§4.2 对偶与鞍点问题

考虑一般约束优化问题(4.1.1), 即:

$$\begin{aligned} \min \quad & f(x) \\ \text{s.t.} \quad & c_i(x) \geq 0, \quad \forall i \in I = \{1, 2, \ldots, p\}, \\ & c_i(x) = 0, \quad \forall i \in E = \{p+1, \ldots, m\}. \end{aligned} \tag{4.2.1}$$

设$\mathscr{F}$为该问题的可行域，且广义Lagrange函数: $L(x, \lambda) = f(x) - \sum_{i \in I \cup E} \lambda_i c_i(x)$. 称问题(4.2.1)为原问题，其对偶问题的形式定义为:

$$\begin{aligned} \max \quad & \theta(\lambda) \\ \text{s.t.} \quad & \lambda_i \geq 0, \ \forall i \in I, \end{aligned} \tag{4.2.2}$$

其中$\theta(\lambda) := \inf L(x, \lambda)$, 称其为原问题(4.2.1)的Lagrange对偶函数. 如果广义Lagrange函数无下界，那么$\theta(\lambda) = -\infty$. 值得注意的是: 对任意的$x \in \mathbb{R}^n$，广义Lagrange 函数$L(x, \cdot)$关于变量$\lambda$是线性函数. 于是, Lagrange对偶函数$\theta(\lambda)$是一个凹函数，即可说明对偶问题(4.2.2)是一个凸规划问题.

4.2 对偶与鞍点问题

原问题(4.2.1)和对偶问题(4.2.2) 的目标函数值具有如下关系.

定理 4.2.1 *(弱对偶定理)* 设x和λ分别是原问题*(4.2.1)*和对偶问题*(4.2.2)*任意的可行解，则

$$f(x) \geq \theta(\lambda).$$

证明 由于$x \in \mathbb{R}^n$和$\lambda = (\lambda_1, \ldots, \lambda_m)^\top$分别是原问题(4.2.1) 和对偶问题(4.2.2)的可行解，所以

$$\theta(\lambda) = \inf L(x, \lambda) = \inf \left\{ f(x) - \sum_{i \in I \cup E} \lambda_i c_i(x) \right\} \leq f(x) - \sum_{i \in I \cup E} \lambda_i c_i(x) \leq f(x).$$

定理得证. □

4.2 对偶与鞍点问题

推论 4.2.1 设$\mathscr{F}$和$\mathscr{D}$分别是原问题*(4.2.1)*和对偶问题*(4.2.2)*的可行域，则下列结论成立：

(i) 若原问题*(4.2.1)*和对偶问题*(4.2.2)*可行，即$\mathscr{F} \neq \emptyset$ 且$\mathscr{D} \neq \emptyset$，则

$$\min\{f(x) \mid x \in \mathscr{F}\} \geq \max\{\theta(\lambda) \mid \lambda \in \mathscr{D}\}.$$

(ii) 若原问题*(4.2.1)*的最优值$f_{\min} = -\infty$，则对任意的$\lambda \in \mathscr{D}$，都有

$$\theta(\lambda) = -\infty.$$

(iii) 若对偶问题*(4.2.2)*的最优值$\theta_{\max} = \infty$，则原问题无可行解，即$\mathscr{F} = \emptyset$.

4.2 对偶与鞍点问题

把原问题(4.2.1)的最优值和对偶问题(4.2.2)的最优值之间的差值称为**对偶间隙(或对偶间隔)**. 对于一般的约束优化问题，对偶间隙不一定为零，那么，在什么条件下对偶间隙为零呢？下面先给出鞍点的概念，然后再给出其等价条件. 在本文中, 以下记$\mathbb{R}^p_+ := \{x \in \mathbb{R}^p \mid x \geq 0\}$.

定义 4.2.1 对于一般约束优化问题*(4.2.1)*, 若存在$x^* \in \mathbb{R}^n$和$\lambda^* \in \mathbb{R}^p_+ \times \mathbb{R}^{m-p}$, 使得对任意的$x \in \mathbb{R}^n$和$\lambda \in \mathbb{R}^p_+ \times \mathbb{R}^{m-p}$, 都有

$$L(x^*, \lambda) \leq L(x^*, \lambda^*) \leq L(x, \lambda^*),$$

则称点(x^*, λ^*)为该约束优化问题广义*Lagrange*函数的鞍点.

由鞍点的定义，直接验证可得下面的引理.

