KKT and duality

Ковалев Алексей

1.

$$x^2 + 1 \to \min_{x \in \mathbb{R}}$$
$$\text{s.t. } (x-2)(x-4) \leqslant 0$$

Feasible set – $[2; 4]$, минимум функции равен $p^* = 5$, достигается на $x^* = 2$.

$$L(x,\, \lambda) = x^2 + 1 + \lambda(x-2)(x-4)$$
$$g(\lambda) = \inf_{x \in R} L(x,\, \lambda)$$

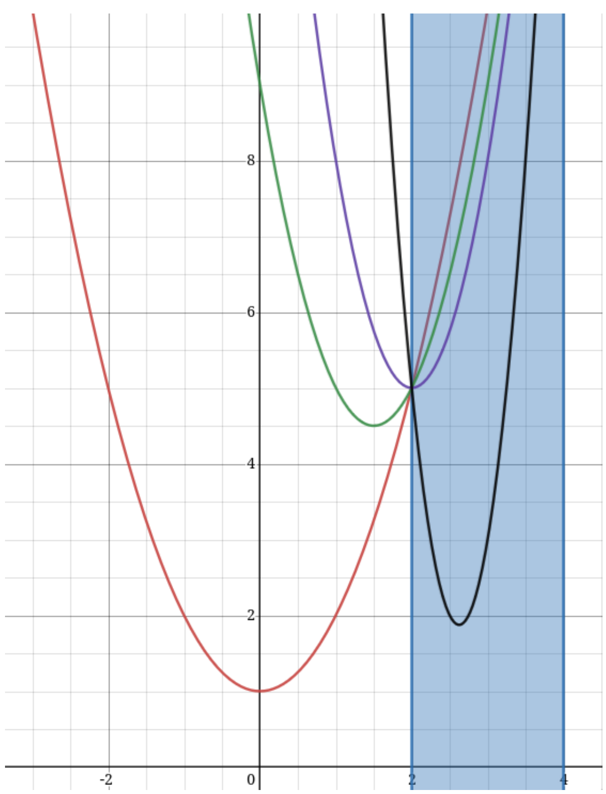
В силу выпуклости лагранжиана получаем, что $\inf\limits_{x \in R} L(x,\, \lambda)$ достигается на $x^* : \nabla_x L(x^*,\, \lambda) = 0$.

$$2x^* + \lambda(2x^* - 6) = 0$$

Тогда двойственная задача

$$\frac{9\lambda^2}{(1+\lambda)^2} + 1 - \lambda \cdot \frac{(\lambda-2)(\lambda+4)}{(1+\lambda)^2} \to \max_{\lambda \in \mathbb{R}}$$
$$\text{s.t. } \lambda \geqslant 0$$

Вогнутость этой задачи непосредственно проверяется дифференциальным критерием выпуклости первого порядка при $\lambda \geqslant 0$. Также из условия $\nabla g(\lambda^*) = 0$ получаем, что $\lambda^* = 2$, $d^* = 5$, то есть сильная двойственность присутствует.
Красным на рисунке x^2+1, зеленым, фиолетовым, черным – лагранжиан при разных λ (1, 2, 7 соответственно)



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

$$\frac{1}{2}\|y-b\|_2^2 + \frac{\lambda}{2}\|x\|_2^2 \to \min_{x\in\mathbb{R}^n,\, y\in\mathbb{R}^m}$$
$$\text{s.t. } y = Ax$$

$$L(x,\, y,\, \mu) = \frac{1}{2}\|y-b\|_2^2 + \frac{\lambda}{2}\|x\|_2^2 + \mu^\top(y - Ax)$$
$$g(\mu) = \inf_{x\in\mathbb{R}^n,\, y\in\mathbb{R}^m} L(x,\, y,\, \mu)$$

$L(x,\, y,\, \mu)$ – выпуклая по x и y функция. Значит $\inf\limits_{x\in\mathbb{R}^n,\, y\in\mathbb{R}^m} L(x,\, y,\, \mu)$ достигается при x^*, y^*, таких что $\nabla_{(x,y)}L(x^*,\, y^*,\, \mu) = 0$.

