2nd Order KKT conditions

$$\min \quad f(x)$$
$$g_i(x) \le 0 \quad i=1,\dots,m \qquad \text{(NLP)}.$$
$$h_j(x) = 0 \quad j=1,\dots,p$$

First order conditions:

i) $\nabla f(\bar{x}) + \sum_{i=1}^{m} \bar{u}_i \nabla g_i(\bar{x}) + \sum_{j=1}^{p} \bar{v}_j \nabla h_j(\bar{x}) = 0$

ii) $\bar{u} \ge 0, \; g_i(\bar{x}) \le 0, \; \bar{u}_i g_i(\bar{x}) = 0$

iii) $h_j(\bar{x}) = 0$

In unconstrained case,

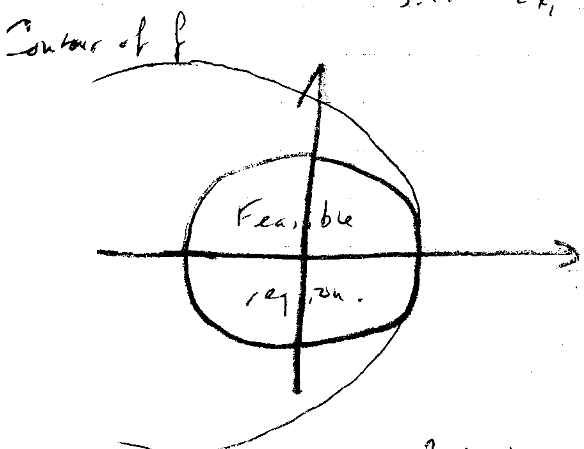
$\bar{x}$ local min $\Rightarrow$ $\nabla^2 f(\bar{x})$ p.s.d., $\nabla f(\bar{x}) = 0$

$\nabla^2 f(\bar{x})$ p.d. $\Rightarrow$ $\bar{x}$ local min
$\nabla f(\bar{x}) = 0$

Consider the problem

Eg 1: $\min \; -\frac{1}{2}(x_1+1)^2 - \frac{1}{2}x_2^2$ Opt point : $\bar{x} = \begin{pmatrix}1\\0\end{pmatrix}$

s.t. $\frac{1}{2}x_1^2 + \frac{1}{2}x_2^2 \le \frac{1}{2}$



$\nabla f(x) = \begin{pmatrix}-x_1-1\\-x_2\end{pmatrix}$ $\nabla f(\bar{x}) = \begin{pmatrix}-2\\0\end{pmatrix}$

$\nabla g(x) = \begin{pmatrix}x_1\\x_2\end{pmatrix}$ $\nabla g(\bar{x}) = \begin{pmatrix}1\\0\end{pmatrix}$

$u = 2$ (KKT multiplier)

$\nabla^2 f(x) = \begin{pmatrix}-1 & 0\\0 & -1\end{pmatrix}$ So $\nabla^2 f(x)$ n.d.

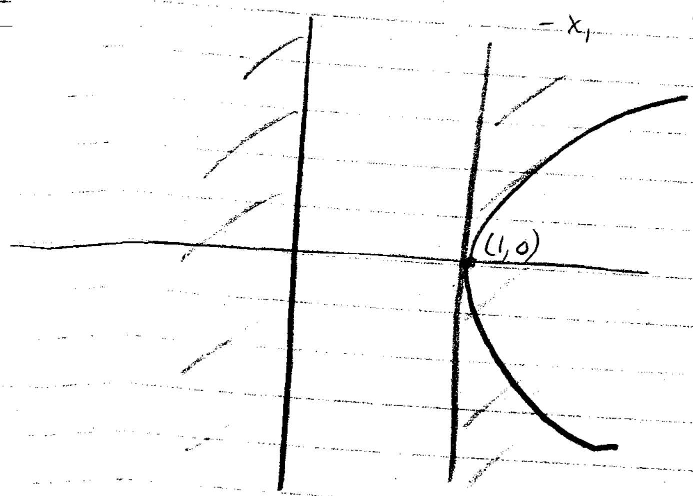
$\nabla^2 g(x) = \begin{pmatrix}1 & 0\\0 & 1\end{pmatrix}$, $\nabla^2 f(x) + u\nabla^2 g(x) = \begin{pmatrix}1 & 0\\0 & 1\end{pmatrix}$ p.d.

