NAME: DHRW NILESH PATEL SFUID: 301471961

1. a) given  $x \in \mathbb{R}$ , prove  $\text{RELU}(x) = \max(x, 0)$  is convex function.

According to convex properties.

$$f(\lambda \overrightarrow{x} + (1-\lambda)\overrightarrow{y}) \leq xf(\overrightarrow{x}) + (1-\lambda)f(\overrightarrow{y}) - 0$$

so, if RELU(x) statisfy above property it will a convex quinc

Let assume a function, RELU(x) = h(x) = max(x, 0),  $\forall x \in \mathbb{R}$ .

Putting h(x) in egh 1. on LHS side, we get

$$h(\lambda x + (1-\lambda)y) = max(\lambda x + (1-\lambda)y, 0)$$

$$\leq \lambda \max(x,0) + (1-\lambda) \max(y,0) - 2$$
  
 $\leq \lambda h(x) + (1-\lambda) h(x)$ 

Also, since at step ), we know, man (x+5, t+4) will always less than max (r, s) + max (t, u)

So can say  $h(\lambda x + (1-\lambda)y) \leq \lambda h(x) + (1-\lambda)h(y)$ 

Hence it proves RELU(x) = max (x0) is a convex quiction.

1 b) @ f(x) = ||Ax + 6 ||2 + \ ||x||

we know norm function are convex function. So splitting them and treating as separate function.

let h(x) be ||AT+10||2 and l(x) is \||x||2

To prove h(x) is convex.  $h(\lambda x + (1-\lambda)y) \leq 2(||A(\lambda x + (1-\lambda)y) + ||b|||_2)$ 

$$\leq \lambda \|Ax\| + ((-\lambda)\|Ay\| + \lambda b + (1-\lambda)b$$
  
 $\leq \lambda \|Ax\| + \lambda b + (1-\lambda)(|Ay + b||_2)$ 

$$\leq \lambda h(x) + (1-\lambda) h(y)$$
 Since, it follows the convex property. Similar to prove  $l(x)$ 

$$\ell(\lambda x + (1-\lambda)y) \leq \delta(\lambda |x| + (1-\lambda)|y|)$$
,  $\lambda$  is a tre constant  $\leq \lambda \ell(x) + (1-\lambda)\ell(y)$  as given  $\lambda > 0$ 

similar l(x) is always a convex question.

sum of two convex function is always convex function.

So we prove that 12 norm and 12 norms are convex using nowers homogenity and triangular inequality.

To prove  $f(x) = \frac{1}{1+e^{-x}}$  is neither convex nor concave.

$$f'(x) = \frac{d(1+e^{-x})^{-1}}{dx}$$
= -1 \* (1+e^{-x})^{-2} \* -e^{-x}

$$= -1 \times (1 + e^{-x})^{-2} \times -e^{-x}$$

$$= \frac{1}{e^{+x} (1 + e^{-x})^{+2}} \Rightarrow \frac{e^{x}}{(1 + e^{-x})^{+2}}$$

$$\frac{1+e^{-x}-1}{(1+e^{-x})^{+2}} = \frac{(1+e^{-x})}{(1+e^{-x})^2} - \frac{1}{(1+e^{-x})^2}$$

$$\frac{(1+e^{-x})^{+2}}{(1+e^{-x})^{+2}}$$

$$\stackrel{\Rightarrow}{=} \frac{1}{(1+e^{-x})} \left( \frac{1}{(1+e^{-x})} \right)$$

$$f'(x) = f(x) \left(1 - f(x)\right)$$

$$f''(x) \text{ or } H(x)$$

$$= f'(x) (1 - f(x)) + f(x) = f(x) (1 - f(x))^{2} + f(x) - (f(x) (1 - f(x)))$$

$$\Rightarrow f(x) (1 - f(x))^{2} - f(x)^{2} (1 - f(x))$$

in order to prove the wen 
$$H(x) \ge 0$$
, so quality the  $eg^n$  to  $\ge 0$ 

$$f(x) (1-f(x))^2 - f(x)^2 (1-f(x)) \ge 0$$

$$f(x) (1-f(x)) [(1-f(x))-f(x)) \ge 0$$

$$f(n) \left(1 - f(n)\right) \left[ (1 - f(n)) - f(n) \right] \ge$$

$$1 - 2f(n) \ge 0$$

$$\frac{1}{2} \le f(n) \Rightarrow \frac{1}{2} \le \frac{1}{1 + 2}$$

$$\frac{1}{2} \leqslant f(n) \Rightarrow \frac{1}{2} \leqslant \frac{1}{1 + e^{-x}}$$

$$e^{-x} \ge 1, \text{ [naing log on both sid]}$$

$$\frac{2}{1+e^{-x}}$$

$$e^{-x} \ge 1, \text{ [using log on both side]}$$

 $-x \ge 0$ , since only for -ve value of x. This egn is tome. Hence for is not a convex function.

