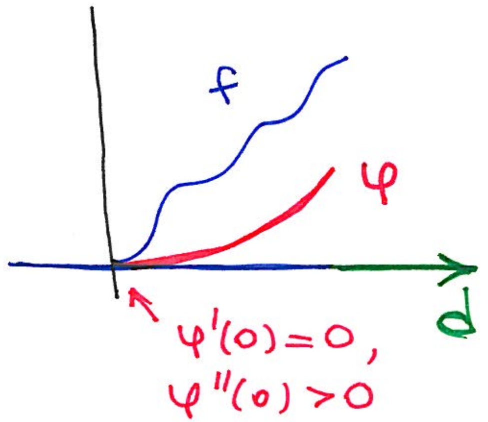
Sufficient conditions (cont'd)

$d^T \nabla^2_{xx} L(\bar{x},\bar{u},\bar{v})\, d > 0$, $\forall$ "almost feasible" $d \neq 0$.

"Almost feasible" d : $\begin{cases} \nabla g_i(\bar{x})^T d \le 0, & \forall i \in I(\bar{x}) \\ \nabla h_j(\bar{x})^T d = 0, & \forall j. \end{cases}$

Geometry of the sufficient condition:

- $\varphi(t) = L(\bar{x} + t d, \bar{u}, \bar{v})$,
- $\varphi'(0) = \underbrace{\nabla_x L(\bar{x},\bar{u},\bar{v})^T}_{=0 \text{ (KKT)}} d = 0$,
- $\varphi''(0) = d^T \nabla^2_{xx} L(\bar{x},\bar{u},\bar{v})\, d > 0$.



$\Rightarrow$ strict loc. min.

$L(\bar{x},\bar{u},\bar{v}) = f(\bar{x})$,

$L(x,\bar{u},\bar{v}) \le f(x)$!!

- It is possible to refine further the sufficient condition.

Take a particular g_k, $k \in I(\bar{x})$:

$\{\nabla g_k(\bar{x})^T d \le 0\} = \{\nabla g_k(\bar{x})^T d < 0\} \cup \{\nabla g_k(\bar{x})^T d = 0\}$

"almost feasible" = strictly feasible + tangent feasible.

KKT: $\nabla f(\bar{x}) + \sum_{i \in I} \bar{u}_i \nabla g_i(\bar{x}) + \sum_{j=1}^{\ell} \bar{v}_j \nabla h_j(\bar{x}) = 0 \Rightarrow$

$\Rightarrow \nabla f(\bar{x})^T d + \sum_{i \in I} \overset{\ge 0}{\bar{u}_i} \underbrace{\nabla g_i(\bar{x})^T d}_{\le 0} + \sum_{j=1}^{\ell} \bar{v}_j \underbrace{\nabla h_j(\bar{x})^T d}_{=0} = 0$

If $\exists k \in I : \bar{u}_k > 0$ and $\nabla g_k(\bar{x})^T d < 0$ then

$\nabla f(\bar{x})^T d > 0 \Rightarrow$ strict loc. min. along $d \Rightarrow$

$\Rightarrow$ no need to check $d^T \nabla^2_{xx} L\, d > 0$.

Define $I^+(\bar{x}) = \{ i \in I(\bar{x}) \mid \bar{u}_i > 0 \}$ and

$I^0(\bar{x}) = \{ i \in I(\bar{x}) \mid \bar{u}_i = 0 \}$.

Then the modified sufficient condition is

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

Remark: often $I^0 = \emptyset \Rightarrow$ only "=".

Ex (from Lecture 9, p. 10)

$$\min(8x_1x_2 + 7x_3) \mid x_1^2 + x_2^2 + x_3^3 \le 2, \; x_3 \ge 0.$$

KKT points: $(0,0,0)$ and $\pm(1,-1,0)$.

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

$$= 8x_1x_2 + 7x_3 + u_1(x_1^2 + x_2^2 + x_3^3 - 2) - u_2 x_3 \Rightarrow$$

$$\Rightarrow \nabla_x L = \begin{bmatrix} 8x_2 + 2u_1x_1 \\ 8x_1 + 2u_1x_2 \\ 7 + 3u_1x_3^2 - u_2 \end{bmatrix}, \quad \nabla^2_{xx} L = \begin{bmatrix} 2u_1 & 8 & 0 \\ 8 & 2u_1 & 0 \\ 0 & 0 & 6u_1x_3 \end{bmatrix}.$$