$$\nabla_x L(x^*,\, y^*,\, \mu) = \lambda x^* - A^\top\mu = 0;\ \nabla_y L(x^*,\, y^*,\, \mu) = y^* - b + \mu = 0$$
$$g(\mu) = \frac{1}{2}\|\mu\|_2^2 + \frac{1}{2\lambda}\|A^\top\mu\|_2^2 + \mu^\top\left(b - \mu - \frac{1}{\lambda}AA^\top\mu\right)$$

Ответ:

$$\frac{1}{2}\|\mu\|_2^2 + \frac{1}{2\lambda}\|A^\top\mu\|_2^2 + \mu^\top\left(b - \mu - \frac{1}{\lambda}AA^\top\mu\right) \to \max_{\mu\in\mathbb{R}^m}.$$

3. $A \in \mathbb{R}^{m\times n}$, $\lambda > 0$.

$$\langle \mathbf{1},\, t\rangle + \frac{\lambda}{2}\|x\|_2^2 \to \min_{x\in\mathbb{R}^n,\, t\in\mathbb{R}^m}$$
$$\text{s.t. } Ax \succeq \mathbf{1} - t,$$
$$t \succeq 0$$

$$L(x,\, t,\, \mu,\, \nu) = \mathbf{1}^\top t + \frac{\lambda}{2}\|x\|_2^2 + \mu^\top(\mathbf{1} - t - Ax) - \nu^\top t$$
$$g(\nu, \lambda) = \inf_{x\in\mathbb{R}^n,\, t\in\mathbb{R}^m} L(x,\, t,\, \mu,\, \nu)$$

Функция $L(x,\, t,\, \mu,\, \nu)$ выпукла по $(x,\, t)$. Значит $\inf\limits_{x\in\mathbb{R}^n,\, t\in\mathbb{R}^m} L(x,\, t,\, \mu,\, \nu)$ достигается на x^*, t^*, таких что $\nabla_{(x,\,t)}L(x^*,\, t^*,\, \mu,\, \nu) = 0$.

$$\nabla_x L(x^*,\, t^*,\, \mu,\, \nu) = \lambda x^* - A^\top\mu = 0;\ \nabla_t L(x^*,\, t^*,\, \mu,\, \nu) = \mathbf{1} - \mu - \nu = 0$$
$$x^* = \frac{1}{\lambda}A^\top\mu;\ \mathbf{1} - \mu - \nu = 0$$

Отсюда получаем

$$g(\mu,\, \nu) = (\mathbf{1} - \mu - \nu)^\top t + \frac{\lambda}{2}\|x^*\|_2^2 + \mathbf{1}^\top\mu - \mu^\top Ax^* = -\frac{1}{2\lambda}\|A^\top\mu\|_2^2 + \mathbf{1}^\top\mu$$

Ответ:

$$\mathbf{1}^\top\mu - \frac{1}{2\lambda}\|A^\top\mu\|_2^2 \to \max_{\mu\in\mathbb{R}^m,\, \nu\in\mathbb{R}^m}$$
$$\text{s.t. } \lambda \succeq 0$$

4.

$$
\begin{gathered}
c^\top x \to \min_{x \in \mathbb{R}^n} \\
\text{s.t. } \mathbf{1}^\top x = 1, \\
x \succeq 0
\end{gathered}
$$

$$
L(x,\, \lambda,\, \nu) = c^\top x + \nu(\mathbf{1}^\top x - 1) - \lambda^\top x
$$

В данном случае выполняется условие регулярности LCQ, так как $\mathbf{1}^\top x - 1$ и x – аффинные функции. Значит ККТ – необходимые и достаточные условия. То есть

$$
\begin{gathered}
\nabla_x L(x^*,\, \lambda^*,\, \nu^*) = c + \nu^* \mathbf{1} - \lambda = 0 \\
\nabla_\nu L(x^*,\, \lambda^*,\, \nu^*) = \mathbf{1}^\top x^* - 1 = 0 \\
\lambda_i^* \geqslant 0,\, i = 1,\, \ldots,\, n \\
\lambda_i^* x_i^* = 0,\, i = 1,\, \ldots,\, n \\
x_i \geqslant 0,\, i = 1,\, \ldots,\, n
\end{gathered}
$$

