(i) Consider a differentiable, scalar fun^n $f: \mathbb{R}\to\mathbb{R}$, $\min_{x\in I} f(x)$, $I\in[a,b]$ is closed Interval. what are the necessary conditions

(a) $I=(a,b)$, Interior

necessary optimality conditions

First order: $\nabla f(x^*)=0$

second order: $\nabla^2 f(x^*)\geq 0$

sufficient optimality conditions

First order: $\nabla f(x^*)=0$

second order: $\nabla^2 f(x^*)>0$

(b) On the boundary ∂I:

→ It has to be satisfy the variational inequality. $f'(x^*)(x-x^*)\geq 0$

→ for minimum on a:

$f(a+t(x-a))\geq f(a)$, $\forall x\in I$, $t>0$ is small enough

$f(a+t(x-a))-f(a)\geq 0$

$$\lim_{t\to 0}\frac{f(a+t(x-a))-f(a)}{t}=f'(a)(x-a)\geq 0 \quad \forall x\in I. \qquad f'(b)(x-b)$$

→ for Maximum on a

$f(a+t(x-a))\leq f(a)$

$f(a+t(x-a))-f(a)\leq 0$

$$\lim_{t\to 0}\frac{f(a+t(x-a))-f(a)}{t}=f'(a)(x-a)\leq 0 \quad \forall x\in I \qquad f'(b)(x-b) \to \text{similarly}$$

Show that a local solution x^* of a convex programme is also a global sol^n.

→ Condition: $f(x^*)\leq f(x)\ \forall x\in X$

→ Local optimality

→ convexity of the fun^n:

(i) If I take an element from the neighborhood of $B_\varepsilon(x^*)$ e.g $\lambda x+(1-\lambda)x^*\in B_\varepsilon(x^*)$ for $\lambda\in(0,1)$ small enough $x\in X$ is arbitrary.

$$f(x^*)\leq f(\lambda x+(1-\lambda)x^*)$$

(ii) Convexity of the fun^n:

$f(\lambda x_1+(1-\lambda)x_2)\leq\lambda f(x_1)+(1-\lambda)f(x_2)$

then I have

$$f(x^*)\underset{\text{due to local optimality}}{\leq} f(\lambda x+(1-\lambda)x^*)\underset{\text{due to convexity of }x}{\leq}\lambda f(x)+(1-\lambda)f(x^*)$$

$f(x^*)\leq\lambda f(x)+f(x^*)-\lambda f(x^*)\ \forall x\in X$

x^* is a global solution.

$f(x)=\frac12 x^TAx-b^Tx+c$, $\nabla f(x)^Td<0$

(a) $t_{min}=-\dfrac{\nabla f(x)^T d}{d^TAd}$ gives $f(x+t_{min}d)\leq f(x+td)$

→ $\nabla f(x)=?$ $g(x)=$

$g(x)=x^TAx=\sum_i\sum_j a_{ij}x_ix_j = \sum_{i=1}^n a_{ii}x_ix_i+\sum_{i=1}^n a_{ij}x_ix_j+\sum_{i=2}^n\sum_{j=2}^n a_{ij}x_ix_j$

$$\frac{\partial g}{\partial x_1}=\sum_{i=1}^n a_{1i}x_i+\sum_{j=1}^n a_{1j}x_j=2\sum_{i=1}^n a_{1i}x_i \quad (\because A \text{ is symmetric})$$

$$\nabla g(x)=\begin{pmatrix}2\sum_{i=1}^n a_{1i}x_i\\ 2\sum_{i=1}^n a_{2i}x_i\\ \vdots\\ 2\sum_{i=1}^n a_{ni}x_i\end{pmatrix}=2\begin{pmatrix}a_{11}&a_{12}&\cdots&a_{1n}\\ a_{21}&a_{22}&\cdots&a_{2n}\\ \vdots&&&\vdots\\ a_{n1}&a_{n2}&\cdots&a_{nn}\end{pmatrix}\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}=2Ax$$

Now

$\Rightarrow \nabla f(x)=\frac12\cdot 2Ax-b^T=Ax-b$

$\Rightarrow \nabla^2 f(x)=A \Rightarrow \nabla^3 f(x)=0$

the exact taylor expansion

$$f(x+t\cdot d)=f(x)+t\cdot\nabla f(x)^T\cdot d+\frac12 t^2 d^T\nabla^2 f(x)\cdot d$$

Now, we define

$\Phi(t)=f(x+t\cdot d)\qquad \Phi'(t)=\nabla f(t)^T\cdot d$

and get

$\Phi(t)=\Phi(0)+t\Phi'(0)+\frac12 t^2\Phi''(0)$

If we search stationary point, we get

$0\overset{!}{=}\Phi'(t_{min})$

$\Phi'(t_{min})=\Phi'(0)+t_{min}\Phi''(0)$

$0=\Phi'(0)+t_{min}\Phi''(0)$

$$\therefore t_{min}=-\frac{\Phi'(0)}{\Phi''(0)}\Rightarrow-\frac{\nabla f(x)^T\cdot d}{d^T\cdot\nabla^2 f(x)\cdot d}$$

$$\Rightarrow t_{min}=-\frac{\nabla f(x)^T\cdot d}{d^TA\cdot d}$$

furthermore,

$\Phi''(t_{min})=\Phi''(0)=d^T\nabla^2 f(x)\cdot d = d^T\cdot A\cdot d\geq 0$

Here, A is SPD, then d^TAd is +ve definite. So, t_{min} is indeed a min^m.

(b) $f(x+t_{min}d)\leq f(x)-\theta\left(\dfrac{\nabla f(x)^T\cdot d}{\|d\|}\right)^2$

taylor expansion

$$=f(x)+t_{min}\nabla f(x)^Td+\frac{t_{min}^2}{2}d^T\nabla^2 f(x)\cdot d$$

$$=f(x)+\left(-\frac{\nabla f(x)^Td}{d^TA\cdot d}\right)\nabla f(x)^Td+\left(-\frac{\nabla f(x)^Td}{d^TAd}\right)^2\frac{d^TAd}{2}$$

$$=f(x)-\frac{(\nabla f(x)^T\cdot d)^2}{d^TA\cdot d}+\frac12\frac{(\nabla f(x)^Td)^2}{d^TA\cdot d}$$

$$=f(x)-\frac12\frac{(\nabla f(x)^T\cdot d)^2}{d^TA\cdot d}\qquad(\because d^TAd\leq\lambda_{max}(A)\|d\|^2)$$

$$\leq f(x)-\frac{(\nabla f(x)^T\cdot d)^2}{(2\lambda_{max}(A))\|d\|^2}$$

$$\leq f(x)-\theta\left(\frac{\nabla f(x)^T\cdot d}{\|d\|}\right)^2\qquad\therefore\theta=\frac{1}{2\lambda_{max}(A)}$$

Let $f:\mathbb{R}^n\to\mathbb{R}$ be a quadratic function i.e $f(x)=\frac12 x^TAx-b^Tx+c$, with ...

then show $t_{min}\in T_{wp}(x,d)$

$\Rightarrow f(x)=\frac12 x^TAx-b^Tx+c$

$\nabla f(x)=Ax-b \Rightarrow \nabla^2 f(x)=A \to$ (i)

$t_{min}=-\dfrac{\nabla f(x)^Td}{d^T\cdot A\cdot d}$ ⇝ From ex sheet 3

$t_{min}\in T_{wp}(x,d)$???

$$T_{wp}(x,d)=\left\{t>0: \begin{array}{l}\text{(i) } f(x+td)\leq f(x)+\alpha\cdot t\cdot f'(x)\cdot d\\ \text{(ii) } f'(x+td)\geq\beta f'(x)\cdot d\end{array}\right.$$

$\alpha\in(0,\frac12)$, $\beta\in[\alpha,1)$

(i) $f(x+t_{min}d)\leq f(x)+\alpha\cdot t_{min}f'(x)\cdot d$

using taylor expansion

$f(x+td)=f(x)+t\nabla f(x)^Td+\frac{t^2}{2}d^T\nabla^2 f(x)\cdot d$

$f(x+td)-f(x)=\underset{t_{min}}{t}\nabla f(x)^T\cdot d+\frac{\underset{t_{min}}{t^2}}{2}d^T\nabla^2 f(x)\cdot d \to (*)$

Put $t_{min}=-\dfrac{\nabla f(x)^T\cdot d}{d^T\cdot A\cdot d}$

$$(*)=\left(-\frac{\nabla f(x)^T\cdot d}{d^T\cdot A\cdot d}\right)\nabla f(x)^T\cdot d+\frac12\left(-\frac{\nabla f(x)^Td}{d^TAd}\right)^2 d^T\underbrace{\nabla^2 f(x)}_{A}d$$

$$=-\frac{(\nabla f(x)^Td)^2}{d^T\cdot A\cdot d}+\frac12\frac{(\nabla f(x)^Td)^2}{d^T\cdot A\cdot d}$$

$$\leq\alpha\left(-\frac{\nabla f(x)^T\cdot d}{d^TAd}\right)\nabla f(x)^Td$$

$\Rightarrow -\frac12\leq-\alpha \Rightarrow 0<\alpha<\frac12$

(ii) $f'(x+t_{min}\cdot d)\geq\beta\cdot f'(x)\cdot d$?

$\nabla f(x+td)^T\cdot d=(A(x+td)-b)^T\cdot d$

$\nabla f(x+td)^T\cdot d=(Ax-b)^Td+t\cdot d^TA\cdot d$

$\nabla f(x)^T\cdot d=(Ax-b)^T\cdot d$

Now insert t_{min} in condition (ii)

$(Ax-b)^Td+t_{min}d^T\cdot A\cdot d\overset{?}{\geq}\beta(Ax-b)^T\cdot d$

$(Ax-b)^Td+\left(-\dfrac{\nabla f(x)^Td}{d^TAd}\right)d^TAd\geq\beta(Ax-b)^T\cdot d$

$-\underbrace{(Ax-b)^Td}_{<0}\geq(\beta-1)(Ax-b)^Td$

$-1\leq\beta-1$

$0\leq\beta \Rightarrow \beta\geq 0$

since, we have $\alpha>0$. we get $\beta\in[\alpha,1]$ both conditions satisfied for t_{min}.

Show that if the sequence (t_k) of Armijo-Goldstein steplengths fulfills $t_k\geq-\alpha\dfrac{\nabla f(x^k)^Td^k}{\|d^k\|^2}$ with $\alpha>0$ then t_k is efficient.

