凸优化 第四次作业

1. 定义：给定一个函数 $f : \mathbb{R}^n \to \mathbb{R}$，其共轭函数（也称作对偶函数）定义为：

$$f^*(y) = \sup_{x \in \mathbb{R}^n} (\langle y, x \rangle - f(x))$$

其中 $\langle y, x \rangle$ 是 y 和 x 的内积。计算函数 $f(x) = \frac{1}{2}\|x\|^2$ 的共轭函数 $f^*(y)$。

证明　计算 $f(x) = \frac{1}{2}\|x\|^2$ **的共轭函数** $f^*(y)$

对于函数 $f(x) = \frac{1}{2}\|x\|^2$，共轭函数的定义是：

$$f^*(y) = \sup_{x \in \mathbb{R}^n} \left(\langle y, x \rangle - \frac{1}{2}\|x\|^2 \right)$$

展开目标函数：

$$\langle y, x \rangle - \frac{1}{2}\|x\|^2 = \langle y, x \rangle - \frac{1}{2}\sum_{i=1}^{n} x_i^2$$

对 x 求导并令其为零：

$$\frac{\partial}{\partial x} \left(\langle y, x \rangle - \frac{1}{2}\|x\|^2 \right) = y - x = 0$$

因此，最优解为 $x = y$。

将 $x = y$ 代入原目标函数：

$$f^*(y) = \langle y, y \rangle - \frac{1}{2}\|y\|^2 = \frac{1}{2}\|y\|^2$$

因此，$f^*(y) = \frac{1}{2}\|y\|^2$。

2. 将下列非标准的几何规划问题转化为标准的几何规划问题:

$$\begin{aligned} \min f_0(x,y) &= x/y \\ \text{subject to}\, 1 &\le x \le 4, \\ x^2 &\le y^2, \\ x/y^3 &= 3. \end{aligned}$$

证明

$$\begin{aligned} \min f_0(x,y) &= x^{-1}y \\ \text{subject to}\, x^{-1} &\le 1, \\ \frac{1}{4}x &\le 1, \\ x^2y^{-2} &\le 1, \\ \frac{1}{3}xy^{-3} &= 1. \end{aligned}$$

3. 用线性规划来描述下列问题:

(a) 给定 $A \in R^{m\times n}$, $b \in R^m$,

$$\min \sum_{i=1}^{m} \max\{0, a_i^T x + b_i\},$$

其中变量 $x \in R^n$。

(b) 给定 $p+1$ 个矩阵 $A_0, A_1, \ldots, A_p \in R^{m\times n}$, 寻找向量 $x \in R^p$ 来最小化

$$\max_{||y||_1=1} ||(A_0 + x_1A_1 + \cdots + x_pA_p)y||_1.$$

证明 (a) 线性规划形式化问题

给定矩阵 $A \in \mathbb{R}^{m\times n}$ 和向量 $b \in \mathbb{R}^m$, 目标是最小化以下函数:

$$\min \sum_{i=1}^{m} \max\{0, a_i^T x + b_i\}$$

其中, $x \in \mathbb{R}^n$ 是优化变量。

线性规划转化: 为了将这个问题转化为线性规划形式, 我们首先定义新的辅助变量 $t_i \geq 0$ 来表示 $\max\{0, a_i^T x + b_i\}$, 即:

$$t_i = \max\{0, a_i^T x + b_i\}$$

这个定义可以分解为两个约束:

$$t_i \geq a_i^T x + b_i \quad (\text{对于 } t_i \geq 0 \text{ 的情况})$$
$$t_i \geq 0 \quad (\text{确保 } t_i \geq 0)$$

目标函数变为:

$$\min \sum_{i=1}^{m} t_i$$

因此, 这个问题可以转化为以下线性规划形式:

$$\begin{aligned} \min \quad & \sum_{i=1}^{m} t_i \\ \text{subject to} \quad & t_i \geq a_i^T x + b_i, \quad i = 1, \ldots, m \\ & t_i \geq 0, \quad i = 1, \ldots, m \end{aligned}$$

