SOLUTIONS OF HOMEWORK I

1) i) $f(x_1, x_2) = e^{x_2} x_1^2 + \sin(x_1)\, x_2^2$, $x_1, x_2 \in \mathbb{R}$

$$\nabla f = \begin{bmatrix} 2x_1 e^{x_2} + x_2^2 \cos(x_1) \\ x_1^2 e^{x_2} + 2x_2 \sin(x_1) \end{bmatrix}$$

$$\nabla^2 f = \begin{bmatrix} 2e^{x_2} - \sin(x_1)\, x_2^2 & 2x_1 e^{x_2} + 2x_2 \cos(x_1) \\ 2x_1 e^{x_2} + 2x_2 \cos(x_1) & x_1^2 e^{x_2} + 2\sin(x_1) \end{bmatrix}$$

ii) $f: \mathbb{R}^n \to \mathbb{R}$, $f(x) = x^T x = x_1^2 + x_2^2 + \cdots x_n^2$

$$\Rightarrow \nabla f = \begin{bmatrix} 2x_1 \\ 2x_2 \\ \vdots \\ 2x_n \end{bmatrix} , \quad \nabla^2 f = \begin{bmatrix} 2 & 0 & \cdots & 0 \\ 0 & 2 & \cdots & 0 \\ \vdots & & \ddots & \\ 0 & \cdots & & 2 \end{bmatrix} = 2 . I_{n \times n}$$

iii) $f: \mathbb{R}^n \to \mathbb{R}$, $f(x) = (x^T x) . e^{x^T x} \Rightarrow$

$$f(x) = (x_1^2 + x_2^2 + \cdots x_n^2)\, e^{(x_1^2 + x_2^2 \cdots + x_n^2)}$$

$$\nabla f = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{bmatrix} = \begin{bmatrix} 2x_1 e^{(x_1^2 + \cdots x_n^2)} + (x_1^2 + \cdots x_n^2)\, 2x_1 e^{(x_1^2 + \cdots x_n^2)} \\ 2x_2 e^{(x_1^2 + \cdots x_n^2)} + (x_1^2 + \cdots x_n^2)\, 2x_2 e^{(x_1^2 + \cdots x_n^2)} \\ \vdots \\ 2x_n e^{(x_1^2 + \cdots + x_n^2)} + (x_1^2 + \cdots + x_n^2)\, 2x_n e^{(x_1^2 + \cdots + x_n^2)} \end{bmatrix}$$

$$\nabla^2 f = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_2 \partial x_1} & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_1} \\ \vdots & \frac{\partial^2 f}{\partial x_2^2} & & \vdots \\ \frac{\partial^2 f}{\partial x_1 \partial x_n} & \cdots & & \frac{\partial^2 f}{\partial x_n^2} \end{bmatrix}$$

$$\frac{\partial^2 f}{\partial x_j^2} = 2e^{\|x\|_2^2} + 4x_j^2 e^{\|x\|_2^2} + (6x_j^2 + 2x_2 + \dots 2x_n) e^{\|x\|_2^2} + \|x\|_2^2 4x_j^2 e^{\|x\|_2^2}$$

$$\frac{\partial f}{\partial x_i \partial x_j} = 4x_i x_j e^{\|x\|_2^2} + 4x_i x_j e^{\|x\|_2^2} + 4x_i x_j \|x\|_2^2 e^{\|x\|_2^2}$$

iv) $f: \mathbb{R}^n \to \mathbb{R}$, $f(x) = b^T x$, $b \in \mathbb{R}^n$

$f(x) = b_1x_1 + b_2x_2 + \dots b_nx_n$, then $\nabla f = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} = b$

$\nabla^2 f = 0_{n\times n}$

v) $f: \mathbb{R}^n \to \mathbb{R}$, $f(x) = \frac{1}{2} x^T A x + b^T x + c$,

where $b \in \mathbb{R}^n$, $A_{n\times n}$ and $c \in \mathbb{R}$.

