DCS440 最优化理论 第二次作业: 凸优化问题与对偶理论

胡瑞康

22336087

推导线性规划问题的对偶问题和KKT条件：

原始问题（Primal Problem）：

$$\begin{aligned}\min_{x} \quad & c^\top x \\ \text{s.t.} \quad & Gx \leq h \\ & Ax = b\end{aligned}$$

一、对偶问题（Dual Problem）的推导

1. **引入拉格朗日乘子（Lagrange Multipliers）**：
   - 对于不等式约束 $Gx \leq h$，引入非负乘子 $\lambda \geq 0$。
   - 对于等式约束 $Ax = b$，引入乘子 ν。
2. **构造拉格朗日函数（Lagrangian）**：

$$L(x, \lambda, \nu) = c^\top x + \lambda^\top (Gx - h) + \nu^\top (Ax - b)$$

3. **对 x 求极小**:

$$g(\lambda, \nu) = \inf_{x} L(x, \lambda, \nu)$$

   为了使 $g(\lambda, \nu)$ 有界，需满足:

$$c + G^\top \lambda + A^\top \nu = 0$$

   此时:

$$g(\lambda, \nu) = -\lambda^\top h - \nu^\top b$$

   否则，$g(\lambda, \nu) = -\infty$。

4. **形成对偶问题**:

$$\begin{aligned}\max_{\lambda, \nu} \quad & -\lambda^\top h - \nu^\top b \\ \text{s.t.} \quad & G^\top \lambda + A^\top \nu = -c \\ & \lambda \geq 0\end{aligned}$$

   或者，调整符号后得到常见形式:

$$\begin{aligned}\max_{\lambda, \nu} \quad & h^\top \lambda + b^\top \nu \\ \text{s.t.} \quad & G^\top \lambda + A^\top \nu = c \\ & \lambda \geq 0\end{aligned}$$

二、KKT条件的推导

1. 构造拉格朗日函数

对于原始问题：

$$\begin{aligned} \min_x \quad & c^\top x \\ \text{s.t.} \quad & Gx \leq h \\ & Ax = b \end{aligned}$$

我们引入拉格朗日乘子 $\lambda \geq 0$ 和 ν，构造拉格朗日函数：

$$L(x, \lambda, \nu) = c^\top x + \lambda^\top (Gx - h) + \nu^\top (Ax - b)$$

2. Stationarity（驻点条件）

对 x 求导并令导数为零，得到驻点条件：

$$\nabla_x L(x, \lambda, \nu) = 0$$

计算梯度：

$$\nabla_x L(x, \lambda, \nu) = c + G^\top \lambda + A^\top \nu = 0$$

这意味着：

$$c + G^\top \lambda + A^\top \nu = 0$$

3. Primal Feasibility（原始可行性）

原始问题的可行解需满足：

$$\begin{aligned} Gx \leq h \\ Ax = b \end{aligned}$$

4. Dual Feasibility（对偶可行性）

拉格朗日乘子对应于对偶变量，需满足：

$$\lambda \geq 0$$

5. Complementary Slackness（互补松弛条件）

互补松弛条件描述了原始约束和对偶变量之间的关系：

$$\lambda_i (Gx - h)_i = 0, \quad \forall i$$

这意味着对于每个 i：

- 如果 $\lambda_i > 0$，则对应的约束必须紧（即 $Gx = h$）。
- 如果 $Gx_i < h_i$，则对应的 $\lambda_i = 0$。

6. 综合KKT条件

综合以上条件，KKT条件如下：

1. **原始可行性**：

$$Gx \leq h$$
$$Ax = b$$

2. **对偶可行性**：

$$\lambda \geq 0$$

3. **驻点条件**：

$$c + G^{\top}\lambda + A^{\top}\nu = 0$$

4. **互补松弛条件**：

$$\lambda_i (Gx - h)_i = 0, \quad \forall i$$

推导以下问题的对偶问题:

$$\min_{x} \frac{1}{2}\|x - x_0\|_2^2 + \sum_{i=1}^{N} \|A_i x + b_i\|_2$$

其中 $A_i \in \mathbb{R}^{m_i \times n}, b_i \in \mathbb{R}^{m_i}$,且 $x_0 \in \mathbb{R}^n$ 。(提示: 引入新的变量 $y_i \in \mathbb{R}^{m_i}$ 以及等式约束 $y_i = A_i x + b_i$,将原无约束优化问题转化为约束优化问题后,再推导其对偶问题。)

原问题：

$$\min_{x} \frac{1}{2}\|x - x_0\|_2^2 + \sum_{i=1}^{N} \|A_i x + b_i\|_2$$

其中，$A_i \in \mathbb{R}^{m_i \times n}$，$b_i \in \mathbb{R}^{m_i}$，$x_0 \in \mathbb{R}^n$。

首先，将原问题改写为带约束的形式：

$$\begin{aligned} \min_{x, y_i} \quad & \frac{1}{2}\|x - x_0\|_2^2 + \sum_{i=1}^{N} \|y_i\|_2 \\ \text{subject to} \quad & y_i = A_i x + b_i, \quad i = 1, 2, \ldots, N \end{aligned}$$

引入拉格朗日乘子 $\lambda_i \in \mathbb{R}^{m_i}$，构造拉格朗日函数：

$$L(x, y_i, \lambda_i) = \frac{1}{2}\|x - x_0\|_2^2 + \sum_{i=1}^{N} \|y_i\|_2 + \sum_{i=1}^{N} \lambda_i^T (y_i - A_i x - b_i)$$

对偶函数定义为：

$$g(\lambda) = \inf_{x, y_i} L(x, y_i, \lambda_i)$$

首先，对 x 最小化：

$$\nabla_x L = (x - x_0) - \sum_{i=1}^{N} A_i^T \lambda_i = 0 \quad \Rightarrow \quad x = x_0 + \sum_{i=1}^{N} A_i^T \lambda_i$$

然后，对每个 y_i 最小化：

$$\min_{y_i} \|y_i\|_2 + \lambda_i^T y_i$$

根据优化理论，该最小值为：

$$\min_{y_i} \|y_i\|_2 + \lambda_i^T y_i = \begin{cases} 0 & \text{如果 } \|\lambda_i\|_2 \leq 1 \\ -\infty & \text{如果 } \|\lambda_i\|_2 > 1 \end{cases}$$

因此，对偶函数 $g(\lambda)$ 为：

$$g(\lambda) = \begin{cases} -\frac{1}{2}(\sum_{i,j} \lambda_i^T A_i A^T \lambda_j + 2\sum_{i=1}^N \lambda_i A_i x_0) - \sum_{i=1}^N \lambda_i^T b_i & \text{如果 } \|\lambda_i\|_2 \leq 1\ \forall i \\ -\infty & \text{否则} \end{cases}$$

最终的对偶问题是：

$$\begin{aligned} \max_{\lambda_i} \quad & -\frac{1}{2}(\sum_{i,j} \lambda_i^T A_i A^T \lambda_j + 2\sum_{i=1}^N \lambda_i A_i x_0) - \sum_{i=1}^N \lambda_i^T b_i \\ \text{subject to} \quad & \|\lambda_i\|_2 \leq 1, \quad i = 1, 2, \ldots, N \end{aligned}$$