这就是原问题的线性规划形式。

(b) 线性规划形式化问题

给定 $p+1$ 个矩阵 $A_0, A_1, \ldots, A_p \in \mathbb{R}^{m\times n}$, 我们需要寻找向量 $x \in \mathbb{R}^p$ 来最小化以下目标函数:

$$\min \max_{\|y\|_1=1} \|(A_0 + x_1A_1 + \cdots + x_pA_p)y\|_1$$

其中，$y \in \mathbb{R}^m$ 满足 $\|y\|_1 = 1$（即 y 的 1-范数为 1）。

线性规划转化: 首先，定义目标函数中的项:

$$T(x) = \max_{\|y\|_1=1} \|(A_0 + x_1A_1 + \cdots + x_pA_p)y\|_1$$

为了将这个问题转化为线性规划形式，我们引入一个新的辅助变量 $t \geq 0$ 来表示 $\|(A_0 + x_1A_1 + \cdots + x_pA_p)y\|_1$。即:

$$t = \|(A_0 + x_1A_1 + \cdots + x_pA_p)y\|_1$$

在 y 的 1-范数约束下，可以通过如下的线性约束表示这个目标函数:

$$t \geq |(A_0 + x_1A_1 + \cdots + x_pA_p)y|_i \quad \forall i = 1, \ldots, m$$
$$\sum_{i=1}^{m} |y_i| = 1$$

因此，问题可以转化为以下的线性规划问题:

$$\begin{aligned} \min \quad & t \\ \text{subject to} \quad & t \geq (A_0 + x_1A_1 + \cdots + x_pA_p)y_i, \quad i = 1, \ldots, m \\ & \sum_{i=1}^{m} |y_i| = 1 \\ & y \in \mathbb{R}^m, x \in \mathbb{R}^p \end{aligned}$$

这就是问题 (b) 的线性规划形式化描述。

4. 请使用线性规划来解决以下问题：给定两个多面体

$$\mathcal{P}_1 = \{x|Ax \le b\}, \mathcal{P}_2 = \{x|Cx \le d\},$$

试证明 $\mathcal{P}_1 \subseteq \mathcal{P}_2$ 或者找到一个在 $\mathcal{P}_1$ 但不在 $\mathcal{P}_2$ 中的点。矩阵 $A \in R^{m\times n}$ 和 $C \in R^{p\times n}$，向量 $c \in R^m$ 和 $d \in R^p$ 是给定的。

解 为了回答这个问题，我们可以将其转化为一个线性规划的可行性问题。具体地，我们需要判断是否对于所有 $x \in \mathcal{P}_1$，都有 $x \in \mathcal{P}_2$。也就是判断是否每个满足 $Ax \le b$ 的点都满足 $Cx \le d$。

1. 判断 $\mathcal{P}_1 \subseteq \mathcal{P}_2$ 我们需要验证是否对于所有 $x \in \mathcal{P}_1$，都有 $Cx \le d$。这等价于检验以下线性约束系统是否存在解：

$$Ax \le b,$$
$$Cx \le d.$$

这是一个标准的线性规划问题。如果该问题有解，则说明 $\mathcal{P}_1 \subseteq \mathcal{P}_2$。

2. 找到一个在 $\mathcal{P}_1$ 但不在 $\mathcal{P}_2$ 中的点如果我们发现该线性规划无解，则说明 $\mathcal{P}_1 \not\subseteq \mathcal{P}_2$，并且存在至少一个点 $x \in \mathcal{P}_1$ 使得 $Cx > d$。为了找到这个点，我们可以通过求解以下优化问题来寻找一个在 $\mathcal{P}_1$ 但不在 $\mathcal{P}_2$ 中的点：

$$\text{maximize} \quad 1^T x \quad \text{subject to} \quad Ax \le b, Cx > d.$$

通过求解这个线性规划问题，我们可以找到一个点 x，使得 $x \in \mathcal{P}_1$ 且 $Cx > d$，从而证明 $\mathcal{P}_1 \not\subseteq \mathcal{P}_2$。

总结: - 如果线性规划问题有解，则 $\mathcal{P}_1 \subseteq \mathcal{P}_2$。- 如果线性规划问题无解，则 $\mathcal{P}_1 \not\subseteq \mathcal{P}_2$，并且可以通过线性规划找到一个在 $\mathcal{P}_1$ 但不在 $\mathcal{P}_2$ 中的点。