Отсюда получаем, что

$$
\begin{gathered}
c_i + \nu^* = \lambda_i^* \geqslant 0,\, i = 1,\, \ldots,\, n \\
\nu^* \geqslant -\min_{i=1,\,\ldots,\,n} c_i,\, i = 1,\, \ldots,\, n
\end{gathered}
$$

Пусть $j = \operatorname*{arg\,min}_{i=1,\,\ldots,\,n} c_i$. Тогда $\nu^* = -c_j;\; \lambda^* = c + \nu^* \mathbf{1};\; x_j = 1;\; x_i = 0,\, i \neq j$ удовлетворяет всем условия ККТ, а значит является решением задачи. Других решений нет, так как если $\nu^* \neq -c_j$, то все $\lambda_i^* \neq 0$, а значит все $x_i^* = 0$ и второе условие ККТ не выполняется.

Ответ: $x^* = (0,\, \ldots,\, 1,\, \ldots,\, 0)^\top$, где 1 стоит на j-том месте, $j = \operatorname*{arg\,min}_{i=1,\,\ldots,\,n} c_i$.

5. $C \in \mathbb{S}^n_{++}$, $0 \neq a \in \mathbb{R}^n$.

$$
\begin{gathered}
\langle C^{-1},\, X \rangle - \log\det X \to \min_{X \in \mathbb{R}^{n \times n}} \\
\text{s.t. } \langle Xa,\, a \rangle \leqslant 1
\end{gathered}
$$

$$
L(X,\, \lambda) = \langle C^{-1},\, X \rangle - \log\det X + \lambda\big(\langle Xa,\, a \rangle - 1\big)
$$

Неравенство $\langle Xa,\, a \rangle \leqslant 1$ можно представить в виде $\langle X,\, aa^\top \rangle \leqslant 1$, что является аффинной функцией, а значит справедливо условие регулярности LCQ. Тогда ККТ – необходимые и достаточные условия. Получаем

$$
\nabla_x L(X^*,\, \lambda^*) = C^{-1} - {X^*}^{-\top} + \lambda^* aa^\top = 0
$$

$$
\begin{gathered}
\lambda^* \geqslant 0 \\
\lambda^*\big(\langle X^*,\, aa^\top \rangle - 1\big) = 0 \\
\langle X^*,\, aa^\top \rangle \leqslant 1
\end{gathered}
$$

Воспользовавшись формулой для обращения матрицы с поправкой ранга 1 получаем

$$
X^* = C - \frac{\lambda^* C aa^\top C}{1 + \lambda^* a^\top C a}
$$

Третье уравнение распадается на два варианта:

- $\lambda^* = 0$. Из соотношения выше $X^* = C$ единственно.

- $\langle X^*,\, aa^\top \rangle = a^\top X^* a = 1$. Домножив первое уравнение скалярно на a справа и на $a^\top$ слева получаем

$$1 = a^\top X^* a = a^\top Ca - \frac{\lambda^*}{1+\lambda^* a^\top Ca} a^\top Caa^\top Ca$$

$$1 + \lambda^* a^\top Ca = a^\top Ca$$

$$\lambda^* = \frac{a^\top Ca - 1}{a^\top Ca},\ a^\top Ca \geqslant 1$$

$$X^* = \frac{a^\top CaCa^\top Ca - a^\top CaCaa^\top C + Caa^\top C}{a^\top Caa^\top Ca},\ a^\top Ca \geqslant 1$$

  Причем X^* снова единственно.