$$f(x) = \frac{1}{2} \sum_{j=1}^{n} \sum_{i=1}^{n} a_{ij} x_i x_j + b^T x + c$$

$$\Rightarrow \nabla f(x) = \frac{1}{2}(A + A^T)x + b$$

$$\Rightarrow \nabla^2 f(x) = \frac{1}{2}(A + A^T)$$

2- $f(x) = x_1^4 + x_1 x_2 + (1+x_2)^2$ 3

$\nabla f(x) = \begin{bmatrix} 4x_1^3 + x_2 \\ x_1 + 2(1+x_2) \end{bmatrix}$, $\nabla^2 f(x) = \begin{bmatrix} 12x_1^2 & 1 \\ 1 & 2 \end{bmatrix}$

a) Since $\exists$ points x, y (on the neigbourhood of $(0,0)$) for all neigbourhood of $(0,0)$ such that $f(x) < f(0,0) < f(y)$. (we can deduce it from contour diagram) ;

b) Since $\nabla^2 f(0) = \begin{bmatrix} 0 & 1 \\ 1 & 2 \end{bmatrix} \not\succ 0$

So $\exists\ p \in \mathbb{R}^2$ s.t $p^T \nabla^2 f(0) p \leq 0$

c) Matlab

3- $\phi(c_1,c_2)(t) = c_1 e^{c_2 t}$

$\phi(c_1,c_2)(1) = c_1 e^{c_2} = 10$

$\phi(c_1,c_2)(2) = c_1 e^{2c_2} = 50$

let $\underbrace{f(c)}_{\text{objective function}} = \underbrace{(c_1 e^{c_2} - 10)^2}_{(r_{(c_1,c_2)}(1))^2} + \underbrace{(c_1 e^{2c_2} - 50)^2}_{(r_{(c_1,c_2)}(2))^2}$, $c = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}$

where $r(c_1,c_2)(t) = |y_t - \phi(c_1,c_2)(t)|$

where y_t is measured displacement at time t

b) $\nabla f(c) = \begin{bmatrix} 2\left(e^{c_2}(c_1e^{c_2}-10) + e^{2c_2}(c_1e^{2c_2}-50)\right) \\ 2\left(c_1e^{c_2}(c_1e^{c_2}-10) + 2c_1e^{2c_2}(c_1e^{2c_2}-50)\right) \end{bmatrix}$

If c^* is local minimizer then

$\nabla f(c^*) = 0$, i.e

$$2e^{c_2^*}(c_1^*e^{c_2^*}-10 + e^{c_2^*}(c_1^*e^{2c_2^*}-50) = 0$$

$$2c_1e^{c_2}(c_1^*e^{c_2^*}-10 + e^{c_2^*}(c_1^*e^{2c_2^*}-50) = 0$$

4 –

$$\|v\|_\infty = \max_{1\le i\le n}|v_i|$$

$$\|v\|_1 = \sum_{i=1}^{n}|v_i|$$

a) If $\|v\|_\infty = \max_{1\le i\le n}|v_i| = 0$

i) then $|v_i| = 0$ for $i = 1 \dots n$, i.e $v = \vec{0}$

If $v \neq \vec{0}$ then $\exists i$ s.t $v_i \neq 0$, $(i = 1 \dots n)$

since $\|v\|_\infty = \max_{1\le i\le n}|v_i|$, Hence $\|v\|_\infty > 0$

Therefore positivity holds

ii) Let $c \in \mathbb{R}$, $\|cv\|_\infty = \max_{1\le i\le n}|cv_i| = \max_{1\le i\le n}|c||v_i|$

$= |c| \max |v_i| = |c| \, \|v\|_\infty$, Homogeneity holds

iii) $\|v+w\|_\infty = \max\limits_{1\le i\le n} |v_i + w_i|$

Since $|v_i + w_i| \le |v_i| + |w_i|$, then

$\max\limits_{1\le i\le n} |v_i + w_i| \le \max\limits_{1\le i\le n} (|v_i| + |w_i|) = \max\limits_{1\le i\le n} |v_i| + \max\limits_{1\le i\le n} |w_i|$

So, $\|w+v\|_\infty \le \|w\|_\infty + \|v\|_\infty$

Hence triangular inequality holds.

therefore $\| \ \|_\infty$ is a vector norm

b) i) Let $\|v\|_1 = \sum\limits_{i=1}^{n} |v_i| = 0$, then $|v_i| = 0$ for all $i = 1, \dots, n$

i.e $v = \vec{0}$, Let $v \neq \vec{0}$ then $\exists\, i \in 1, \dots, n$

s.t $|v_i| \neq 0$ which shows that $\|v\|_1 > 0$ ✓

(positivity)

ii) Let $c \in \mathbb{R}$, $\|cv\|_1 = \sum\limits_{i=1}^{n} |cv_i| = \sum\limits_{i=1}^{n} |c||v_i| = |c| \sum\limits_{i=1}^{n} |v_i|$