⇒ Efficiency equation

$$f(x+t\cdot d)\leq f(x)-\theta(f'(x)\cdot d)^2$$

Armijo Goldstein equ^n

$f(x+td)\leq f(x)+t_k\alpha f'(x)\cdot d \to (1)$

steplength (t_k) is given, Put in (1).

$$f(x^k+t_kd^k)\leq f(x^k)+\alpha\left[\frac{-\alpha\nabla f(x^k)^T\cdot d^k}{\|d^k\|^2}\right]f'(x^k)d^k$$

$$\leq f(x^k)-\varepsilon\frac{\nabla f(x^k)^T\cdot d^k}{\|d^k\|^2}\nabla f(x^k)^T\cdot d^k$$

$$\leq f(x^k)-\underset{\varepsilon>0}{\varepsilon}\underbrace{\left[\frac{\nabla f(x^k)^Td^k}{\|d^k\|}\right]^2}_{>0}\quad\because\varepsilon=\alpha\alpha$$

same as efficiency eq^n. This shows efficiency.

Show that Armijo steplength is not efficient.

$\Rightarrow f(x)=\dfrac{x^2}{8}$, $d^k=-2^{-k}f'(x^k)=-2^{-(k+2)}\cdot x^k$

Let $x^0>0$, then we see that with a generalised by Armijo, $\alpha>0$

$x^{k+1}-x^k=x^k+\alpha_kd^k-x^k$

$=-\alpha_k\cdot 2^{-(k+2)}x_k$

$x^{k+1}=\left(1-\alpha_k 2^{-(k+2)}\right)x^k$

$x^{k+1}=\left(1-\dfrac{\alpha_{k-1}}{2^{k+1}}\right)x^k$

$x^k=\left(1-\dfrac{\alpha_{k-1}}{2^{k+1}}\right)x^{k-1}$, $x^1=\left(1-\dfrac{\alpha_0}{2^2}\right)x^0$

$$x^{k+1}=x_0\prod_{n=1}^{k}\left(1-\frac{\alpha_{n-1}}{2^{n+1}}\right)$$

$$\geq x_0\prod_{n=1}^{k-1}\left(1-\frac{1}{2^{n+1}}\right)\geq x_0\prod_{n=1}^{\infty}\left(1-\frac{1}{2^{n+1}}\right)$$

$x^k\geq\frac{x_0}{2}$ ↓ since x^k is monotonically decreasing. we have $x^k\leq x_0$

$f(x)=\frac{x^2}{8}$ is increasing. i.e

for $0<a<b$, $0<f(a)<f(b)$

Note, $f(a)=0 \Leftrightarrow a=0$.

$0<f(\frac{x^0}{2})<f(x^k)\leq f(x^0)\ \forall k\in\mathbb{N}$

$\Rightarrow f(x^k)\searrow f(\bar{x})\geq f(\frac{x^0}{2})>0$

$\bar{x}$ is not a stationary point.

Use Newton's method to calculate the extremum of $g(x)=2x_1^2+x_2^2-4x_1+x_2$ with $x^0=(0,0)^T$.

$\Rightarrow x^1=x^0-[\text{Hess}\, g(x^0)]^{-1}\nabla g(x^0)$

$\nabla g(x)=\begin{bmatrix}4x_1-4\\2x_2+1\end{bmatrix}\Rightarrow\nabla g(x^0)=\begin{bmatrix}-4\\1\end{bmatrix}$

$\text{Hess}\, g(x)=\begin{bmatrix}4&0\\0&2\end{bmatrix}\Rightarrow[\text{Hess}\, g(x^0)]^{-1}=\begin{bmatrix}1/4&0\\0&1/2\end{bmatrix}$

$x^1=\begin{bmatrix}0\\0\end{bmatrix}-\begin{bmatrix}1/4&0\\0&1/2\end{bmatrix}\begin{bmatrix}-4\\1\end{bmatrix}=\begin{bmatrix}1\\-1/2\end{bmatrix}$ $\quad x^2=\begin{bmatrix}0\\0\end{bmatrix}$

Let $p>2$, $x^0>0$ arbitrary and ...

$|x^{k+1}-\bar{x}|=\left(1-\dfrac{1}{p-1}\right)|x^k-\bar{x}|$.

⇒ given $f(x)=|x|^p$

we can rewrite $|x|^p$ as

$f(x)=(\text{sign}(x)\cdot x)^p$

$f'(x)=p(\text{sign}(x)\cdot x)^{p-1}\,\text{sign}(x)$.

$f''(x)=\left[p(p-1)(\text{sign}(x)\cdot x)^{p-2}\,\text{sign}\,x\right]\text{sign}\,x$

$=p(p-1)(\text{sign}(x)\cdot x)^{p-2}$

$=p(p-1)|x|^{p-2}$

By applying newton's equation for all k

$$x^{k+1}=x^k-\frac{f'(x^k)}{f''(x^k)}$$

$$=x^k-\frac{p|x^k|^{p-1}\,\text{sign}(x^k)}{p(p-1)|x^k|^{p-2}}$$

$$=x^k-\frac{1}{(p-1)}\left[|x^k|\,\text{sign}(x^k)\right]$$

$$=x^k-\frac{1}{(p-1)}\left[\text{sign}(x)^k\cdot x^k\cdot\text{sign}\,x^k\right]$$

$$=x^k-\frac{1}{(p-1)}[x^k]$$

By induction, we prove that $0<x^{k+1}<x^k\ \forall k$

given $x^0>0$

$x^1=x^0\left(1-\frac{1}{p-1}\right)<x^0$, $x^2=x^1\left[1-\frac{1}{p-1}\right]>0$

since $0<x^k<x^{k+1}$

$x^{k+1}=\left(1-\frac{1}{p-1}\right)x^k<x^k$

$x^{k+1}=\left(1-\frac{1}{p-1}\right)x^k>0$

convergence by applying limit

$$\lim_{k\to\infty}x^k=\lim_{k\to\infty}\left[1-\frac{1}{p-1}\right]x^{k-1}$$

$=0=\bar{x}$

$|x^{k+1}-\bar{x}|=|x^{k+1}|$

$=\left[\left(1-\frac{1}{p-1}\right)x^k\right]\Rightarrow 1-\frac{1}{p-1}|x^k|$

$\Rightarrow 1-\frac{1}{p-1}|x^k-\bar{x}|$

In this case, Hess $f(\bar{x})>0$ is violated Because Hess $f(\bar{x})=p(p-1)|\bar{x}|^{p-2}=0$.

Proof the Lemma of Banach, which...

(i) $A+B$ is invertible.

⇒ Neumann result: Let $A\in\mathbb{R}^{n\times n}$ be invertible matrix and $\rho(A)$ is spectral radius of A.

If $\rho(A)<1$, then $(I-A)^{-1}$ matrix exists and $(I_n-A)^{-1}=\sum_{k=0}^{\infty}A^k$ Neumann series

Now, we are given $A\in\mathbb{R}^{n\times n}$ & $B\in\mathbb{R}^{n\times n}$ are invertible & $\|A^{-1}B\|<1$. we set $M=-A^{-1}B$

$S=\sum_{k=0}^{\infty}M^k$, $S_n=\sum_{k=0}^{n}M^k$

For the convergence of Neumann series we look at:

$$\|S-S_n\|=\left\|\sum_{k=0}^{\infty}M^k-\sum_{k=0}^{n}M^k\right\|$$

$$=\left\|\sum_{k=n+1}^{\infty}M^k\right\|$$

$$\leq\sum_{k=n+1}^{\infty}\|M^k\|\quad(\because\text{triangle inequality})$$

$$\leq\sum_{k=n+1}^{\infty}\|M\|^k\quad(\because\|a\cdot b\|\leq\|a\|\|b\|)$$

$$=\|M\|^{n+1}\sum_{k=0}^{\infty}\|M\|^k$$

$$=\|M\|^{n+1}\times\frac{1}{1-\|M\|}>0$$

If $n\to\infty$ $\|S-S_n\|=0$

$(I-M)S_n=(I-M)\sum_{k=0}^{n}M^k$

$\Rightarrow\sum_{k=0}^{n}M^k-\sum_{k=0}^{n}M^{k+1}=M^0-M^{n+1}$

$\Rightarrow I-M^{n+1}=M^0-M^{n+1}$

$\sum_{k=0}^{n}M^k-\sum_{k=0}^{n}M^k\cdot M$

$=\sum_{k=0}^{n}M^k(I-M)=S_n(I-M)$

$\to(I-M)S=I=S(I-M)$

⇒ for $n\to\infty$, we get

$(I-M)\cdot S=I\Rightarrow(I-M)^{-1}=S$

It means, $I-M$ is invertible

$A+B=A(I-M)$

$=A+AA^{-1}B$ $(\because M=-A^{-1}B)$

$=A+B$

since, A & $(I-M)$ are both invertible.

we have

$(A+B)^{-1}=(A(I-M))^{-1}=(I-M)^{-1}A^{-1}=SA^{-1}$

(ii) $\|(A+B)^{-1}\|=\|SA^{-1}\|$

$\leq\|S\|\,\|A^{-1}\|$

$=\|A^{-1}\|\left\|\sum_{k=0}^{\infty}M^k\right\|$

$=\|A^{-1}\|\sum_{k=0}^{\infty}\|M\|^k$

$$=\frac{\|A^{-1}\|}{1-\|A^{-1}B\|}$$

If f is twice differentiable at x and $f''(x)<0$ then f has a local maximum at x. If f is twice differentiable at x and $f''(x)>0$ then f has a local minimum at x.

Let $f: \mathbb{R} \to \mathbb{R}$, $f(x) = -\frac{1}{4}x^4$ with ...
starting point $x^0 \in \mathbb{R}$ arbitrary. ...
what kind of extremum is x^*?

$> f(x) = -\frac{1}{4}x^4$

$f'(x) = -x^3 \Rightarrow f''(x) = -3x^2$

$$x^{k+1} = x^k - \frac{f'(x^k)}{f''(x^k)}$$
$$= x^k - \frac{(-x^{3k})}{(-3x^{2k})} = \frac{2x^k}{3}$$

$$\lim_{k\to\infty} x^k = \lim_{k\to\infty} \frac{2}{3}x^{k-1}$$
$$= \lim_{k\to\infty} \left(\frac{2}{3}\right)^k x^0 = 0 = x^*$$

x^* is a maximum.
" " the extremum
(i) $f''(x^*) > 0 \Rightarrow x^*$ is min^m
(ii) $f''(x^*) < 0 \Rightarrow x^*$ is max^m
But if $f''(x^*) = f'(0) = 0$
then we can't decide
Here, $f(x) = -\frac{1}{4}x^4 \le 0$, $\forall x \in \mathbb{R}$

Given the matrices $U, V \in \mathbb{R}^{n\times m}$ ...
show that $(I_{dm} - V^TA^{-1}U)$ is invertible
then $M = A - UV^T$ is invertible with $M^{-1} = A^{-1} + A^{-1}U(I_{dm} - V^TA^{-1}U)^{-1}V^TA^{-1}$.

