THE CHINESE UNIVERSITY OF HONG KONG
Department of Mathematics
MATH4230 2024-25
Lecture 5
January 21, 2025 (Tuesday)

1 Recall

In the previous lecture, we introduce a problem (P) and the *feasible* set K as follows:

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

where $f, g_i, h_j \in C^1$, and

$$K = \{x \in \mathbb{R}^n : g_i(x) \le 0,\ h_j(x) = 0,\ i = 1,\dots,\ell,\ j = 1,\dots,m\}$$

We have the following theorems related to KKT condition and proved in the lecture 2.

Theorem 1. *Assume that $x^* \in K$ is an optimal solution to (P), then there exists $p_0 \ge 0,\ p_1,\dots,p_\ell \ge 0,\ q_1,\dots,q_m \in \mathbb{R}$ such that the following holds:*

$$\begin{cases} (p_0, p_1, \dots, p_\ell, q_1, \dots, q_m) \ne \mathbf{0} \\ \displaystyle\sum_{i=1}^{\ell} p_i g_i(x^*) = 0 \iff p_i g_i(x^*) = 0,\ \forall i = 1,2,\dots,\ell \\ \displaystyle p_0 \nabla f(x^*) + \sum_{i=1}^{\ell} p_i \nabla g_i(x^*) + \sum_{j=1}^{m} q_j \nabla h_j(x^*) = \mathbf{0} \end{cases}$$

Also, we discuss the **Mangasarian Fromovitz Qualification condition** last week:

① the family of vectors $(\nabla h_1(x), \dots, \nabla h_m(x))$ is linearly independent.

② there exists a vector $v \in \mathbb{R}^n$ satisfying

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

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 $x \in K$.

Therefore, together with optimal solution $\boldsymbol{x^*}$ and the **qualification**, we have the following:

Let $x^* \in K$ be a solution to (P) and assume that K is **qualified** at x^*. Then there exists $\lambda_1, \cdots, \lambda_\ell \ge 0$ and $\mu_1, \cdots, \mu_m \in \mathbb{R}$ such that

$$\begin{cases} \displaystyle\sum_{i=1}^{\ell} \lambda_i g_i(x^*) = 0 \\ \displaystyle\nabla f(x^*) + \sum_{i=1}^{\ell} \lambda_i \nabla g_i(x^*) + \sum_{j=1}^{m} \mu_j \nabla h_j(x^*) = \mathbf{0} \end{cases}$$

Remember that we still not yet finished the proof mentioned in the claim 2 of the lecture 2. We will discuss the proof for the claim now. Let us simply recall the setting.

Define

$$f_N(x) := f(x) + \|x - x^*\|^2 + \frac{N}{2}\left(\sum_{i=1}^{\ell} \underbrace{\max(0, g_i(x))^2}_{g_i^+(x)} + \sum_{j=1}^{m} h_j^2(x)\right)$$

Then, by computation, we can check that

- $f_N(x^*) = f(x^*)$
- $f_N(x) \geq f(x), \ \forall x \neq x^*$

Next, we claim that:

> There exists $\varepsilon_0 > 0$ such that for all $0 < \varepsilon < \varepsilon_0$, there exists $N_\varepsilon \in \mathbb{N}$ satisfying $f_{N_\varepsilon}(x) > f_{N_\varepsilon}(x^*) = f(x^*)$ for all $x \in \{z : \|z - x^*\| = \varepsilon\}$.

