Lagrangian Duality

- 原问题为带约束的最优化问题

$$\min_{x\in R^n} f(x)$$
$$c_i(x) \le 0, i=1,2,\dots,k$$
$$h_j(x)=0, j=1,2,..,l$$

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

$$L(x,\alpha,\beta)=f(x)+\sum_{i=1}^{k}\alpha_i c_i(x)+\sum\nolimits_{j=1}^{l}\beta_j h_j(x)$$
$$\text{其中}x\in R^n,\alpha_i,\beta_j\text{是拉格朗日乘子，且要求}\alpha_i\ge 0$$

- 原问题变化

$$\min_{x\in R^n} f(x) \quad ==> \quad \min_x \theta_P(x)$$
$$\theta_P(x)=\max_{\alpha\ge 0,\beta} L(x,\alpha,\beta)$$
$$\text{原问题转化成}\min_x \max_{\alpha\ge 0,\beta} L(x,\alpha,\beta)$$

证明可以从x满足约束情况下与不满足情况下分类证明

- 拉格朗日乘子法将带约束的最优化问题转化成无约束最优化问题
- 通常先求解原问题关于α,β的最大值不太好求，所有转化成对偶问题求解
- 定义原始问题的对偶问题

$$\max_{\alpha>0,\beta}\theta_D(\alpha,\beta)=\max_{\alpha>0,\beta}\min_x L(x,\alpha,\beta)$$

- 定理（弱对偶性 Weak Duality）：如果原始问题和对偶问题都有最优值，则

$$\theta_D(\alpha,\beta)\le\theta_P(x)$$
$$\max_{\alpha>0,\beta}\theta_D(\alpha,\beta)\le\min_x\theta_P(x)$$
$$d^*=\max_{\alpha>0,\beta}\min_x L(x,\alpha,\beta)\le\min_x\max_{\alpha\ge 0,\beta}L(x,\alpha,\beta)=p^*$$

- 推论：设x^*,α^*,β^*分别是原始问题和对偶问题的可行解，如果d^*=p^*，那么x^*,α^*,β^*分别是原始问题和对偶问题的最优解。

Slater条件

- 定理(强对偶性)：对于原始问题和对偶问题，假设函数f(x)和c_i(x)是凸函数， h_j(x)是仿射函数（即由一阶多项式构成的函数）；并且假设**不等式约束c_i(x)是严格可行的**，即存在x，对所有i有c_i(x) <0，则存在x^*,α^*,β^*，使得x^*是原始问题的最优解， α^*,β^*是对偶问题的最优解，**并且**

$$d^*=p^*=L(x^*,\alpha^*,\beta^*)$$

KKT条件

- 定理：对于原始问题和对偶问题，假设函数f(x)和c_i(x)是凸函数， h_j(x)是仿射函数（即由一阶多项式构成的函数）；并且假设不等式约束c_i(x)是严格可行的，即存在x，对所有i有c_i(x)<0，则x^∗,α^∗,β^∗分别是原始问题和对偶问题的最优解的充分必要条件是x^∗,α^∗,β^∗满足下面的Karush-Kuhn-Tucker(KKT)条件：

$$
\begin{gathered}
\nabla_x L(x^*, \alpha^*, \beta^*) = 0 \\
\alpha_i^* c_i(x^*) = 0, i = 1, 2, \ldots, k \qquad (KKT\text{对偶互补条件}) \\
c_i(x^*) \leq 0, i = 1, 2, \ldots, k \\
\alpha_i^* \geq 0, i = 1, 2, \ldots, k \\
h_j(x^*) = 0, j = 1, 2, .., l
\end{gathered}
$$

- 通过KKT条件可以将求解原始问题的最优解，转化成其对偶问题的最优解
- 线性规划的对偶形式可由拉格朗日乘子法推导