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 $g_k \in C^1(X)$, $k = 1, 2, \dots, m$.

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

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

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

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

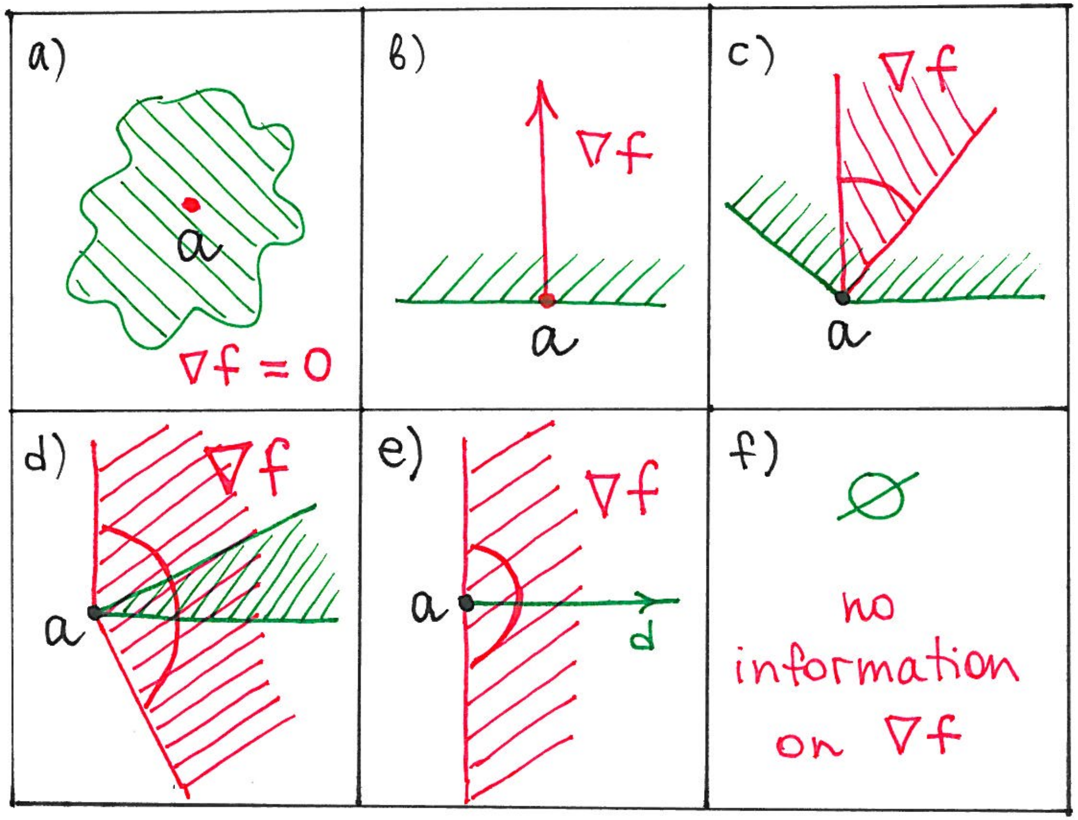
The necessary condition from Lemma:

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

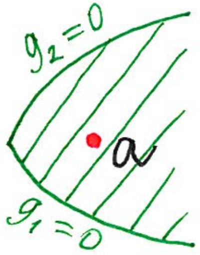
more feasible directions d $\Rightarrow$

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

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



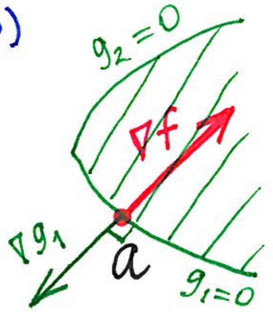
a)



<u>No active g_k:</u>

$$\boxed{\nabla f = 0} \quad \text{(at a)}$$

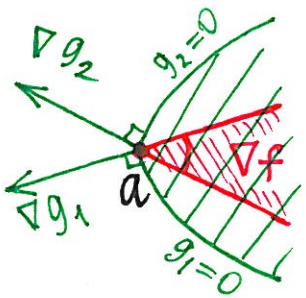
b)



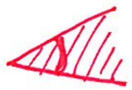
<u>g_1 is active</u>: $\nabla f \uparrow\downarrow \nabla g_1 \iff$

$$\iff \boxed{\begin{cases} \nabla f + u_1 \nabla g_1 = 0 \\ u_1 \geq 0 \\ g_1 = 0 \end{cases}}$$

c)–d)



<u>g_1 and g_2 are active</u>:

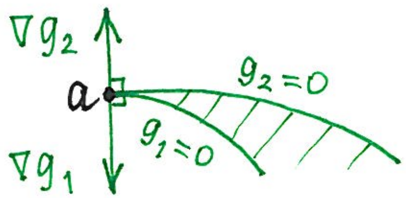
$\nabla f \in$  $\iff$

$$\iff \boxed{\begin{cases} \nabla f + u_1 \nabla g_1 + u_2 \nabla g_2 = 0 \\ u_1 \geq 0, \quad u_2 \geq 0 \\ g_1 = 0, \quad g_2 = 0 \end{cases}}$$

e)

