THE CHINESE UNIVERSITY OF HONG KONG
Department of Mathematics
MATH4230 2024-25
Lecture 17
March 19, 2025 (Wednesday)

1 Recall

Recall that the primal problem (P) that we discussed yesterday:

$$P := \inf_{x\in K} f(x) \quad \text{subject to} \quad K := \begin{cases} g_i(x) \le 0, & i = 1,\dots,m \\ h_j(x) = 0, & j = 1,\dots,\ell \end{cases} \tag{P}$$

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 $j = 1,\dots,m$, and **the Lagrange dual problem** is defined as:

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

where $d(\lambda,\mu) := \inf_{x\in\mathbb{R}^n} L(x,\lambda,\mu)$.

Theorem 1. *Under Slater condition (i.e. there exists $x \in K$ such that $g_i(x) < 0,\ i = 1,\dots,m$ and $h_j(x) = 0,\ j = 1,\dots,\ell$), one has*

1. *$P = D$.*

2. *there exists $(\lambda^*,\mu^*) \in \mathbb{R}_m^+ \times \mathbb{R}^\ell$ such that $P = D = d(\lambda^*,\mu^*)$.*

Also, we discussed two systems of constraints as

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

$$d(\lambda,\mu) \ge C, \quad \text{for}\ \ (\lambda,\mu) \in \mathbb{R}_+^m \times \mathbb{R}^\ell \tag{II}$$

and a proposition:

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

Proof of Theorem. 1. Weak Duality: $P \ge D$.

2. Set $C = P$ in both (I) and (II) so that (I) is insolvable.
   By the proposition, (II) is solvable, so there exists $(\lambda^*,\mu^*) \in \mathbb{R}_+^m \times \mathbb{R}^\ell$ such that $D \ge d(\lambda^*,\mu^*) \ge C = P$.
   Thus, we have $P = d(\lambda^*,\mu^*) = D$.

□

Theorem 3. *(Saddle Point)*

(i) Let $(x^*, (\lambda^*, \mu^*)) \in \mathbb{R}^n \times \mathbb{R}^m_+ \times \mathbb{R}^\ell$ *be a saddle point in the sense that*

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

Then x^* *is an optimizer of* (P).

(ii) Let x^* *be an optimizer of* (P) *and Slater condition holds. Then there exists* $(\lambda^*, \mu^*) \in \mathbb{R}^m_+ \times \mathbb{R}^\ell$ *such that* $(x^*, (\lambda^*, \mu^*))$ *is a saddle point of*

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

Proof of (ii). By Duality Theorem, there exists $(\lambda^*, \mu^*) \in \mathbb{R}^m_+ \times \mathbb{R}^\ell$ such that

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

This implies that

$$\begin{aligned} f(x^*) = P &= \inf_{x \in \mathbb{R}^n} L(x, \lambda^*, \mu^*) \\ &\leq L(x^*, \lambda^*, \mu^*) \\ &= f(x^*) + \sum_{i=1}^m \lambda_i^* g_i(x^*) + \sum_{j=1}^\ell \mu_j^* h_j(x^*) \end{aligned}$$

Since $x^* \in K$, we know

$$\begin{cases} g_i(x) \leq 0, \ i = 1, \ldots, m \\ h_j(x^*) = 0, \ j = 1, \ldots, \ell \end{cases} \quad \Longrightarrow \sum_{i=1}^m \lambda_i^* g_i(x^*) \leq 0 \quad \text{and} \quad \sum_{j=1}^\ell \mu_j^* h_j(x^*) = 0.$$

From the above, we can deduce that

$$\begin{aligned} f(x^*) \leq f(x^*) + \underbrace{\sum_{i=1}^m \lambda_i^* g_i(x^*)}_{\leq 0} &\implies \sum_{i=1}^m \lambda_i^* g_i(x^*) = 0 \\ &\implies \lambda_i^* g_i(x^*) = 0, \quad \forall i = 1, \ldots, m \end{aligned}$$

Hence, we have

$$\begin{aligned} L(x, \lambda^*, \mu^*) &\geq L(x^*, \lambda^*, \mu^*), \ \forall x \in \mathbb{R}^n \\ &= f(x^*) \\ &\geq f(x^*) + \sum \underbrace{\lambda_i g_i(x^*)}_{\leq 0} + \underbrace{\sum \mu_j h_j(x^*)}_{=0} \\ &= L(x^*, \lambda, \mu) \end{aligned}$$

for any $(\lambda, \mu) \in \mathbb{R}^m_+ \times \mathbb{R}^\ell$. □

2 Qualification

Recall that, we discuss the **Mangasarian Fromovitz** qualification condition before.

Proposition 4. *Let $x \in K$ satisfies the followings (**Mangasarian Fromovitz condition**)*

1. *the family of vectors $(\nabla h_1(x), \ldots, \nabla h_m(x))$ is linearly independent.*
2. *there exists a vector $v \in \mathbb{R}^n$ satisfying*

$$\langle \nabla h_j(x), v \rangle = 0,\ \forall j = 1, \ldots, m$$

and

$$\langle \nabla g_i(x), v \rangle < 0,\ \forall i \in I(x) := \{k : g_k(x) = 0\}.$$

Then the constraint K is qualified at $x \in K$.

Proposition 5. *Let $g_i(x)$ be a C^1 convex function, and $h_j(x) = A_j^T x + b_j$. Assume that*

1. *$\{A_1, \ldots, A_\ell\}$ is linearly independent.*
2. *There exists $\hat{x} \in \mathbb{R}^n$ such that*

$$\begin{aligned} g_i(\hat{x}) < 0, &\quad i = 1, \cdots, m \\ h_j(\hat{x}) = 0, &\quad j = 1, \ldots, \ell \end{aligned}$$

Then, qualification condition holds for all $x \in K$.

Proof.

1. To apply the Mangasarian Fromovitz condition, we check that

$$\{\nabla h_1(x), \cdots, \nabla h_\ell(x)\} = \{A_1, \cdots, A_\ell\}$$

is linearly independent.

2. For all $x \neq \hat{x}$, and satisfying $g_i(x) \leq 0, h_j(x) = 0$ and $g_i(\hat{x}) < 0,\ h_j(\hat{x}) = 0$. Put $v := \hat{x}$-x, then

$$0 = h_j(x) - h_j(\hat{x}) = A_j^T(\hat{x} - x) = \langle \nabla h_j(x), v \rangle$$

Further, when $g_i(x) = 0$, we have

$$\langle \nabla g_i(x), v \rangle = \langle \nabla g_i(x), \hat{x} - x \rangle \leq g_i(\hat{x}) - g_i(x) = g_i(\hat{x}) < 0$$

Thus, applying the Mangasarian Fromovitz condition, the proof is finished. □

— End of Lecture 17 —