引理 4.2.1 点(x^*, λ^*)为一般约束优化问题广义*Lagrange*函数的鞍点的充分必要条件为

$$\sup_{\lambda \in \mathbb{R}^p_+ \times \mathbb{R}^{m-p}} L(x^*, \lambda) = L(x^*, \lambda^*) = \inf_{x \in \mathbb{R}^n} L(x, \lambda^*).$$

4.2 对偶与鞍点问题

定理 4.2.2 *(强对偶定理或鞍点定理)* 对于问题*(4.2.1)*，对应的广义*Lagrange*函数$L(\cdot,\cdot)$存在鞍点(x^*,λ^*)当且仅当x^*和λ^*分别是原问题*(4.2.1)*和对偶问题*(4.2.2)*的最优解，且对偶间隙为零，即

$$\theta(\lambda^*) = f(x^*).$$

证明 先证必要性. 假设点(x^*,λ^*)是约束优化问题(4.2.1)对应广义Lagrange函数的鞍点. 由鞍点的定义及其等价条件，则对任意的$x \in \mathbb{R}^n$和$\lambda \in \mathbb{R}^p_+ \times \mathbb{R}^{m-p}$, 都有

$$\begin{aligned}\inf_{x\in\mathbb{R}^n} \sup_{\lambda\in\mathbb{R}^p_+\times\mathbb{R}^{m-p}} L(x,\lambda) &\le \sup_{\lambda\in\mathbb{R}^p_+\times\mathbb{R}^{m-p}} L(x^*,\lambda) = L(x^*,\lambda^*) = \inf_{x\in\mathbb{R}^n} L(x,\lambda^*) \\ &\le \sup_{\lambda\in\mathbb{R}^p_+\times\mathbb{R}^{m-p}} \inf_{x\in\mathbb{R}^n} L(x,\lambda). \qquad (4.2.3)\end{aligned}$$

由于$\sup_{\lambda\in\mathbb{R}^p_+\times\mathbb{R}^{m-p}} \inf_{x\in\mathbb{R}^n} L(x,\lambda) \le \inf_{x\in\mathbb{R}^n} \sup_{\lambda\in\mathbb{R}^p_+\times\mathbb{R}^{m-p}} L(x,\lambda)$成立, 根据(4.2.3)式, 有

$$\sup_{\lambda\in\mathbb{R}^p_+\times\mathbb{R}^{m-p}} \inf_{x\in\mathbb{R}^n} L(x,\lambda) = \inf_{x\in\mathbb{R}^n} \sup_{\lambda\in\mathbb{R}^p_+\times\mathbb{R}^{m-p}} L(x,\lambda).$$

再结合(4.2.3)式, 可得x^*是原约束优化问题(4.2.1)的最优解，λ^*是对偶问题(4.2.2)的最优解, 且对偶间隙为零.

4.2 对偶与鞍点问题

反之，若x^*和λ^*分别是原问题(4.2.1) 和对偶问题(4.2.2)的最优解, 则

$$\sup_{\lambda\in\mathbb{R}_+^p\times\mathbb{R}^{m-p}} \inf_{x\in\mathbb{R}^n} L(x,\lambda) = \inf_{x\in\mathbb{R}^n} L(x,\lambda^*)$$

且

$$\inf_{x\in\mathbb{R}^n} \sup_{\lambda\in\mathbb{R}_+^p\times\mathbb{R}^{m-p}} L(x,\lambda) = \sup_{\lambda\in\mathbb{R}_+^p\times\mathbb{R}^{m-p}} L(x^*,\lambda).$$

由原问题(4.2.1)和对偶问题(4.2.2)的对偶间隙为零，故

$$\sup_{\lambda\in\mathbb{R}_+^p\times\mathbb{R}^{m-p}} \inf_{x\in\mathbb{R}^n} L(x,\lambda) = \inf_{x\in\mathbb{R}^n} \sup_{\lambda\in\mathbb{R}_+^p\times\mathbb{R}^{m-p}} L(x,\lambda).$$

此外，由于$\inf_{x\in\mathbb{R}^n} L(x,\lambda^*) \le L(x^*,\lambda^*) \le \sup_{\lambda\in\mathbb{R}_+^p\times\mathbb{R}^{m-p}} L(x^*,\lambda)$ 成立，再结合上式，可得

$$\sup_{\lambda\in\mathbb{R}_+^p\times\mathbb{R}^{m-p}} L(x^*,\lambda) = L(x^*,\lambda^*) = \inf_{x\in\mathbb{R}^n} L(x,\lambda^*).$$

由引理4.2.1得: 点(x^*,λ^*)是广义Lagrange函数$L(\cdot,\cdot)$的鞍点. □

4.2 对偶与鞍点问题 — 作业

4.17 对于二次规划问题

$$
\begin{aligned}
\min \quad & g(x) = -x_1^2 - x_2^2 - x_3^2 + 4x_1 + 6x_2 \\
\text{s.t.} \quad & x_1 + x_2 \le 2, \\
& 2x_1 + 3x_2 \le 12, \\
& x_1, x_2, x_3 \ge 0.
\end{aligned}
$$

(1) 写出该问题所对应的Lagrange函数，并求出该函数的鞍点.

(2) 写出该问题的K-T条件，并求出相应的K-T点.

(3) 用罚函数法求解该二次规划问题.