$\Rightarrow MM^{-1}$ doing multiplication any step

$= AA^{-1} + AA^{-1}U(I_{dm} - V^TA^{-1}U)^{-1}V^TA^{-1}$
$- UV^TA^{-1} - UV^TA^{-1}U(I_{dm} - V^TA^{-1}U)^{-1}V^TA^{-1}$

$= I_{dn} + U(I_{dm} - V^TA^{-1}U)^{-1}V^TA^{-1} - UV^TA^{-1}$
$- UV^TA^{-1}U(I_{dm} - V^TA^{-1}U)^{-1}V^TA^{-1}$

$= I_{dn} + U[(I_{dm} - V^TA^{-1}U)^{-1} - I_{dm}$
$- V^TA^{-1}U(I_{dm} - V^TA^{-1}U)^{-1}]V^TA^{-1}$

$= I_{dn} + U[\underbrace{(I_{dm} - V^TA^{-1}U)(I_{dm} - V^TA^{-1}U)^{-1} - I_{dm}}_{=0}]V^TA^{-1}$

$= I_{dn}$

$M^{-1}M$ doing multi^n any step

$= A^{-1}A - A^{-1}UV^T + A^{-1}U(I_{dm} - V^TA^{-1}U)^{-1}V^TA^{-1}A - A^{-1}U(I_{dm} - V^TA^{-1}U)^{-1}V^TA^{-1}UV^T$

$= I_{dn} - A^{-1}UV^T + A^{-1}U(I_{dm} - V^TA^{-1}U)^{-1}V^T$
$- A^{-1}U(I_{dm} - V^TA^{-1}U)^{-1}V^TA^{-1}UV^T$

$= I_{dn} + A^{-1}U(-I_{dm} + (I_{dm} - V^TA^{-1}U)^{-1}$

$= I_{dn} + A^{-1}U[-I_{dm} + (I_{dm} - \ldots)^{-1} = 0$
$(V^TA^{-1}U)^{-1}(I_{dm} - V^TA^{-1}U)]V^T$

$= I_{dn}$

$MM^{-1} = M^{-1}M = I_{dn}$

(b) s.t $M = A - uv^T$
$$M^{-1} = A^{-1} + \frac{A^{-1}uv^TA^{-1}}{1 - v^TA^{-1}u}$$

$\Rightarrow M = A - uv^T$, $u, v \in \mathbb{R}^n$

$\rightsquigarrow M$ is invertible $\Leftrightarrow 1 - v^TA^{-1}u \neq 0$

$(\Leftarrow:)$ let $1 - v^TA^{-1}u \neq 0$, then from (a) with $m=1$ (which gives $I_{dm} - V^TA^{-1}U = 1 - v^TA^{-1}u$)
It follows that M is invertible with the given M^{-1}.

$(:\Rightarrow)$ let $M = A - uv^T$ is invertible
Assume that
$1 - v^TA^{-1}u = 0$
$\Rightarrow AA^{-1}u(1 - v^TA^{-1}u) = 0$
$\Rightarrow AA^{-1}u - uv^TA^{-1}u = 0$
$\Rightarrow (A - uv^T)A^{-1}u = 0$
$\Rightarrow A(A - uv^T)^{-1}(A - uv^T)A^{-1}u = 0$
If $u = 0 \Rightarrow 1 - \underbrace{v^TA^{-1}u}_{0} = 0$
$1 = 0$ contradiction
Hence, $(1 - v^TAu) \neq 0$ must hold true.

S.t the simplified Newton's method
$||x^{k+1} - x^*|| \le c||x^k - x^*||$ $\forall k$, $c<1$.

$\Rightarrow$ Simplified Newton's method
$x^{k+1} = x^k - \text{Hess}(f(x^0))^{-1}\nabla f(x^k)$
$||x^{k+1} - x^*|| = ||x^k - \text{Hess}(f(x^0))^{-1}\nabla f(x^k) - x^*||$
(∵ subtract x^* both side)
$= ||-\text{Hess}(f(x^0))^{-1}\nabla f(x^k) + x^k - x^*||$
$= ||\text{Hess}(f(x^0))^{-1}[-\nabla f(x^k) + \text{Hess}(f(x^0))(x^k - x^*)]||$
$= ||\text{Hess}(f(x^0))^{-1}[\underbrace{\nabla f(x^*)}_{0} - \nabla f(x^k) + \text{Hess}(f(x^0))(x^k - x^*)]||$
$= ||\text{Hess}(f(x^0))^{-1}[\int_0^1 \text{Hess}(f(x^k + s(x^* - x^k)))ds(x^* - x^k) + \text{Hess}(f(x^0))\cdot(x^k - x^*)]||$

$\le ||\text{Hess}(f(x^0))^{-1}||\,||\int_0^1 \text{Hess}(f(x^k + s(x^* - x^k))ds(x^* - x^k) - \text{Hess}(f(x^0))(x^* - x^k)||$ (∵ $||a \cdot b|| \le ||a||\,||b||$)

$\le ||\text{Hess}(f(x^0))^{-1}||\,||\int_0^1 \text{Hess}(f(x^k + s(x^* - x^k)))ds - \text{Hess}(f(x^0))||\,||x^* - x^k||$

$\le ||\text{Hess}(f(x^0))^{-1}||\,||\int_0^1 \text{Hess}(f(x^k + s(x^* - x^k))) - \text{Hess}(f(x^0))ds||\,||x^k - x^*||$

$\le ||\text{Hess}(f(x^0))^{-1}|| \int_0^1 ||\text{Hess}(f(x^k + s(x^* - x^k)) - \text{Hess}(f(x^0))||ds\,||x^k - x^*||$

$= ||\text{Hess}(f(x^0))^{-1}|| \int_0^1 \varepsilon \cdot ds\,||x^k - x^*|| \Rightarrow ||x^{k+1} - x^*|| \le \frac{c}{\varepsilon} \times \varepsilon\,||x^k - x^*||$

$\Rightarrow ||x^{k+1} - x^*|| = c||x^k - x^*|| \Rightarrow ||x^{k+1} - x^*|| \le c||x^k - x^*||$.

Show that the local Newton's method for optimization problems ...
$x^0 = Ay^0 + b \Rightarrow x^k = Ay^k + b$, $\forall k \in \mathbb{N}$
$\Rightarrow x = Ay + b$, $A \in \mathbb{R}^{n\times n}$ invertible $b \in \mathbb{R}^n$
$g(y) = f(Ay + b)$
Newton applied to $f(x) = x^k$
" " " $g(y) = y^k$
To show $x^0 = Ay^0 + b \Rightarrow x^k = Ay^k + b$ $\forall k \in \mathbb{N}$
By the Newton's method for $g(y)$ we have
$y^{k+1} = y^k - \text{Hess}_y(g(y^k))^{-1}\nabla_y g(y^k)$
$\nabla_y g(y^k) = \nabla_y f(Ay^k + b)$
$= \nabla_x f(x^k) \cdot A$
$= A^T \cdot \nabla_x f(x^k) \longrightarrow (1)$
$\text{Hess}_y(g(y^k))^{-1} = \frac{\partial}{\partial y}(A^T \nabla_x f(x^k))^{-1}$
$= (A^T \cdot \text{Hess}_x(f(x^k)) \cdot A)^{-1}$
$= A^{-1}\text{Hess}_x(f(x^k))^{-1}(A^T)^{-1} \to (2)$

Proof by induction
$k=0 \Rightarrow x^0 = Ay^0 + b$
let, $x^k = Ay^k + b$ hold true.
then, $y^{k+1} = y^k - \text{Hess}_y(g(y^k))^{-1}\nabla_y(g(y^k))$
$= y^k - (A^T \cdot \text{Hess}_x(f(x^k))A)^{-1}(A^T\nabla_x f(x^k))$
$= y^k - (A^{-1}\text{Hess}_x(f(x^k))^{-1}(A^T)^{-1}A^T\nabla_x f(x^k))$
$= y^k - A^{-1}\text{Hess}_x(f(x^k))^{-1}\nabla_x f(x^k)$
$\rightsquigarrow A \cdot y^{k+1} + b$
$= A(y^k - A^{-1}\text{Hess}_x(f(x^k))^{-1}\nabla_x f(x^k)) + b$
$= \underbrace{Ay^k + b}_{x^k \text{ we consider}} - \text{Hess}_x(f(x^k))^{-1}\nabla_x f(x^k)$
$= x^k - \text{Hess}_x(f(x^k))^{-1}\nabla_x f(x^k)$
$= x^{k+1}$ (∵ Newton's method - Formula)

Find the global minimum of ...
$$\min_{x\in\mathbb{R}^2} f(x) = (1-x_1)(x_2-2)^2 - x_2 \quad \begin{pmatrix} -x_1 \\ x_1 - 1 \end{pmatrix} \le 0$$

$\Rightarrow -x_1 \le 0 \Rightarrow x_1 \ge 0$
$x_1 - 1 \le 0 \Rightarrow x_1 \le 1$ } $x_1 \in [0,1]$, $0 \le x_1 \le 1$

$(x_1 + x_2 - 1)^2(x_1 - x_2 - 1)^2 \le 0$
case: 01 $x_2 = 1 - x_1$ case: 02 $x_2 = x_1 - 1$

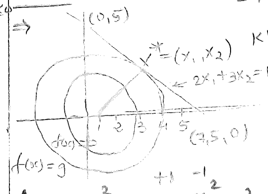
case: 01
$f(x_1) = (1-x_1)[(1-x_1) - 2]^2 - (1-x_1)$
$= -x_1^3 - x_1^2 + 2x_1$
$\nabla f(x_1^*) = 0 \Rightarrow -3x_1^2 - 2x_1 + 2 = 0$
$\Delta = b^2 - 4ac = 28$
$x_{1,2} = -\frac{1}{3} \pm \frac{\sqrt{7}}{3}$ $x_1 = -\frac{1}{3} + \frac{\sqrt{7}}{3} \in [0,1]$
all constraints satisfied $x_2 = \frac{4}{3} - \frac{\sqrt{7}}{3}$
$x^* = (x_1, x_2) = \left(-\frac{1}{3} + \frac{\sqrt{7}}{3}, \frac{4}{3} - \frac{\sqrt{7}}{3}\right)$
$f(x^*) = 0.6311$

case: B $x_2 = x_1 - 1$
$f(x_1) = (1-x_1)[(x_1-1) - 2]^2 - (x_1 - 1)$
$= -x_1^3 + 7x_1^2 - 16x_1 + 10$
$\nabla f(x_1^*) = 0 \Rightarrow -3x_1^2 + 14x_1 - 16 = 0$
$x_1' = 2 \notin [0,1]$ $x_1'' = \frac{-7-1}{-3} \notin [0,1]$

* check at the boundary values
$x_1 \in [0,1] \Rightarrow x_1 = 0$ $x_1 = 1$
$x_1 = 0 \Rightarrow x_2 = 1 - 0$ $x_1 = 0 \Rightarrow x_2 = 0 - 1$
$x_2 = 1$ $x_2 = -1$
The last one is violated. No solution
$x_1 = 1$, $x_2 = 0 \Rightarrow$ all constraint satisfied.
$f(1,0) = 0$. $\Rightarrow f(x^*) = 0$ global solution.