Таким образом, единственность X^* доказана.
Ответ:

$$X^* = \begin{cases} C, & a^\top Ca \leqslant 1 \\ \frac{a^\top CaCa^\top Ca - a^\top CaCaa^\top C + Caa^\top C}{a^\top Caa^\top Ca}, & a^\top Ca > 1 \end{cases}$$

6. $A \in \mathbb{S}^n_{++}$, $c \neq 0$, $x_c \in \mathbb{R}^n$

$$c^\top x \to \min_{x \in \mathbb{R}^n}$$
$$\text{s.t. } (x - x_c)^\top A(x - x_c) \leqslant 1$$

$$L(x,\, \lambda) = c^\top x + \lambda\big((x-x_c)^\top A(x-x_c) - 1\big)$$

В данном случае справедливо условие регулярности SC, так как при $x = x_c$ неравенство становится строгим и все функции выпуклы (для строгости можно считать, что есть аффинные равенства вида $0 \cdot x = 0$, но они никак не повлияют на решение). Значит ККТ – необходимые и достаточные условия. Запишем их

$$\nabla_x L(x^*,\, \lambda^*) = c + 2\lambda A(x^* - x_c) = 0$$

$$\lambda^* \geqslant 0$$
$$\lambda^*\big((x^*-x_c)^\top A(x^*-x_c) - 1\big) = 0$$
$$(x^*-x_c)^\top A(x^*-x_c) \leqslant 1$$

Отсюда получаем

$$x^* = x_c - \frac{1}{2\lambda^*} A^{-1} c$$

Третье уравнение ККТ дает два варианта

- $\lambda^* = 0$. Тогда $c = 0$, что противоречит условию.
- $(x^*-x_c)^\top A(x^*-x_c) = 1$. Тогда $c^\top A^{-1} c = 4\lambda^{*2}$, то есть $\lambda^* = \frac{1}{2}\sqrt{c^\top A^{-1} c}$.

Итого получаем

$$x^* = x_c - \frac{1}{\sqrt{c^\top A^{-1} c}} A^{-1} c$$

Ответ:

$$x^* = x_c - \frac{1}{\sqrt{c^\top A^{-1} c}} A^{-1} c;\ p^* = c^\top x^*$$

7. $A \in \mathbb{R}^{m\times n}$, $\operatorname{rk} A = n$, $C \in \mathbb{R}^{k\times n}$, $\operatorname{rk} C = k$.

$$\|Ax - b\|_2^2 \to \min_{x\in\mathbb{R}^n}$$
$$\text{s.t. } Cx = d$$

$$L(x,\, \nu) = \|Ax - b\|_2^2 + \nu^\top (Cx - d)$$

Единственное ограничение – аффинная функция, значит справедливо условие регулярности LCQ. Тогда воспользуемся ККТ

$$\nabla_x L(x^*,\, \nu^*) = 2A^\top (Ax^* - b) + C^\top \nu^* = 0$$
$$\nabla_\nu L(x^*,\, \nu^*) = Cx^* - d = 0$$

Из этих условий получаем

$$x^* = C^\top \left(CC^\top\right)^{-1} d$$
$$x^* = \left(A^\top A\right)^{-1} A^\top b - \frac{1}{2}\left(A^\top A\right)^{-1} C^\top \nu^*$$
$$\left(A^\top A\right)^{-1} C^\top \nu^* = 2\left(A^\top A\right)^{-1} A^\top b - 2C^\top \left(CC^\top\right)^{-1} d$$
$$C^\top \nu^* = 2A^\top b - 2A^\top A C^\top \left(CC^\top\right)^{-1} d$$
$$\nu^* = \left(CC^\top\right)^{-1} C \left(2A^\top b - 2A^\top A C^\top \left(CC^\top\right)^{-1} d\right)$$

Ответ: $x^* = C^\top \left(CC^\top\right)^{-1} d$; $\nu^* = \left(CC^\top\right)^{-1} C \left(2A^\top b - 2A^\top A C^\top \left(CC^\top\right)^{-1} d\right)$.

