Optimization Methods **Fall 2023**

Homework 3

Instructor: Lijun Zhang *Name:* 季千F, *StudentId:* 221300066

Notice

- The submission email is: **optfall2023@163.com**.
- Please use the provided LaTeX file as a template.
- If you are not familiar with LaTeX, you can also use Word to generate a **PDF** file.

Problem 1: Linear optimization problems

Formulate the following problems as LPs. Explain in detail the relation between the optimal solution of each problem and the solution of its equivalent LP.

(1) Minimize $\|Ax-b\|_\infty$ (ℓ_∞-norm approximation).

(2) Minimize $\|Ax-b\|_1$ (ℓ_1-norm approximation).

(3) Minimize $\|Ax-b\|_1$ subject to $\|x\|_\infty \le 1$.

Solution. (1) 最小化 $\|Ax-b\|_\infty$ 等价于最小化 t，t 是一个新的变量，满足 $-t \le a_i^T x - b_i \le t$，对于所有的 $i=1,\dots,m$。这个线性规划问题的标准型为：

$$\begin{aligned} \min \quad & t \\ s.t. \quad & -t \le a_i^T x - b_i \le t, \quad i=1,\dots,m \end{aligned}$$

此问题的最优解与其等价的线性规划问题的最优解之间的关系是：如果 (t^*, x^*) 是线性规划问题的最优解，那么 x^* 就是原问题的最优解，且 $t^* = \|Ax^*-b\|_\infty$。

(2) 最小化 $\|Ax-b\|_1$ 等价于最小化 $t_1+\cdots+t_m$，$t_1,\dots,t_m$ 是新变量，满足 $-t_i \le a_i^T x - b_i \le t_i$，对于所有的 $i=1,\dots,m$。这个线性规划问题的标准型为：

$$\begin{aligned} \min \quad & t_1+\cdots+t_m \\ s.t. \quad & -t_i \le a_i^T x - b_i \le t_i, \quad i=1,\dots,m \end{aligned}$$

该问题的最优解与其等价的线性规划问题的最优解之间的关系是：如果 $(t_1^*,\dots,t_m^*, x^*)$ 是线性规划问题的最优解，那么 x^* 就是原问题的最优解，且 $t_1^*+\cdots+t_m^* = \|Ax^*-b\|_1$。

(3) 最小化 $\|Ax-b\|_1$，同时满足 $\|x\|_\infty \leq 1$，等价于最小化 $t_1+\cdots+t_m$，$t_1,\ldots,t_m$ 是新变量，满足 $-t_i \leq a_i^T x - b_i \leq t_i$，对于所有的 $i=1,\ldots,m$，以及 $-1 \leq x_j \leq 1$，对于所有的 $j=1,\ldots,n$。这个线性规划问题的标准型为：

$$\begin{aligned} \min \quad & t_1+\cdots+t_m \\ s.t. \quad & -t_i \leq a_i^T x - b_i \leq t_i, \quad i=1,\ldots,m \\ & -1 \leq x_j \leq 1, \quad j=1,\ldots,n \end{aligned}$$

该问题的最优解与其等价的线性规划问题的最优解之间的关系是：如果 $(t_1^*,\ldots,t_m^*,x^*)$ 是线性规划问题的最优解，那么 x^* 就是原问题的最优解，且 $t_1^*+\cdots+t_m^* = \|Ax^*-b\|_1$。

□

Problem 2: Negative-entropy Regularization

Please show how to compute

$$\underset{x\in\Delta^n}{\operatorname{argmin}} \quad b^\top x + c\cdot\sum_{i=1}^n x_i \ln x_i$$

where $\Delta^n = \{x \mid \sum_{i=1}^n x_i = 1, x_i \geq 0, i=1,\cdots,n\}$, $b\in\mathbb{R}^n$ and $c>0$.