Derive the necessary optimality conditions
$$\min_{x\in\mathbb{R}^2} x_1^2 - 2x_1 + x_2^2 - 3 \text{ s.t } 2x_1 + 3x_2 - 15 = 0.$$

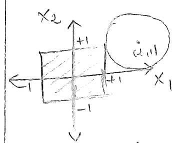
$\Rightarrow$



$f(x) = x_1^2 - 2x_1 + x_2^2 - 3$
$= (x_1 - 1)^2 + x_2^2 - 4$
$f(x) = 0 \Rightarrow (x_1 - 1)^2 + x_2^2 - 4 = 0$
$(x_1 - 1)^2 + x_2^2 = 2^2$ centre
This is Circle, with radius 2, (1,0)
x^* satisfies the contraint $h(x^*) = 0$
$\therefore 2x_1^* + 3x_2^* - 15 = 0$
moreover, x^* must be closest point to the circle.
$$l = \frac{|2\times1 + 3\times0 - 15|}{\sqrt{4+9}} = \frac{13}{\sqrt{13}} = \sqrt{13}$$
$(x_1 - 1)^2 + x_2^2 - 4 = 9$
$f(x) = 9$
$\rightarrow x^*$ solves $h(x^*) = 0$ & $f(x^*) = 9$
$2x_1^* + 3x_2^* - 15 = 0$
$x_2^* = \frac{15 - 2x_1^*}{3}$, Put x_2^* in $f(x^*) = 9$
$(x_1^* - 1)^2 + x_2^{*2} - 4 = 9$
$(x_1^2 - 6x_1 + 9) = 0$
$x_1^* = 3$ $x_2^* = 3$ $x^* = (3,3)$
Now, we have to look at the optimality conditions:
$\nabla f(x^*) = \mu^* \cdot \nabla h(x^*)$
$\begin{pmatrix}4\\6\end{pmatrix} = \mu^*\begin{pmatrix}2\\3\end{pmatrix}$ $\therefore \mu^* = 2$
the necessary condition fullfilled.
Now, we calculate the solution
$L(x, \mu) = f(x) - \mu h(x)$
$\nabla L(x,\mu) = \begin{pmatrix} 2x_1 - 2 - 2\mu \\ 2x_2 - 3\mu \end{pmatrix} \overset{!}{=} \begin{pmatrix}0\\0\end{pmatrix} \begin{matrix}\to (1) \\ \to (2)\end{matrix}$
solve ... $2x_1 + 3x_2 - 15 = 0 \to (3)$ ... KKT ... $\left(\begin{pmatrix}3\\3\end{pmatrix}, 2\right)$

For the second order condition
$F_1(x^*) = \{d : \nabla h(x^*)^T d = 0\}$
$= \{d : \begin{pmatrix}2\\3\end{pmatrix}^T d = 0\}$
$= \{d : 2d_1 + 3d_2 = 0\}$
$= \{d : d_2 = -\frac{2}{3}d_1\}$
$F_2(x^*) = \{\omega \in F_1(x^*) : \omega^T g_i(x_i) = 0$ $\forall$ active, with $\lambda_i^* > 0\}$
no $g(x) = \{\omega \in F_1(x^*)\}$
$= \{\omega : \omega_2 = -\frac{2}{3}\omega_1\}$
Sufficient 2nd order condition
$\omega^T \nabla_x^2 L(x, \mu)\omega > 0$, $\forall \omega \in F_2(x^*) \setminus \{0\}$
$\begin{pmatrix}\omega_1 \\ -\frac{2}{3}\omega_1\end{pmatrix}^T \begin{pmatrix}2 & 0 \\ 0 & 2\end{pmatrix}\begin{pmatrix}\omega_1 \\ -\frac{2}{3}\omega_1\end{pmatrix}$
$2\omega_1^2 + \frac{8}{9}\omega_1^2 > 0$, $\omega_1 \neq 0$
we found a local sol^n
KKT $\left(\begin{pmatrix}3\\3\end{pmatrix}, 2\right)$

Derive the ...
$\min_{x\in C} (x_1 - 2)^2 + (x_2 - 1)^2$ $C = [-1,1]\times[-1,1]$
$\Rightarrow$ (i) $x_1 \ge -1$; $x_1 + 1 \ge 0$ —(1)
(a) ≤ 1 ; $-x_1 + 1 \ge 0$ —(2)
$x_2 \ge -1$; $x_2 + 1 \ge 0$ —(3)
≤ 1 ; $-x_2 + 1 \ge 0$ —(4)



(b) $L(x,\lambda) = f(x) - \sum_{i=1}^{n} \lambda_i g_i(x)$ for all i active
$\nabla_x L(x,\lambda) = \begin{pmatrix} 2x_1 - 4 - \lambda_1 + \lambda_2 \\ 2x_2 - 2 - \lambda_3 + \lambda_4 \end{pmatrix} = \begin{pmatrix}0\\0\end{pmatrix}$
$\lambda_1(x_1 + 1) = 0$, $\lambda_2(-x_1 + 1) = 0$
$\lambda_3(x_2 + 1) = 0$, $\lambda_4(-x_2 + 1) = 0$.
(i) $(-1,-1) = \left(\begin{pmatrix}-1\\-1\end{pmatrix}, -6, 0, -4, 0\right)$
$\left(\begin{pmatrix}1\\1\end{pmatrix}, 0, 2, 0, 0\right) \left(\begin{pmatrix}1\\-1\end{pmatrix}, 0, 2, -4, 0\right)$
$\left(\begin{pmatrix}-1\\1\end{pmatrix}, 0, -6, 0, 0\right)$
$f(-1,1) = 9$, $f(1,-1) = 5$
$f(-1,-1) = 13$, $f(1,1) = 1$ ✓
KKT $\left(\begin{pmatrix}1\\1\end{pmatrix}, 0, 2, 0, 0\right)$

1) Second order condition

$F_1(x^*) = \{d : \nabla h(x^*)^T d = 0, \nabla g_i(x^*)^T d \ge 0$ $\forall$active, $\lambda_i^* > 0\}$

Active i ⇒ (i) $-x_1+1 \ge 0$ (ii) $-x_2+1 \ge 0$

$\{d: \nabla g_1(x^*)^T d \ge 0,\ \nabla g_2(x^*)^T d \ge 0\}$

$\{d: (-1\ 0)\binom{d_1}{d_2} \ge 0 \quad (0\ -1)\binom{d_1}{d_2} \ge 0\}$

$\{d: -d_1 \ge 0 \quad -d_2 \ge 0\}$

$d_1 \le 0 \quad d_2 \le 0$

$F_2(x^*) = \{\omega \in F_1(x^*) : \omega^T \nabla g_i(x^*) = 0$ $\forall$active i $\lambda_i^* > 0\}$

$\{\omega: \omega_1 \le 0, \omega_2 \le 0, (-\omega_1\ -\omega_2)\binom{-1}{0} = 0\}$

$\omega_1 + 0 = 0 \Rightarrow \omega_1 = 0$

sufficient second order condition

$\omega^T \cdot \nabla_x^2 L(x,\lambda) \cdot \omega > 0 \quad \forall \omega \in F_2(x^*)$

$\nabla_x^2 L(x,\lambda) = \begin{pmatrix}2&0\\0&2\end{pmatrix}$

$\binom{0}{\omega_2}^T \begin{pmatrix}2&0\\0&2\end{pmatrix}\binom{0}{\omega_2} > 0$

$(0\ \omega_2)\binom{0}{2\omega_2} > 0 \Rightarrow 2\omega_2^2 > 0 \quad \forall \omega_2 \ne 0$

we found a local solution.

Let $f: R^2 \to R$ be defined by

$f(x) = \begin{cases} x_1 & \text{if } x_1 = x_2 \\ 0 & \text{otherwise} \end{cases}$ ... is not Gâteaux differentiable at (0,0).

$\frac{\partial f}{\partial x_1}\Big|_{(0,0)} = \lim_{t\to 0} \frac{f(0+t,0) - f(0,0)}{t} = \lim_{t\to 0} \frac{0-0}{t} = 0$

$\frac{\partial f}{\partial x_2}\Big|_{(0,0)} = \lim_{t\to 0} \frac{f(0,0+t) - f(0,0)}{t} = \lim_{t\to 0} \frac{0-0}{t} = 0$

Proof by contradiction. Assume that f is Gâteaux differentiable at (0,0). Then the Gâteaux derivative of f at (0,0) must be equal to zero. i.e. $(\delta f)(0,0) = 0$

Since f Gâteaux differentiable at $x_0 \in R^2$ it hold for any $U \in R^2$, $t \in R$

$f(x_0 + tU) - f(x_0) - t(\delta f)(x_0)U \in o(|t|)$

Let $\hat e = \frac{1}{\sqrt 2}(1,1)$ then $U = \hat e$ & $x_0 = (0,0)$

$f(t,\hat e) - f(0,0) \in o(|t|)$

$= \frac{1}{\sqrt 2} t \notin (|t|)$

→ This is contradiction, so, f is not Gâteaux diff^le at (0,0).

Let A be an n×n matrix & let $f: R^n \to R$ be func $f(x) = x^TAx$. Prove that $(\delta f)x_0 = x_0^T(A+A^T)$.

⇒ $\delta f(x_0)U$

$= \lim_{t\to 0} \frac{f(x_0+tU) - f(x_0)}{t}$

$= \lim_{t\to 0} \frac{(x_0+tU)^T A (x_0+tU) - x_0^TAx_0}{t}$

$= \lim_{t\to 0} \frac{x_0^TAx_0 + x_0^TA(tU) + (tU)^TAx_0 + (tU)^TA(tU) - x_0^TAx_0}{t}$

$= \lim_{t\to 0} \frac{x_0^TAtU + U^TtAx_0 + \cancel{U^TtAtU}}{t}$

$= x_0^TAU + U^TAx_0$

$= x_0^TAU + x_0^TA^TU = x_0^T(A+A^T)U$

∴ $(\delta f)x_0(U) = x_0^T(A+A^T)$ (∵ U is direction)

It is valid for all directions

Let $f: R^2 \to R$ defined by

$f(x,y) = \begin{cases} 0 & (x,y) = (0,0) \\ \frac{x^4y}{x^6+|y|^3} & x^2+y^2 > 0 \end{cases}$

Prove that f has 0 as it Gâteaux derivative at the origin but fails to be continuous there.

⇒ $f: R^2 \to R$

f is not continuous at (0,0) since for $\varepsilon \in R$ we have

$f(\varepsilon, \varepsilon^2) = \frac{\varepsilon^6}{\varepsilon^6 + \varepsilon^6} = \frac{1}{2} \ne 0 = f(0,0)$

⇒ f is Gâteaux differentiable at (0,0) with $(\delta f)(0,0) = 0$ since for all $U \in R^2$

$|f(t \cdot U_x, tU_y) - f(0,0) - t(\delta f)(0,0)U| \in O(t)$

$\left|\frac{t^4U_x^4 \cdot t \cdot U_y}{t^6U_x^6 + t^3U_y^3}\right| \le \left|\frac{t^5U_x^4U_y}{t^3U_y^3}\right| = \frac{t^2|U_x^4U_y|}{|U_y^3|} \in O(t)$

f is Gâteaux diff^ble but Not continuous

$\begin{pmatrix} 3x_1^2 \\ x_3^2 - 1 \\ 2x_2x_3 \\ \cos x_4 \end{pmatrix}$

$\hat f(x) = f(x) + \epsilon_f(x)$, (a) Approximate the gradient and the hessian of f with central difference.