8. $y \in \mathbb{R}^n$, $s \in \mathbb{R}^n$, $y^\top s = 1$.

$$\operatorname{tr} X - \log\det X \to \min_{X\in\mathbb{S}^n_{++}}$$
$$\text{s.t. } Xs = y$$

$$L(x,\, \nu) = \operatorname{tr} X - \log\det X + \nu^\top (Xs - y)$$

Единственное ограничение – аффинная функция, значит справедливо условие регулярности LCQ. Запишем ККТ

$$\nabla_X L(X^*,\, \nu^*) = I - X^{*-\top} + \nu^* s^\top = 0$$
$$\nabla_\nu L(X^*,\, \nu^*) = X^* s - y = 0$$

Проверим, что

$$X^* = I + yy^\top - \frac{1}{s^\top s} ss^\top$$

дает оптимальное решение поставленной задачи.

$$s + yy^\top s - \frac{1}{s^\top s} ss^\top s - y = s + y - s - y = 0$$

$$I - \left(I + yy^\top - \frac{1}{s^\top s} ss^\top\right)^{-\top} + \nu^* s = I - \left(I + yy^\top\right)^{-1} + \frac{\left(I + yy^\top\right)^{-1} ss^\top \left(I + yy^\top\right)^{-1}}{s^\top s - s^\top \left(I + yy^\top\right)^{-1} s} + \nu^* s^\top = 0$$

$$\left(I + yy^\top\right)^{-1} = I - \frac{yy^\top}{1 + y^\top y}$$

$$-\nu^* s^\top = \frac{yy^\top}{1 + y^\top y} + \frac{1}{s^\top s - s^\top \left(I - \frac{yy^\top}{1+y^\top y}\right) s} \left(I - \frac{yy^\top}{1 + y^\top y}\right) ss^\top \left(I - \frac{yy^\top}{1 + y^\top y}\right)$$

$$\nu^* = -\frac{yy^\top y}{1 + y^\top y} - \frac{1}{s^\top s - s^\top \left(I - \frac{yy^\top}{1+y^\top y}\right) s} \left(I - \frac{yy^\top}{1 + y^\top y}\right) ss^\top \left(I - \frac{yy^\top}{1 + y^\top y}\right) y$$

ν^* существует и единственно, второе соотношение из ККТ также выполняется. Значит

$$X^* = I + yy^\top - \frac{1}{s^\top s} ss^\top$$

действительно решение исходной задачи. □

9. Дана выпуклая задача

$$f_0(x) \to \min_{x \in \mathbb{R}^n}$$
$$\text{s.t. } f_i(x) \leqslant 0,\ i = 1,\ \ldots,\ n$$

$\exists x^* \in \mathbb{R}^n$, $\exists \mu^* \in \mathbb{R}^m$:

$$\nabla_x L(x^*,\ \mu^*) = \nabla f_0(x^*) + \sum_{i=1}^{m} \mu_i^* \nabla f_i(x^*) = 0$$
$$\mu_i^* \geqslant 0,\ i = 1,\ \ldots,\ m$$
$$\mu_i^* f_i(x^*) = 0,\ i = 1,\ \ldots,\ m$$
$$f_i(x^*) \leqslant 0,\ i = 1,\ \ldots,\ m$$

Рассмотрим

$$\nabla f_0^\top(x^*)(x - x^*) = -\sum_{i=1}^{m} \mu_i^* \nabla f_i^\top(x^*)(x - x^*)$$

В силу выпуклости всех f_i получаем

$$f_i(x) \geqslant f_i(x^*) + \nabla f_i^\top(x^*)(x - x^*)$$
$$\nabla f_0^\top(x^*)(x - x^*) \geqslant -\sum_{i=1}^{m} \mu_i^* \left(f_i(x) - f_i(x^*)\right) = \sum_{i=1}^{m} \mu_i^* \left(f_i(x^*) - f_i(x)\right) = -\sum_{i=1}^{m} \mu_i^* f_i(x) \geqslant 0$$

Последнее неравенство справедливо, так как $f_i(x) \leqslant 0$ и $\mu_i^* \geqslant 0$. $\square$

10. $0 \neq c \in \mathbb{R}^n$.