Proof. We prove by contradiction. If the claim is not true, then there exists $\varepsilon > 0$ such that for all $N > 0$, there exists some $x_N \in \{z : \|z - x^*\| = \varepsilon\}$ satisfies:

$$f_N(x_N) \leq f_N(x^*) = f(x^*) < +\infty$$

Note that $\{z : \|z - x^*\| = \varepsilon\}$ is a compact set in $\mathbb{R}$, then we choose sequence $\{x_N\}_{N \geq 1}$ in the compact set, there exists a subsequence $\{x_{N_k}\}_{k \in \mathbb{N}}$ converges to $\bar{x} \in \{z : \|z - x^*\| = \varepsilon\}$, and

$$\sup_{N_k}\left(C + \frac{N_k}{2}\left(\sum_{i=1}^{\ell}(g_i^+(x_{N_k}))^2 + \sum_{j=1}^{m} h_j^2(x_{N_k})\right)\right) \leq \sup_{N_k} f_{N_k}(x_{N_k}) \leq f(x^*) < +\infty$$

This implies that

$$\limsup_{N \to +\infty} N \cdot \ell_N(x) < +\infty \implies \lim_{N \to +\infty} \ell_N(x) = 0$$

So, we have

$$\sum_{i=1}^{\ell}(g_i^+(\bar{x}))^2 + \sum_{j=1}^{m} h_j^2(\bar{x}) = 0$$

and hence $\bar{x} \in K$. So from the above inequality, taking limit $N \to +\infty$ yields:

$$\lim_{N \to +\infty} f_N(x_N) = f(\bar{x}) + \underbrace{\|\bar{x} - x^*\|}_{\varepsilon} \leq f(x^*)$$

This is a contradiction to the fact that x^* is the minimizer, i.e.

$$f(\bar{x}) \geq f(x^*), \quad \forall \bar{x} \in K$$

□

2 Cone

Definition 1. Let $x \in K$, we define the **tangent cone** of K at point x by

$$T_K(x) := \left\{ v \in \mathbb{R}^n : \begin{array}{c} \text{there exists sequence } (s_k, v_k)_{k\geq 1} \subset \mathbb{R}_+ \times \mathbb{R}^n \text{ such that} \\ s_k \searrow 0^+,\ v_k \to 0 \text{ and } x + s_k v_k \in K,\ \forall k \geq 1 \end{array} \right\}$$

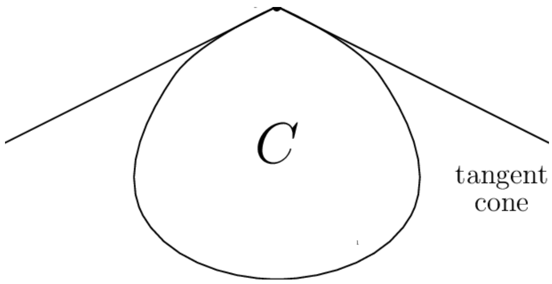

Figure 1: Tangent Cone

Example 1. In $\mathbb{R}^2$, and $K = \{x : \|x\| \leq 1\}$.

- If $x \notin \partial K$, then $T_k(x) = \mathbb{R}^2$.
- If $x \in \partial K$, then $T_k(x) = \{\text{all vectors in } \mathbb{R}^2 \text{ towards the interior of } K\}$

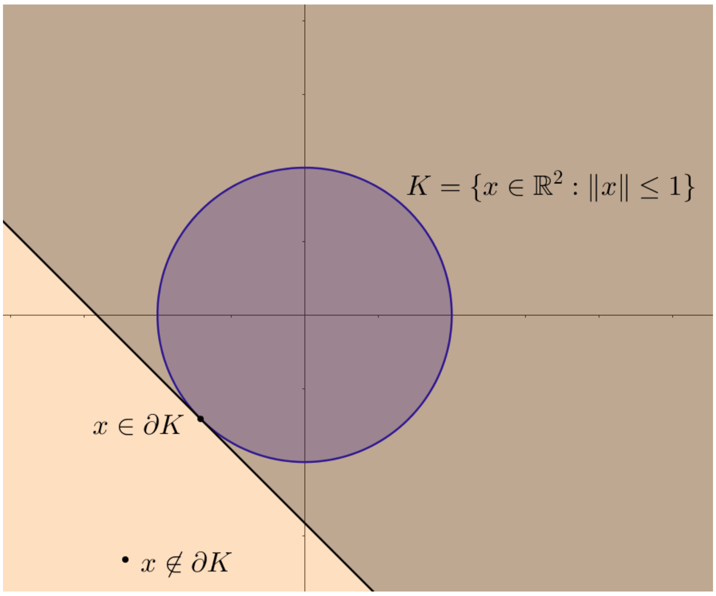

Figure 2: Example 1

After we discuss the **tangent cone** of a set K, we introduce the following lemma.

