THE CHINESE UNIVERSITY OF HONG KONG
Department of Mathematics
MATH4230 2024-25
Lecture 16
March 18, 2025 (Tuesday)

1 Recall

In the previous lecture, recall that the **primal problem** as follows:

$$\min_{x\in\mathbb{R}^n} f(x) \quad \text{subject to} \quad \begin{cases} g_i(x) \le 0, & i = 1,\dots,\ell \\ h_j(x) = 0, & j = 1,\dots,m \end{cases} \tag{P}$$

where $f, g_i : \mathbb{R}^n \to \mathbb{R}$ are convex differentiable functions while $h_j : \mathbb{R}^n \to \mathbb{R}$ are affine functions, i.e. $h_j(x) = A_j^T x + b_j$.

and we define the **Lagrangian function** L by

$$L(x,\lambda,\mu) := f(x) + \sum_{i=1}^{\ell} \lambda_i g_i(x) + \sum_{j=1}^{m} \mu_j h_j(x)$$

for $\lambda_i \ge 0$, $\mu_j \in \mathbb{R}$ for all $i = 1,\dots,\ell$ and $j = 1,\dots,m$. Also **the Lagrange dual problem** is defined as:

$$\max_{\substack{\lambda_i\in\mathbb{R}_+,\, i=1,\dots,\ell \\ \mu_j\in\mathbb{R},\, j=1,\dots,m}} d(\lambda,\mu) \tag{D}$$

where $d(\lambda,\mu) := \inf_{x\in\mathbb{R}^n} L(x,\lambda,\mu)$.
Also, from the weak duality, we have $P \ge D$ since $\inf_x \sup_{(\lambda,\mu)} L(x,\lambda,\mu) \ge \sup_{(\lambda,\mu)} \inf_x L(x,\lambda,\mu)$.
Consider a system of constraints on x as

$$\begin{cases} f(x) & < C, \\ g_i(x) & \le 0, \quad i = 1,\dots,m \\ h_j(x) & = 0, \quad j = 1,\dots,\ell \end{cases} \tag{I}$$

along with the system of constraints on λ:

$$\begin{cases} d(\lambda,\mu) & \ge C \\ \lambda_i & \ge 0, \quad i = 1,\dots,m \\ \mu_j & \in \mathbb{R}, \quad j = 1,\dots,\ell \end{cases} \tag{II}$$

From last lecture, we have the following propositions:

Proposition 1. *If (II) is solvable, (that is $\exists(\lambda,\mu)$ satisfying (II)), then (I) is insolvable.*

Proposition 2. *If (I) is insolvable and Slater condition holds, then (II) is solvable.*

Proof of Proposition 2. Let

$$S := \left\{(u_0, u, v) : u_0 \in \mathbb{R}, u \in \mathbb{R}^m, v \in \mathbb{R}^\ell, u_0 < c, u_i \leq 0,\ v_j = 0\right\}$$

and

$$T := \left\{(u_0, u, v) : \exists x \in \mathbb{R}^n \text{ s.t. } \substack{f(x) \leq u_0 \\ g_i(x) \leq u_i,\, i=1,\ldots,m \\ h_j = v_j,\, j=1,\ldots,\ell}\right\}$$

Then S is non-empty and convex, and T is non-empty and convex with $S \cap T = \emptyset$, otherwise (I) is solvable.
By the Separation theorem, there exists $w \in \mathbb{R}^{1+m+\ell}$ such that

$$\inf_{(u_0,u,v)\in T} \langle w, (u_0, u, v)\rangle \geq \sup_{(u_0,u,v)\in S} \langle w, (u_0, u, v)\rangle$$

Then, since

$$+\infty > \inf_{(u_0,u,v)\in T} \langle w, (u_0, u, v)\rangle = \inf_{\substack{u_0 \geq f(x) \\ u_i \geq g_i(x)\ ,x\in\mathbb{R}^n \\ v_j = h_j(x)}} \left(w_0 u_0 + \sum_{i=1}^m w_i u_i + \sum_{j=1}^\ell w_{m+j} v_j\right)$$

and

$$\begin{aligned}
\sup_{(u_0,u,v)\in S} \langle w, (u_0, u, v)\rangle &= \sup_{\substack{u_0 < c \\ u_i \leq 0 \\ v_j = 0}} \left(w_0 u_0 + \sum_{i=1}^m w_i u_i + \cancelto{0}{\sum_{j=1}^\ell w_{m+j} v_j}\right) \\
&= \sup_{\substack{u_0 < c \\ u_i \leq 0}} \left(w_0 u_0 + \sum_{i=1}^m w_i u_i\right)
\end{aligned}$$

Therefore, we have

$$w_0 C = \sup_{\substack{u_0 < c \\ u_i \leq 0}} \left(w_0 u_0 + \sum_{i=1}^m w_i u_i\right) < +\infty \implies w_i \geq 0,\ \forall i = 1, \ldots, m \quad \text{and} \quad w_0 \geq 0$$