$\nabla_h f(x) = \frac{\hat f(x+h) - \hat f(x-h)}{2h}$ →(*) ... error holds $\epsilon_{\nabla f} = O((\epsilon_{\nabla f})^{2/3}) = O((\epsilon_f)^{4/9})$.

our goal is to approximate ∇f & Hess(f) with (*)

$\hat f(x+h) = f(x+h) + \epsilon_f(x+h)$

$= f(x) + h\nabla f(x) + \frac{h^2}{2} hess(f(x)) + O(h^3) + \epsilon_f(x+h)$ →(1)

$\hat f(x-h) = f(x) - h\nabla f(x) + \frac{h^2}{2} hess(f(x)) + O(h^3) + \epsilon_f(x-h)$ →(2)

① & ② in (*)

$\nabla_h f(x) = \frac{2h\nabla f(x) + O(h^3) + \epsilon_f(x+h) - \epsilon_f(x-h)}{2h}$

$= \nabla f(x) + O(h^2) + O(\epsilon_f/h) \Rightarrow \nabla f(x) + O(h^2 + \epsilon_f/h)$

Balancing the error terms gives $h_{opt} = \epsilon_f^{1/3}$ (∵ $\frac{\epsilon_f}{h} = h^2 \Rightarrow h_{opt} = (\epsilon_f)^{1/3}$)

$= \nabla f(x) + O(\epsilon_f^{2/3})$ →(3)

By using the taylor expansion, we have

$\hat f(x+h) = \hat f(x) + h\nabla \hat f(x) + \frac{h^2}{2} Hess(\hat f(x)) + O(h^3)$

$\hat f(x-h) = \hat f(x) - h\nabla \hat f(x) + \frac{h^2}{2} Hess(\hat f(x)) + O(h^3)$

$\hat f(x+h) - \hat f(x-h) = 2h\nabla \hat f(x) + O(h^3)$

$\nabla \hat f(x) = \frac{\hat f(x+h) - \hat f(x-h) + O(h^3)}{2h}$

$\nabla \hat f(x) = \nabla_h f(x) + O(h^2)$

$\nabla \hat f(x) = \nabla f(x) + O(\epsilon_f^{2/3}) + O(h^2)$ (∵③)

$\nabla \hat f(x) = \nabla f(x) + O(\epsilon_f^{2/3})$

$\nabla \hat f(x) = \nabla f(x) + \epsilon_{\nabla f}(x)$ (∵ By Def^n)

∴ $\epsilon_{\nabla f}(x) = O(\epsilon_f^{2/3})$

Approximate the hessian by

$\nabla_h^2 f(x) = \frac{\nabla \hat f(x+h) - \nabla \hat f(x-h)}{2h}$ →(4)

$Hess(\hat f(x)) = Hess(f(x)) + \epsilon_{\nabla f}(x)$

$\nabla \hat f(x+h) = \nabla f(x+h) + \epsilon_{\nabla f}(x+h)$

$= \nabla f(x) + h \cdot Hess(f(x)) + \frac{h^2}{2} f'''(x) + O(h^3) + \epsilon_{\nabla f}(x+h)$ →(A)

$\nabla \hat f(x-h) = \nabla f(x) - h \cdot Hess(f(x)) + \frac{h^2}{2} f'''(x) + O(h^3) + \epsilon_{\nabla f}(x-h)$ →(B)

Ⓐ & Ⓑ in 4

$\nabla_h^2 f(x) = \frac{2h Hess(f(x)) + O(h^3) + \epsilon_{\nabla f}(x+h) - \epsilon_{\nabla f}(x-h)}{2h}$

$= Hess(f(x)) + O(h^2) + O\left(\frac{\epsilon_{\nabla f}}{h}\right)$

$= Hess(f(x)) + O\left(h^2 + \frac{\epsilon_{\nabla f}}{h}\right)$

Balancing the error $h_{opt} = \epsilon_{\nabla f}^{1/3}$

$\nabla_h^2 f(x) = Hess f(x) + O(h^2)$

By Taylor expansion

$\nabla \hat f(x+h) = \nabla \hat f(x) + h \cdot Hess(\hat f(x)) + \frac{1}{2} \hat f'''(x)h^2 + O(h^3)$

$\nabla \hat f(x-h) = \nabla \hat f(x) - h \cdot Hess(\hat f(x)) + \frac{1}{2} \hat f'''(x)h^2 + O(h^3)$

So, $\nabla \hat f(x+h) - \nabla \hat f(x-h) = 2h Hess(\hat f(x)) + O(h^3)$

→ devide by 2h both side

$\nabla_h^2 f(x) = Hess(\hat f(x)) + O(h^2)$ -(5)

(4) & (5) same

$Hess \hat f(x) = Hess(f(x)) + O\left(h^2 + \frac{\epsilon_{\nabla f}}{h}\right) + O(h^2)$

$= Hess(f(x)) + O(\epsilon_{\nabla f}^{2/3})$

but $\epsilon_{\nabla f} = \epsilon_f^{2/3}$

This means

$\boxed{\epsilon_{\nabla f} = O(\epsilon_{\nabla f}^{2/3}) = O(\epsilon_f^{4/9})}$

Condition for convexity

$f(\lambda x_1 + (1-\lambda)x_2) \le \lambda f(x_1) + (1-\lambda) f(x_2)$

$f(x) = x^2$, 3rd step

$\lambda^2 x_1^2 + 2\lambda(1-\lambda)x_1x_2 + x_2^2 - 2\lambda x_2^2 + \lambda^2 x_2^2$

$\le \lambda x_1^2 + x_2^2 - \lambda x_2^2$ (as left side)

multiplication

$(\lambda - \lambda^2)(x_1 - x_2)^2 \ge 0, \lambda \in [0,1]$

$h(x) = ||x||$

$||\lambda x_1 + (1-\lambda)x_2|| \le \lambda ||x_1|| + (1-\lambda)||x_2||$

$x_1 = 1$, $x_2 = 2$, $\lambda = 1/2$ $\quad \frac{3}{2} \le \frac{3}{2}$

∴ Not strictly convex but convex

$g(x) = x^3 \Rightarrow g'(x) = 3x^2 \Rightarrow g''(x) = 6x$

we can't decide bcz x is +ve or -ve.

x>0 (convex) x<0 (concave/Not convex)

$x_1 = -2 \quad x_2 = -3$

$g(-2\lambda + 3\lambda - 3) \le \lambda g(-2) + (1-\lambda) g(-3)$

$g(\lambda - 3) \le \lambda(-2)^3 + (1-\lambda)(-3)^3$

$\lambda = 0.5$ bcz $\lambda \in [0,1]$

$-15.625 \not\le -17.5$ (Not convex)

Let $f: X \to R$ be defined on the convex set $X \subset R^n$. Prove that f is convex, Iff its epigraph: $epi(f) = \{(x,\alpha) \in X \times R;\ f(x) \le \alpha\}$ is convex set.

⇒ Let f is convex and (x,a) $(y,b) \in epi(f)$

then f is convex ⇒

$f(\lambda x + (1-\lambda)y) \le \lambda f(x) + (1-\lambda)f(y)$ $\quad \lambda \in [0,1]$

$(\lambda x + (1-\lambda)y, \lambda a + (1-\lambda)b) \in epi f$

$\lambda(x,a) + (1-\lambda)(y,b) \in epi f$ → epi f is convex

⇐ let epigraph of be convex.

$\forall (x,a), (y,b) \in epi(f) \quad \forall \lambda \in [0,1]$

$(\lambda x + (1-\lambda)y, \lambda a + (1-\lambda)b) \in epi(f)$

Again let $x, y \in epi(f)$, $\lambda \in [0,1]$

by definition of convexity of epi(f)

$a \ge f(x)$, $b \ge f(y)$

$\lambda a + (1-\lambda)b \ge \lambda f(x) + (1-\lambda)f(y)$

$\lambda a + (1-\lambda)b \ge f(\lambda x + (1-\lambda)y)$

→ f is convex.

Rosenbrock fun^n, $f: R^2 \to R$

$f(x) = 100(x_2 - x_1^2)^2 + (1-x_1)^2$

has a global minimum, ...

→ $100x_2^2 - 200x_2x_1^2 + 100x_1^4 + 1 - 2x_1 + x_1^2$

$\frac{\partial f}{\partial x_1} = 400x_1^3 - 400x_2x_1 + 2x_1 - 2$

$\frac{\partial f}{\partial x_2} = 200x_2 - 200x_1^2 = 0 \Rightarrow x_2 = x_1^2$

$\frac{\partial f}{\partial x_1} = 0 \Rightarrow \cancel{400x_1^3} - \cancel{400x_1^3} + 2x_1 - 2 = 0 \quad \boxed{x_1 = 1}$

∴ $x_1 = x_2 = 1$

$f_{min}(1,1) = 100(1-1)^2 + (1-1)^2 = 0$

$\frac{\partial^2 f}{\partial x_1^2}\Big|_{(1,1)} = 1200x_1^2 - 400x_2 + 2 = 802$

$\frac{\partial^2 f}{\partial x_2^2}\Big|_{(1,1)} = 200 \qquad H(A) = \begin{bmatrix} 802 & -400 \\ -400 & 200 \end{bmatrix}$

$\frac{\partial^2 f}{\partial x_1 \partial x_2} = -400$

The orthogonal projection $P_X: R^n \to X$ for $X = \prod_{i=1}^n [a_i, b_i]$ with $a_i < b_i$ $(1 \le i \le n)$ is given as $P_X(x) = (\max(a_i, \min(x_i, b_i)))_{i=1}^n$

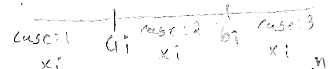
⇒ P_X is an orthogonal projection iff

⇔ $((x - P_X(x), y - P_X(x)) \le 0$, $\forall y \in X$, $x \in [a,b]$

case:1 $x_i < a_i$

case:2 $x_i \in (a_i, b_i)$

case:3 $x_i \ge b_i$



$P_X(x) = (\max(a_i, \min(x_i, b_i)))_{i=1}^n$

case:01 $P_X(x) = a_i$

〃 02 〃 $= x_i$

〃 03 〃 $= b_i$

case:01 $(x_i - a_i, y_i - a_i) \le 0$ $\quad \le 0 \quad \ge 0$

case:02 $(x_i - x_i, y_i - x_i) = 0$ $\quad = 0$

case:03 $(x_i - b_i, y_i - b_i) \le 0$ $\quad \ge 0 \quad \le 0$

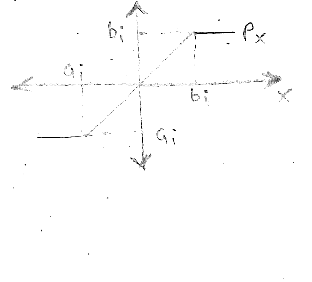
→ In all cases the components are non-positive and therefore for any $y \in X$ we get

$(x - P_X(x), y - P_X(x)) \le 0$

$\sum_{i=1}^n (x_i - P_X(x_i), y_i - P_X(x_i))$

$P_X(x)$ is an orthogonal projection

P_X is not differentiable due to the min & max operators.



