拉格朗日对偶性

> 在约束最优化问题中，可以利用**拉格朗日对偶性**（Lagrange duality）将原始问题转换为对偶问题，通过求解对偶问题得到原始问题的解

1. 原始问题

- 假设$f(x), c_i(x), h_j(x)$是定义在$\mathbb{R}^n$上的连续可微函数，考虑约束最优化问题：

$$\begin{aligned} \min_{P \in \mathcal{C}} \quad & f(x) & & \\ \text{s.t.} \quad & c_i(x) & \leq 0, \quad & i = 1, 2, \ldots, k \\ & h_j(x) & = 0 \quad & j = 1, 2, \ldots, l \end{aligned}$$

称此约束最优化问题为**原始最优化问题**或**原始问题**；

- **广义拉格朗日函数**（generalized Lagrange function）：

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

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

- 考虑x的函数（下标P表示原始问题）：

$$\theta_P(x) = \max_{\alpha, \beta; \alpha_i \geq 0} L(x, \alpha, \beta) \tag{2}$$

可以证明：

$$\theta_P(x) = \begin{cases} f(x), & x\text{满足原始问题约束} \\ +\infty, & \text{其他} \end{cases}$$

- 所有再考虑其极小化问题：

$$\min_x \theta_P(x) = \min_x \max_{\alpha,\beta;\alpha_i \geq 0} L(x,\alpha,\beta) \tag{3}$$

与原始最优化问题等价（有相同的解），问题$\min_x \max_{\alpha,\beta;\alpha_i \geq 0} L(x,\alpha,\beta)$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：

$$p^* = \min_x \theta_P(x) \tag{4}$$

2. 对偶问题

- 定义α,β的函数（下标D表示对偶问题）：

$$\theta_D(\alpha,\beta) = \min_x L(x,\alpha,\beta) \tag{5}$$

- 再考虑其极大化：

$$\max_{\alpha,\beta;\alpha_i \geq 0} \theta_D(\alpha,\beta) = \max_{\alpha,\beta;\alpha_i \geq 0} \min_x L(x,\alpha,\beta) \tag{6}$$

问题$\max_{\alpha,\beta;\alpha_i \geq 0} \min_x L(x,\alpha,\beta)$称为**广义拉格朗日函数的极大极小问题**

- 广义拉格朗日函数的极大极小问题可以表示为约束最优化问题：

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

称为原始问题的**对偶问题**；

- 定义对偶问题的最优值：

$$d^* = \max_{\alpha,\beta;\alpha_i \geq 0} \theta_D(\alpha,\beta) \tag{7}$$

3. 原始问题和对偶问题的关系

- **定理1**：可证明如果原始问题和对偶问题都有最优值，则

$$d^* = \max_{\alpha,\beta;\alpha_i \geq 0} \theta_D(\alpha,\beta) \quad \leq \quad \min_x \theta_P(x) = p^* \tag{8}$$

- **推论1**：设x^*和α^*, β^*分别是原始问题和对偶问题的*可行解*，且$d^* = p^*$，则x^*和α^*, β^*分别是原始问题和对偶问题的最优解；（某些条件下原始问题和对偶问题的最优值相等，可以*用解对偶问题替代解原始问题*）；

- **定理2**：假设函数$f(x)$和$c_i(x)$是凸函数，$h_j(x)$是仿射函数；且不等式约束$c_i(x)$是严格可行（存在x，有$c_i(x) < 0, \forall i$），则存在x^*, α^*, β^*，使得x^*是原始问题的解，α^*, β^*是对偶问题的解，且：

$$d^* = p^* = L(x^*, \alpha^*, \beta^*) \tag{9}$$

- **定理3**：假设函数$f(x)$和$c_i(x)$是凸函数，$h_j(x)$是仿射函数；且不等式约束$c_i(x)$是严格可行，则使得x^*是原始问题的解，α^*, β^*是对偶问题的解的*充分必要条件*是满足下面的Karush-Kuhn-Tucker（KKT）条件：

$$\begin{aligned}
\nabla_x L(x^*, \alpha^*, \beta^*) = 0 \\
\nabla_\alpha L(x^*, \alpha^*, \beta^*) = 0 \\
\nabla_\beta L(x^*, \alpha^*, \beta^*) = 0 \\
\alpha_i^* c_i(x^*) = 0, i = 1, 2, \ldots, k \\
c_i(x^*) \leq 0, i = 1, 2, \ldots, k \\
\alpha_i^* \geq 0, i = 1, 2, \ldots, k \\
h_j(x^*) = 0, j = 1, 2, \ldots, l
\end{aligned}$$

其中$\alpha_i^* c_i(x^*) = 0, i = 1, 2, \ldots, k$成为KKT的*对偶互补条件*，由此条件可知：如果$\alpha_i^* > 0$，则$c_i(x^*) = 0$