$$c^\top x \to \min_{x \in \mathbb{R}^n}$$
$$\text{s.t. } f(x) \leqslant 0$$

$$L(x,\ \lambda) = c^\top x + \lambda f(x)$$
$$g(\lambda) = \inf_{x \in \mathbb{R}^n} L(x,\ \lambda) = \inf_{x \in \mathbb{R}^n} \left(c^\top x + \lambda f(x)\right) = -\sup_{x \in \mathbb{R}^n} \left(-\langle c,\ x\rangle - \lambda f(x)\right)$$
$$\frac{g(\lambda)}{\lambda} = -\sup_{x \in \mathbb{R}^n} \left(-\left\langle \frac{1}{\lambda} c,\ x\right\rangle - f(x)\right)$$
$$f^*\left(-\frac{1}{\lambda} c\right) = \sup_{x \in \operatorname{dom} f} \left(\left\langle x,\ -\frac{1}{\lambda} c\right\rangle - f(x)\right)$$
$$g(\lambda) = \lambda f^*\left(-\frac{1}{\lambda} c\right)$$

Тогда двойственная задача имеет вид

$$\lambda f^*\left(-\lambda^{-1} c\right) \to \max_{\lambda \in \mathbb{R}}$$
$$\text{s.t. } \lambda \geqslant 0$$

Можем эквивалентно переформулировать эту задачу в виде

$$-\lambda f^*\left(-\lambda^{-1} c\right) \to \min_{\lambda \in \mathbb{R}}$$
$$\text{s.t. } \lambda \geqslant 0$$

Функция $g(\lambda)$ является вогнутой, потому что является инфимумом линейных по λ функций. Значит полученная задача минимизации является выпуклой, так как $-g(\lambda)$ в таком случае выпукла и ограничения выпуклы.
Ответ:

$$\lambda f^*\left(-\lambda^{-1} c\right) \to \max_{\lambda \in \mathbb{R}}$$
$$\text{s.t. } \lambda \geqslant 0$$

11. To be done...

12.

$$-\sum_{i=1}^{m} \log\left(b_i - a_i^\top x\right) \to \min_{x \in \mathbb{R}^n}$$

с областью определения $\{x:\ a_i^\top x < b_i,\ i = 1, \ldots, n\}$.
Введем новую переменную $y_i = b_i - a_i^\top x$ и сформулируем новую задачу

$$-\sum_{i=1}^{m} \log y_i \to \min_{y \in \mathbb{R}^n_{++},\, x \in \mathbb{R}^n}$$
$$\text{s.t. } y_i = b_i - a_i^\top x$$

Для этой задачи лагранжиан выглядит так (здесь y – вектор из y_i, A – матрица из a_i)

$$L(x,\, y,\, \nu) = -\sum_{i=1}^{m} \log y_i + \nu^\top (y - b + Ax)$$

$$g(\nu) = \inf_{y \in \mathbb{R}^n_{++},\, x \in \mathbb{R}^n} L(x,\, y,\, \nu)$$

Он является выпуклой по $(x,\, y)$ функцией, поэтому $\inf\limits_{y \in \mathbb{R}^n_{++},\, x \in \mathbb{R}^n} L(x,\, y,\, \nu)$ достигается на $x^*,\, y^*$, таких что $\nabla_{(x,\, y)} L(x^*,\, y^*,\, \nu) = 0$.

$$\nabla_x L(x^*,\, y^*,\, \nu) = A^\top \nu = 0;\ \nabla_y L(x^*,\, y^*,\, \nu) = -\left(\frac{1}{y_1^*},\, \ldots,\, \frac{1}{y_n^*}\right)^\top + \nu = 0$$

$$g(\nu) = \sum_{i=1}^{n} \log \nu_i + m - \nu^\top b + \nu^\top A x = \sum_{i=1}^{n} \log \nu_i + m - \nu^\top b$$

Ответ:

$$\sum_{i=1}^{n} \log \nu_i + m - \nu^\top b \to \max_{\nu \in \mathbb{R}^n}$$
$$\text{s.t. } A^\top \nu = 0$$