Take $\bar{x} = (1,-1,0)$ with $\bar{u}_1 = 4$, $\bar{u}_2 = 7 \Rightarrow$

$$\Rightarrow \nabla^2_{xx} L = \begin{bmatrix} 8 & 8 & 0 \\ 8 & 8 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \nabla g_1(\bar{x}) = \begin{bmatrix} 2\bar{x}_1 \\ 2\bar{x}_2 \\ 3\bar{x}_3^2 \end{bmatrix} = \begin{bmatrix} 2 \\ -2 \\ 0 \end{bmatrix}, \quad \nabla g_2(\bar{x}) = \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix}.$$

Let's try the original version of the theorem:

$$d \neq 0: \begin{cases} 2d_1 - 2d_2 \le 0 & (\text{since } g_1(\bar{x}) = 0) \\ -d_3 \le 0 & (\text{since } g_2(\bar{x}) = 0) \end{cases} \Rightarrow$$

$$\Rightarrow d \neq 0: d_1 \le d_2, \; d_3 \ge 0 \Rightarrow d^T \nabla^2_{xx} L d =$$

$$= [d_1 \; d_2 \; d_3] \begin{bmatrix} 8 & 8 & 0 \\ 8 & 8 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} d_1 \\ d_2 \\ d_3 \end{bmatrix} = 8(d_1 + d_2)^2 \not> 0, \text{ e.g.}$$

$d_1 = -1$, $d_2 = 1$, $d_3 = 0$ ger $= 0$.

The original th. gives no information.

Try now the modified version:

$$d \neq 0: \begin{cases} 2d_1 - 2d_2 = 0 & (\bar{u}_1 = 4 > 0 \Rightarrow 1 \in I^+) \\ -d_3 = 0 & (\bar{u}_2 = 7 > 0 \Rightarrow 2 \in I^+) \end{cases} \Rightarrow$$

$$\Rightarrow d_1 = d_2, \; d_3 = 0 \Rightarrow d = t(1,1,0), \; t > 0 \Rightarrow$$

$$\Rightarrow d^T \nabla^2_{xx} L d = 8t^2(1+1)^2 = 32t^2 > 0 \Rightarrow$$

$\Rightarrow (1,-1,0)$ is a local minimum.

Remark: in the constrained minimization

the Lagrange function L replaces f:

	$S = \underline{X}$ - open	$S = \{g \le 0,\ h = 0\}$
General	Loc. min. $\Rightarrow \nabla f = 0$	Loc. min. $\Rightarrow \nabla_x L = 0$ (or $\underline{CQ}$ point)
	$\nabla f = 0$, $\nabla^2 f$ pos. def. $\Rightarrow$ loc. min.	$\nabla_x L = 0$, $d^T \nabla^2_{xx} L d > 0$, $\forall$ "almost feasible" $d \neq 0$ $\Rightarrow$ loc. min.
Convex	$\nabla f = 0 \Rightarrow$ glob. min.	$\nabla_x L = 0 \Rightarrow$ glob. min.

7.5. Quadratic Programming
7.6. Some applications.

} read yourself

Ch. 8. Saddle point and duality.

$$\min_{x \in S} f(x) \qquad S = \{x \in \underline{X} \mid g(x) \le 0,\ h(x) = 0\}.$$

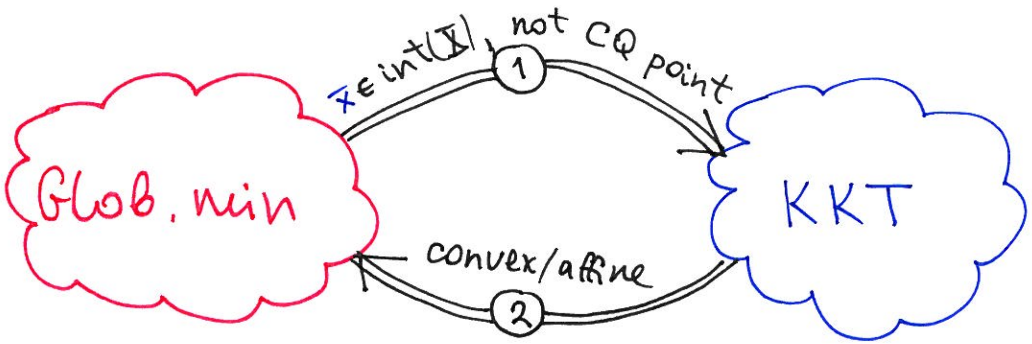
(no longer assumed to be open)

1. Th. 4, p. 251:

$\bar{x} \in \mathrm{int}(\underline{X})$ - glob. min. $\Longrightarrow \bar{x}$ - CQ/KKT point.

2. Corollary to Th. 5, p. 265:

KKT + "convex/affine" $\Longrightarrow$ glob. min.



Let us recall the proof of 2.

For $(\bar{x},\bar{u},\bar{v})$ - KKT, $u \geq 0$, $x \in S$:

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

$$\leq$$

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

$$\leq$$

$L(x,\bar{u},\bar{v})$ since $\nabla_x L(\bar{x},\bar{u},\bar{v}) = 0 \Rightarrow$ glob. min.

