THE CHINESE UNIVERSITY OF HONG KONG
Department of Mathematics
MATH4230 2024-25
Lecture 3
January 14, 2025 (Tuesday)

1 Recall

From Week 1, we discuss the follows:

<u>Euler's first order condition</u>
If $f(\mathbf{x})$ is **continuously** differentiable, $\emptyset \neq K$ is an open set in $\mathbb{R}^n$ and $\mathbf{x}^* \in K$ is an optimal solution to (P), then

$$\nabla f(\mathbf{x}^*) = \mathbf{0}$$

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

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

Also, we have the following theorems related to KKT condition and proved in the last lecture.

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:*

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

2 Qualification Condition

Definition 1. We say the constraints K is **qualified** at $x \in K$ if $p_i \ge 0$ and $q_j \in \mathbb{R}$ satisfy

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

then it implies that $p_1 = \cdots = p_\ell = q_1 = \cdots = q_m = 0$.

Now, when (P) has an optimal solution $x^* \in K$ and constraints K is **qualified**, the following theorem is introduced.

Theorem 2. *Let $x^* \in K$ be a solution to (P) and assume that K is **qualified** at x^*. Then there exists $\lambda_1, \cdots, \lambda_\ell \geq 0$ and $\mu_1, \cdots, \mu_m \in \mathbb{R}$ such that*

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

Proof. From Theorem 1, if $x^* \in K$ is a solution to (P) and K is **qualified** at x^*, there exist

$$(p_0, p_1, \ldots, p_\ell, q_1, \ldots, q_m) \neq \mathbf{0}$$

such that

$$\begin{cases} \sum_{i=1}^{\ell} p_i g_i(x^*) = 0 \\ p_0 \nabla f(x^*) + \sum_{i=1}^{\ell} p_i \nabla g_i(x^*) + \sum_{j=1}^{m} q_j \nabla h_j(x^*) = \mathbf{0} \end{cases} \qquad (*)$$

Now, if $p_0 = 0$, then $(*)$ becomes

$$\begin{cases} \sum p_i g_i(x^*) = 0 \\ \sum p_i \nabla g_i(x^*) + \sum q_j \nabla h_j(x^*) = \mathbf{0} \end{cases}$$

By the Qualification condition, this implies that $p_1 = \cdots = p_\ell = q_1 = \cdots = q_m = 0$ and hence

$$(p_0, p_1, \ldots, p_\ell, q_1, \ldots, q_m) = \mathbf{0}$$

Contradiction arises! Thus, we have $p_0 > 0$.
Dividing the second equality of $(*)$ by p_0 gives

$$\nabla f(x^*) + \sum_{i=1}^{\ell} \frac{p_i}{p_0} \nabla g_i(x^*) + \sum_{j=1}^{m} \frac{q_j}{p_0} \nabla h_j(x^*) = \mathbf{0}$$

Thus, it is natural to put $\lambda_i = \frac{p_i}{p_0} \geq 0$ for $i = 1, \ldots, \ell$ and $\mu_j = \frac{q_j}{p_0} \in \mathbb{R}$ for $j = 1, \ldots, m$. □

Let's see a simple example for the importance of qualification condition.

Example 1. Solve the following problem

$$\min_{x\in\mathbb{R}} x, \quad \text{subject to} \quad x^2 = 0$$

Solution. As there is one variable problem, so putting $n = 1$, $f(x) = x$. Since there is no inequality constraint, so $\ell = 0$. There is one equality constraint, so $m = 1$, and let $h_1(x) = x^2$.
It is easy to see that $x^* = 0$ is an optimal solution to the problem as $\{x \in \mathbb{R} : x^2 = 0\} = \{0\}$.
From Theorem 1, there exists $(p_0, q_1) \neq \mathbf{0}$ such that

$$p_0 \underbrace{\nabla f(x^*)}_{f'(0)=1} + q_1 \underbrace{\nabla h_1(x^*)}_{h_1'(0)=0} = 0$$

Thus, $p_0 = 0$, which means the qualification condition of the Theorem 2 is not satisfied. ◀