2KKT

Eg 2: $\min \; -\frac{1}{2}x_1^2 + \frac{1}{2}x_2^2$

s.t. $x_1 - 1 \le 0$

$-x_1 \le 0$



Optimal point $\bar{x} = \begin{pmatrix}1\\0\end{pmatrix}$

$Df = \begin{pmatrix}-x_1\\x_2\end{pmatrix}$ $Df(\bar{x}) = \begin{pmatrix}-\\0\end{pmatrix}$

$Dg_1(x) = \begin{pmatrix}1\\0\end{pmatrix}$ $Dg_2(x) = \begin{pmatrix}-\\0\end{pmatrix}$

So $u_1 = 1$ $(u_2 = 0)$

$D^2 f = \begin{pmatrix}-1 & 0\\0 & 1\end{pmatrix}$ $D^2 g_1 = \begin{pmatrix}0 & 0\\0 & 0\end{pmatrix}$

So $D^2 f + u D^2 g_1 = \begin{pmatrix}-1 & 0\\0 & +1\end{pmatrix}$

not psd.

Bad direction is $\begin{pmatrix}-1\\0\end{pmatrix}$

But this moves off $g_1(x) = 0$ straight away.

So maybe only want $D^2 f + \sum u_i D^2 g_i$ p.s.d. on the directions which have $d^T Dg_I(\bar{x}) = 0$.

2KKT3

(2nd order nec condns)

Let $\bar{x}$ be a local minimum for (NLP).

Assume $\{\nabla g_i\} : i \in I\}$, $\{\nabla h_j\}$ are lin indep.

Then there exist ~~—~~ $\bar{u}, \bar{v}$ s.t. first order conditions hold.

In addition, if

$$\left.\begin{array}{l} d^T \nabla g_i(\bar{x}) = 0 \quad \forall i \in I \\ d^T \nabla h_j(\bar{x}) = 0 \quad \forall j \end{array}\right\} \text{ then } d^T \nabla^2 f(\bar{x}) d + \sum u_i d^T \nabla^2 g_i d + \sum v_j d^T \nabla^2 h_j d \geq 0.$$

Proof

For case with no equality constraints.

Proof by contradiction, so assume $\exists \bar{d}$ s.t. $\bar{d}^T \nabla g_i(\bar{x}) = 0 \ \forall i \in I$, $\underbrace{\bar{d}^T \nabla^2 f(\bar{x}) \bar{d} + \sum u_i \bar{d}^T \nabla^2 g_i(\bar{x}) \bar{d}}_{=: c} < 0$

Will construct $x(\lambda)$, curve, such that $x(0) = \bar{x}$,

$x(\lambda)$ feasible and $f(x(\lambda)) < f(\bar{x})$ for λ suff small.

Set $x(\lambda) = \bar{x} + \lambda \bar{d} + \frac{1}{2}\lambda^2 w$ for some fixed w

Pick w to ensure feasibility, and also, from KKT 1st order conditions, improvement (*) obj.

