Q^T x) = b^T \cdot y$, т.е. $Q^T c = b$; $Q^T x = y$. В силу невырожденности матрицы Q оптимизация по x и по введенному вектору y равносильны. Используя (*), понимаем как изменится ограничение:
$x^T A x = x^T \left(\sum_{k=1}^n \lambda_k u_k u_k^T \right) x = \sum_{k=1}^n \lambda_k (x^T u_k) u_k^T x = \sum_{k=1}^n \lambda_k (x^T u_k)^2 = \sum_{k=1}^n \lambda_k y_k^2$, поскольку $y^T = x^T Q = (x^T u_1; \cdots ; x^T u_n)$

Получена эквивалентная задача
$$\begin{cases} b^T y \to min \\ y^T diag(\lambda_1, \cdots, \lambda_n) y \le 1 \end{cases} = \begin{cases} \sum_i b_i y_i \to min \\ \sum_i \lambda_i y_i^2 \le 1 \end{cases}$$

Случай, когда все $\lambda_i > 0$ был рассмотрен выше.
Если $\exists q : \lambda_q \le 0, b_q \ne 0$, то y_q можно взять бесконечно большим с нужным знаком чтобы получить $\sum_i b_i y_i = -\infty$
Если $\forall q : \lambda_q \le 0, b_q = 0$, задача сводится к случаю $\forall k \; \lambda_k > 0$, но меньшей размерности (все такие q-ые координаты удалим).

(№5)
Решается аналогично пункту задачи 4, где $A \in S_{++}^n$. Укажу отличия:

$$L = \langle c, x \rangle > +\lambda \langle x - x_c, A(x - x_c) \rangle, \; \lambda \in R$$

$$\begin{cases} L_x = c + 2\lambda A(x - x_c) = 0 \\ L_\lambda = \langle x - x_c, A(x - x_c) \rangle - 1 = 0 \end{cases}$$ Повторим те же вычисления после замены $y = x - x_c$

Находим $y' = x' - x_c = \dfrac{-A^{-1} c}{\sqrt{c^T A^{-1} c}} \Rightarrow min(c^T x') = c^T (x_c - \dfrac{A^{-1} c}{\sqrt{c^T A^{-1} c}}) = c^T x_c - \sqrt{c^T A^{-1} c}$

(№6)

$$L = x^T B x + \lambda x^T A x - \lambda$$

Так как B - положительно полуопределенная, то $x^T B x \ge 0$, т.е. минимум равен 0. Он достигается при $x = 0$, который является допустимым ($x^T A x \le 1$).

(№7)

$$L = \langle Ax - b, Ax - b \rangle + \lambda^T (Gx - d) = x^T A^T A x + \langle G^T \lambda - 2A^T b, x \rangle - \langle \lambda, d \rangle$$

Условия ККТ для задачи с ограничением типа равенство:
$$\begin{cases} L_x = 2A^T(Ax-b)G^T\lambda = 0 \\ L_\lambda = Gx - d = 0 \\ y^T(L_{xx})y \geq 0 \ \forall y : Gy = 0 \end{cases}$$
Тут $L_{xx} = 2A^TA$

Из первого уравнения находим: $x = (A^TA)^{-1}(A^Tb - 0.5G^T\lambda)$
Подставляем это во второе уравнение и выражаем
$\lambda^* = -2[G(A^TA)^{-1}G^T]^{-1} \cdot [d - G(A^TA)^{-1}A^Tb]$
Теперь получаем из последних двух формул
$x^* = (A^TA)^{-1}A^Tb + (A^TA)^{-1}G^T[G(A^TA)^{-1}G^T]^{-1} \cdot [d - G(A^TA)^{-1}A^Tb]$

(№8)

Есть три условия ККТ для оптимальности решения, два из которых получим из лагранжиана:
$L = trX - log(det\ X) + \lambda^T(Xs - y)$

Т.к. $d\langle z, Xs\rangle = d\langle s, Xz\rangle \Rightarrow d\langle z, Xs\rangle = 0.5(d\langle z, Xs\rangle + d\langle s, Xz\rangle) = 0.5(sz^T + zs^T)$, так как $\langle z, Xs\rangle = \sum_{i,j} x_{ij}z_is_j \Rightarrow \partial\langle z, Xs\rangle/\partial x_{kl} = z_ks_l$

$L_x = I - X^{-1} + 0.5(sz^T + zs^T) = 0$
$L_\lambda = Xs - y = 0$
Еще нужно добавить $X \succ 0$.

Для проверки выразим X^{-1} через y и s и домножим на предложенное решение X чтобы получить I (если решение верно). Для этого выразим z через y и s.