(iii) Let $f(x)=x_1^2+3x_2^2-4$ &

$h(x)=x_1^2-x_2-2$

(a) KKT

$\Rightarrow L(x,\lambda,\mu)=f(x)-\mu h(x)-\sum_{i=1}^{n}\lambda_i g_i$

$=(x_1^2+3x_2^2-4)-\mu(x_1^2-x_2-2)$

* KKT conditions

(i) $\nabla_x L(x,\mu,\lambda)=0$

(ii) $\lambda_i g_i(x)=0$ $\forall i$ active

(iii) $h(x)=0$ (iv) $\lambda_i\geq 0$ for all i

(i) $\nabla_x L(x,\mu,\lambda)=\begin{pmatrix}2x_1-\mu(2x_1)\\ 6x_2-\mu(-1)\end{pmatrix}=\begin{pmatrix}0\\0\end{pmatrix}$

(ii) $x_1^2-x_2-2=0$ case 1 $x_1=0$ case 2 $\mu=1$

$2x_1(1-\mu)=0\Rightarrow$

case: 01, $x_1=0$, $x_2=-2$

$6x_2-\mu(-1)=0$

$6x_2+\mu=0$

$\mu=-12$ $\left(\begin{pmatrix}0\\-2\end{pmatrix},12\right)$

case: 02 $\mu=1$ તો $x_2=-1/6$

$x_1^2-(-1/6)-2=0\Rightarrow$

$\Rightarrow x_1=\pm\sqrt{11/6}$ $\left(\begin{pmatrix}\pm\sqrt{11/6}\\-1/6\end{pmatrix},1\right)$

2nd order condition

$F_1(x^*)=\{d:\nabla h(x^*)^T d=0\}$

$=\{d:\begin{pmatrix}2x_1\\-1\end{pmatrix}^T\begin{pmatrix}d_1\\d_2\end{pmatrix}=0\}$

$2x_1d_1-d_2=0$

$d_2=2x_1d_1$

$F_2(x^*)=\{\omega\in F_1(x^*)\}$

$=\{\omega:\omega_2=2x_1\omega_1\}$

Sufficient 2nd order condition:

$\omega^T\nabla_x^2L(x,\mu)\omega>0$, $\forall\omega\in F_2(x^*)\setminus\{0\}$

$\begin{pmatrix}\omega_1\\2x_1\omega_1\end{pmatrix}^T\begin{pmatrix}2-2\mu&0\\0&6\end{pmatrix}\begin{pmatrix}\omega_1\\2x_1\omega_1\end{pmatrix}>0$

$(\omega_1\ \ 2x_1\omega_1)\begin{pmatrix}\omega_1(2-2\mu)\\12x_1\omega_1\end{pmatrix}>0$

$\omega_1^2(2-2\mu)-24x_1^2\omega_1^2>0$

when, $\mu<1$ then this condⁿ satisfied.

$\min_{x\in R^2} -2x_1+x_2$ & $\begin{cases}(1-x_1)^3-x_2\geq 0\\ x_2+\frac14x_1^2-1\geq 0\end{cases}$

$x^*=(0,1)^T$ ← Linear Independent constraint Qualification.

(a) check LICQ at x^*

⇒ Condition for LICQ:

If $[\nabla g(x)\ \ \nabla h(x)]$ has a full ... then LICQ is ...

$\nabla g_1(x)=\begin{bmatrix}3(1-x_1)^2(-1)\\-1\end{bmatrix}\Rightarrow\nabla g_1(0,1)^T=\begin{bmatrix}-3\\-1\end{bmatrix}$

$\nabla g_2(x)=\begin{bmatrix}x_1/2\\1\end{bmatrix}\Rightarrow\nabla g_2(0,1)^T=\begin{bmatrix}0\\1\end{bmatrix}$

If $[\nabla g(x)\ \ \nabla h(x)]\in R^{n\times(p^*+\#\text{active inequalities})}$ — equality cons; No of variables in objective funⁿ

$A=\begin{bmatrix}-3&0\\-1&1\end{bmatrix}\in R^{2\times(0+2)}$

Matrix A has full rank Matrix.

∴ LICQ is fullfilled.

(b) KKT satisfied at $(0,1)^T$

$L(x,\mu,\lambda)=-2x_1x_2-\lambda_1[(1-x_1)^3-x_2]-\lambda_2[x_2+\frac14x_1^2-1]$

$\nabla L(x,\mu,\lambda)=0$

$\begin{pmatrix}-2+\lambda_1(3-6x_1+3x_1^2)-\frac{\lambda_2 x_1}{2}\\ 1+\lambda_1-\lambda_2\end{pmatrix}=\begin{pmatrix}0\\0\end{pmatrix}$ ① ②

$\lambda_1\{(1-x_1)^3-x_2\}=0$ ③

$\lambda_2\{x_2+\frac14x_1^2-1\}=0$ ④

Put $x^*=(0,1)^T$ $x_1=0$ $x_2=1$ (from ① & ②)

so, $\lambda_1=\frac23$, $\lambda_2=\frac53$

Let, $f:[a,b]\to R$ be continuous,

$\min_{x\in[a,b]} f(x)$ (a) s.t Problem (1) admits a solution.

⇒ f(x) is continuous & bounded, $f:[a,b]\to IR$

$\forall\varepsilon>0\ \exists P\ \forall x$ in $[a,b]$

$|f(x)-P(x)|<\varepsilon$

i.e admits the theorem of weirstaβ

→ A continuous funⁿ defined over a compact set admits a global-minimum.

(b) Derive necessary optimality conditions for local minᵐ x^* of (P)1.

⇒ for x^* to be a local minᵐ then $f'(x^*)=0$ & $f''(x^*)\geq 0$

(i) checking for the interior betⁿ boundary a & b.

1st order: for $x^*\in[a,b]$ $f'(x^*)=0$

2nd order $f''(x^*)\geq 0$

(ii) Checking at the boundary

$I\in[a,b]$

* for a minimum in a

$f(a+t(x-a))\geq f(a)$ $\forall x\in I$, t is small enough.

$\frac{f(a+t(x-a))-f(a)}{t}\geq 0$

$\lim_{t\to 0}\frac{f(a+t(x-a))-f(a)}{t}=f'(a)\cdot(x-a)\geq 0$

(c) Let f in addition be strictly convex. s.t (1) admits a unique solⁿ ... [a,b].

⇒ For unique solⁿ

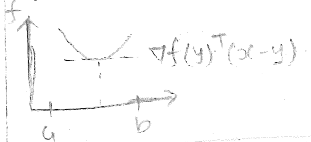
→ Let x & y be 2 solⁿ of f.

$\lambda x+(1-\lambda)y\in[a,b]$ for $\lambda\in[0,1]$

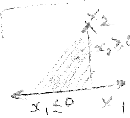
By defⁿ: $f(x)-f(y)>\nabla f(y)^T(x-y)$

1st order necessary condition $\nabla f(y)^T=0$

$f(x)-f(y)\geq 0\Rightarrow f(x)>f(y)$.



Calculate the tangential cone $T(X,x^*)$ for $X:=\{x\in R^2:-x_1\geq 0,\ x_2\geq 0\}$ in $x^*=(0,0)^T$



$-x_1\geq 0\Rightarrow x_1\leq 0$, $x_2\geq 0$

$T(X,x^*)=$

$\left\{d\in R^n:\lim_{K\to\infty}x^K-x^*\ \exists\mu_K>0\right\}$

$\lim_{K\to\infty}d^K=\mu_K(x^K-x^*)=d,\ x^K\in X$

$\lim_{K\to\infty}d^K=\mu_K\left(\begin{pmatrix}x_1\\x_2\end{pmatrix}-\begin{pmatrix}0\\0\end{pmatrix}\right)=d$

$\lim_{K\to\infty}d^K=\mu_K\begin{pmatrix}x_1\\x_2\end{pmatrix}=d$

$T(X,x^*)=\left\{\mu_K\begin{pmatrix}d_1\\d_2\end{pmatrix}:d_1\leq 0,\ d_2\geq 0,\ \mu_K>0\right\}$ → (1)

Tangential cone.

(b) ACQ fullfilled in x^*?

⇒ ACQ (Abadie Contraint Qualification.)

$T(X,x^*)=F_1(x^*)$

$F_1(x^*)=\{d\in R^n:\nabla g_i(x^*)^Td\geq 0,\ \nabla h(x)^Td=0\}$

$\nabla g_1(x^*)=\begin{pmatrix}-1\\0\end{pmatrix}$ $\nabla g_2(x^*)=\begin{pmatrix}0\\1\end{pmatrix}$

$F_1(x^*)=\left\{d\in IR^2:(-1,0)\begin{pmatrix}d_1\\d_2\end{pmatrix}\geq 0,\ (0,1)\begin{pmatrix}d_1\\d_2\end{pmatrix}\geq 0\right\}$

$-d_1\geq 0$ $d_2\geq 0$

$d_1\leq 0$

$F_1(x^*)=\{d:d_1\leq 0,\ d_2\geq 0\}$

$T(X,x^*)=F_1(x^*)$

Hence, ACQ is fullfilled.

(c) LICQ = ?

$\min_{x\in R^2} f(x)=2x_1^2+x_2^2+x_1$ s.t $x_1\geq 0$

(a) S.t $x^*=(0,0)^T$ with $\lambda^*=1$ is global solution.

$\Rightarrow L(x,\lambda)=2x_1^2+x_2^2+x_1-\lambda x_1$

KKT $\nabla_x L(x,\lambda)=\begin{pmatrix}4x_1+1-\lambda\\2x_2\end{pmatrix}=\begin{pmatrix}0\\0\end{pmatrix}$

$\lambda x_1=0$

$x_2=0$

$\lambda x_1=0\Rightarrow\lambda=0$ $x_1=0$.

case: 01 $\lambda=0$ | $\left(\begin{pmatrix}-1/4\\0\end{pmatrix},0\right)$

case: 02 $x_1=0$ | $\left(\begin{pmatrix}0\\0\end{pmatrix},1\right)$

for global minᵐ solution

$f(x^*)\leq f(x)$ $\forall x\in X$

$x^*=\begin{pmatrix}0\\0\end{pmatrix}\Rightarrow f(x^*)=0$ since $x\in IR^2$, $x_1\geq 0$, $f(x)>0$

$f(x^*)<f(x)$ Hence, x^* is the global minᵐ solⁿ with $\lambda^*=1$.