3 Formal Justification of Theorem 2

For the expansion of Theorem 2, it looks quite similar as "Lagrangian". Indeed, we can rewrite the problem $\min\limits_{\substack{g(x)\leq 0\\h(x)=0}} f(x)$ as follows:

$$\begin{aligned}\min_{g(x)\leq 0, h(x)=0} f(x) &= \min_{x\in\mathbb{R}^k}\left(\sup_{\lambda\geq 0,\ \mu\in\mathbb{R}} f(x) + \lambda\cdot g(x) + \mu\cdot h(x)\right)\\ &\overset{(?)}{=} \sup_{\lambda\geq 0, \mu\in\mathbb{R}}\left(\min_{x\in\mathbb{R}} f(x) + \lambda\cdot g(x) + \mu\cdot h(x)\right)\\ &= \min_{x\in\mathbb{R}} f(x) + \lambda^* g(x) + \mu^* h(x)\end{aligned}$$

Denote $L(x, \lambda, \mu) = f(x) + \lambda\cdot g(x) + \mu\cdot h(x)$. If the feasible set K is qualified at x^*, then Theorem 2 implies $\nabla L(x^*, \lambda^*, \mu^*) = 0$, and such λ^*, μ^* are called the **Language Multiplier**.

Now, let us do some exercises together on minimization problems subject to different constraints.

4 Exercises

Exercise 1. Solve the following problem

$$\min_{x^2+y^2=1} (2x + y)$$

Solution. As there are two variables, so putting $n = 2$, and let $f(x) = 2x + y$.
Moreover, there is no inequality constraint, so $\ell = 0$.
There is one equality constraint, so $m = 1$ and let $h(x, y) = x^2 + y^2 - 1$.
If $(x^*, y^*) \in K$ is an optimal solution (let us assume that K is qualified at (x^*, y^*) without checking), then by Theorem 2, there exists $\mu_1 \in \mathbb{R}$ such that

$$\nabla f(x^*, y^*) + \mu_1 \nabla h(x^*, y^*) = \mathbf{0}$$
$$\begin{pmatrix}2\\1\end{pmatrix} + \mu_1\begin{pmatrix}2x^*\\2y^*\end{pmatrix} = \begin{pmatrix}0\\0\end{pmatrix}$$

Therefore, we have

$$\begin{cases}1 + \mu_1 x^* = 0\\ 1 + 2\mu_1 y^* = 0\\ (x^*)^2 + (y^*)^2 = 1\end{cases}$$

Note that $\mu_1 \neq 0$, and from the first two equations, we have

$$x^* = -\frac{1}{\mu_1} \quad \text{and} \quad y^* = -\frac{1}{2\mu_1}$$

Putting into the third equation yields:

$$\begin{aligned}\left(-\frac{1}{\mu_1}\right)^2 + \left(-\frac{1}{2\mu_1}\right)^2 &= 1\\ \mu_1^2 &= \frac{5}{4}\\ \mu_1 &= \pm\frac{\sqrt{5}}{2}\end{aligned}$$

So, we have $(x^*, y^*) = \left(\frac{2}{\sqrt{5}}, \frac{1}{\sqrt{5}}\right)$ or $\left(-\frac{2}{\sqrt{5}}, -\frac{1}{\sqrt{5}}\right)$.
It remains to compare the values of $f\left(\frac{2}{\sqrt{5}}, \frac{1}{\sqrt{5}}\right)$ and $f\left(-\frac{2}{\sqrt{5}}, -\frac{1}{\sqrt{5}}\right)$ to get the minimum.
By direct computation, we find that $(x^*, y^*) = \left(-\frac{2}{\sqrt{5}}, -\frac{1}{\sqrt{5}}\right)$ and $f(x^*, y^*) = -\sqrt{5}$. ◀

Exercise 2. Solve the problem $\min\limits_{x^2+y^2\leq 1} x \cdot y$.