$-\nabla f \notin \text{cone}\{\nabla g_1, \nabla g_2\}$

f)



no info on ∇f

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

We have there $\nabla g_1 \uparrow\downarrow \nabla g_2 \iff$

$\iff \exists\, \lambda_1 \geq 0,\ \lambda_2 \geq 0$, not all $\lambda_k = 0$:

$$\lambda_1 \nabla g_1 + \lambda_2 \nabla g_2 = 0.$$

To avoid this possibility, consider

<u>CQ condition at a:</u> (CQ = Constraint Qualification)

for active g_k the gradients ∇g_k are positively linearly independent:

$$\boxed{\sum_{\substack{\text{active}\\ g_k}} \lambda_k \nabla g_k(a) = 0,\ \lambda_k \geq 0 \implies \text{all } \lambda_k = 0 \qquad \underline{\underline{\text{CQ}}}}$$

- active $g_k \iff g_k(a) = 0$
- $\displaystyle\sum_{\substack{\text{active}\\ g_k}} \lambda_k \nabla g_k = \sum_{k=1}^{m} \lambda_k \nabla g_k$

$\boxed{\text{if } \lambda_k = 0 \text{ for } k \text{ with } g_k(a) \neq 0} \iff \underline{\underline{\lambda_k g_k(a) = 0.}}$

Thus CQ condition at a becomes

$$\left\{\begin{array}{l} \sum \lambda_k \nabla g_k(a) = 0 \\ \lambda_k \geq 0, \ \forall k = 1,..,m \\ \lambda_k g_k(a) = 0, \ \forall k = 1,\dots,m \\ g_k(a) \leq 0 \end{array}\right. \Longrightarrow \text{all } \lambda_k = 0 \qquad \text{CQ}$$

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

$$\Longrightarrow -\nabla f \in \underset{\substack{\text{active}\\ g_k}}{\text{cone}} \{\nabla g_k\} \overset{*)}{\Longrightarrow}$$

$$\Longrightarrow \exists\, u_1, u_2, \dots, u_m \text{ such that } \left\{\begin{array}{l} \nabla f(a) + \sum_{k=1}^{m} u_k \nabla g_k(a) = 0 \\ u_k \geq 0, \ k = 1,\dots,m \\ u_k g_k(a) = 0, \ k = 1,\dots,m \\ g_k(a) \leq 0 \end{array}\right. \qquad \text{KKT}$$

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

*) same trick with adding zeros as above.

$a \overset{\text{def}}{=}$ KKT point if KKT holds.

$a \overset{\text{def}}{=}$ CQ point if CQ does not hold.

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 (**) $\left\{\begin{array}{l} A^T y = 0 \\ y \geq 0. \end{array}\right.$

Then $\exists$ solution to (*)

$$\Updownarrow$$

$\nexists$ non-trivial solution to (**).

Proof of the lemma:

$$\exists x: Ax<0 \iff \exists (x,w): \begin{cases} Ax \leqslant -\begin{bmatrix}1\\1\\ \vdots \\1\end{bmatrix} w \\ w>0 \end{cases} \iff$$

"$\Leftarrow$" trivial

"$\Rightarrow$" $Ax=b<0 \Rightarrow$ take $w=\min_k |b_k|$

$$\iff \exists (x,w): \begin{cases} \left[\begin{array}{c|c} \begin{matrix}1\\1\\ \vdots\\1\end{matrix} & A \end{array}\right] \begin{bmatrix} w\\x\end{bmatrix} \leqslant 0 \\ \begin{bmatrix}1 & 0\end{bmatrix}\begin{bmatrix} w\\x\end{bmatrix} >0 \end{cases} \iff$$

Farkas

$$\iff \nexists y: \begin{cases} \left[\begin{array}{c} 1\,1\dots 1 \\ \hline A^T \end{array}\right] y = \begin{bmatrix}1\\0\end{bmatrix} \\ y \geqslant 0 \end{cases} \iff$$

$$\iff \nexists y: \begin{cases} y_1+y_2+\dots+y_n=1 \\ A^T y = 0 \\ y\geqslant 0 \end{cases} \iff \nexists y\neq 0: \begin{cases} A^T y = 0 \\ y \geqslant 0 \end{cases}$$

can always scale to get $\sum y_k=1$

■

Suppose $g_1, g_2, \dots, g_p$ are active at a. Then

$\nabla g_k^T d<0 \Rightarrow d$ feasible $\Rightarrow \nabla f^T d \geqslant 0$. Thus
$(k=1,2,\dots,p)$ Lecture 8, p. 7

$$\nexists d: \begin{bmatrix} \nabla f^T \\ \nabla g_1^T \\ \vdots \\ \nabla g_p^T \end{bmatrix} d<0 \iff \exists u\neq 0: \begin{cases} [\nabla f \ \nabla g_1 \dots \nabla g_p] u = 0 \\ u\geqslant 0 \end{cases}.$$