Lemma 2. *$T_K(x)$ is a closed cone.*

Proof. 1. To prove that $T_K(x)$ is a cone, it is sufficient to prove that

$$\boxed{\text{If } v \in T_K(x), \text{ then } \lambda v \in T_K(x) \text{ for all } \lambda \geq 0.}$$

- **Case 1:** $\lambda = 0$ or $v = 0$, then $\lambda v = 0 \in T_K(x)$.
- **Case 2:** $\lambda \neq 0$ and $v \neq \mathbf{0}$
  Then there exists $(s_k, v_k)_{k\geq 1} \subset \mathbb{R}_+ \times \mathbb{R}^n$ satisfies the conditions in the definition of $T_K(x)$. We define $\bar{s_k} := \dfrac{s_k}{\lambda}$, $\bar{v_k} := \lambda v_k$. Then, it follows that

$$\begin{cases} (\bar{s_k}, \bar{v_k}) \to (0^+, \lambda v) \\ x + \bar{s_k}\bar{v_k} = x + s_k v_k \in K \end{cases} \implies \lambda v \in T_K(x)$$

2. Next, it remains to prove $T_K(x)$ is closed. We take $(v_n)_{n\geq 1} \subset T_K(x)$, $v_n \to v$. Then there exists $(s_{n,k_n}, v_{n,k_n})_{n\geq 1}$ such that $s_{n,k_n} < \dfrac{1}{n}$ and $\|v_{n,k_n} - v_n\| < \dfrac{1}{n}$ such that $x + s_{n,k_n} v_{n,k_n} \in K$. Then, together with $v_{n,k_n} \to v$ and $s_{n,k_n} \to 0$ as $n \to +\infty$, we can prove $v \in T_K(x)$.

Thus, $T_K(x)$ is closed cone and the proof is finished. □

Lemma 3. *Let $K = \{x \in \mathbb{R}^n : g_i(x) \leq 0,\ h_j(x) = 0,\ i = 1, \ldots, \ell,\ j = 1, \ldots, m\}$ and $x \in K$. Then*

$$T_K(x) \subseteq \left\{ v \in \mathbb{R}^n : \begin{array}{c} \langle \nabla h_j(x), v\rangle = 0, \quad \forall j = 1, \ldots, m, \\ \langle \nabla g_i(x), v\rangle \leq 0, \ \forall i = 1, \ldots, \ell \quad \textit{satisfying} \quad g_i(x) = 0 \end{array} \right\} = J$$

Proof. Let $v \in T_K(x)$. Then by definition, there exists sequence $(s_k, v_k) \in \mathbb{R}_+ \times \mathbb{R}^n$ such that $s_k \searrow 0^+$, $v_k \to v$ and $x + s_k v_k \in K$.

1. For all $i = 1, \ldots, \ell$ satisfying $g_i(x) = 0$, one has $g_i(x + s_k v_k) \leq 0$. Using taylor expansion, we have

$$\underbrace{g_i(x)}_{=0} + s_k \langle \nabla g_i(x), v_k\rangle + s_k O(s_k) \leq 0$$

   Dividing $s_k > 0$ on both sides then taking limit $k \to +\infty$, we have

$$\langle \nabla g_i(x), v\rangle \leq 0$$

2. For all $j = 1, \ldots, m$, one has $h_j(x + s_k v_k) = 0$. Using taylor expansion again yields:

$$\underbrace{h_j(x)}_{=0} + s_k \langle \nabla h_j(x), v_k\rangle + s_k O(s_k) = 0$$

   Dividing $s_k > 0$ and taking limit $k \to +\infty$, we have

$$\lim_{k\to+\infty} \left(\underbrace{h_j(x)}_{=0} + \langle \nabla h_j(x), v_k\rangle + O(s_k) \right) = 0 \implies \langle \nabla h_j(x), v\rangle = 0$$

