7.2. Special case of 7.3 (read yourself)

7.3. Necessary condition for a minimum.

7.3.1. Only inequality constraints.

Let $X \subset \mathbb{R}^n$ be an open set (think for now $X = \mathbb{R}^n$ to simplify) and $g_k \in C^1(X)$, $k = 1, 2, \ldots, m$.

$S = \{ x \in X \mid g_1(x) \leq 0, g_2(x) \leq 0, \ldots, g_m(x) \leq 0 \}$

The problem: $\min_{x \in S} f(x)$

We will see how the necessary condition for min from Lemma looks like for this particular $S$.

Def. $g_k \overset{\text{def}}{=} \text{active at } a \in S \text{ if } g_k(a) = 0$.

The necessary condition from Lemma:

$a$-loc. min. $\Rightarrow \nabla f(a)^T d > 0$, $\forall$ feasible $d$

More feasible directions $d \Rightarrow$

$\Rightarrow$ more information about $\nabla f(a)$.

Ex: $a) \rightarrow f)$: less $d$

---

(a) $\nabla f = 0$

(b) $\nabla f$

(c) $\nabla f$

(d) $\nabla f$

(e) $\nabla f$

(f) $\emptyset$ no information on $\nabla f$
a) No active $g_k$:
\[
\nabla f = 0 \quad \text{ (at a)}
\]

b) $g_1$ is active: $\nabla f \uparrow \nabla g_1$ \iff
\[
\begin{align*}
\nabla f + u_1 \nabla g_1 &= 0 \\
u_1 &> 0 \\
g_1 &= 0
\end{align*}
\]

c) $g_1$ and $g_2$ are active:
\[
\nabla f \in \text{cone} \{ \nabla g_1, \nabla g_2 \} \quad \text{no info on } \nabla f
\]

d) $g_2$ is active: $\nabla f \uparrow \nabla g_2$ \iff
\[
\begin{align*}
\nabla f + u_1 \nabla g_1 + u_2 \nabla g_2 &= 0 \\
u_1 &> 0, \; u_2 &> 0 \\
g_1 &= 0, \; g_2 &= 0
\end{align*}
\]

To avoid this possibility, consider the Constraint Qualification (CQ) condition at $a$:

- for active $g_k$ the gradients $\nabla g_k$ are positively linearly independent:
\[
\sum_{\text{active } g_k} \lambda_k \nabla g_k(a) = 0, \; \lambda_k > 0 \implies \text{all } \lambda_k = 0
\]

- active $g_k \iff g_k(a) = 0$

- $\sum_{\text{active } g_k} \lambda_k \nabla g_k = \sum_{k=1}^{\infty} \lambda_k \nabla g_k$

  if $\lambda_k = 0$ for $k$ with $g_k(a) \neq 0$ \iff $\lambda_k g_k(a) = 0$. 

Look first at "bad" cases e), f):

We have there $\nabla g_1 \uparrow \nabla g_2$ \iff
\[
\exists \lambda_1 > 0, \lambda_2 > 0, \text{ not all } \lambda_k = 0:
\]
\[
\lambda_1 \nabla g_1 + \lambda_2 \nabla g_2 = 0.
\]
Thus CQ condition at $a$ becomes

$$\begin{align*}
\sum \lambda_k \nabla g_k(a) &= 0 & \text{(CQ)} \\
\lambda_k &> 0, \, \forall k=1, \ldots, m \\
\lambda_k g_k(a) &= 0, \, \forall k=1, \ldots, m \\
g_k(a) &\leq 0
\end{align*}$$

\[ \Rightarrow \text{all } \lambda_k = 0 \]

Now look at "good" cases a)-d) $\Rightarrow$

$$\Rightarrow -\nabla f \in \text{cone active } g_k$$

$\exists u_1, u_2, \ldots, u_m$ such that

$$\begin{align*}
\nabla f(a) + \sum_{k=1}^m u_k \nabla g_k(a) &= 0 \\
u_k &> 0, \, k=1, \ldots, m \\
u_k g_k(a) &= 0, \, k=1, \ldots, m \\
g_k(a) &\leq 0
\end{align*}$$

\[ \Rightarrow \text{KKT} \]

It's called Karush-Kuhn-Tucker condition (or KKT) at $a$.