By Lemma

Adding non-active constraints with zeros (as above) we get:

$\exists u_0, u_1, u_2, \dots, u_m$, <u>not all zero</u>:

$$\begin{cases} u_0 \nabla f + \sum_{k=1}^m u_k \nabla g_k = 0 \\ u_k \geqslant 0, \quad k=0,1,2,\dots,m \\ u_k g_k = 0, \quad k=1,2,\dots,m \quad \leftarrow *) \\ g_k \leqslant 0, \quad k=1,2,\dots,m \end{cases}$$

- $u_0=0 \Rightarrow$ CQ point.
- $u_0>0 \Rightarrow$ KKT point (divide by u_0).

■

---

*) $u_k g_k = 0$ – <u>complementary slackness</u>.

<u>Remark</u>: CQ/KKT necessary condition is used similar to $\nabla f = 0$, i.e.

$$\min \Rightarrow CQ/KKT, \text{ but } \not\Leftarrow$$

Existence of min is important!

<u>How to use the necessary condition:</u>

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.

Ex $\min(8x_1x_2 + 7x_3) \;\Big|\; \begin{array}{l} x_1^2 + x_2^2 + x_3^3 \leq 2 \\ x_3 \geq 0 \end{array}$

<u>Define</u>: $f(x) = 8x_1x_2 + 7x_3$,

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

1. $f \in C(\mathbb{R}^3)$

$\left.\begin{array}{l} 0 \leq x_3 \leq 2,\ x_1^2 + x_2^2 \leq 2 \Rightarrow S \text{ bounded} \\ \text{Only nonstrict inequal.} \Rightarrow S \text{ closed} \end{array}\right\} \Rightarrow$

$\Rightarrow$ compact $\Rightarrow \exists$ min by Weierstraß' th.

2. <u>CQ points</u>:

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

not possible.

$$\lambda_1 \nabla g_1 = \lambda_1 \begin{bmatrix} 2x_1 \\ 2x_2 \\ 3x_3^2 \end{bmatrix} = 0 \Rightarrow x = (0,0,0) \text{ or } \underline{\lambda_1 = 0}.$$

CQ condition holds $\Rightarrow$ no CQ points.

2) Only g_2 is active: $x_3 = 0$.

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

3) Both g_1 & g_2 are active:

$$x_3 = 0, \quad x_1^2 + x_2^2 = 2.$$

not possible

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

No CQ points for all cases.

③ <u>KKT points</u>:

$$\begin{cases} 8x_2 + u_1 2x_1 = 0 & (1) \\ 8x_1 + u_1 2x_2 = 0 & (2) \\ 7 + u_1 3x_3^2 - u_2 = 0 & (3) \\ u_1(x_1^2 + x_2^2 + x_3^3 - 2) = 0 & (4) \\ u_2(-x_3) = 0 & (5) \\ \text{all } u_k \geq 0 & (6) \\ x_1^2 + x_2^2 + x_3^3 \leq 2, \quad x_3 \geq 0 & (7)-(8) \end{cases}$$

1) $u_2 = 0 \overset{(3)}{\Rightarrow} 7 + \text{"}\geq 0\text{"} = 0 \Rightarrow$ impossible!

2) $u_2 > 0 \overset{(5)}{\Rightarrow} x_3 = 0 \overset{(3)}{\Rightarrow} u_2 = 7$: (6) ok

- $u_1 = 0 \overset{(1,2)}{\Rightarrow} x_1 = x_2 = 0 \overset{(7)}{\Rightarrow} x = (0,0,0)$ KKT

- $u_1 > 0 \overset{(4)}{\Rightarrow} x_1^2 + x_2^2 = 2$ (NB: $x_3 = 0$)

$$(1,2) \Rightarrow \begin{bmatrix} u_1 & 4 \\ 4 & u_1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 0 \text{ and } (x_1, x_2) \neq 0 \Rightarrow$$

$$\Rightarrow \det \begin{bmatrix} u_1 & 4 \\ 4 & u_1 \end{bmatrix} = 0 \Rightarrow u_1^2 - 16 = 0 \Rightarrow u_1 = \pm 4.$$

But $u_1 = -4$ contradicts (6) $\Rightarrow u_1 = 4. \Rightarrow$

$$\Rightarrow \begin{bmatrix} 4 & 4 \\ 4 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 0 \Rightarrow x_1 + x_2 = 0 \Rightarrow x_2 = -x_1 \Rightarrow$$

$$\Rightarrow x_1^2 + x_1^2 = 2 \Rightarrow x_1 = \pm 1, \ x_2 = \mp 1 \Rightarrow$$

$$\Rightarrow x = \pm(1, -1, 0) \text{ KKT}$$

④ $f(0,0,0) = \boxed{0}$, $f(1,-1,0) = f(-1,1,0) = \boxed{-8} \leftarrow$ min!