$= |c| \, \|v\|_1$ ✓ (homogeneity

iii) Let $v, w \in \mathbb{R}^n$, then $\|v+w\|_1 = \sum\limits_{i=1}^{n} |v_i + w_i|$

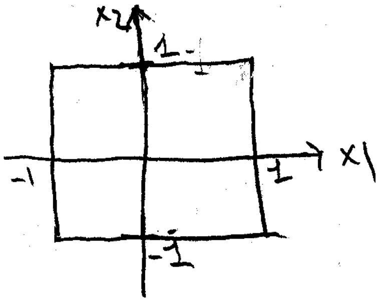
$\le \sum\limits_{i=1}^{n} |v_i| + |w_i| = \sum\limits_{i=1}^{n} |v_i| + \sum\limits_{i=1}^{n} |w_i| = \|v\|_1 + \|w\|_1$ ✓

(Triangle inequality)

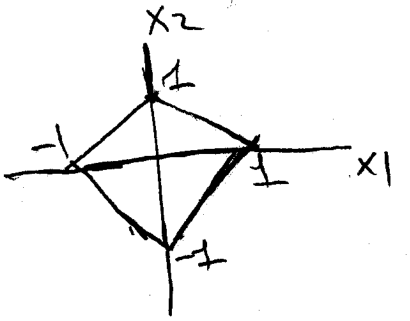
So, $\|\cdot\|_1$ is a vector norm

c) ⑥

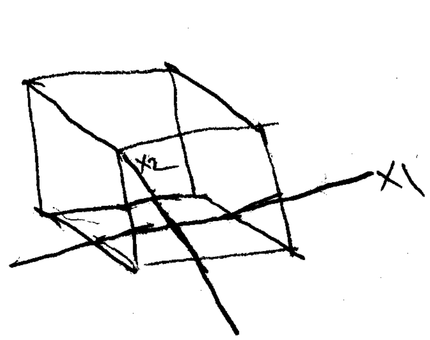
$S_2^{\infty} = \{v \in \mathbb{R}^2 : \|v\|_\infty = 1\}$ is square with the side length 2 and origin $(0,0)$ in $\mathbb{R}^2$



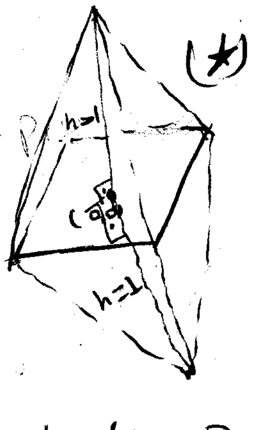
$S_2^1 = \{v \in \mathbb{R}^2 : \|v\|_1 = 1\}$ is the square with the length side $\sqrt{2}$ and origin $(0,0)$ in $\mathbb{R}^2$



$S_3^{\infty} = \{v \in \mathbb{R}^3 : \|v\|_\infty = 1\}$ is the cube with length side 2 and origin $(0,0)$ in $\mathbb{R}^3$



$S_3^1 = \{v \in \mathbb{R}^3 : \|v\|_1 = 1\}$



is the shape (*) with hight 1 in both direction $\downarrow\uparrow$

also, base is square with the side length 2 and

5- $l(\alpha) = f(x+\alpha p)$, $L: \mathbb{R} \to \mathbb{R}^2$, $L(\alpha) = \nabla f(x+\alpha p)$

$f: \mathbb{R}^2 \to \mathbb{R}$, $f(x) = [x_1\ x_2] \underbrace{\begin{bmatrix}2 & -1\\ 1 & 3\end{bmatrix}}_{A} \begin{bmatrix}x_1\\x_2\end{bmatrix} + \underbrace{[4\ -2]}_{b} \begin{bmatrix}x_1\\x_2\end{bmatrix}$

$l'(\alpha) = \nabla f(x+\alpha p)^T p$, $\nabla f(x) = (A+A^T)x + b$

$$\Rightarrow \nabla f(x+\alpha p)^T = \left[\left(\begin{bmatrix}2&-1\\1&3\end{bmatrix} + \begin{bmatrix}2&1\\-1&3\end{bmatrix}\right)(x+\alpha p) + \begin{bmatrix}4\\-2\end{bmatrix}\right]^T$$

$$\Rightarrow l'(\alpha) = \left((x+\alpha p)^T \begin{bmatrix}4&0\\0&6\end{bmatrix} + [4\ -2]\right) p$$

