1

7.3.2. Inequalities and equalities

$$S = \{ x \in X \subset \mathbb{R}^n \mid g_1(x) \le 0, \dots, g_m(x) \le 0,$$
$$h_1(x) = 0, \dots, h_\ell(x) = 0 \}.$$

Denote $g = \begin{bmatrix} g_1 \\ \vdots \\ g_m \end{bmatrix}$, $h = \begin{bmatrix} h_1 \\ \vdots \\ h_\ell \end{bmatrix}$ $\Rightarrow$

$\Rightarrow S = \{ x \in X \mid g(x) \le 0, h(x) = 0 \}$.

<u>Ex</u> $X = \mathbb{R}^2$, only one equality $h_1(x) = 0$.

Let $a \in S$ be loc. min. for $\min_{x \in S} f(x)$.

- if $\nabla h_1(a) \neq 0$ then $h_1(x_1, x_2) = 0$ can locally near a be parameterized as $x(t) = (x_1(t), x_2(t))$ with $a = x(0)$ by implicit function th. $\Rightarrow$ $\Rightarrow F(t) = f(x(t))$, $G_k(t) = g_k(x(t))$ $\Rightarrow$ $\Rightarrow$ new problem $\{\min F \mid G_k \le 0\}$ has $t = 0$ a loc. min. $\Rightarrow$ apply 7.3.1.
- if $\nabla h_1(a) = 0$ then "Bad" point.

2

<u>In general</u>: $h_1(x) = 0, \dots, h_\ell(x) = 0$.

- $\nabla h_1(a), \dots, \nabla h_\ell(a)$ are linearly independent $\Rightarrow$ $\Rightarrow \exists$ local parametrization $x = x(t)$ for $x \in \mathbb{R}^n \sim t \in \mathbb{R}^{n-\ell}$ and the necessary condition for loc. min. can be checked without equalities:

  $F(t) = f(x(t))$, $G_k(t) = g_k(x(t))$.
- $\nabla h_1(a), \dots, \nabla h_\ell(a)$ are lin. dep. $\Rightarrow$ "Bad" point.

Doing that one gets

<u>CQ condition at a</u>:

$$\boxed{\begin{cases} \sum\limits_{\substack{\text{active} \\ g_k}} \lambda_k \nabla g_k + \sum\limits_{j=1}^{\ell} \mu_j \nabla h_j = 0 \\ \lambda_k \ge 0 \end{cases} \Rightarrow \begin{array}{l} \text{all } \lambda_k = 0, \\ \text{all } \mu_j = 0. \end{array} \quad \text{CQ}}$$

In particular, all ∇h_j are <u>lin. indep.</u>

KKT condition at a:

$$
\begin{cases}
\nabla f + \sum_{k=1}^{m} u_k \nabla g_k + \sum_{j=1}^{\ell} v_j \nabla h_j = 0 \\
u_k \geq 0, \quad k = 1, \dots, m \\
u_k\, g_k = 0, \quad k = 1, \dots, m \\
g \leq 0, \quad h = 0
\end{cases}
$$

Th ($\approx$ Th. 4, p. 251)

a – local min for $\min_{x \in S} f(x) \Rightarrow$

$\Rightarrow$ a – CQ point or KKT point.

<u>Remark</u>: as usual

- a $\overset{\text{def}}{=}$ CQ point if CQ not satisfied.
- a $\overset{\text{def}}{=}$ KKT point if KKT satisfied.
- $L(x,u,v) = f(x) + u^T g(x) + v^T h(x) \Rightarrow$

$\Rightarrow$ KKT: $\underline{\underline{\nabla_x L = 0}}$.

