Homework 4

Name: 方嘉聪 ID: 2200017849

Problem 1. 对于如下原问题和对偶问题, 其中 $f(x), g(x)$ 为凸函数, $h(x)$ 为线性函数.

$$
(\mathbf{P}) \quad \begin{aligned} &\min_x f(x) \\ &\text{s.t. } g_i(x) \le 0 \\ &\qquad h_i(x) = 0 \end{aligned} \qquad \text{and} \qquad (\mathbf{D}) \quad \begin{aligned} &\max_{\lambda,\mu} \mathcal{L}(\varphi(\lambda,\mu);\lambda,\mu) \\ &\text{s.t. } \lambda \ge 0. \end{aligned}
$$

其中

$$
\mathcal{L}(x;\lambda,\mu) := f(x) + \sum_i \lambda_i g_i(x) + \sum_i \mu_i h_i(x). \quad \varphi(\lambda,\mu) := \operatorname*{argmin}_x \mathcal{L}(x;\lambda,\mu).
$$

以下为关于 x^*, λ^*, μ^* 的 KKT condition:

$$
\textbf{1.Stationary:} \qquad \left.\frac{\partial \mathcal{L}}{\partial x}\right|_{x^*,\lambda^*,\mu^*} = 0 \tag{1}
$$

$$
\textbf{2.Primal feasible:} \qquad g_i(x^*) \le 0, \quad h_i(x^*) = 0 \tag{2}
$$

$$
\textbf{3.Dual feasible:} \qquad \lambda_i^* \ge 0 \tag{3}
$$

$$
\textbf{4.Complementary slackness:} \qquad \lambda_i^* g_i(x^*) = 0. \tag{4}
$$

请证明:

1. x^*, λ^*, μ^* 满足 KKT condition 是 x^*, λ^*, μ^* 为原问题和对偶问题的最优解的必要条件. 即若 x^*, λ^*, μ^* 为原问题和对偶问题的最优解, 则 x^*, λ^*, μ^* 满足 KKT condition.
2. (Optional) KKT condition 是 x^*, λ^*, μ^* 为原问题和对偶问题的最优解的充分条件.

◀

Solution. 1. (必要条件) 若 x^*, λ^*, μ^* 为原问题和对偶问题的最优解, (2) 和 (3) 显然成立. 对于 (1), 结合课上已经完成的推导, 最优解满足 $x^* = \varphi(\lambda^*, \mu^*) = \operatorname{argmin}_x \mathcal{L}(x;\lambda^*,\mu^*)$, 即

$$
\left.\frac{\partial \mathcal{L}}{\partial x}\right|_{x^*,\lambda^*,\mu^*} = 0.
$$

对于 (4), 记 Lagrange dual function 为

$$
s(\lambda,\mu) := \inf_x \mathcal{L}(x;\lambda,\mu) = \inf_x \left(f(x) + \sum_i \lambda_i g_i(x) + \sum_i \mu_i h_i(x) \right).
$$

由于 $f(x), g(x)$ 为凸函数, $h(x)$ 为线性函数, 故满足强对偶性, 即 $f(x^*) = s(\lambda^*, \mu^*)$. 进一步有

$$
\begin{aligned}
f(x^*) = s(\lambda^*,\mu^*) &= \inf_x \left(f(x) + \sum_i \lambda_i^* g_i(x) + \sum_i \mu_i^* h_i(x) \right) \\
&\le f(x^*) + \sum_i \lambda_i^* g_i(x^*) + \sum_i \mu_i^* h_i(x^*) \\
&\le f(x^*).
\end{aligned}
$$

故有

$$\sum_i \lambda_i^* g_i(x^*) + \sum_i \mu_i^* h_i(x^*) = 0. \implies \sum_i \lambda_i^* g_i(x^*) = 0.$$

由于 $\forall i, \lambda_i^* g_i(x^*) \leq 0$, 故 $\lambda_i^* g_i(x^*) = 0, \forall i$. 综上, x^*, λ^*, μ^* 满足 KKT condition.

2. (充分条件) 由于 x^*, λ^*, μ^* 满足 (2) 和 (3). 故 x^* 和 λ^*, μ^* 分别为原问题和对偶问题的可行解. 设原问题的最优解为 x_{opt}, 记 $f(x_{\text{opt}}) = f^*$. 可以证明 (即弱对偶性), 对于任意 $\lambda \geq 0$ 和 μ, 有

$$s(\lambda, \mu) \leq f^*. \tag{5}$$

取 $\lambda = \lambda^*, \mu = \mu^*$, 有

$$s(\lambda^*, \mu^*) \leq f^*.$$

而由于 x^*, λ^*, μ^* 满足 (1), 那么

$$x^* = \varphi(\lambda^*, \mu^*) = \operatorname*{argmin}_x \mathcal{L}(x; \lambda^*, \mu^*).$$

故

$$\begin{aligned} s(\lambda^*, \mu^*) &= \inf_x \mathcal{L}(x; \lambda^*, \mu^*) \\ &= f(x^*) + \sum_i \lambda_i^* g_i(x^*) + \sum_i \mu_i^* h_i(x^*) \\ &= f(x^*) \leq f^*. \end{aligned}$$

由于 x^* 为原问题的可行解, 故 $f(x^*) \geq f^*$, 故 $f(x^*) = f^*$. 即 x^* 为原问题的最优解. 结合 (5), 可知 λ^*, μ^* 为对偶问题的最优解.

综上, x^*, λ^*, μ^* 为原问题和对偶问题的最优解.

$\lhd$