《Convex Optimization》 4.1, 4.3, 4.9, 4.12, 4.13, 4.19. 4.23

4.1 Consider the optimization problem

$$\begin{array}{ll} \text{minimize} & f_0(x_1, x_2) \\ \text{subject to} & 2x_1 + x_2 \geq 1 \\ & x_1 + 3x_2 \geq 1 \\ & x_1 \geq 0, \quad x_2 \geq 0 \end{array}$$

Make a sketch of the feasible set. For each of the following objective functions, give the optimal set and the optimal value. (a) $f_0(x_1, x_2) = x_1 + x_2$. (b) $f_0(x_1, x_2) = -x_1 - x_2$. (c) $f_0(x_1, x_2) = x_1$. (d) $f_0(x_1, x_2) = \max\{x_1, x_2\}$. (e) $f_0(x_1, x_2) = x_1^2 + 9x_2^2$.

Solution. The feasible set is the convex hull of $(0, \infty), (0, 1), (2/5, 1/5), (1, 0), (\infty, 0)$. (a) $x^\star = (2/5, 1/5)$. (b) Unbounded below. (c) $X_{\text{opt}} = \{(0, x_2) \mid x_2 \geq 1\}$. (d) $x^\star = (1/3, 1/3)$. (e) $x^\star = (1/2, 1/6)$. This is optimal because it satisfies $2x_1 + x_2 = 7/6 > 1, x_1 + 3x_2 = 1$, and

$$\nabla f_0(x^\star) = (1, 3)$$

is perpendicular to the line $x_1 + 3x_2 = 1$.

4.3 Prove that $x^\star = (1, 1/2, -1)$ is optimal for the optimization problem

$$\begin{array}{ll} \text{minimize} & (1/2)x^T P x + q^T x + r \\ \text{subject to} & -1 \leq x_i \leq 1, \quad i = 1, 2, 3 \end{array}$$

where

$$P = \begin{bmatrix} 13 & 12 & -2 \\ 12 & 17 & 6 \\ -2 & 6 & 12 \end{bmatrix}, \quad q = \begin{bmatrix} -22.0 \\ -14.5 \\ 13.0 \end{bmatrix}, \quad r = 1.$$

Solution. We verify that $x^\star$ satisfies the optimality condition (4.21). The gradient of the objective function at $x^\star$ is

$$\nabla f_0\left(x^\star\right) = (-1, 0, 2)$$

Therefore the optimality condition is that

$$\nabla f_0\left(x^\star\right)^T (y - x) = -1\left(y_1 - 1\right) + 2\left(y_2 + 1\right) \geq 0$$

for all y satisfying $-1 \leq y_i \leq 1$, which is clearly true.

4.9 Square LP. Consider the LP

$$
\begin{array}{ll}
\text{minimize} & c^T x \\
\text{subject to} & Ax \preceq b
\end{array}
$$

with A square and nonsingular. Show that the optimal value is given by

$$
p^\star = \begin{cases} c^T A^{-1} b & A^{-T} c \preceq 0 \\ -\infty & \text{otherwise.} \end{cases}
$$

Solution. Make a change of variables $y = Ax$. The problem is equivalent to

$$
\begin{array}{ll}
\text{minimize} & c^T A^{-1} y \\
\text{subject to} & y \preceq b.
\end{array}
$$

If $A^{-T} c \preceq 0$, the optimal solution is $y = b$, with $p^\star = c^T A^{-1} b$. Otherwise, the LP is unbounded below.

4.12 Network flow problem. Consider a network of n nodes, with directed links connecting each pair of nodes. The variables in the problem are the flows on each link: x_{ij} will denote the flow from node i to node j. The cost of the flow along the link from node i to node j is given by $c_{ij}x_{ij}$, where c_{ij} are given constants. The total cost across the network is

$$C = \sum_{i,j=1}^{n} c_{ij}x_{ij}$$

Each link flow x_{ij} is also subject to a given lower bound l_{ij} (usually assumed to be nonnegative) and an upper bound u_{ij}. The external supply at node i is given by b_i, where $b_i > 0$ means an external flow enters the network at node i, and $b_i < 0$ means that at node i, an amount $|b_i|$ flows out of the network. We assume that $1^T b = 0$, i.e., the total external supply equals total external demand. At each node we have conservation of flow: the total flow into node i along links and the external supply, minus the total flow out along the links, equals zero. The problem is to minimize the total cast of flow through the network, subject to the constraints described above. Formulate this problem as an LP.