Similar to Lecture 9, p. 8, we define a "unified" CQ/KKT point as

$\exists\, u_0, u_1, \dots, u_m, v_1, \dots, v_\ell$, not all zeros:

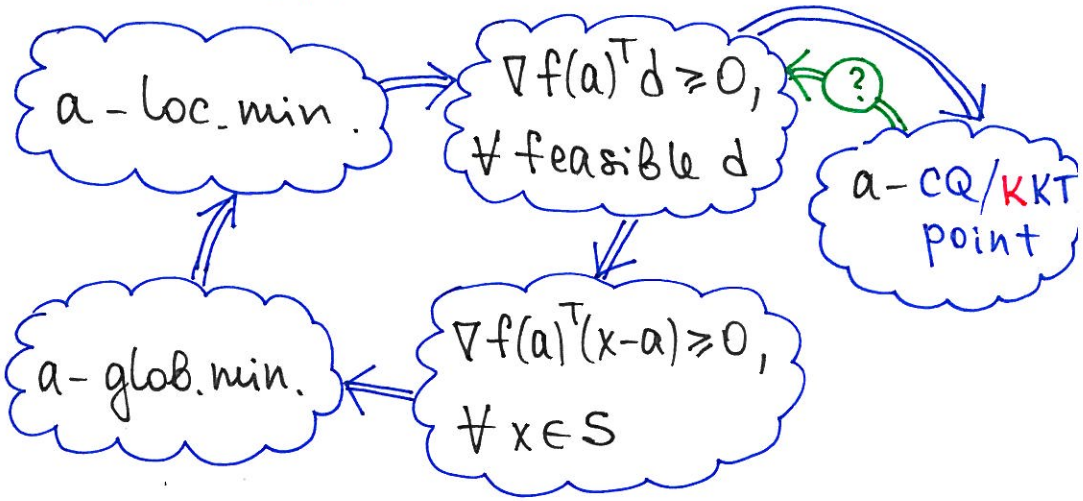
$$
\begin{cases}
u_0 \nabla f + \sum_{k=1}^{m} u_k \nabla g_k + \sum_{j=1}^{\ell} v_j \nabla h_j = 0 \\
u_k \geq 0, \quad k = 0, 1, \dots, m \\
u_k \cdot g_k = 0, \quad k = 1, \dots, m \\
g \leq 0, \quad h = 0.
\end{cases}
$$

- We have in general:

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

$\Rightarrow$ a – CQ/KKT point ("unified").

- for convex f we have:



a - loc. min. $\Rightarrow$ necessary condition

sufficient condition $\Rightarrow$ a - loc. min.

7.4. Sufficient conditions for minimum.

$$\begin{cases} \min f(x) \\ x \in S \end{cases} \qquad \begin{array}{l} X \subset \mathbb{R}^n \text{ - open}, \ f: X \to \mathbb{R} \\ S = \{x \in X \mid g(x) \leq 0, \ h(x) = 0\} \\ g: X \to \mathbb{R}^m, \ h: X \to \mathbb{R}^l. \end{array}$$

Ex. $S = X$. a - loc. min. $\Rightarrow \nabla f(a) = 0$

$a \in S$, $\nabla f(a) = 0$, $\nabla^2 f(a)$ pos. def. $\Rightarrow$ a - loc. min.

For $x \in S$ denote $I(x) \overset{def}{=} \{i \mid g_i(x) = 0\}$, i.e. indices for active constraints at x.

Th (Th. 5, p. 264)

$\bar{x} \in S$ is KKT point, $f, g_i, i \in I(\bar{x})$ are locally convex at $\bar{x}$ and h is locally affine at $\bar{x} \Rightarrow \bar{x}$ - loc. min.

Remark: locally = in some ball around the point $\bar{x}$.

Proof: $\bar{x}$ - KKT point $\Rightarrow \exists \bar{u}, \bar{v}$ from KKT condition. Define

$$x \mapsto L(x, \bar{u}, \bar{v}) = f(x) + \bar{u}^T g(x) + \bar{v}^T h(x).$$

The function is locally convex at $\bar{x}$.

KKT $\Rightarrow \nabla_x L(\bar{x}, \bar{u}, \bar{v}) = 0 \Rightarrow \bar{x}$ - stationary point $\Rightarrow$

$\Rightarrow$ loc. min. for $L(x, \bar{u}, \bar{v})$ at $\bar{x}$, i.e.

$$L(\bar{x}, \bar{u}, \bar{v}) \leq L(x, \bar{u}, \bar{v}), \ \forall x \in S \text{ near } \bar{x}.$$

$$L(\bar{x}, \bar{u}, \bar{v}) = f(\bar{x}) + \underbrace{\bar{u}^T g(\bar{x})}_{= 0 \ (CSP)} + \underbrace{\bar{v}^T h(\bar{x})}_{= 0 \text{ in } S} = f(\bar{x})$$

$$L(x, \bar{u}, \bar{v}) = f(x) + \underbrace{\overset{\geq 0}{\bar{u}^T} \overset{\leq 0}{g(x)}}_{\leq 0 \text{ in } S} + \underbrace{\bar{v}^T h(x)}_{= 0 \text{ in } S} \leq f(x)$$

!