L convex

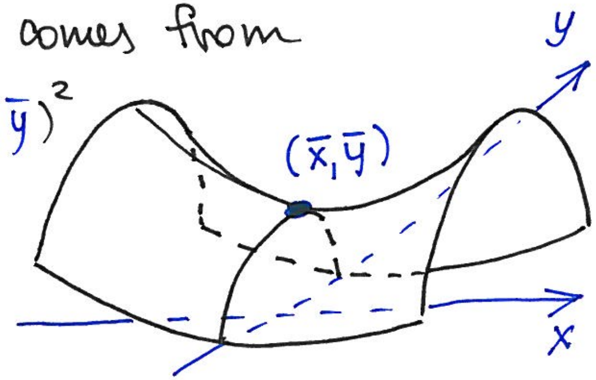
<u>Def</u> Let $\bar{x} \in \underline{X}$, $\bar{u} \geq 0$.

$(\bar{x},\bar{u},\bar{v})$ is called a saddle point of L if

$$L(\bar{x},u,v) \leq L(\bar{x},\bar{u},\bar{v}) \leq L(x,\bar{u},\bar{v}),\ \forall x \in \underline{X},\ \forall u \geq 0, v$$

<u>Remark:</u> the name comes from

$f(x,y) = (x-\bar{x})^2 - (y-\bar{y})^2$

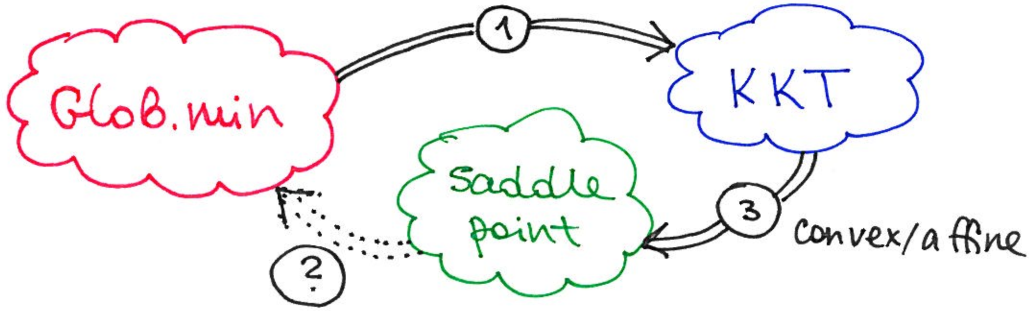


$$f(\bar{x},y) \leq \underbrace{f(\bar{x},\bar{y})}_{=0} \leq f(x,\bar{y})$$

Thus, we have got

③ Th. 2, p. 296:

KKT + "convex/affine" $\Rightarrow$ saddle point.



<u>Lemma:</u> (Lemma 1, p. 294)

$(\bar{x},\bar{u},\bar{v})$ - saddle point $\iff$

$$\iff \begin{cases} L(\bar{x},\bar{u},\bar{v}) = \min\limits_{x \in \underline{X}} L(x,\bar{u},\bar{v}) & (1) \\ \bar{u}_k \underbrace{g_k(\bar{x}) = 0}_{c)},\ \underbrace{g(\bar{x}) \leq 0}_{b)},\ \underbrace{h(\bar{x}) = 0}_{a)}. & (2) \end{cases}$$

<u>Proof:</u> $L(\bar{x},u,v) \underset{\mathrm{I}}{\leq} L(\bar{x},\bar{u},\bar{v}) \underset{\mathrm{II}}{\leq} L(x,\bar{u},\bar{v})$

$\mathrm{II} \underset{\text{trivial}}{\iff} (1)$. Prove now that $\mathrm{I} \iff (2)$.

$\boxed{\text{I} \Leftarrow (2)}$ Easy:

$$L(\bar{x},u,v) = f(\bar{x}) + \underbrace{u^T g(\bar{x})}_{\le 0 \ \ 2b)} + v^T \underbrace{h(\bar{x})}_{= 0 \ \ 2a)} \le f(\bar{x}) =$$

$$= f(\bar{x}) + \underbrace{\bar{u}^T g(\bar{x})}_{= 0 \ \ 2c)} + \bar{v}^T \underbrace{h(\bar{x})}_{= 0 \ \ 2a)} = L(\bar{x},\bar{u},\bar{v}), \ \forall u \ge 0, v.$$

$\boxed{\text{I} \Rightarrow (2)}$ $L(\bar{x},u,v) - L(\bar{x},\bar{u},\bar{v}) \le 0 \iff$

$$\iff (u-\bar{u})^T g(\bar{x}) + (v-\bar{v})^T h(\bar{x}) \le 0, \ \forall u \ge 0, v.$$

a) Take $u = \bar{u}$, $v = \bar{v} + h(\bar{x}) \Rightarrow \|h(\bar{x})\|^2 \le 0 \Rightarrow$

$\Rightarrow h(\bar{x}) = 0$.