(where $p = \begin{bmatrix}1\\-1\end{bmatrix}$)

$$= [x_1+\alpha, x_2-\alpha]\begin{bmatrix}4\\-6\end{bmatrix} + 6 = 4x_1 - 6x_2 + 10\alpha + 6$$

$$l''(\alpha) = p^T \nabla^2 f(x+\alpha p)\, p = p^T (A+A^T)\, p$$

$$= [1\ -1]\begin{bmatrix}4&0\\0&6\end{bmatrix}\begin{bmatrix}1\\-1\end{bmatrix} = 10$$

$$L'(\alpha) = \nabla f(x+\alpha p)' = \nabla^2 f(x+\alpha p)\, p = (A+A^T)\, p$$

$$= \begin{bmatrix}4&0\\0&6\end{bmatrix}\begin{bmatrix}1\\-1\end{bmatrix} = \begin{bmatrix}4\\-6\end{bmatrix}$$

⑧

6- $f: \mathbb{R}^n \to \mathbb{R}$ be a twice cont. diff

let $\ell(\alpha): \mathbb{R} \to \mathbb{R}$ s.t $\ell(\alpha) = f(x + \alpha p)$

then $\ell(1) = \ell(0) + \ell'(0) + \frac{1}{2}\ell''(t)$

$t \in (0, 1)$

by Taylor's Theorem.

Equivalently,

$$f(x+p) = f(x) + \nabla f(x)^T p + \frac{1}{2} p^T \nabla f(x+tp) p$$

as we desire.

7-

$$f(x_1, x_2) = x_1^2 + 4x_2^2 - 4x_1x_2 + x_1 + x_2 + 4$$

a) $f(x_1, x_2) = \frac{1}{2} x^T \begin{bmatrix} 2 & -4 \\ -4 & 8 \end{bmatrix} x + [1 \; 1] x + 4$

b) $\nabla f(x_1, x_2) = Ax + b = \begin{bmatrix} 2 & -4 \\ -4 & 8 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

$$= \begin{bmatrix} 2x_1 - 4x_2 + 1 \\ -4x_1 + 8x_2 + 1 \end{bmatrix}$$

Let $\nabla f(x) = 0$, then $\left.\begin{matrix} 2x_1 - 4x_2 + 1 = 0 \\ -4x_1 + 8x_2 + 1 = 0 \end{matrix}\right\} \Rightarrow 2 = 1 \Rightarrow$ contradiction ... has no stationary

$g(x_1,x_2) = 2x_1^2+3x_2^2+2x_1x_2+4x_1-x_2+3$

a) $g(x_1,x_2) = \frac{1}{2}x^T \underbrace{\begin{bmatrix} 4 & 2 \\ 2 & 6 \end{bmatrix}}_{A} x + \underbrace{[4 \ -1]}_{b} x + 3$

b) $\nabla g(x_1,x_2) = Ax+b = \begin{bmatrix} 4 & 2 \\ 2 & 6 \end{bmatrix} x + \begin{bmatrix} 4 \\ -1 \end{bmatrix} = \begin{bmatrix} 4x_1+2x_2+4 \\ 2x_1+6x_2-1 \end{bmatrix}$

If $\nabla g(x) = 0$, then $\left.\begin{matrix} 4x_1+2x_2+4=0 \\ 2x_1+6x_2-1=0 \end{matrix}\right\} \Rightarrow -10x_2=-6$

$x_2 = \frac{3}{5}$

$x_1 = -\frac{13}{10}$

$\bar{x} = \begin{bmatrix} -13/10 \\ 3/5 \end{bmatrix}$ is stationary point of g

c) Since $\nabla^2 g(x) = \begin{bmatrix} 4 & 2 \\ 2 & 6 \end{bmatrix} \succ 0$

so, $\bar{x}$ is local minimizer.