Thus $f(\bar{x}) \leq f(x)$, $\forall x \in S$ near $\bar{x}$. ■

Remark: replace all word "locally" with "globally" $\Rightarrow$ the result holds.

Th (Th. 6, p. 266)

Let $\bar{x}$ be a KKT point that satisfies

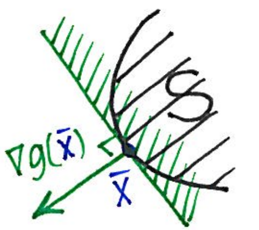
$$d^T \nabla^2_{xx} L(\bar{x},\bar{u},\bar{v}) d > 0, \ \forall d \underset{0}{\neq} : \begin{cases} \nabla g_i(\bar{x})^T d \le 0, \ \forall i \in I(\bar{x}) \\ \nabla h_j(\bar{x})^T d = 0, \ \forall j \end{cases}$$

$\Rightarrow \bar{x}$ is a strict local minimum. *)

Remark: all such $d = \emptyset \Rightarrow$ OK (th. is true).

Interpretation: $\nabla^2_{xx} L$ is positive along all "almost feasible" directions d.

Ex $S = \{x \mid g(x) \le 0\}$ – only one constraint.



"Almost feasible" d =

= feasible + limiting feasible =

$= \{ d \mid \nabla g(\bar{x})^T d \le 0 \}$.

Proof: By contradiction. Assume *) false $\Rightarrow$

$\Rightarrow \exists\, x_k \in S, \ x_k \to \bar{x} : f(x_k) \le f(\bar{x}), \ \forall k$

(1) Technical step: set $\lambda_k = \|x_k - \bar{x}\|$ and

$d_k = \dfrac{x_k - \bar{x}}{\lambda_k} \Rightarrow \|d_k\| = 1$ and

By Bolzano–Weierstraß theorem (p. 117)

$\exists$ lim. point $d \neq 0$: subsequence $d_k \to d$.

Thus $x_k = \bar{x} + \lambda_k d_k, \ d_k \to d, \ \lambda_k \searrow 0, \ x_k \to \bar{x}$.

(2) $\forall i \in I(\bar{x})$ we take the Taylor expansion

$$\underset{\overset{\wedge|}{0}}{g_i(x_k)} = \underset{\overset{||}{0}}{g_i(\bar{x})} + \nabla g_i(\bar{x})^T (x_k - \bar{x}) + \|x_k - \bar{x}\|^2 \cdot B(x_k) =$$

$$= \nabla g_i(\bar{x})^T \lambda_k d_k + \lambda_k^2 \cdot B(x_k) \Rightarrow \begin{bmatrix} \text{divide} \\ \text{by } \lambda_k > 0 \end{bmatrix}$$

$\Rightarrow 0 \ge \nabla g_i(\bar{x})^T d_k + \lambda_k \cdot B(x_k) \Rightarrow$ let $k \to \infty$

$$\Rightarrow \nabla g_i(\bar{x})^T d \le 0.$$

(3) Similarly, we get $\nabla h_j(\bar{x})^T d = 0$

(4) $\underset{\overset{\wedge|}{f(x_k) \text{ by } !}}{L(x_k,\bar{u},\bar{v})} = \underset{\overset{||}{f(\bar{x})}}{L(\bar{x},\bar{u},\bar{v})} + \underset{\overset{||}{0 \text{ by KKT}}}{\nabla_x L(\bar{x},\bar{u},\bar{v})^T} (x_k - \bar{x}) +$

$+\frac{1}{2}(x_k - \bar{x})^T \nabla^2_{xx} L(\bar{x},\bar{u},\bar{v})(x_k - \bar{x}) + \|x_k - \bar{x}\|^3 \cdot B(x_k) \Rightarrow$