Solution. This can be formulated as the LP

$$\begin{array}{rl} \text{minimize} & C = \sum_{i,j=1}^{n} c_{ij}x_{ij} \\ \text{subject to} & b_i + \sum_{j=1}^{n} x_{ij} - \sum_{j=1}^{n} x_{ji} = 0, \quad i = 1, \dots, n \\ & l_{ij} \le x_{ij} \le u_{ij} \end{array}$$

4.13 Robust LP with interval coefficients. Consider the problem, with variable $x \in \mathrm{R}^n$,

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax \preceq b \text{ for all } A \in \mathcal{A} \end{array}$$

where $\mathcal{A} \subseteq \mathrm{R}^{m \times n}$ is the set

$$\mathcal{A} = \left\{ A \in \mathrm{R}^{m \times n} \mid \bar{A}_{ij} - V_{ij} \leq A_{ij} \leq \bar{A}_{ij} + V_{ij}, i = 1, \ldots, m, j = 1, \ldots, n \right\}$$

(The matrices $\bar{A}$ and V are given.) This problem can be interpreted as an LP where each coefficient of A is only known to lie in an interval, and we require that x must satisfy the constraints for all possible values of the coefficients. Express this problem as an LP. The LP you construct should be efficient, i.e., it should not have dimensions that grow exponentially with n or m.

Solution. The problem is equivalent to

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & \bar{A}x + V|x| \preceq b \end{array}$$

where $|x| = (|x_1|, |x_2|, \ldots, |x_n|)$. This in turn is equivalent to the LP

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & \bar{A}x + Vy \preceq b \\ & -y \preceq x \preceq y \end{array}$$

with variables $x \in \mathrm{R}^n, y \in \mathrm{R}^n$.

4.19 Consider the problem

$$\begin{array}{ll} \text{minimize} & \|Ax-b\|_1/\left(c^Tx+d\right) \\ \text{subject to} & \|x\|_\infty \leq 1 \end{array}$$

where $A \in \mathrm{R}^{m\times n}, b \in \mathrm{R}^m, c \in \mathrm{R}^n$, and $d \in \mathrm{R}$. We assume that $d > \|c\|_1$, which implies that $c^Tx + d > 0$ for all feasible x. (a) Show that this is a quasiconvex optimization problem. (b) Show that it is equivalent to the convex optimization problem

$$\begin{array}{ll} \text{minimize} & \|Ay-bt\|_1 \\ \text{subject to} & \|y\|_\infty \leq t \\ & c^Ty+dt=1 \end{array}$$

with variables $y \in \mathrm{R}^n, t \in \mathrm{R}$.

Solution. (a) $f_0(x) \leq \alpha$ if and only if

$$\|Ax-b\|_1 - \alpha\left(c^Tx+d\right) \leq 0,$$

which is a convex constraint. (b) Suppose $\|x\|_\infty \leq 1$. We have $c^Tx+d > 0$, because $d > \|c\|_1$. Define

$$y = x/\left(c^Tx+d\right), \quad t = 1/\left(c^Tx+d\right).$$

Then y and t are feasible in the convex problem with objective value

$$\|Ay-bt\|_1 = \|Ax-b\|_1/\left(c^Tx+d\right).$$

Conversely, suppose y, t are feasible for the convex problem. We must have $t > 0$, since $t = 0$ would imply $y = 0$, which contradicts $c^Ty + dt = 1$. Define

$$x = y/t.$$

Then $\|x\|_\infty \leq 1$, and $c^Tx + d = 1/t$, and hence

$$\|Ax-b\|_1/\left(c^Tx+d\right) = \|Ay-bt\|_1.$$

4.23 ℓ_4-norm approximation via $QCQP$. Formulate the ℓ_4-norm approximation problem

$$\text{minimize} \quad \|Ax - b\|_4 = \left(\sum_{i=1}^{m} \left(a_i^T x - b_i\right)^4 \right)^{1/4}$$

as a QCQP. The matrix $A \in \mathrm{R}^{m \times n}$ (with rows a_i^T) and the vector $b \in \mathrm{R}^m$ are given.

Solution.

$$\begin{array}{ll} \text{minimize} & \sum_{i=1}^{m} z_i^2 \\ \text{subject to} & a_i^T x - b_i = y_i, i = 1, \ldots, m \\ & y_i^2 \leq z_i, \quad i = 1, \ldots, m \end{array}$$