DATE

Q a) $\|Ux\|^2 = (Ux)^T(Ux) = (x^T U^T)(Ux)$

$= x^T \underbrace{(U^T U)}_{= I} x$ {for orthogonal matrix}
& U is given to be orthogonal

$\therefore = x^T I x$
$= x^T x$
$= \|x\|^2$

$\therefore \|Ux\|^2 = \|x\|^2$
$\because \|a\|$ is +ve
$\therefore \|Ux\| = \|x\|$

~~b) For any 2x2 orthogonal matrix (let say Q)~~
~~$Q = [q_1 \ q_2]$ where $q_1, q_2 \in \mathbb{R}^2$ has following properties.~~
~~$\langle q_1, q_2 \rangle = 0$ & $\|q_1\| = 1$ & $\|q_2\| = 1$~~

b) The property of 2x2 orthogonal matrices is
$\|$Column vector$\| = 1$
&

Let $\begin{bmatrix} a & c \\ b & d \end{bmatrix}$ be a orthogonal matrix

Then $\begin{bmatrix} a \\ b \end{bmatrix}^T \cdot [a \ b]^T = 1$ ie $a^2 + b^2 = 1$ —(1)

||ly $c^2 + d^2 = 1$ —(2)

& $u_1^T u_2 = 0$
ie $ac + bd = 0$ —(3)

For eq(1) → Let $a = \cos\alpha$ & $b = \sin\alpha$
For eq(2) c & d can be $\pm\cos\beta$ & $\pm\sin\beta$ {in any order}.
But from eq (3) $ac + db = 0$
$\therefore c = \pm\sin\alpha$ & $d = \mp\cos\alpha$

or the two possible matrices are
$\begin{bmatrix} \cos\alpha & \sin\alpha \\ \sin\alpha & -\cos\alpha \end{bmatrix}$ or $\begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix}$

Typo

DATE

Q1 c) Equation of a rotated ellipse can be written as.

$$(x-c)^T R D R^T (x-c) = 1$$

where.

$$R = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$ θ is rotation in counter clockwise direction wrt x axis.

& $D = \begin{bmatrix} 1/a^2 & 0 \\ 0 & 1/b^2 \end{bmatrix}$ a^2, b^2 are squares of lenght of major & minor axis

$$\therefore \begin{pmatrix} x-3 & y+1 \end{pmatrix} \begin{pmatrix} \cos\pi/6 & -\sin\pi/6 \\ \sin\pi/6 & \cos\pi/6 \end{pmatrix} \begin{pmatrix} 1/1.5^2 & 0 \\ 0 & 1/0.5^2 \end{pmatrix} \begin{pmatrix} \cos\pi/6 & -\sin\pi/6 \\ \sin\pi/6 & \cos\pi/6 \end{pmatrix}^T \begin{pmatrix} x-3 \\ y+1 \end{pmatrix}$$

$$\Rightarrow \begin{pmatrix} x-3 & y+1 \end{pmatrix} \begin{pmatrix} \sqrt{3}/2 & -1/2 \\ 1/2 & \sqrt{3}/2 \end{pmatrix} \begin{pmatrix} 1/2.25 & 0 \\ 0 & 1/0.25 \end{pmatrix} \begin{pmatrix} \sqrt{3}/2 & -1/2 \\ 1/2 & \sqrt{3}/2 \end{pmatrix}^T \begin{pmatrix} x-3 \\ y+1 \end{pmatrix}$$

DATE

Q2 $A = U \Lambda U^T$ {when A is a symmetric matrix}

~~A is +ve PSD if $x^T A x \geq 0 \;\forall x \neq 0$~~

i.e. $x^T A x$

$= x^T U \Lambda U^T x$

Now, $\because$ A is symmetric

$$= x^T \begin{bmatrix} | & & | \\ u_1 & \cdots & u_n \\ | & & | \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & & \\ 0 & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix}_{n \times n} \begin{bmatrix} — u_1^T — \\ — u_2^T — \\ \vdots \\ — u_n^T — \end{bmatrix}_{n \times n} x$$

$$= x^T \begin{bmatrix} | & & | \\ u_1 & \cdots & u_n \\ | & & | \end{bmatrix} \begin{bmatrix} — \lambda_1 u_1^T — \\ \vdots \\ — \lambda_n u_n^T — \end{bmatrix} x$$

$$= x^T \{ \lambda_1 u_1 u_1^T + \lambda_2 u_2 u_2^T + \cdots + \lambda_n u_n u_n^T \} x$$

$$= \lambda_1 x^T u_1 u_1^T x + \lambda_2 x^T u_2 u_2^T x + \cdots + \lambda_n x^T u_n u_n^T x$$

$$= \lambda_1 (u_1^T x)^T (u_1^T x) + \lambda_2 (u_2^T x)^T (u_2^T x) + \cdots + \lambda_n (u_n^T x)^T (u_n^T x)$$

$$x^T A x = \lambda_1 \underbrace{\| u_1^T x \|^2}_{\text{always +ve.}} + \lambda_2 \| u_2^T x \|^2 + \cdots + \lambda_n \| u_n^T x \|^2 \quad — (1)$$

a) A is PSD iff $\lambda_i \geq 0$ for each i

If $\lambda_i \geq 0$ for each i

RHS of eq 1 ≥ 0

or $x^T A x \geq 0$

ie A is PSD

b) If $\lambda_i > 0$ for each i

RHS of eq (1) > 0

or $x^T A x > 0$

or A is PD

Typo

Scanned by CamScanner

DATE

Q3 a) Given : f is strictly convex
Show: f has atmost one global minimizer.