Первые два условия ККТ перепишем $X^{-1} = I + 0.5(sz^T + zs^T)$ (2) и $Xs = y$. Домножим первое на второе слева. Поскольку $s^Ty = 1$ а $(sz^T)y = \sum_j s_iz_jy_j = s(z^Ty)$, то
$s = X^{-1}y = y + 0.5(zs^Ty + (sz^T)y) = y + 0.5(z + (z^Ty)s)$ (1).
Домножаем скалярно на y:
$1 = y^Ts = y^Ty + 0.5[y^Tz + (z^Ty)y^Ts] = y^Ty + 0.5[y^Tz + z^Ty] = y^Ty + y^Tz \Rightarrow y^Tz = 1 - y^Ty$.
Подставим это в (1): $s = y + 0.5(z + (1 - y^Ty)s) \Rightarrow z = (1 + y^Ty)s - 2y$
Это подставим в (2):
$X^{-1} = I + 0.5(-2ys^T - sy^T + 2(1 + y^Ty)ss^T) = I + (1 + y^Ty)ss^T - ys^T - sy^T$
Проверим, что $X^{-1}X^* = I$:

$X^{-1}X^* = [I + (1 + y^T y)ss^T - ys^T - sy^T] \cdot [I + yy^T - \frac{1}{s^T s}ss^T] = (I + yy^T - \frac{1}{s^T s}ss^T) + (ss^T - ss^T + sy^T) - (ys^T - ys^T + yy^t) - (sy^T(y^T y)sy^T - - \frac{1}{s^T s}ss^T) = I$
Значит, действительно, X^* - оптимальная точка, поскольку она удовлетворяет условиям ККТ.

(№9)
Пусть x - произвольный вектор, удовлетворяющий ограничениям задачи. Поскольку задача выпуклая, то ограничения задания выпуклыми функциями, значит, по определению, в котором y=x^*:
$f_i(x^*) + [\bigtriangledown f_i(x^*)]^T(x - x^*) \leq f_i(x) \leq 0 \forall i$

Используя это и условия ККТ преобразуем:
$\bigtriangledown f_0(x^*)(x - x^*) = -\sum_i \lambda_i^* \bigtriangledown^T f_i(x^*)(x - x^*) = -\sum_i \lambda_i^* \bigtriangledown f_i(x^*) - \sum_i \lambda_i^* f_i(x^*) = -\sum_i \lambda_i [f_i(x^*) + \bigtriangledown^T f_i(x^*)(x - x^*) \geq 0$

General optimization problems

(№1)
Составим лагранжиан для этой задачи:
$L = c^T x + \lambda f(x),\ \lambda \geq 0$
Тогда целевая функция в двойственной задаче:
$g(\lambda) = inf_x\, L(x, \lambda) = inf_x(c^T x + \lambda f(x)) = -sup_x(-c^T - \lambda f(x)) = -\lambda f^* \left(-\frac{c}{\lambda}\right)$

Тогда получена двойственная задача:
$$\begin{cases} -\lambda f^* \left(-\frac{c}{\lambda}\right) \to max \\ \lambda \geq 0 \end{cases} \Leftrightarrow \begin{cases} \lambda f^* \left(-\frac{c}{\lambda}\right) \to min \\ \lambda \geq 0 \end{cases}$$

Эта задача минимизации выпуклая задача оптимизации, так как сопряженная функция всегда выпукла как поточенный супремум и ограничения типа неравенств заданы выпуклыми функциям.

(№2)
Преобразуем ограничение:
$a_i^T X a_i = \sum_{p,q} a_i^p a_i^q x_{pq} = tr(a_i a_i^T X)$
Тогда
$L = ln(det\ X^{-1}) + \sum_i \lambda_i (tr(a_i a_i^T X) - 1) = ln(det\ X^{-1}) + tr(\sum_i \lambda_i a_i a_i^T X) - 1^T \lambda$
Целевая функция есть инфимум лагранжиана по X:

$L_X = -X^{-1} + \sum_i \lambda_i(a_i a_i^T) = 0 \Rightarrow X_*^{-1} = \sum_i \lambda_i(a_i a_i^T) \Rightarrow inf_X\ L(X, \lambda) = L(\lambda, X_*) = ln\ det(\sum_i \lambda_i(a_i a_i^T)) + tr(I) - 1^T\lambda = ln\ det(\sum_i \lambda_i(a_i a_i^T)) + n - 1^T\lambda = g(\lambda)$

Тогда двойственная задача:
$\begin{cases} ln\ det(\sum_i \lambda_i(a_i a_i^T)) + n - 1^T\lambda \to max \\ \forall i\ \lambda_i \geq 0 \end{cases}$ (аргумент логарифма положительный должен быть)

Сильная двойственность есть, поскольку есть допустимая точка $\lambda = \overline{1}$.

(№3)

Из необходимого условия минимума для функции ϕ получаю: $\bigtriangledown\phi(\tilde{x}) = 0 =$

$= \bigtriangledown f_0(\tilde{x}) + 2\alpha A^T(A\tilde{x} - b) = 0$, но поскольку для функции $\varphi = f_0(x) + 2\alpha\langle A\tilde{x} - b, Ax - b\rangle$, градиент такой же, то $min\ \varphi = min\ \phi = f_0(\tilde{x}) + 2\alpha\|A\tilde{x} - b\|_2^2$.

Заметим, что лагранжиан исходной задачи $L = f_0(x) + \langle\lambda, Ax - b\rangle$, и целевая функция в задаче, двойственной к исходной это $g(\lambda) = inf\ L = inf_x(f_0(x) + \langle\lambda, Ax - b\rangle)$. При $\lambda = 2\alpha(A\tilde{x} - b) \hookrightarrow L = \varphi \Rightarrow g = f_0(\tilde{x}) + 2\alpha\|A\tilde{x} - b\|_2^2$. Тогда $\forall x : Ax = b \hookrightarrow f_0(x) \geq f_0(\tilde{x}) + 2\alpha\|A\tilde{x} - b\|_2^2$.