Solution. 使用拉格朗日乘子法来计算

首先，构造拉格朗日函数：

$$L(x,\lambda) = b^\top x + c\cdot\sum_{i=1}^n x_i \ln x_i + \lambda\left(\sum_{i=1}^n x_i - 1\right)$$

其中，λ 是拉格朗日乘子，用于引入等式约束 $\sum_{i=1}^n x_i = 1$。

然后令拉格朗日函数对 x_i 和 λ 的偏导数为零，得到：

$$\frac{\partial L}{\partial x_i} = b_i + c + c\ln x_i + \lambda = 0, \quad i=1,\ldots,n\frac{\partial L}{\partial \lambda} = \sum_{i=1}^n x_i - 1 = 0$$

解得：

$$x_i = e^{-\frac{b_i+c+\lambda}{c}}, \quad i=1,\ldots,n\sum_{i=1}^n e^{-\frac{b_i+c+\lambda}{c}} = 1$$

由于 λ 是未知的，使用用牛顿法来迭代求解：

$$\lambda_{k+1} = \lambda_k - \frac{f(\lambda_k)}{f'(\lambda_k)}$$

其中，

$$f(\lambda) = \sum_{i=1}^n e^{-\frac{b_i+c+\lambda}{c}} - 1 f'(\lambda) = -\frac{1}{c}\sum_{i=1}^n e^{-\frac{b_i+c+\lambda}{c}}$$

当 λ 收敛时，就得到了 x 的最优解。

□

Problem 3: KKT conditions

Consider the problem

$$
\begin{aligned}
\min_{x\in\mathbb{R}^2} \quad & x_1^2 + x_2^2 \\
\text{s.t.} \quad & (x_1-1)^2 + (x_2-1)^2 \le 2 \\
& (x_1-1)^2 + (x_2+1)^2 \le 2
\end{aligned}
$$

where $x = \begin{bmatrix} x_1 & x_2 \end{bmatrix}^\top \in \mathbb{R}^2$.

(1) Write the Lagrangian for this problem.

(2) Does strong duality hold in this problem?

(3) Write the KKT conditions for this optimization problem.

Solution. (1) 这个问题的拉格朗日函数是

$$L(x,\lambda,\mu) = x_1^2 + x_2^2 + \lambda\left((x_1-1)^2 + (x_2-1)^2 - 2\right) + \mu\left((x_1-1)^2 + (x_2+1)^2 - 2\right)$$

其中，λ 和 μ 是拉格朗日乘子，用来引入不等式约束。

(2) 强对偶性在这个问题中不一定成立。因为这个问题的目标函数和约束函数都是非凸的，所以不能保证满足斯莱特条件或其他正则性条件。如果存在一个可行点 x，使得 $(x_1-1)^2 + (x_2-1)^2 < 2$ 且 $(x_1-1)^2 + (x_2+1)^2 < 2$，那么斯莱特条件成立，从而强对偶性成立。否则，强对偶性可能不成立。

(3) KKT 条件是这个问题的最优解必须满足的必要条件，即是：

$$
\begin{aligned}
\frac{\partial L}{\partial x_1} &= 2x_1 + 2\lambda(x_1-1) + 2\mu(x_1-1) = 0 \\
\frac{\partial L}{\partial x_2} &= 2x_2 + 2\lambda(x_2-1) + 2\mu(x_2+1) = 0 \\
\lambda &\ge 0 \\
\mu &\ge 0 \\
\lambda\left((x_1-1)^2 + (x_2-1)^2 - 2\right) &= 0 \\
\mu\left((x_1-1)^2 + (x_2+1)^2 - 2\right) &= 0 \\
(x_1-1)^2 + (x_2-1)^2 &\le 2 \\
(x_1-1)^2 + (x_2+1)^2 &\le 2
\end{aligned}
$$

□

Problem 4: Equality Constrained Least-squares

Consider the equality constrained least-squares problem

$$
\begin{aligned}
\min \quad & \frac{1}{2}\|Ax-b\|_2^2 \\
\text{s.t.} \quad & Gx = h
\end{aligned}
$$

where $A \in \mathbf{R}^{m\times n}$ with $\textbf{rank}\ A = n$, and $G \in \mathbf{R}^{p\times n}$ with $\textbf{rank}\ G = p$.

(1) Derive the Lagrange dual problem with Lagrange multiplier vector v.

(2) Derive expressions for the primal solution $x^\star$ and the dual solution $v^\star$.