Choose w to satisfy

$$\nabla g_i(\bar{x})^T w = -\bar{d}^T \nabla^2 g_i(\bar{x}) \bar{d} - \varepsilon \quad \text{(can do } \cancel{i = 1, \ldots, m} \ i \in I \text{ since } \nabla g_i \text{ lin indep)}$$

where $\varepsilon > 0$, $\varepsilon < \dfrac{-c}{\sum_{i=1}^{m} u_i}$ ($= +\infty$ if $u_i = 0 \ \forall i$)

Then $x(\lambda)$ is feasible for suff small λ:

~~Let $\gamma_i(\lambda) = g_i(x(\lambda))$~~

~~$g_i(x(\lambda)) = g_i(\bar{x}) + \lambda \bar{d}^T$~~

$$g_i(x(\lambda)) = g_i(\bar{x}) + \lambda \bar{d}^T D g_i(\bar{x}) + \tfrac{1}{2}\lambda^2 \omega^T D g_i(\bar{x}) + \tfrac{1}{2}\lambda^2 \bar{d}^T D^2 g_i(\bar{x})$$
$$+ \lambda^2 \alpha(\bar{x}; \lambda d + \tfrac{1}{2}\lambda^2 \omega)$$

Let $\gamma_i(\lambda) = g_i(x(\lambda)) \Leftarrow$

Then $\frac{d\gamma_i}{d\lambda}(0) = \lim_{\lambda \to 0} \frac{g_i(x(\lambda)) - g_i(\bar{x})}{\lambda} = 0. \Leftarrow$

$$\frac{d^2\gamma_i}{d\lambda^2}(0) = \tfrac{1}{2}\omega^T D g_i(\bar{x}) + \tfrac{1}{2}\bar{d}^T D^2 g_i(\bar{x})\bar{d}$$
$$= -\tfrac{1}{2}\varepsilon \quad \text{from } (*)$$
$$< 0 \Leftarrow$$

In addition, $x(\lambda)$ is better than $\bar{x}$ for λ suff small $\Leftarrow$

Let $\phi(\lambda) = f(x(\lambda)) \Leftarrow$

So $\phi(\lambda) = \phi(0) + \lambda \bar{d}^T D f(\bar{x}) + \tfrac{1}{2}\lambda^2 \bar{d}^T D^2 f(\bar{x})\bar{d}$
$$+ \tfrac{1}{2}\lambda^2 \omega^T D f(\bar{x})$$
$$+ \lambda^2 \alpha(\cancel{\bar{x}; \lambda d + \ldots})$$

Thus $\phi'(0) = 0 \Leftarrow$ since $Df(\bar{x}) = -\sum u_i D g_i(\bar{x})$, $\bar{d}^T D g_i(\bar{x})$.

$$\phi''(0) = \tfrac{1}{2}\bar{d}^T D^2 f(\bar{x})\bar{d} + \tfrac{1}{2}\omega^T D f(\bar{x})$$
$$= -\tfrac{1}{2}\sum u_i \omega^T D g_i(\bar{x}) + \tfrac{1}{2}\bar{d}^T D^2 f(\bar{x})\bar{d}$$
$$= +\tfrac{1}{2}\varepsilon \sum u_i + \tfrac{1}{2}\sum u_i \bar{d}^T D^2 g_i(\bar{x})\bar{d} + \tfrac{1}{2}\bar{d}^T D^2 f(\bar{x})\bar{d}$$
$$= \tfrac{1}{2}\varepsilon \sum u_i - \tfrac{1}{2}c$$
$$< 0 \Leftarrow$$

//

Sufficient conditions?

? As for nec, except now

$$\left.\begin{array}{l} d^T Dg_i(\bar{x}) = 0 \ \forall i \\ d^T Dh_j(\bar{x}) = 0 \\ d \neq 0 \end{array}\right\} \Rightarrow d^T D^2 L(\bar{x})\, d > 0 \ ?$$

~~Next recall~~:

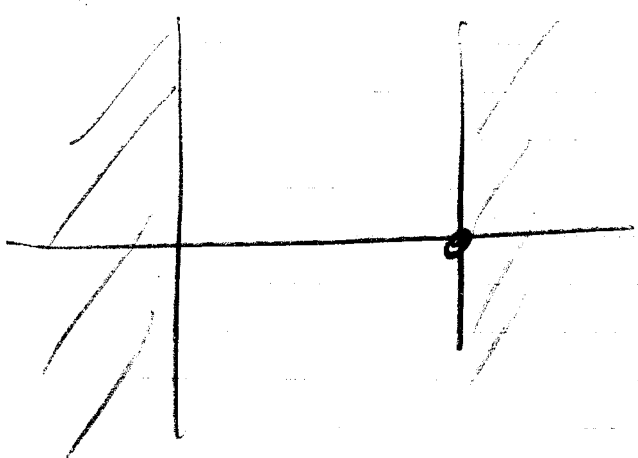
Consider

(Ex 3 illustrates how the necessary conditions can be used to determine that a point is <u>not</u> a local minimizer.)

Ex 3: $\min \ -\frac{1}{2}(x_1+1)^2 - \frac{1}{2}x_2^2$

$$x_1 - 1 \leq 0$$

$$-x_1 \leq 0$$



Consider $\bar{x} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$

KKT point, but <u>NOT</u> optimal

(increase or decrease x_2).

$$Df = \begin{pmatrix} -x_1 - 1 \\ -x_2 \end{pmatrix} \qquad Df(\bar{x}) = \begin{pmatrix} -2 \\ 0 \end{pmatrix}$$

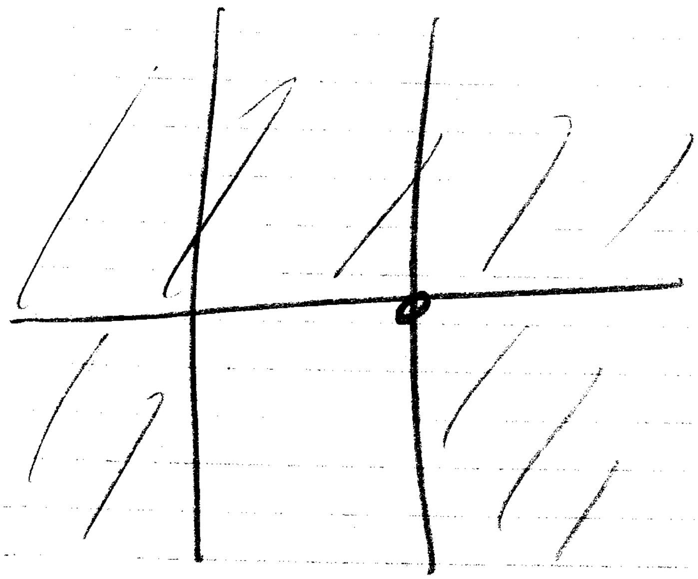
$$Dg_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad u_1 = 2.$$

$$D^2 f = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \qquad D^2 g_1 = 0$$

So direction $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ violates 2nd order nec conditions.

So $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ not optimum.

However, now consider adding constraint $g_3(x) = x_2 \le 0$:



$\nabla g_3 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$

KKT with $u_1 = 2$, $u_3 = 0$

However,

$d^T \nabla g_1(\bar{x}) = 0$, $d^T \nabla g_3(\bar{x}) = 0$, d

does not have a soln,

so vacuously satisfy $d^T(\nabla^2 f + \sum u_i \nabla^2 g_i) d > 0$

if $d^T \nabla g_1(\bar{x}) = 0$, $d^T \nabla g_3(\bar{x}) = 0$, $d \ne 0$.

Problem is caused by zero KKT multiplier.

Theorem (2nd order suff conds):

Suppose $\bar{x}, \bar{u}, \bar{v}$ satisfy KKT conditions

i) $\nabla f + \sum \bar{u}_i \nabla g_i + \sum \bar{v}_j \nabla h_j = 0$

ii) $\bar{u} \ge 0$, $g \le 0$, $\bar{u}_i g_i(\bar{x}) = 0$

iii) $h_j(\bar{x}) = 0$

NB: this is less restrictive than hypothesised restrictions: $d^T \nabla h_j = 0$, $d^T \nabla g_I = 0$.

