最优化算法作业 1

陈文轩

更新：March 24, 2025

1. 考虑以下问题：

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

其中 $x = (x_1, x_2)^\top \in \mathbb{R}^2$。

(a). 这是一个凸优化问题吗？

(b). 写出此问题的拉格朗日函数。使用 Slater 条件验证在这个问题中是否存在强对偶性。

(c). 写出这个优化问题的 KKT 条件。求出 KKT 点和最优点。

解

(a). 由于目标函数与约束函数均为二次函数，且 Hessian 矩阵为 $2I_2$，故是凸优化问题。

(b). 引入非负 Lagrange 乘子 $\lambda_1, \lambda_2 \ge 0$，可以得到问题的 Lagrange 函数为：

$$L(x, \lambda_1, \lambda_2) = x_1^2 + x_2^2 + \lambda_1((x_1-1)^2 + (x_2-1)^2 - 2) + \lambda_2(x_1-1)^2 + (x_2+1)^2 - 2$$

取 $x = (1, 0)^\top$，显然这个 x 满足严格不等式，故 Slater 条件成立，即具有强对偶性。

(c). KKT 条件为：

$$\begin{cases}\nabla_x L(x, \lambda_1, \lambda_2) = 2(x_1 + (\lambda_1+\lambda_2)(x_1-1), x_2 + (\lambda_1+\lambda_2)(x_2-1)) = 0 \\ \lambda_1((x_1-1)^2 + (x_2-1)^2 - 2) = 0, \lambda_2((x_1-1)^2 + (x_2+1)^2 - 2) = 0 \\ (x_1-1)^2 + (x_2-1)^2 \le 2, (x_1-1)^2 + (x_2+1)^2 \le 2 \\ \lambda_1 \ge 0, \lambda_2 \ge 0\end{cases}$$

解得 $x_1 = x_2 = \lambda_1 = \lambda_2 = 0$，即 $(0, 0)^\top$ 是 KKT 点，也是最优点。

2. 考虑一个凸的分段线性最小化问题，变量为 $x \in \mathbb{R}^n$

$$\min \max_{i=1,\cdots,m} (a_i^\top x + b_i)$$

其中 $a_i \in \mathbb{R}^n, n_i \in \mathbb{R}$。

(a). 考虑原问题的如下等价问题2.1，推导2.1的对偶问题

$$\begin{aligned} \min\ \max_{i=1,\cdots,m} \quad & y_i \\ s.t. \quad & a_i^\top x + b_i \le y_i, i = 1, \cdots, m \end{aligned} \tag{2.1}$$

变量为 $x \in \mathbb{R}^n, y \in \mathbb{R}^m$。

(b). 假设我们通过平滑函数逼近目标函数, 逼近函数为：

$$f_0(x) = \log \sum_{i=1}^{m} \exp(a_i^\top x + b_i)$$

现在我们考虑无约束问题，即 $\min f_0(x)$。证明该问题的对偶问题如下：

$$\begin{aligned} \max \quad & b^\top z - \sum_{i=1}^{m} z_i \log z_i \\ s.t. \quad & A^\top z = 0, \mathbb{1}^\top z = 1, z \succeq 0. \end{aligned} \tag{2.2}$$

其中 $\mathbb{1}$ 表示全是 1 的向量。

(c). 设问题2.1的最优函数值是 p_1^*，问题2.2的最优函数值是 p_2^*，证明 $0 \le p_2^* - p_1^* \le \log m$。

解

(a). 先转化为线性规划问题：

$$\begin{aligned} \min \quad & t \\ s.t. \quad & a_i^\top x + b_i \le t, i = 1, \cdots, m \end{aligned}$$

Lagrange 函数为 $L(x, t, \lambda) = t + \sum\limits_{i=1}^{m} \lambda_i (a_i^\top x + b_i - t), \lambda_i \ge 0$，目标是 $\max\limits_{\lambda} \min\limits_{x,t} L(x, t, \lambda)$。

$\nabla_t L(x, t, \lambda) = 1 - \sum\limits_{i=1}^{m} \lambda_i = 0 \Rightarrow \mathbb{1}^\top \lambda = 1, \nabla_x L(x, t, \lambda) = \sum\limits_{i=1}^{m} \lambda_i a_i = 0$，

此时 $L(x, y, \lambda) = \sum\limits_{i=1}^{m} \lambda_i b_i$。因此对偶问题是：

$$\begin{aligned} \max \quad & \sum_{i=1}^{m} \lambda_i b_i \\ s.t. \quad & \mathbb{1}^\top \lambda = 1, \sum_{i=1}^{m} \lambda_i a_i = 0 \\ & \lambda_i \ge 0, i = 1, \cdots, m \end{aligned}$$

(b). 记 $A = (a_i, \cdots, a_m), z_j = \dfrac{\exp(a_j^\top x + b_j)}{\sum_{i=1}^{m} \exp(a_i^\top x + b_i)}$，则 $\sum\limits_{j=1}^{m} z_j = 1, z_j \ge 0$。

此时 $f_0(x) = \log \sum\limits_{i=1}^{m} \exp(a_i^\top x + b_i) = \sum\limits_{i=1}^{m} z_i (a_i^\top x + b_i) - \sum\limits_{i=1}^{m} z_i \log z_i := L(x, z)$。

$\nabla_x L(x, z) = \sum\limits_{i=1}^{m} z_i a_i = 0 \Rightarrow A^\top z = 0$，此时 $L(x, z) = b^\top z - \sum\limits_{i=1}^{m} z_i \log z_i$。

因此对偶问题是：

$$\begin{aligned} \max \quad & b^\top z - \sum_{i=1}^{m} z_i \log z_i \\ s.t. \quad & A^\top z = 0, \mathbb{1}^\top z = 1, z \succeq 0. \end{aligned}$$

(c). 记 $M = \max_i (a_i^\top x + b_i)$，则 $\exp(M) \le \sum_{i=1}^m \exp(a_i^\top x + b_i) \le m\exp(M)$。

取对数，即有 $M \le \log \sum_{i=1}^m \exp(a_i^\top x + b_i) \le M + \log m, \forall x$。

对问题2.2的最优解 x_2^*，有 $p_2^* = \log \sum_{i=1}^m \exp(a_i^\top x_2^* + b_i)$，由上有 $\max_i (a_i^\top x_2^* + b_i) \le p_2^*$。

又 $p_1^* = \min_x \max_i (a_i^\top x + b_i) \le \max_i (a_i^\top x_2^* + b_i) \le p_2^*$，故 $p_1 \le p_2$。

对问题2.1的最优解 x_1^*，有 $\max_i (a_i^\top x_1^* + b_i) = p_1^*$。由上有 $\log \sum_{i=1}^m \exp(a_i^\top x_1^* + b_i) \le p_1^* + \log m$。又 $p_2^* = \min_x \log \sum_{i=1}^m \exp(a_i^\top x + b_i) \le p_1^* + \log m$，故 $p_2^* \le p_1^* + \log m$。

综上所述有 $0 \le p_2^* - p_1^* \le \log m$。