拉格朗日乘子法

假设$f\left(\mathbf{x}\right), c_i\left(\mathbf{x}\right), h_j\left(\mathbf{x}\right)$是定义在$\mathbb{R}^n$上的连续可微函数。
称约束最优化问题

$$\begin{aligned} \min_{\mathbf{x}\in\mathbb{R}^n} \quad & f\left(\mathbf{x}\right) \\ s.t. \quad & c_i\left(\mathbf{x}\right) \leq 0, \quad i=1,2,\cdots,k \\ & h_j\left(\mathbf{x}\right) = 0, \quad j=1,2,\cdots,l \end{aligned}$$

为原始最优化问题或原始问题。

引入拉格朗日函数

$$L\left(\mathbf{x},\alpha,\beta\right) = f\left(\mathbf{x}\right) + \sum_{i=1}^{k} \alpha_i c_i\left(\mathbf{x}\right) + \sum_{j=1}^{l} \beta_j h_j\left(\mathbf{x}\right)$$

其中，$\mathbf{x} = \left(x^{(1)}, x^{(2)}, \cdots, x^{(n)}\right)^T \in \mathbb{R}^n, \alpha_i, \beta_j$是拉格朗日乘子，$\alpha_i \geq 0$。

构建关于$\mathbf{x}$的函数

$$\theta_P\left(\mathbf{x}\right) = \max_{\alpha,\beta;\alpha_i\geq 0} L\left(\mathbf{x},\alpha,\beta\right)$$

假设给定某个违反原始问题约束条件的$\mathbf{x}$，即存在某个i使得$c_i\left(\mathbf{x}\right) > 0$或$h_j\left(\mathbf{x}\right) \neq 0$。若$c_i\left(\mathbf{x}\right) > 0$，可令$\alpha_i \to +\infty$，使得$\theta_P\left(\mathbf{x}\right) = +\infty$；若$h_j\left(\mathbf{x}\right) \neq 0$，可令$\beta_j$使$\beta_j h_j\left(\mathbf{x}\right) \to +\infty$，使得$\theta_P\left(\mathbf{x}\right) = +\infty$。将其余$\alpha_i, \beta_j$均取值为0。
即

$$\theta_P\left(\mathbf{x}\right) = \max_{\alpha,\beta;\alpha_i\geq 0} \left[f\left(\mathbf{x}\right) + \sum_{i=1}^{k} \alpha_i c_i\left(\mathbf{x}\right) + \sum_{j=1}^{l} \beta_j h_j\left(\mathbf{x}\right) \right] = +\infty$$

假设给定某个符合原始问题约束条件的$\mathbf{x}$，即$c_i\left(\mathbf{x}\right) \leq 0$且$h_j\left(\mathbf{x}\right) = 0$，
则

$$\theta_P\left(\mathbf{x}\right) = \max_{\alpha,\beta;\alpha_i\geq 0} \left[f\left(\mathbf{x}\right) + \sum_{i=1}^{k} \alpha_i c_i\left(\mathbf{x}\right) + \sum_{j=1}^{l} \beta_j h_j\left(\mathbf{x}\right) \right] = f\left(\mathbf{x}\right)$$

由以上，得

$$\theta_P\left(\mathbf{x}\right) = \begin{cases} f\left(\mathbf{x}\right), \text{当}\mathbf{x}\text{满足原始问题约束} \\ +\infty, \text{否则} \end{cases}$$

则极小化问题

$$\min_{\mathbf{x}} \theta_P\left(\mathbf{x}\right) = \min_{\mathbf{x}} \max_{\alpha,\beta;\alpha_i\geq 0} L\left(\mathbf{x},\alpha,\beta\right)$$

与原始最优化问题等价，即有相同的解。

$$\min_{\mathbf{x}} \max_{\alpha,\beta;\alpha_i \geq 0} L(\mathbf{x}, \alpha, \beta)$$