$h(x_1,x_2,x_3) = x_1^2-2x_2^2+x_3^2+3x_1+4x_2-x_3+8$

a) $h(x_1,x_2,x_3) = \frac{1}{2}x^T \underbrace{\begin{bmatrix} 2 & 0 & 0 \\ 0 & -4 & 0 \\ 0 & 0 & 2 \end{bmatrix}}_{A} + \underbrace{[3 \ 4 \ -1]}_{b} x + 8$

b) $\nabla h(x) = Ax+b = \begin{bmatrix} 2 & 0 & 0 \\ 0 & -4 & 0 \\ 0 & 0 & 2 \end{bmatrix} x + \begin{bmatrix} 3 \\ 4 \\ -1 \end{bmatrix}$

$\nabla h(x) = 0 \Rightarrow \left.\begin{matrix} 2x_1 + 3 = 0 \\ -4x_2 + 4 = 0 \\ 2x_3 - 1 = 0 \end{matrix}\right\} \Rightarrow x_3 = \frac{1}{2}, x_2 = 1, x_1 = -\frac{3}{2}$

$\bar{x} = \begin{bmatrix} -\frac{3}{2} \\ 1 \\ \frac{1}{2} \end{bmatrix}$ is stationary point of h

c) Since $\nabla^2 h(x) = A = \begin{bmatrix} 2 & 0 & 0 \\ 0 & -4 & 0 \\ 0 & 0 & 2 \end{bmatrix}$ is negative definite $\bar{x}$ is not local min of h.

8- $f(x) = (x_2 - x_1^2)^2 + x_1^5$

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

$\nabla f(x) = \begin{bmatrix} 5x_1^4 - 4x_1(x_2 - x_1^2) \\ 2(x_2 - x_1^2) \end{bmatrix} = 0 \Rightarrow x_2 = x_1^2$

let $x_1 \neq 0$ then $5x_1^4 - 4x_1 \underbrace{(x_2 - x_1^2)}_{0} = 5x_1^4 \neq 0$

which is contradiction, Therefore $x_1 = 0$ and $x_2 = 0^2 = 0$

Hence $(0,0)$ is unique stationary point of f

$\nabla^2 f(x) = \begin{bmatrix} 20x_1^3 + 12x_1^2 - 4x_2 & -4x_1 \\ -4x_1 & 2 \end{bmatrix} \Rightarrow \nabla^2 f(0,0) = \begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix}$

$\nabla^2 f(0,0) \not\succ$ but $\nabla^2 f(0,0) \succeq 0$, $(0,0)$ can be local minimizer, we cannot say anything from second order sufficient conditions.

$$\nabla g(x) = \begin{bmatrix} -400x_1(x_2 - x_1^2) - 2(1-x_1) \\ 200(x_2 - x_1^2) \end{bmatrix} = 0$$

$$\Rightarrow x_2 = x_1^2 \Rightarrow -400x_1 \underbrace{(x_2 - x_1^2)}_{0} - 2(1-x_1) = 0$$

$$\Rightarrow 1 - x_1 = 0 \Rightarrow x_1 = 1 \Rightarrow x_2 = 1$$

$\bar{x} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ is unique stationary point of g.

$$\nabla^2 g(x) = \begin{bmatrix} 1200x_1^2 - 400x_2 + 2 & -400x_1 \\ -400x_1 & 200 \end{bmatrix}$$

$$\nabla^2 g(1,1) = \begin{bmatrix} 802 & -400 \\ -400 & 200 \end{bmatrix} \succ 0$$

Hence $\bar{x} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ is local minimizer of $g(x)$

9- By assumption $\exists\ x^*, p^* \in \mathbb{R}^n$ such that $f(x) = f(x+p)$ and $\nabla^2 f(x+\alpha p) \succ 0$ for all $\alpha \in (0,1)$

$l(\alpha) = f(x+\alpha p)$ and we know that $f: \mathbb{R}^n \to \mathbb{R}$ twice continuously differentiable

$l(1) - l(0) = l'(t)$ for some $t \in (0,1)$ (By M.V.T)

equivalently $f(x+p) - f(x) = l'(t)$ for some $t \in (0,1)$

let $x = x^*$ and $p = p^*$, then $f(x^*+p) - f(x) = 0 = l'(t)$

t is stationary point of l.

$l''(t) = p^T \nabla^2 f(x+tp) p \succ 0$ by assumption for $t \in (0,1)$

so t is local minimizer of $l(\alpha)$

Since $l(t) = f(x+tp)$, can we say $(x+tp)$ is local minimizer of f?

Since p is a specific direction, The answer is NO. If $f(x) = f(x+p)$ for all $p \in \mathbb{R}^n$ given, Then we can say $(x+tp)$ is local min of f