Together with 1 and 2, we prove $v \in J$. Since $v \in T_K(x)$ is arbitrary, so this proves $T_K(x) \subseteq J$. □

3 Checking Qualification

Lemma 4. ***(Farkas's lemma)*** *Let $c \in \mathbb{R}^n$, $(c_i)_{i=1,\dots,k} \subseteq \mathbb{R}^n$ and $(d_j)_{j=1,\dots,m} \subseteq \mathbb{R}^n$. Assume that for all $v \in \mathbb{R}^n$, we have*

$$\begin{cases} \langle d_j, v\rangle = 0, & j = 1,\dots,m \\ \langle c_i, v\rangle \leq 0, & i = 1,\dots,k \end{cases} \implies \langle c, v\rangle \leq 0$$

Then there exist $\lambda_i \geq 0,\ i = 1,\dots,k$ and $\mu_j \in \mathbb{R},\ j = 1,\dots,m$ such that $c = \sum_{i=1}^{k} \lambda_i c_i + \sum_{j=1}^{m} \mu_j d_j$.

Remarks. Let us accept it for the moment because the proof of this lemma requires the convex duality. We may discuss this afterwards.

Based on the **Farka's lemma**, we introduce the following theorem.

Theorem 5. *Let $x^* \in K$ be a solution to (P). Assume that*

$$T_K(x^*) = \left\{ v \in \mathbb{R}^n : \begin{array}{c} \langle \nabla g_i(x^*), v\rangle \leq 0,\ \forall i = 1,\dots,\ell \ \text{ satisfying } \ g_i(x^*) = 0 \\ \langle \nabla h_j(x^*), v\rangle = 0,\ \ \forall j = 1,\dots,m \end{array} \right\}$$

Then, there exists $\lambda_1, \cdots, \lambda_\ell \geq 0$ and $\mu_1, \cdots, \mu_m \in \mathbb{R}$ such that

$$\begin{cases} \displaystyle\sum_{i=1}^{\ell} \lambda_i g_i(x^*) = 0 \\ \displaystyle\nabla f(x^*) + \sum_{i=1}^{\ell} \lambda_i \nabla g_i(x^*) + \sum_{j=1}^{m} \mu_j \nabla h_j(x^*) = \mathbf{0} \end{cases}$$

*and this is called the **Abadie's qualification condition**.*

Proof. Note that $x^* \in K$ is the minimizer, so for all $v \in T_K(x^*)$, then

$$f(x^*) \leq f(x^* + s_k v_k) = f(x^*) + s_k \langle \nabla f(x^*), v_k\rangle + s_k O(s_k)$$

dividing both sides by $s_k > 0$ and taking $k \to +\infty$ yields $\langle -\nabla f(x^*), v\rangle \leq 0$. By the Farkas' lemma, there exists $\lambda_i \geq 0$ for all $i = 1,\dots,\ell$ satisfying $g_i(x^*) = 0$, and $\mu_j \in \mathbb{R}$ for $j = 1,\dots,m$ such that

$$-\nabla f(x^*) = \sum_{i : g_i(x^*) = 0} \lambda_i \nabla g_i(x^*) + \sum_{j=1}^{m} \mu_j \nabla h_j(x^*)$$

On the other hand, for those i such that $g_i(x^*) \neq 0$, we set $\lambda_i = 0$ so that

$$\lambda_i g_i(x^*) = 0,\ \ \forall i = 1,\dots,\ell$$

and

$$\nabla f(x^*) + \sum_{i=1}^{\ell} \lambda_i \nabla g_i(x^*) + \sum_{j=1}^{m} \mu_j \nabla h_j(x^*) = 0.$$

□

Remarks. Note that the Abadie's qualification condition is much general than the Mangasarian Fromovitz Qualification condition. If K satisfies the **Mangasarian-Fromovitz condition** at $x \in K$, then it satisfies **Abadie's condition**.

— End of Lecture 5 —