(b) For $\alpha>0$ find the global minᵐ $x(\alpha)$ of the penalty problem (P_α).

$\min_{x\in R^2}P_\alpha(x)=f(x)+\frac{\alpha}{2}\left[(\min[0,g(x)])^2+h(x)^2\right]$

$=2x_1^2+x_2^2+x_1+\frac{\alpha}{2}\left[(\min(0,x_1))^2\right]$

$\nabla P_\alpha(x)=\begin{pmatrix}4x_1+1+\alpha(\min(0,x_1))\\2x_2\end{pmatrix}=\begin{pmatrix}0\\0\end{pmatrix}$

$x_1<0$: $\begin{pmatrix}4x_1+1+\alpha x_1\\2x_2\end{pmatrix}=\begin{pmatrix}0\\0\end{pmatrix}$ | $x_1>0$: $\begin{pmatrix}4x_1+1\\2x_2\end{pmatrix}=\begin{pmatrix}0\\0\end{pmatrix}$

$x_2=0\Rightarrow x_1=-\frac{1}{4+\alpha}$

$x(\alpha)=\begin{pmatrix}-\frac{1}{4+\alpha}\\0\end{pmatrix}$ $\forall\alpha>0$

x(α) is not admissible set Bcz it doesn't satisfy the contraint $x_1\geq 0$

(c) $\lim_{\alpha\to\infty}x(\alpha)=x^*$, $-\lim_{\alpha\to\infty}\alpha\cdot\min\{0,g(x(\alpha))\}=\lambda^*$

$\Rightarrow\lim_{\alpha\to\infty}\begin{pmatrix}-\frac{1}{4+\alpha}\\0\end{pmatrix}=\begin{pmatrix}0\\0\end{pmatrix}=x^*$

$-\lim_{\alpha\to\infty}\alpha\left(\frac{-1}{4+\alpha}\right)\Rightarrow 1=\lambda^*$.

(d) Cond (Hess $(P_\alpha(x(\alpha)))$) = ?

$\Rightarrow$ Hess$(P_\alpha(x(\alpha)))=\begin{pmatrix}4+\alpha&0\\0&2\end{pmatrix}$

$\|\text{Hess}(P_\alpha(x(\alpha)))^{-1}\|=\begin{pmatrix}\frac{2}{8+2\alpha}&0\\0&\frac{4+\alpha}{8+2\alpha}\end{pmatrix}\Rightarrow\left\|\begin{matrix}\frac{1}{4+\alpha}&0\\0&\frac12\end{matrix}\right\|$

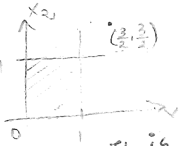
cond$(A)=\|A\|\,\|A^{-1}\|=(4+\alpha)\left(\frac12\right)$ as $\lim_{\alpha\to\infty}$ the number very large so bad condition matrix

$\min_{x\in R^2}(x_1-2)^2+(x_2-1)^2$, $x_2-x_1^2\geq 0$, $x_1-x_2=0$

(a) Draw admissible set.

$\Rightarrow x_1=x_2$, $x_2(1-x_2)\geq 0$ $0\leq x_1\leq 1$

$x_2\geq 0$ $1-x_2\geq 0$ $0\leq x_2\leq 1$



$f(x_1)=(x_1-2)^2+(x_1-1)^2$

$f(x_1)=2x_1^2-6x_1+5$

$\nabla f(x_1)=4x_1-6$

$\nabla f(x_1)=0\Rightarrow x_1=\frac32$ $\left(\frac32,\frac32\right)$

It is out of contraints so It is not solution. The solution is out of contraints so

→ Now we have to check at boundary.

$f(0,0)=5$ | $f(1,1)=1$ ✓ we have solⁿ at (1,1) which satisfies all contraints & minimizes the funⁿ.

$f(1,0)=2$ | $f(0,1)=4$

(b) KKT ? $L(x,\mu,\lambda)=f(x)-\mu h(x)-\sum_{i=1}^{n}\lambda_i g_i(x)$

$=x_1^2-2x_1+4+x_2^2-2x_2+1-\mu(x_1-x_2)-\lambda_1(x_2-x_1^2)$

KKT conditions

(i) $\nabla_x L(x,\mu,\lambda)=0$

(ii) $\lambda_i\cdot g_i(x_i)=0$ $\forall i$ active

(iii) $h(x)=0$, (iv) $\lambda_i\geq 0$ for all i.

(i) $\begin{pmatrix}2x_1-4-\mu+2x_1\lambda_1\\2x_2-2+\mu-\lambda_1\end{pmatrix}=\begin{pmatrix}0\\0\end{pmatrix}$

$\lambda_1(x_2-x_1^2)=0$, $x_1-x_2=0\Rightarrow x_1=x_2$

$\lambda_1(x_1-x_1^2)=0$

$\lambda_1=0$ | $x_1(1-x_1)=0$

$x_1=0=x_2$ $x_1=1=x_2$

Case: A

$\lambda_1=0$ & $x_1=x_2$, [illegible] $x_1=x_2=\frac32$ [illegible] $0\leq x_2\leq 1$ [illegible] violate [illegible] (doesn't exist)

Case: B(i) $x_1=x_2=0$, [illegible] $\lambda_1=-6$ [illegible] $\lambda_1\geq 0$ [illegible] violate [illegible]. doesn't exist

Case B(ii) $x_1=x_2=1$

KKT $\left(\begin{pmatrix}1\\1\end{pmatrix},\overset{\mu}{2},\overset{\lambda}{2}\right)$

$\min_{x\in R^2} x_1^2+4x_2+x_2^2$ s.t $x_2\geq 0$

$\Rightarrow L(x,\lambda)=x_1^2+4x_2+x_2^2-\lambda x_2$

$\nabla_x L(x,\lambda)=\begin{pmatrix}2x_1\\4+2x_2-\lambda\end{pmatrix}=\begin{pmatrix}0\\0\end{pmatrix}$

$\lambda x_2=0$

* 2 cases

(i) $\lambda=0\Rightarrow 4+2x_2-0=0$ $x_2=-2$ $\left(\begin{pmatrix}0\\-2\end{pmatrix},0\right)$

(ii) $x_2=0\Rightarrow 4-\lambda=0$ $\lambda=4$ $\left(\begin{pmatrix}0\\0\end{pmatrix},4\right)$

(2) $\min_{x\in R^2}P_\alpha(x)=x_1^2+4x_2+x_2^2+\frac{\alpha}{2}\left[(\min(0,g(x)))^2+h(x)^2\right]$

$\Rightarrow\nabla P_\alpha(x)=\begin{pmatrix}2x_1\\4+2x_2+\alpha\cdot\min(0,x_2)\end{pmatrix}=\begin{pmatrix}0\\0\end{pmatrix}$

$=\begin{pmatrix}2x_1\\4+2x_2+\alpha x_2\end{pmatrix}=\begin{pmatrix}0\\0\end{pmatrix}$ $x_2<0$

$\min_{x\in R^2} x_1^2+\frac34x_2$ s.t $x_1\geq 0$, $x_2\geq 0$.

(a) Find $x(\alpha)$ that solves the associated logarithmic barrier Problem (P_α) & calculate $\lambda_1(\alpha)$ & $\lambda_2(\alpha)$. where $\lambda_i(\alpha)=\alpha/x_i(\alpha)$.

$\Rightarrow$ (P) $\min_{x\in\Omega}f(x)$, $\Omega=\{x\in R^n, g(x)\geq 0; h(x)=0\}$

Define Barrier funⁿ

$\beta(x)=-\ln g(x)$, $g(x)>0$

$P_\alpha=\min_{x\in R^n}\{f(x)+\alpha\cdot\beta(x)\}$ for $\alpha>0$

$P_\alpha=x_1^2+\frac34x_2+\alpha[-\ln x_1-\ln x_2]$

$\nabla P_\alpha=\begin{pmatrix}2x_1-\frac{\alpha}{x_1}\\ \frac34-\frac{\alpha}{x_2}\end{pmatrix}=\begin{pmatrix}0\\0\end{pmatrix}$

$x_1=\pm\sqrt{\frac{\alpha}{2}}\Rightarrow x_1=\sqrt{\frac{\alpha}{2}}$

(∵ − is Not admissible)

$x_2=\frac{4\alpha}{3}$, we have $\lambda_i(\alpha)=\frac{\alpha}{x_i(\alpha)}$

$\therefore\lambda_1(\alpha)=\boxed{\sqrt{2\alpha}}$ & $\lambda_2(\alpha)=\boxed{\frac34}$

(b) KKT

$L(x,\lambda)=x_1^2+\frac34x_2-\lambda_1x_1-\lambda_2x_2$

$\nabla_x L(x,\lambda)=\begin{pmatrix}2x_1-\lambda_1\\ \frac34-\lambda_2\end{pmatrix}=\begin{pmatrix}0\\0\end{pmatrix}$

$\lambda_1x_1=0$

$\lambda_2x_2=0$

$\left(\begin{pmatrix}0\\0\end{pmatrix},0,\frac34\right)$

$\lim_{\alpha\to 0}x_1(\alpha)=\sqrt{\frac{\alpha}{2}}=0$

$x_2(\alpha)=\frac{4\alpha}{3}=0$ $\Big\}x^*$

$\lambda_1(\alpha)=\sqrt{2\alpha}=0=\lambda_1^*$

$\lambda_2(\alpha)=\frac34=\frac34=\lambda_2^*$

(c) $x^*=\begin{pmatrix}0\\0\end{pmatrix}$ satisfy $\min f(x)=0$ $\forall x\geq 0$

$\lambda_1^*=0$ $\lambda_i=0$

$\lambda_2^*=\frac34$ $\lambda_i>0$ $\Big\}\lambda_i\geq 0$

Hence $\left(\begin{pmatrix}0\\0\end{pmatrix},0,\frac34\right)$ systems satisfy all KKT conditions.

$x(\alpha)=\begin{pmatrix}0\\-4/2+\alpha\end{pmatrix}$

$\lim_{\alpha\to\infty}x(\alpha)\Rightarrow\lim_{\alpha\to\infty}\begin{pmatrix}0\\ \frac{-4}{2+\alpha}\end{pmatrix}=\begin{pmatrix}0\\0\end{pmatrix}=x^*$

$\lambda(\alpha)=\lim_{\alpha\to\infty}-\alpha\min\{0,g(x(\alpha))\}$

$=\lim_{\alpha\to\infty}-\alpha\left(\frac{-4}{2+\alpha}\right)$

$=4$

$=\lambda^*$ of (P)

Exercise 1: (4 points (1+1+2)) Consider the following questions related to the lectures of Professor Sanjay.

a) Let $f(x) := x^{-\alpha}$. For which $\alpha \in \mathbb{R}$ is f an element of $L^p((0,1))$, where $1 \leq p < \infty$ is fix?

b) Give $f \in L^1_{loc}((0,1))$, which is not an element of $L^1((0,1))$.

c) Show that $f(x) := \max\{0, x\}$ is weakly differentiable as a function on $\mathbb{R}$ and give its weak derivative.