Also, we consider

$$\inf_{\substack{x\in\mathbb{R}^n \\ u_0 \geq f(x) \\ u_i(x) \geq g(x) \\ v_j = h_j(x)}} \left(w_0 u_0 + \sum_{i=1}^m w_i u_i + \sum_{j=1}^\ell w_{m+j} v_j\right) = \inf_{x\in\mathbb{R}^n} \left(w_0 f(x) + \sum_{i=1}^m w_i g_i(x) + \sum_{j=1}^\ell w_{m+j} h_j(x)\right) \geq w_0 C$$

If $w_0 = 0$, then

$$0 = w_0 C \leq \inf_{x\in\mathbb{R}^n} \left(w_0 f(x) + \underbrace{\sum w_i g_i(x) + \sum w_{m+j} h_j(x)}_{<0}\right) < 0$$

it is a contradiction. So, we deduce that $w_0 > 0$.
From the above, dividing $w_0 > 0$ on both sides yields:

$$\inf_{x\in\mathbb{R}^n} \left(f(x) + \sum_{i=1}^m \left(\frac{w_i}{w_0}\right) g_i(x) + \sum_{j=1}^\ell \left(\frac{w_{m+j}}{w_0}\right) h_j(x)\right) \geq C \iff d\left(\frac{w_i}{w_0}, \frac{w_{m+j}}{w_0}\right) \geq C$$

Thus, this proves that (I) is solvable. $\square$

Theorem 3. *Assume that $P > -\infty$ and there exists $x \in \mathbb{R}^n$ such that*

$$\begin{cases} g_i(x) < 0, & i = 1, \dots, m \\ h_j(x) = 0, & j = 1, \dots, \ell \end{cases} \tag{Slater Condition}$$

Then $P = D$.

Proof. Let $C := P = \inf_{x \in K} f(x)$ in both (I) and (II), then (I) is insolvable.
Together with the Slater condition, it follows by proposition 2 that (II) is solvable, thus there exists (λ, μ) such that $\lambda_i \geq 0,\ \mu_j \in \mathbb{R}$ and $d(\lambda, \mu) \geq P$. This follows that

$$D \geq d(\lambda, \mu) \geq P \implies D = P.$$

□

2 Saddle Points

Theorem 4. *(i) Let $x^* \in \mathbb{R}^n,\ \lambda^* \in \mathbb{R}^m_+,\ \mu^* \in \mathbb{R}^\ell$ be such that*

$$L(x, \lambda^*, \mu^*) \geq L(x^*, \lambda^*, \mu^*) \geq L(x^*, \lambda, \mu)$$

for all $x \in \mathbb{R}^n,\ (\lambda, \mu) \in \mathbb{R}^m_+ \times \mathbb{R}^\ell$. Then x^* is solution to (P):

$$\inf_{x \in \mathbb{R}^n} f(x) \quad \textit{subject to} \quad \begin{matrix} g_i(x) \leq 0 \\ h_j(x) = 0 \end{matrix}$$

(ii) Let x^* be a solution to (P) and (P) satisfies Slater condition ($\exists x$ such that $g_i(x) < 0,\ h_j(x) = 0$). Then there exists $(\lambda^*, \mu^*) \in \mathbb{R}^m_+ \times \mathbb{R}^\ell$ such that

$$L(x, \lambda^*, \mu^*) \geq L(x^*, \mu^*, \lambda^*) \geq L * (x^*, \mu, \lambda)$$

for all $x \in \mathbb{R}^n$ and $(\lambda, \mu) \in \mathbb{R}^m_+ \times \mathbb{R}^\ell$.

Proof. We first prove part (i). Note that

$$L(x^*, \lambda^*, \mu^*) \geq L(x^*, \lambda, \mu),\ \forall (\lambda, \mu) \in \mathbb{R}^m_+ \times \mathbb{R}^\ell$$

implies that

$$\begin{aligned} L(x^*, \lambda^*, \mu^*) &\geq \sup_{(\lambda,\mu) \in \mathbb{R}^m_+ \times \mathbb{R}^\ell} \Big(f(x^*) + \sum \lambda_i g_i(x^*) + \sum \mu_j h_j(x^*) \Big) \\ &= \begin{cases} f(x^*), & x* \in K = \{x : g_i(x) \leq 0,\ h_j(x) = 0\} \\ +\infty, & x^* \notin K \end{cases} \end{aligned}$$

Since $L(x^*, \lambda^*, \mu^*) < +\infty$, then $x^* \in K$ i.e. $g_i(x^*) \leq 0,\ h_j(x^*) = 0$ and $L(x^*, \lambda^*, \mu^*) = f(x^*)$. This implies that

$$f(x^*) = L(x^*, \lambda^*, \mu^*) \leq \inf_{x \in \mathbb{R}^n} L(x, \lambda^*, \mu^*) = d(\lambda^*, \mu^*)$$

$$\implies f(x^*) \leq \sup_{(\lambda,\mu)} d(\lambda, \mu) = D \overset{\text{by weak duality}}{\leq} P.$$

□

— End of Lecture 16 —