$\because$ f is strictly convex

$\therefore \quad f(tx + (1-t)y) < tf(x) + (1-t)\cancel{f}(y)$

$\forall\, x \neq y$ & $t \in (0,1)$

Suppose x^* & y^* ~~local min strit not global~~ are two global minimizer

$\therefore$ By convexity, $\forall\, t \in (0,1)$ we have

$$f(tx^* + (1-t)y^*) < tf(x^*) + (1-t)f(y^*)$$

Now, $\because$ x^* & y^* are both minimizer

$\therefore \quad f(x^*) = f(y^*) = \min(f(x))$

$$\therefore \quad f(tx^* + (1-t)y^*) < tf(x^*) + (1-t)f(x^*)$$
$$< f(x^*)$$

Let $t = 1/2$ then

$$f\left(\frac{x^* + y^*}{2}\right) < f(x^*)$$

This contridicts the defination of a global minimizer which states that x^* is global minizer of

$$f(x^*) \leq f(x) \quad \forall\, x \in \mathbb{R}^d$$

$\therefore$ There can be only one global minimizer.

Typo

Typo

b) Let h & g be two convex functions

$\therefore \nabla^2 h(x)$ is PSD

& $\nabla^2 g(x)$ is PSD

Let $f(x) = h(x) + g(x)$

$$\therefore \underbrace{\nabla^2 f(x)}_{C} = \nabla^2 (h(x) + g(x))$$
$$= \underbrace{\nabla^2 h(x)}_{A} + \underbrace{\nabla^2 g(x)}_{B}$$

We know $x^T A x \geq 0$ {as A is PSD}

& $x^T B x \geq 0$ {as B is PSD}.

$$\therefore x^T C x = x^T (A+B) x$$
$$= x^T A x + x^T B x$$
$$\geq 0$$

or C is PSD

or $\nabla^2 f(x)$ is PSD

or $\nabla^2 (h(x) + g(x))$ is PSD

or $h(x) + g(x)$ is a convex function

c) $$\nabla^2 f(x) = \begin{bmatrix} \frac{\partial f}{\partial x_1^2} & \cdots & \partial f / \partial x_1 x_d \\ \vdots & & \vdots \\ \frac{\partial f}{\partial x_1 \partial x_d} & & \frac{\partial f}{\partial x_d^2} \end{bmatrix}$$

As shown in Q2 earlier

$$x^T A x = \lambda_1 \|u_1^T x\|^2 + \lambda_2 \|u_2^T x\|^2 \cdots + \lambda_d \|u_d^T x\|^2$$
$$= \lambda_1 (u_{11} x_1 + u_{12} x_2 \cdots + u_{1d} x_d)^2$$
$$+ \lambda_2 (u_{21} x_1 + \cdots + u_{2d} x_d)^2$$
$$\vdots$$
$$+ \lambda_d (u_{d1} x_1 + \cdots + u_{dd} x_d)^2$$

$1\times n \times n\times n \times n\times 1$

Q 3c) $f(x) = \frac{1}{2} x^T A x + b^T x + c$

$$\cancel{x^T A x = \sum x_i (A x)_i}$$

$$= \cancel{[x_1 \dots x_n] A \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}}$$

$$= [x_1 \dots x_n] \begin{bmatrix} a_1^T \\ a_1 \; a_2 \; \dots \; a_n \\ a_d^T \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

$x_1 x_1 a_{11} + x_1 x_2 a_{21} + \dots + x_1 x_n a_{n1}$

$$= \begin{bmatrix} (x_1 \dots) & \dots & (\dots x_n) \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \quad (x_1 a_1^T + x_2 a_2^T \dots + x_d a_d^T) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

$$x^T A x = \sum_{i=1}^{d} x_i (A x_i) = \sum_{i=1}^{d} x_i \left(\sum_{i=1}^{d} A_{ik} x_k \right)$$

$$x^T A x = \sum_{k=1}^{d} \sum_{i=1}^{d} A_{ik} x_i x_k$$

$$= \sum_{i=1}^{d} A_{ii} x_i^2 + \sum_{j=1}^{d} \sum_{\substack{i=1 \\ \& j \neq i}}^{d} (A_{ki}) x_i x_k$$

$\therefore$ ~~$x^T A x$~~ =

Now we know

$$\nabla^2 f(x) = \begin{bmatrix} \frac{\partial f}{\partial x_1^2} & \dots & \frac{\partial f}{\partial x_1 \partial x_n} \\ \vdots & & \vdots \\ \frac{\partial f}{\partial x_1 \partial x_n} & & \frac{\partial f}{\partial x_n^2} \end{bmatrix}$$

$$\frac{\partial f}{\partial x_i \partial x_i} = 2 A_{ii}$$

$$\frac{\partial f}{\partial x_i \partial x_j} = A_{ij} + A_{ji} = 2 A_{ij}$$ as A is symmetric

DATE

$$\therefore \nabla^2 (x^T A x) = \begin{bmatrix} 2A_{11} & 2A_{21} & \cdots & 2A_{n1} \\ 2A_{21} & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ 2A_{n1} & & & 2A_{nn} \end{bmatrix}$$

$$= 2A$$

$$\nabla^2 (b^T x) = \nabla^2 (b_1 x_1 + b_2 x_2 + \dots + b_n x_n)$$

$$= 0$$

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

$$= \frac{1}{\cancel{2}} \cancel{2} A$$

$$= A$$

$\therefore$ f is convex if $\nabla^2 f(x)$ is PSD
ie A is PSD

& f is strictly convex if $\nabla^2 f(x)$ is PD
ie A is PD

Typo