Exercise 2: (4 points) The risk in a portfolio of three stocks is given by

$$R(x_1, x_2, x_3) := \frac{1}{2}x_1^2 + x_2^2 + \frac{1}{2}x_3^2,$$

where $x_j \in [0,1]$ for $j = 1,2,3$ are the shares invested in the j-th stock, such that $x_1 + x_2 + x_3 = 1$ holds. With which shares does one obtain the minimum risk?

Exercise 3: (4 points (3+1)) Let $f(x) := \frac{1}{2}x^t Gx - b^t x + c$ with $G \in \mathbb{R}^{n \times n}$, $b \in \mathbb{R}^n$ and $c \in \mathbb{R}$. Furthermore let $A \in \mathbb{R}^{m \times n}$ have full rank with $m \leq n$ and $d \in \mathbb{R}^m$.

a) Write down the KKT system for

$$\min_{x \in \mathbb{R}^n} f(x), \quad \text{such that} \quad Ax = d. \tag{1}$$

Show that the KKT system to (1) has exactly one solution (x^*, μ^*), if $G \succ 0$ on $\ker(A)$, i.e. if G is symmetric and positive definite on the null space of A.

b) Let (x^*, μ^*) be the solution of the KKT system related to (1). Is x^* the unique solution to problem (1)? Explain your answer.

Exercise 4: (4 points (2+1+1))

a) Calculate the tangential cone $T(X, x^*)$ for $X := \{x \in \mathbb{R}^2 : -x_1 \geq 0, x_2 \geq 0\}$ in $x^* = (0,0)^\top$.

b) Is the (ACQ) fulfilled in x^*?

c) Let $g_1(x) = (1 - x_1)^3 - x_2 \geq 0$ and $g_2(x) = x_2 + \frac{1}{4}x_1^2 - 1 \geq 0$. Is the (LICQ) fulfilled in $\hat{x} = (0,1)^\top$?

Exercise 5:(4 points(2+1+1)) Consider the minimization problem

$$\min_{x \in \mathbb{R}^2} f(x) := x_1^2 + 4x_2 + x_2^2 \quad \text{s.t.} \quad g(x) := x_2 \geq 0.$$

a) Show that $x^* = (0,0)^\top$ with multiplier $\lambda^* = 4$ is the global solution and satisfies the KKT system.

b) For $\alpha > 0$ find the global minimum $x(\alpha)$ of the penalty problem (P_α).

c) Prove that $\lim_{\alpha \to \infty} x(\alpha) = x^*$ and $-\lim_{\alpha \to \infty} \alpha \cdot \min\{0, g(x(\alpha))\} = \lambda^*$.

0.1 0.2 0.3 0.4

0.4 > 0.2

3 grade

Passed

Exercise 1: (4 points (2+1+1))Let $f(x) := x_1^2 + 3x_2^2 - 4$ and $h(x) := x_1^2 - x_2 - 2$. Consider minimizing f under $h = 0$.

a) Give the KKT system of this minimization problem.

b) Compute the critical points of the minimization problem together with the multiplier.

c) Check with the second order sufficient optimality conditions which critical points are locally minimal, and which are locally maximal.

Exercise 2: (4 points) Show that a local solution x^* of a convex programme is also a global solution.

Exercise 3: (4 points (2+1+1)) Let $f : \mathbb{R} \to \mathbb{R}, f(x) := -\frac{1}{4}x^4$ with starting point $x^0 \in \mathbb{R}$ arbitrary.

a) Show that the sequence (x^k) generated by Newton's method converges to $x^* = 0$.

b) What kind of extremum is x^*?

c) Show that the convergence is q-linear.

Exercise 4: (4 points (2+2)) Consider

$$\min_{x \in \mathbb{R}^2} f(x) = -2x_1 + x_2 \quad \text{bei} \quad \begin{cases} (1-x_1)^3 - x_2 & \geq 0, \\ x_2 + \frac{1}{4}x_1^2 - 1 & \geq 0, \end{cases}$$

whose solution is $x^* = (0,1)^T$.

a) Check LICQ at x^*.

b) Check whether the KKT conditions are satisfied at x^*, and give the Lagrange multipliers.

Exercise 5: (4 points (2+2))Consider a twice continuously differentiable, scalar function $f : \mathbb{R} \to \mathbb{R}$ and the minimization problem

$$\min_{x \in I} f(x),$$

where $I = [a, b]$ is a closed interval.

a) What are the necessary conditions for an extremum? Distinguish the cases of extremum in the interior $\mathring{I}$ and on the boundary ∂I and give a graphical sketch.

b) How can you determine whether an extremum $x \in \mathring{I}$ is a minimum or a maximum?

OPTIMIZATION-I

2021 1st

Exercise 1: (4 points (1+2+1)) Let $f : [a, b] \to \mathbb{R}$ be continuous. Consider the minimization problem

$$\min_{x \in [a,b]} f(x) \tag{1}$$

a) Show that problem (1) admits a solution.

b) Derive the necessary optimality conditions for a local minimum x^* of problem (1).

c) Let f in addition be strictly convex. Show that problem (1) admits a unique solution and sketch its possible locations in the interval $[a, b]$.

(16)
3.1 **Exercise 2**: (4 points) Show that a local solution x^* of a convex programme is also a global solution.

Exercise 3: (4 points (3+1)) Let $f(x) := \frac{1}{2}x^t G x - b^t x + c$ with $G \in \mathbb{R}^{n \times n}$, $b \in \mathbb{R}^n$ and $c \in \mathbb{R}$. Furthermore let $A \in \mathbb{R}^{m \times n}$ have full rank with $m \leq n$ and $d \in \mathbb{R}^m$.

a) Write down the KKT system for — Karush-Kuhn-Tucker conditions

$$\min_{x \in \mathbb{R}^n} f(x), \quad \text{such that} \quad Ax = d. \tag{2}$$

Show that the KKT system to (2) has exactly one solution (x^*, μ^*), if $G \succ 0$ on $\ker(A)$, i.e. if G is symmetric and positive definite on the null space of A.

b) Let (x^*, μ^*) be the solution of the KKT system related to (2). Is x^* the unique solution to problem (2)? Explain your answer.

Exercise 4: (4 points (2+1+1))

a) Calculate the tangential cone $T(X, x^*)$ for $X := \{x \in \mathbb{R}^2 : -x_1 \geq 0, x_2 \geq 0\}$ in $x^* = (0, 0)^\top$.

b) Is the (ACQ) fulfilled in x^*?

c) Let $g_1(x) = (1 - x_1)^3 - x_2 \geq 0$ and $g_2(x) = x_2 + \frac{1}{4}x_1^2 - 1 \geq 0$. Is the (LICQ) fulfilled in $\hat{x} = (0, 1)^\top$?

Linear Independent Constraint Qualification

Exercise 5:(4 points(1+1+1+1)) Consider the minimization problem

$$\min_{x \in \mathbb{R}^2} f(x) := 2x_1^2 + x_2^2 + x_1 \quad \text{s.t.} \quad x_1 \geq 0.$$

a) Show that $x^* = (0, 0)^\top$ with multiplier $\lambda^* = 1$ is the global solution. (global minimum)

b) For $\alpha > 0$ find the global minimum $x(\alpha)$ of the penalty problem (P_α).
unconstraining problem.

c) Prove that $\lim_{\alpha \to \infty} x(\alpha) = x^*$ and $-\lim_{\alpha \to \infty} \alpha \cdot \min\{0, g(x(\alpha))\} = \lambda^*$.

d) Calculate cond(Hess($P_\alpha(x(\alpha))$)). What happens if $\alpha \to \infty$?

[11]

OPTIMIZATION - I

Bhavik

21 October 2019

Exercise 1: (4 points (1+2+1)) Let $f : [a, b] \to \mathbb{R}$ be continuous. Consider the minimization problem

$$\min_{x \in [a,b]} f(x) \tag{1}$$

a) Show that problem (1) admits a solution.

b) Derive the necessary optimality conditions for a local minimum x^* of problem (1).

c) Let f in addition be strictly convex. Show that problem (1) admits a unique solution. Sketch all possible locations of this solution in the interval $[a, b]$.

Exercise 2: (4 points (2+2)) Let $f(x) = \frac{1}{2}x^\top Ax - b^\top x + c$ be given with A spd and $d \in \mathbb{R}^n$, such that $\nabla f(x)^\top d < 0$. Show that

a) $t_{\min} := -\frac{\nabla f(x)^\top d}{d^\top Ad}$ gives $f(x + t_{\min} d) \leq f(x + td) \quad \forall t \in \mathbb{R}$,

b) $T(x, d) := \{t_{\min}\}$ is efficient, i.e. there exists $\theta > 0$, such that

$$f(x + t_{\min} d) \leq f(x) - \theta \left(\frac{\nabla f(x)^\top d}{\|d\|} \right)^2.$$

Here you can choose $\theta = \frac{1}{2\,\lambda_{\max}(A)}$.

Exercise 3: (4 points) Show that the local Newton method for optimization problems is invariant under affine-linear transformations $x = Ay + b$, where $A \in \mathbb{R}^{n \times n}$ is invertible and $b \in \mathbb{R}^n$. This means that if the Newton method applied to $f(x)$ generates the sequence (x^k) and the Newton method applied to $g(y) := f(Ay + b)$ generates the sequence (y^k), then

$$x^0 = Ay^0 + b \quad \Rightarrow \quad x^k = Ay^k + b \quad \forall k \in \mathbb{N}.$$

Exercise 4:(4 points (1+3)) Consider the minimization problem

$$\min_{x \in \mathbb{R}^2} f(x) := (x_1 - 2)^2 + (x_2 - 1)^2 \quad \text{s.t.} \quad x_2 - x_1^2 \geq 0,\ x_1 - x_2 = 0.$$

a) Draw the admissible set and find the solution x^* of the problem graphically.
(or) Feasible set

b) Solve the problem analytically, e.g. by using the Karush-Kuhn-Tucker conditions or a reduction technique.

Exercise 5:(4 points (2+1+1)) Consider the minimization problem

$$\min_{x \in \mathbb{R}^2} f(x) := x_1^2 + \frac{3}{4}x_2 \quad \text{s.t.} \quad x_1 \geq 0,\ x_2 \geq 0. \tag{2}$$

Global minimum (0,0)

a) Find $x(\alpha)$ that solves the associated logarithmic barrier problem (B_α) and calculate $\lambda_1(\alpha)$ and $\lambda_2(\alpha)$, where $\lambda_i(\alpha) := \alpha / x_i(\alpha)$. (Interior Point Method)

b) Show that $(x^*, \lambda_1^*, \lambda_2^*) := \lim_{\alpha \to 0} (x(\alpha), \lambda_1(\alpha), \lambda_2(\alpha))$ exists and compute the limit.

c) Does $(x^*, \lambda_1^*, \lambda_2^*)$ satisfy the KKT conditions of the minimization problem (2)?