Solution. Letting $f(x, y) = xy$, $\ell = 1,\ m = 0$ and $g(x) = x^2 + y^2 - 1$.
Assume the feasible set $K := \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq 1\}$ is *qualified*, then by Theorem 2, there exists $\lambda \geq 0$ such that

$$\begin{cases}\lambda \cdot g(x, y) = 0\\ \nabla f(x, y) + \lambda g(x, y) = \mathbf{0}\end{cases} \implies \begin{cases}\lambda(x^2 + y^2 - 1) = 0\\ \begin{pmatrix} y\\ x\end{pmatrix} + \lambda\begin{pmatrix} 2x\\ 2y\end{pmatrix} = \begin{pmatrix} 0\\ 0\end{pmatrix}\end{cases}$$

Now, we separate it into different cases:

- **Case 1:** $\lambda = 0$
  In the second equation, this leads to $(x, y) = (0, 0)$.
- **Case 2:** $\lambda > 0$
  Then, we have

  $$\begin{cases}x^2 + y^2 = 1\\ y + 2\lambda x = 0\\ x + 2\lambda y = 0\end{cases}$$

  Combining the second and the third equations gives $y + 2\lambda(-2\lambda y) = 0 \implies (1 - 4\lambda^2)y = 0$.
  Hence, we have $\lambda = +\frac{1}{2}$ and $y \neq 0$ (*think why?*). As $\lambda = \frac{1}{2}$, this would follows that $x = -y$.
  Putting back to $x^2 + y^2 = 1$, we have $x = \pm\frac{1}{\sqrt{2}}$ and so $y = \mp\frac{1}{\sqrt{2}}$.
  So, we now have $(x, y) = (0, 0), (1/\sqrt{2}, -1/\sqrt{2})$ and $(-1/\sqrt{2}, 1/\sqrt{2})$.

By comparing values on those points to evaluate $x \cdot y$, both $(1/\sqrt{2}, -1/\sqrt{2})$ and $(-1/\sqrt{2}, 1/\sqrt{2})$ are the optimal solutions, and $\min\limits_{x^2+y^2\leq 1} x \cdot y = -\frac{1}{2}$. ◀

Exercise 3. Solve the following problem

$$\min_{\substack{x^2+y^2=1\\ y^2+z^2=4}} (x+z)$$

Solution. As there are three variables, so putting $n = 3$, and $f(x, y, z) = x + z$. Since there is no inequality constraint but 2 equality constraints, so we have $\ell = 0,\ m = 2$ and letting

$$h_1(x, y, z) = x^2 + y^2 - 1$$
$$h_2(x, y, z) = y^2 + z^2 - 4$$

By Theorem 2 (without checking qualification condition), there exists $\mu_1, \mu_2 \in \mathbb{R}$ such that

$$\begin{cases} \nabla f(x,y,z) + \mu_1 \nabla h_1(x,y,z) + \mu_2 \nabla h_2(x,y,z) = \mathbf{0} \\ x^2 + y^2 = 1 \\ y^2 + z^2 = 4 \end{cases}$$

$$\implies \begin{cases} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} + \mu_1 \begin{pmatrix} 2x \\ 2y \\ 0 \end{pmatrix} + \mu_2 \begin{pmatrix} 0 \\ 2y \\ 2z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \\ x^2 + y^2 = 1 \\ y^2 + z^2 = 4 \end{cases}$$

$$\implies \begin{cases} 1 + 2\mu_1 x = 0 \\ 2(\mu_1 + \mu_2) y = 0 \\ 1 + 2\mu_2 z = 0 \\ x^2 + y^2 = 1 \\ y^2 + z^2 = 4 \end{cases}$$

Now, we consider into several cases:

- **Case 1:** $y = 0$
  Then, we have $x = \pm 1$ and $z = \pm 2$. So, the solution are $(x, y, z) = (\pm 1, 0, \pm 2)$.
- **Case 2:** $y \neq 0$
  Then, we have $\mu_1 = -\mu_2$. Plug into the first and the third equations, we have $x = -z = -\dfrac{1}{2\mu_1}$.
  In this case, there exists $\mu_1, \mu_2 \in \mathbb{R} \setminus \{0\}$ such that $x + z \equiv 0$.

By comparing $f(\pm 1, 0, \pm 2)$ and 0, simple calculation gives the optimal solution $(x^*, y^*, z^*) = (-1, 0, -2)$ and $\displaystyle\min_{\substack{x^2+y^2=1\\ y^2+z^2=4}} (x+z) = -3$. ◀

— End of Lecture 3 —