Therefore, $(u-\bar{u})^T g(\bar{x}) \le 0, \ \forall u \ge 0$.

b) If $g_k(\bar{x}) > 0$ then take

$$u = \bar{u} + (0,0,\dots,0,\underset{>0}{\overset{k}{g_k(\bar{x})}},0,\dots,0) \ge 0 \Rightarrow$$

$$\Rightarrow (g_k(\bar{x}))^2 \le 0 \Rightarrow g_k(\bar{x}) = 0. \ ↯$$

The contradiction proves $g(\bar{x}) \le 0$.

c) Take $u = 0 \Rightarrow \bar{u}^T g(\bar{x}) \ge 0$, but

$\bar{u} \ge 0$, $g(\bar{x}) \le 0 \Rightarrow \bar{u}^T g(\bar{x}) \le 0 \Rightarrow$

$\Rightarrow \bar{u}^T g(\bar{x}) = 0$. ■

(4) Th. 1, p. 296:

Saddle point $\Rightarrow$ global minimum.

<u>Proof</u>: trivial from Lemma above:

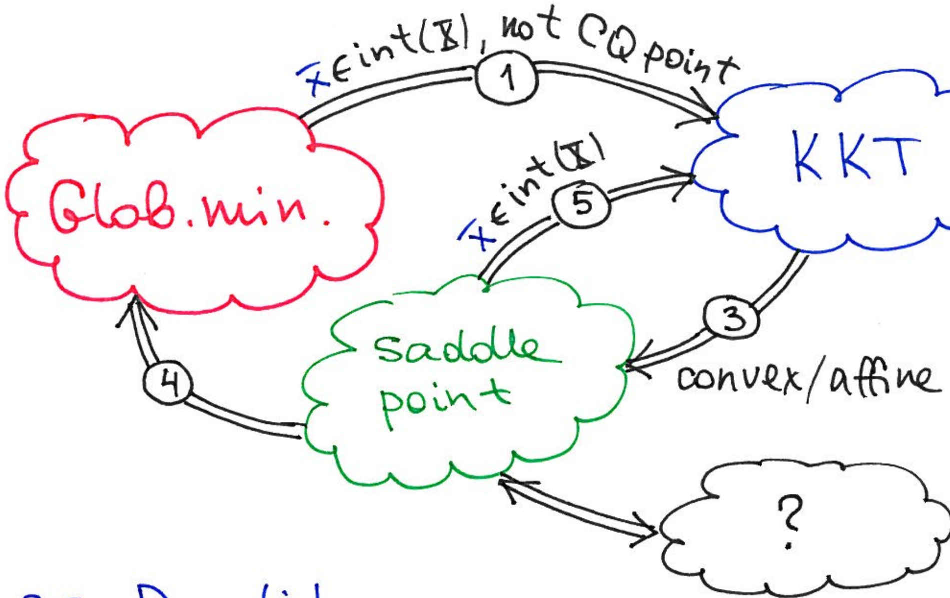
$$f(\bar{x}) = L(\bar{x},\bar{u},\bar{v}) \le L(x,\bar{u},\bar{v}) \le f(x), \ \forall x \in S. \ ■$$

(5) Th. 1, p. 296:

Saddle point $+ \ \bar{x} \in \text{int}(X) \Rightarrow$ KKT.

<u>Proof</u>: trivial from Lemma above

$$\left.\begin{aligned} &L(\bar{x},\bar{u},\bar{v}) = \min_{x \in X} L(x,\bar{u},\bar{v}) \\ &\bar{x} \in \text{int}(X) \end{aligned}\right\} \Rightarrow \nabla_x L(\bar{x},\bar{u},\bar{v}) = 0. \ ■$$



8.2. Duality

Denote $U = \{(u,v) \mid u \geq 0\}$ and

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

We have $\forall x \in S,\ \forall (u,v) \in U$:

$$\inf_{x \in X} L(x,u,v) \leq L(x,u,v) =$$

$$= f(x) + \underbrace{u^T g(x)}_{\leq 0} + v^T \underbrace{h(x)}_{= 0} \leq f(x).$$

Denote $\theta(u,v) = \inf_{x \in X} L(x,u,v)$

Remark: inf = infimum =

= same as minimum, but always $\exists$

and can be $-\infty$.

Th (weak duality)

$$\theta(u,v) \leq f(x),\ \forall x \in S,\ \forall (u,v) \in U.$$

Proof: see above. ∎