人工智能数学原理 作业 4

潘天麟 2023K8009908023

October 28, 2025

T1

将下面的问题转化为线性规划: 给定 $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$,

(a). $\min_{x\in\mathbb{R}^n} \|Ax-b\|_1$, s.t. $\|x\|_\infty \le 1$;

(b). $\min_{x\in\mathbb{R}^n} \|x\|_1$, s.t. $\|Ax-b\|_\infty \le 1$.

(a).

$$
\begin{aligned}
\min_{x\in\mathbb{R}^n} \quad & \sum_{i=1}^{m} t_i \\
\text{s.t.} \quad & -t_i \le A_i x - b_i \le t_i, \quad i = 1,2,...,m; \\
& -1 \le x_j \le 1, \quad j = 1,2,...,n.
\end{aligned} \tag{1}
$$

其中 A_i 表示矩阵 A 的第 i 行, 下同.

(b).

$$
\begin{aligned}
\min_{x\in\mathbb{R}^n} \quad & \sum_{j=1}^{n} s_j \\
\text{s.t.} \quad & -s_j \le x_j \le s_j, \quad j = 1,2,...,n; \\
& -1 \le A_i x - b_i \le 1, \quad i = 1,2,...,m.
\end{aligned} \tag{2}
$$

T2

求解下面的线性规划问题: 给定向量 $c \in \mathbb{R}^n$,

(a). $\min_{x\in\mathbb{R}^n} c^T x$, s.t. $\mathbf{0} \le x \le \mathbf{1}$;

(b). $\min_{x\in\mathbb{R}^n} c^T x$, s.t. $-\mathbf{1} \le x \le \mathbf{1}$.

(a). 首先显然有 $c^T x = \sum_{i=1}^n c_i x_i$. 因为 $0 \le x_i \le 1$, 所以当 $c_i \ge 0$ 时, x_i 取最小值 0; 当 $c_i < 0$ 时, x_i 取最大值 1. 因此当

$$
x = \begin{cases} 0, c_i \ge 0 \\ 1, c_i < 0 \end{cases} \tag{3}
$$

时, $c^T x$ 取得最小值 $\sum_{i=1}^n \min(0, c_i)$.

(b). 首先显然有 $c^T x = \sum_{i=1}^n c_i x_i$. 因为 $-1 \le x_i \le 1$, 所以当 $c_i \ge 0$ 时, x_i 取最小值 -1; 当 $c_i < 0$ 时, x_i 取最大值 1. 因此当

$$x = \begin{cases} -1, c_i \geq 0 \\ 1, c_i < 0 \end{cases} \tag{4}$$

时, $c^T x$ 取得最小值 $-\sum_{i=1}^n |c_i|$.

T3

给定矩阵 $A_i \in \mathcal{S}^m$, $i = 0, 1, ..., m$, 定义线性映射 $A(x) = A_0 + \sum_{i=1}^n x_i A_i$. 令 $\lambda_1(x) \geq \lambda_2(x) \geq ... \geq \lambda_m(x)$ 为 $A(x)$ 的特征值. 将下面的优化问题转化为半定规划问题:

(a). $\min_{x \in \mathbb{R}^n} \lambda_1(x) - \lambda_m(x)$;

(b). $\min_{x \in \mathbb{R}^n} \sum_{i=1}^m |\lambda_i(x)|$.

(a). 我们把 A 正交对角化, 则有 $A = Q\Lambda Q^T$, 其中 $\Lambda = \mathrm{diag}(\lambda_1, \lambda_2, ..., \lambda_m)$. 考虑

$$A - \alpha I = Q(\Lambda - \alpha I)Q^T, \tag{5}$$

那么在正交偏序下, $A - \alpha I \succcurlyeq 0$ 要求 $\Lambda - \alpha I \succcurlyeq 0$, 也就是 Λ 中的每一个特征值都大于等于 α, 即 $\lambda_m \geq \alpha$. 所以我们可以用下面的方式求出 A 的最小特征值:

$$\lambda_m = \max\{\alpha \mid A - \alpha I \succcurlyeq 0\}; \tag{6}$$

同理, 我们可以求出 A 的最大特征值:

$$\lambda_1 = \min\{\beta \mid \beta I - A \succcurlyeq 0\}. \tag{7}$$

所以, 我们可以把原问题转化为下面的半定规划问题:

$$\begin{aligned} \min_{x \in \mathbb{R}^n, \alpha \in \mathbb{R}, \beta \in \mathbb{R}} \quad & \beta - \alpha \\ \text{s.t.} \quad & A(x) - \alpha I \succcurlyeq 0; \\ & \beta I - A(x) \succcurlyeq 0. \end{aligned} \tag{8}$$