称为广义拉格朗日函数的极小极大问题。

定义原始问题的最优值

$$p^* = \min_{\mathbf{x}} \theta_P(\mathbf{x})$$

称为原始问题的值。

构建关于α, β的函数

$$\theta_D(\alpha, \beta) = \min_{\mathbf{x}} L(\mathbf{x}, \alpha, \beta)$$

则极大化问题

$$\max_{\alpha,\beta;\alpha_i \geq 0} \theta_D(\alpha, \beta) = \max_{\alpha,\beta;\alpha_i \geq 0} \min_{\mathbf{x}} L(\mathbf{x}, \alpha, \beta)$$

称为广义拉格朗日函数的极大极小问题。

将广义拉格朗日函数的极大极小问题表示为约束最优化问题

$$\begin{aligned} \max_{\alpha,\beta;\alpha_i \geq 0} \theta_D(\alpha, \beta) = \max_{\alpha,\beta;\alpha_i \geq 0} \min_{\mathbf{x}} L(\mathbf{x}, \alpha, \beta) \\ s.t. \quad \alpha_i \geq 0, \quad i = 1, 2, \cdots, k \end{aligned}$$

称为原始问题的对偶问题。

定义对偶问题的最优值

$$d^* = \max_{\alpha,\beta;\alpha_i \geq 0} \theta_D(\alpha, \beta)$$

称为对偶问题的值。

若原始问题与对偶问题都有最优解，则

$$d^* = \max_{\alpha,\beta;\alpha_i \geq 0} \min_{\mathbf{x}} L(\mathbf{x}, \alpha, \beta) \leq \min_{\mathbf{x}} \max_{\alpha,\beta;\alpha_i \geq 0} L(\mathbf{x}, \alpha, \beta) = p^*$$

对于原始问题及其对偶问题，假设函数$f(\mathbf{x})$和$c_i(\mathbf{x})$是凸函数，$h_j(\mathbf{x})$是仿射函数，且不等式约束$c_i(\mathbf{x})$是严格可行的，即存在$\mathbf{x}$，对所有i有$c_i(\mathbf{x}) < 0$，则存在$\mathbf{x}^*, \alpha^*, \beta^*$，使$\mathbf{x}^*$是原始问题的解，$\alpha^*, \beta^*$是对偶问题的解，并且

$$p^* = d^* = L\left(\mathbf{x}^*, \alpha^*, \beta^*\right)$$

对于原始问题及其对偶问题，假设函数$f(\mathbf{x})$和$c_i(\mathbf{x})$是凸函数，$h_j(\mathbf{x})$是仿射函数，且不等式约束$c_i(\mathbf{x})$是严格可行的，即存在$\mathbf{x}$，对所有i有$c_i(\mathbf{x})<0$，则存在$\mathbf{x}^*, \alpha^*, \beta^*$，使$\mathbf{x}^*$是原始问题的解，$\alpha^*, \beta^*$是对偶问题的解的充分必要条件是$\mathbf{x}^*, \alpha^*, \beta^*$满足下面的Karush-Kuhn-Tucker(KKT)条件：

$$
\begin{aligned}
& \nabla_{\mathbf{x}} L\left(\mathbf{x}^*, \alpha^*, \beta^*\right)=0 \\
& \nabla_\alpha L\left(\mathbf{x}^*, \alpha^*, \beta^*\right)=0 \\
& \nabla_\beta L\left(\mathbf{x}^*, \alpha^*, \beta^*\right)=0 \\
& \alpha_i^* c_i\left(\mathbf{x}^*\right)=0, \quad i=1,2, \cdots, k \\
& c_i\left(\mathbf{x}^*\right) \leq 0, \quad i=1,2, \cdots, k \\
& \alpha_i^* \geq 0, \quad i=1,2, \cdots, k \\
& h_j\left(\mathbf{x}^*\right)=0, \quad j=1,2, \cdots, l
\end{aligned}
$$