Solution. (1) 这个问题的拉格朗日函数是

$$L(x,v) = \frac{1}{2}\|Ax-b\|_2^2 + v^\top(Gx-h)$$

其中，v 是拉格朗日乘子向量，用来引入等式约束 $Gx = h$。

拉格朗日对偶问题是

$$\begin{aligned} \max_{v} \quad & g(v) \\ \text{s.t.} \quad & v \in \mathbb{R}^p \end{aligned}$$

其中，$g(v)$ 是拉格朗日对偶函数，定义为

$$g(v) = \inf_x L(x,v) = \inf_x \left(\frac{1}{2}\|Ax-b\|_2^2 + v^\top(Gx-h)\right)$$

(2) 要求出原始问题的最优解 $x^\star$ 和对偶问题的最优解 $v^\star$，可以令拉格朗日函数对 x 的偏导数为零，得到

$$\frac{\partial L}{\partial x} = A^\top(Ax-b) + G^\top v = 0$$

由于 A 的秩为 n，所以 $A^\top A$ 是可逆的，因此解出

$$x = (A^\top A)^{-1}(A^\top b - G^\top v)$$

代入对偶函数得到

$$g(v) = -\frac{1}{2}(A^\top b - G^\top v)^\top (A^\top A)^{-1}(A^\top b - G^\top v) - v^\top h + \frac{1}{2}b^\top b$$

为求出 $v^\star$，令对偶函数对 v 的偏导数为零得到

$$\frac{\partial g}{\partial v} = -G(A^\top A)^{-1}(A^\top b - G^\top v) - h = 0$$

由于 G 的秩为 p，所以 $G(A^\top A)^{-1}G^\top$ 是可逆的，因此可以解出

$$v = (G(A^\top A)^{-1}G^\top)^{-1}(G(A^\top A)^{-1}A^\top b + h)$$

代入 x 的表达式，得到

$$x = (A^\top A)^{-1}A^\top b - (A^\top A)^{-1}G^\top(G(A^\top A)^{-1}G^\top)^{-1}(G(A^\top A)^{-1}A^\top b + h)$$

则得到了原始问题的最优解 $x^\star$ 和对偶问题的最优解 $v^\star$

$\square$

Problem 5: Determinant optimization

Derive the dual problem of the following problem

$$\begin{aligned} \min \quad & \log\det X^{-1} \\ \text{s.t.} \quad & A_i^T X A_i \preceq B_i, i = 1,\ldots,m \end{aligned}$$

where $X \in \mathbb{S}_{++}^n, A_i \in \mathbb{R}^{n\times k_i}, B_i \in \mathbb{S}_{++}^{k_i}, k_i \in \mathbb{N}_+, i = 1,\ldots,m.$

Solution. 使用拉格朗日对偶理论来求解。首先构造拉格朗日函数：

$$L(X,S) = \log\det X^{-1} + \sum_{i=1}^{m} \mathrm{tr}(S_i(A_i^T X A_i - B_i))$$

其中，$S = (S_1, \ldots, S_m)$ 是拉格朗日乘子矩阵，满足 $S_i \succeq 0$，对于所有的 $i = 1, \ldots, m$。tr 表示矩阵的迹。

拉格朗日对偶函数是

$$g(S) = \inf_{X \succ 0} L(X,S) = -\sup_{X \succ 0} \left(-\log\det X^{-1} - \sum_{i=1}^{m} \mathrm{tr}(S_i(A_i^T X A_i - B_i)) \right)$$

拉格朗日对偶问题是

$$\begin{aligned} \max_{S} \quad & g(S) \\ \text{s.t.} \quad & S_i \succeq 0, \quad i = 1, \ldots, m \end{aligned}$$

为了求出对偶函数的值，令拉格朗日函数对 X 的偏导数为零得到

$$\frac{\partial L}{\partial X} = -X - \sum_{i=1}^{m} A_i S_i A_i^T = 0$$

解得

$$X = \left(\sum_{i=1}^{m} A_i S_i A_i^T \right)^{-1}$$

代入对偶函数得到

$$g(S) = -\log\det \left(\sum_{i=1}^{m} A_i S_i A_i^T \right) + \sum_{i=1}^{m} \mathrm{tr}(S_i B_i)$$

对偶问题的标准形式可得

□