$\Rightarrow \underbrace{f(x_k) - f(\bar{x})}_{\overset{\wedge|}{0} \text{ by assumption}} \ge \frac{1}{2} \lambda_k^2 d_k^T \nabla^2_{xx} L \, d_k + \lambda_k^3 \cdot B(x_k) \Rightarrow$

divide by λ_k^2

$\Rightarrow 0 \ge \frac{1}{2} d_k^T \nabla^2_{xx} L \, d_k + \lambda_k \cdot B(x_k) \Rightarrow$ let $k \to \infty$

$\Rightarrow 0 \geq d^T \nabla^2_{xx} L(\bar{x},\bar{u},\bar{v})\, d.$ ⚡ ∎

<u>Ex</u> (Ex. 9, p. 255 + Ex. 15, p. 268)

$\min (x_1^3 + x_2^2) \mid x_2 \geq 0,\ x_1^2 + x_2^2 = 9.$

KKT points are: $(\pm 3, 0), (0,3), \left(\frac{2}{3}, \frac{\sqrt{77}}{3}\right)$.

Take $\bar{x} = (-3,0)$. We need to know even the solutions $\bar{u} = 0$, $\bar{v} = \frac{9}{2}$ from KKT condition for this $\bar{x}$.

<u>NB</u>: often the triple $(\bar{x},\bar{u},\bar{v})$ is called KKT point instead of just $\bar{x}$.

$f(x) = x_1^3 + x_2^2, \quad g(x) = -x_2, \quad h(x) = x_1^2 + x_2^2 - 9.$

$L(x,u,v) = f(x) + u\,g(x) + v\,h(x) =$

$= x_1^3 + x_2^2 - u x_2 + v(x_1^2 + x_2^2 - 9) \Rightarrow$

$$\Rightarrow \nabla_x L = \begin{bmatrix} 3x_1^2 + 2vx_1 \\ 2x_2 - u + 2vx_2 \end{bmatrix} \Rightarrow \nabla^2_{xx} L = \begin{bmatrix} 6x_1 + 2v & 0 \\ 0 & 2+2v \end{bmatrix}$$

$$\Rightarrow \nabla^2_{xx} L(\bar{x},\bar{u},\bar{v}) = \begin{bmatrix} -18+9 & 0 \\ 0 & 2+9 \end{bmatrix} = \begin{bmatrix} -9 & 0 \\ 0 & 11 \end{bmatrix}.$$

NB: not positive definite!

$$\left.\begin{array}{l} \nabla g(x) = \begin{bmatrix} 0 \\ -1 \end{bmatrix} \Rightarrow \nabla g(\bar{x}) = \begin{bmatrix} 0 \\ -1 \end{bmatrix} \\ \nabla h(x) = \begin{bmatrix} 2x_1 \\ 2x_2 \end{bmatrix} \Rightarrow \nabla h(\bar{x}) = \begin{bmatrix} -6 \\ 0 \end{bmatrix} \end{array}\right\} \Rightarrow$$

$$\Rightarrow \text{admissible } \underset{\neq 0}{d}: \begin{cases} [0\ \ -1]\, d \leq 0 \\ [-6\ \ 0]\, d = 0 \end{cases} \Rightarrow$$

$$\Rightarrow \begin{cases} -d_2 \leq 0 \\ d_1 = 0 \end{cases} \Rightarrow d = \begin{bmatrix} 0 \\ t \end{bmatrix} = t \begin{bmatrix} 0 \\ 1 \end{bmatrix},\ t > 0.$$

$$d^T \nabla^2_{xx} L\, d = t^2 [0\ \ 1] \begin{bmatrix} -9 & 0 \\ 0 & 11 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = 11t^2 > 0$$

$\Rightarrow (-3,0)$ is a <u>local minimum</u>.

Now take $\bar{x} = (3,0)$ with $\bar{u} = 0$, $\bar{v} = -\frac{9}{2} \Rightarrow$

$\Rightarrow \nabla^2_{xx} L = \begin{bmatrix} 9 & 0 \\ 0 & -7 \end{bmatrix}$ and the same $d \Rightarrow$

$\Rightarrow d^T \nabla^2_{xx} L\, d = -7t^2 \not> 0 \Rightarrow$ <u>no information</u>

<u>Remark</u>: the theorem cannot say anything about global min.