(b). 同样, 我们考虑 A 的谱分解 $A = Q\Lambda Q^T$. 这里我们要把 Λ 中的正特征值和负特征值分离开来, 我们定义

$$\begin{aligned} \Lambda_+ &= \mathrm{diag}[\max(0, \lambda_1), \max(0, \lambda_2), ..., \max(0, \lambda_m)], \\ \Lambda_- &= \mathrm{diag}[\min(0, \lambda_1), \min(0, \lambda_2), ..., \min(0, \lambda_m)]. \end{aligned} \tag{9}$$

则我们就有

$$A = Q\Lambda_+ Q^T + Q\Lambda_- Q^T = A_+ + A_-, \tag{10}$$

考虑到

$$\sum_{i=1}^{m} |\lambda_i| = \mathrm{tr}(\Lambda_+) - \mathrm{tr}(\Lambda_-) = \mathrm{tr}(A_+) - \mathrm{tr}(A_-), \tag{11}$$

所以我们可以把原问题转化为下面的半定规划问题:

$$\begin{aligned} \min_{x\in\mathbb{R}^n, A_+\in\mathcal{S}^m, A_-\in\mathcal{S}^m} \quad & \mathrm{tr}(A_+) - \mathrm{tr}(A_-) \\ \text{s.t.} \quad & A_0 + \sum_{i=1}^{n} x_i A_i = A_+ + A_-; \\ & A_+ \succcurlyeq 0; A_- \preccurlyeq 0. \end{aligned} \tag{12}$$

T4

考虑如下模型:

$$\min_{X_1\in\mathbb{R}^{q\times p}, X_2\in\mathbb{R}^{q\times q}} \sum_{i=1}^{m} \|X_2 X_1 a_i - b_i\|_2^2, \tag{13}$$

其中 $a_i \in \mathbb{R}^p$, $b_i \in \mathbb{R}^q$, $i = 1, 2, ..., m$.

(a). 试计算目标函数关于 X_1 和 X_2 的导数.

(b). 给出该问题的最优解.

(a). 目标函数为:

$$\begin{aligned} f(X_1, X_2) &= \sum_{i=1}^{m} \|X_2 X_1 a_i - b_i\|_2^2 \\ &= \sum_{i=1}^{m} (X_2 X_1 a_i - b_i)^T (X_2 X_1 a_i - b_i) \end{aligned} \tag{14}$$

设 $Y = X_2 X_1$, 有

$$\frac{\partial f}{\partial Y} = 2\sum_{i=1}^{m} (Y a_i - b_i) a_i^T \tag{15}$$

于是

$$
\begin{aligned}
\frac{\partial f}{\partial X_1} &= 2\sum_{i=1}^{m} X_2^T (X_2 X_1 a_i - b_i) a_i^T, \\
\frac{\partial f}{\partial X_2} &= 2\sum_{i=1}^{m} (X_2 X_1 a_i - b_i)(X_1 a_i)^T.
\end{aligned} \tag{16}
$$

(b). 令 $A = [a_1, a_2, ..., a_m] \in \mathbb{R}^{p\times m}$, $B = [b_1, b_2, ..., b_m] \in \mathbb{R}^{q\times m}$, $Z = X_2 X_1 \in \mathbb{R}^{q\times p}$, 则原问题可转化为

$$
\min_{Z\in\mathbb{R}^{q\times p}} \quad \|ZA - B\|_F^2. \tag{17}
$$

这其实就是最小二乘问题, 它的最优解为

$$
Z^* = BA^\dagger, \tag{18}
$$

其中 $A^\dagger$ 是 A 的 Moore-Penrose 伪逆. 最优值为

$$
f^* = \|B\big(I - AA^\dagger\big)\|_F^2. \tag{19}
$$