Assume in addition

$d^T \nabla h_j = 0 \ \forall j$
$d^T \nabla g_i \le 0 \ \forall i \in I$
$d^T \nabla g_i = 0$ if $u_i \ne 0$
$d \ne 0$

$\Rightarrow d^T(\nabla^2 f + \sum u_i \nabla^2 g_i + \sum v_j \nabla^2 h_j) d > 0$ (Nor

Then: $\bar{x}$ is a strict local min

$$\min \quad x_2$$

$$\text{s.t.} \quad x_1^2 - x_2 \leq 0$$

$$Df = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \qquad Dg = \begin{bmatrix} 2x_1 \\ -1 \end{bmatrix}$$

$$Df + u\,Dg = 0 \quad \text{if} \quad u = 1,\ x_1 = 0,\ x_2 = 0,\ g_1(x) = 0.$$

$$D^2 f = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \qquad D^2 g = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}.$$

$$d^T Dg\begin{pmatrix} 0 \\ 0 \end{pmatrix} = 0 \Rightarrow d = \begin{bmatrix} a \\ 0 \end{bmatrix}. \Rightarrow d^T D^2 g\, d = 2a^2 > 0 \text{ if } a \neq 0.$$

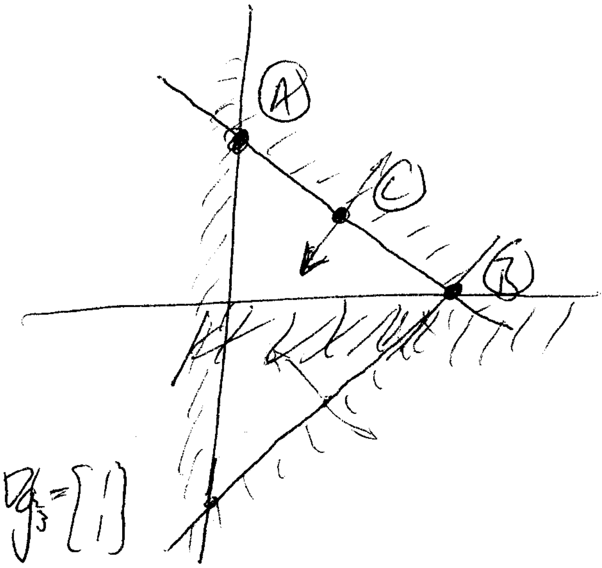
This example is a simple illustration of 2nd order KKT conditions.

min $\quad -\frac{1}{2}x_1^2 - \frac{1}{2}x_2^2$

s.t. $\quad -x_1 \le 0$

$\qquad -x_2 \le 0$

$\qquad x_1 + x_2 - 1 \le 0.$



$Df = \begin{bmatrix} -x_1 \\ -x_2 \end{bmatrix} \quad Dg_1 = \begin{bmatrix} -1 \\ 0 \end{bmatrix} \quad Dg_2 = \begin{bmatrix} 0 \\ -1 \end{bmatrix} \quad Dg_3 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

A: $x_1 = 0, x_2 = 1$, $\quad u_1 = 1, u_2 = 0, u_3 = 1 \qquad d^T Dg_1 = d^T Dg_3 = 0 \Rightarrow d = 0.$

so vacuous at 2nd order

B: $x_1 = 1, x_2 = 0, \quad u_1 = 0, u_2 = 1, u_3 = 1 \qquad$ 〃

C: $x_1 = \frac{1}{2}, x_2 = \frac{1}{2}, \quad u_1 = u_2 = 0, \quad u_3 = \frac{1}{2} \qquad d^T Dg_3 = 0 \Rightarrow d = c\begin{bmatrix} 1 \\ -1 \end{bmatrix}$

$D^2 L = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$

so $d^T D^2 L d = -2c^2.$