*) same trick with adding zeros as above.

\[ \text{a def KKT point if KKT holds.} \]

\[ \text{a def CQ point if CQ does not hold.} \]

\[ \text{(Th) (\approx Th. 3, p. 248)} \]

Let $a$ be a local min for $f$ in $S = \{ x \in X \mid g_k(x) \leq 0 \}$ and $f, g_k \in C^1(X)$.

Then $a$ is CQ point or KKT point.

Remark: It is a generalization of $\nabla f = 0$.

Proof: We need first a lemma.

Lemma: ("almost Farkas")

Consider: $()$ $Ax < 0$ and $()$ $A^T y = 0, y \geq 0$.

Then $\exists$ solution to $(*)$ \[ \iff \]

$\not\exists$ non-trivial solution to $(***)$. 
Proof of the Lemma:

\[ \exists x: Ax < 0 \iff \exists (x, w): \begin{cases} 
    A x \leq \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \\ 1 \end{bmatrix} \begin{bmatrix} w \\ 0 \\ \vdots \\ 0 \end{bmatrix} \\
    w > 0 
\]
Remark: CQ/KKT necessary condition is used similar to $\nabla f = 0$, i.e.

$$\min \Rightarrow CQ/KKT,$$ but $\nRightarrow$

Existence of $\min$ is important!

How to use the necessary condition:

1. Prove that $\exists \min$ (often Weierstrass).
2. Find all CQ points, i.e. exceptional points that do not satisfy CQ condition.
3. Find all KKT points, i.e. exceptional points that satisfy KKT condition.
4. Compare the functional values for all candidates and find the smallest one.

Example: \[
\min (8x_1x_2 + 7x_3) \quad \text{s.t.} \begin{align*}
x_1^2 + x_2^2 + x_3^3 &\leq 2 \\
x_3 &\geq 0
\end{align*}
\]

Define: $f(x) = 8x_1x_2 + 7x_3$,

$g_1(x) = x_1^2 + x_2^2 + x_3^3 - 2$, $g_2(x) = -x_3$.

1. $f \in C(\mathbb{R}^3)$
   
   \(0 \leq x_3 \leq 2, \quad x_1^2 + x_2^2 \leq 2 \Rightarrow S \text{ bounded}\)

   Only nonstrict inequality $\Rightarrow S$ closed
   \(\Rightarrow S \text{ compact} \Rightarrow \exists \min \text{ by Weierstrass' th.}\)

2. CQ points:
   
   1) Only $g_1$ is active: $x_1^2 + x_2^2 + x_3^3 = 2$.

   \(\lambda_1 \nabla g_1 = \lambda_1 \begin{bmatrix} 2x_1 \\ 2x_2 \\ 3x_3^2 \end{bmatrix} = 0 \Rightarrow x = (0,0,0) \text{ or } x = (0,0,0) \text{ not possible.}\)

   $\lambda_1 = 0$.

   CQ condition holds $\Rightarrow$ no CQ points.
2) Only $g_2$ is active: $x_3 = 0$.

$$\lambda_2 \nabla g_2 = \lambda_2 \begin{bmatrix} 0 \\ -1 \end{bmatrix} = 0 \Rightarrow \lambda_2 = 0. \text{ (No CQ points.)}$$

3) Both $g_1$ & $g_2$ are active:

$$x_3 = 0, \quad x_1^2 + x_2^2 = 2.$$  \hspace{1cm} \text{not possible}

$$\lambda_1 \begin{bmatrix} 2x_1 \\ 2x_2 \\ 0 \end{bmatrix} + \lambda_2 \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix} = 0 \Rightarrow \begin{cases} \lambda_1 = 0 \text{ or } x = (0, 0, *) \\ \lambda_2 = 0 \end{cases}$$

No CQ points for all cases.

KKT points:

$$\begin{cases}
8x_2 + u_1 2x_1 = 0 \\
8x_1 + u_1 2x_2 = 0 \\
7 + u_1 3x_3 - u_2 = 0 \\
u_1 (x_1^2 + x_2^2 + x_3^2 - 2) = 0 \\
u_2 (-x_3) = 0 \\
\text{all } u_k \geq 0 \\
x_1^2 + x_2^2 + x_3^3 \leq 2, \quad x_3 \geq 0
\end{cases}$$

4) $f(0,0,0) = 0$, $f(1,-1,0) = f(-1,1,0) = -8 \leq \min$