having the hint (-f(n) is also not convex) we can prove is not concave. so,

Assume 
$$h(n) = -f(n)$$
. Now  $h'(n) = -\frac{1}{1+e^{-x}}$ 

$$\frac{1+e^{-x}}{(1+e^{-x})^2}$$

Now taking 
$$h^{\parallel}(n)$$
 or  $H(n)$ 

Now taking 
$$h^{\parallel}(n)$$
 or  $H(n)$ , we get
$$\frac{d}{dx}\left(\frac{e^{-x}}{(1+e^{-x})^2}\right) = \frac{d}{dx}\left(e^{-x}\right)\left(1+e^{-x}\right)^2 - \frac{d}{dx}\left((1+e^{-x})^2\right)e^{-x}$$

$$\frac{(1+e^{-x})^{2}}{(1+e^{-x})^{2}}^{2}$$

$$= (-e^{-x})(1+e^{-x})^{2} - (-2e^{-x}(1+e^{-x}))e^{-x}$$

$$= (-e^{-x})(1+e^{-x})^{2} - (-2e^{-x}(1+e^{-x}))e^{-x}$$

$$= -\frac{((1+e^{-x})^{2})^{2}}{(1+e^{-x})^{3}} \Rightarrow \frac{(1+e^{-x})^{2}}{e^{x}} + \frac{1}{e^{2x}(1+e^{-x})^{3}}$$

As we know inorder for it to be convex 
$$H(x) \ge 0$$
.  
So,  $e^{x} \ge \frac{-1}{(1+e^{-x})^{-5}}$ ,
$$e^{x} \ge -(1+e^{-x})^{-5}$$
,

 $e^{x} \left(1 + e^{-x}\right)^{5} \leq 1$ This equality does hold for x=1. Hence, it does statisfy for all volues of x.

As negult h(x) i.e -f(x) is not convex function either.

$$h(x,y) = -\cos(x^{2}) + e^{xy} - 2y^{2}$$
(i) gradient of  $x = \nabla h = \left(\frac{\partial}{\partial x} (h(x,y))\right)$ 

$$= \left(\frac{\partial}{\partial x} (h(x,y))\right)$$

$$= \left(\frac{\partial}{\partial x} (h(x,y)\right)$$

$$= \left(\frac{\partial}{\partial x} ($$

2. a) einen xer, yer

$$f(\vec{x_0}) + (\vec{x} - \vec{x_0})^T \vec{q} + \frac{1}{2} (\vec{x} - \vec{x_0})^T H (\vec{x} - \vec{x_0}) \rightarrow \{ex. | x^t = g^t\}.$$

$$f(\vec{x_0}, \vec{y_0}) + [(x_0 - \vec{x_0}) (y - y_0)^T \vec{q} + \frac{1}{2} [x - \vec{x_0}]^T H [x - \vec{x_0} y - \vec{y_0}]$$

$$= -\cos(x_0^2) + e^{x_0 y_0} - 2y_0^2 + [(x - x_0) (y - y_0)]^T [2x_0 \sin(x_0^2) + e^{xy_0}y] + e^{xy_0}x_0 - 4y_0$$

 $= -\cos(\chi_0^2) + e^{\chi_0 y_0} - 2y_0^2 + [(\chi - \chi_0) (y - y_0)]^{T} \left[ 2\chi_0 \sin(\chi_0^2) + e^{\chi y_0} y_0 \right] + e^{\chi_0 y_0} + e^{\chi_0 y_0}$  $\frac{1}{2} \left[ (x_{5} - x_{5}) \quad (y_{5} - y_{5}) \right]^{T} \left[ 2 \sin(x_{5}^{2}) + 4x_{5}^{2} \cos(x_{5}^{2}) + y_{5}^{2} e^{x_{5}y_{5}} \quad x_{5}y_{5} e^{x_{5}y_{5}} + e^{x_{5}y_{5}} \right]$   $x_{5}y_{5} e^{x_{5}y_{5}} + e^{x_{5}y_{5}} \quad x_{5}^{2} e^{x_{5}y_{5}} - 4$ x [(x-x;) (y-y;)]

(iii) When 
$$x_0=0$$
 and  $y_0=0$ , so substituting values in (ii) eqn. we get  $-1+1-0+[\vec{x}-0)(\vec{y}-0)]^T[0]+\frac{1}{2}[\vec{x}-0)(\vec{y}-0)]^T[0]+\frac{1}{2}[\vec{x}-0](\vec{y}-0)$ 

$$\Rightarrow \frac{1}{2} (2\pi \dot{y} - 4\dot{y}^2)$$

$$\Rightarrow \dot{x}\dot{y} - 2\dot{y}^2 = \dot{y}(x - 2\dot{y})$$

$$\Rightarrow \vec{x}\vec{y} - \mathbf{x}y^2 = y(x - 2y)$$

(iv) we need to calculate eigenvalues to determine definiteness for point 
$$(0,0)$$
.

So taking hissian of 
$$k$$
, i.e  $k(x,y) = xy - 2y^2$ 

$$\forall (k(x,y)) = \left[\frac{\partial}{\partial x}(xy - 2y^2)\right] \quad [y]$$

$$\frac{\partial}{\partial x}(x_1,y_1) = \begin{bmatrix} \frac{\partial}{\partial x}(x_1y_1 - 2y_2) \\ \frac{\partial}{\partial y}(x_1y_1 - 2y_2) \end{bmatrix} = \begin{bmatrix} y \\ x_1 - 4y \end{bmatrix}$$

$$H(k(x,y)) = \begin{bmatrix} \frac{\partial^2(k(x,y))}{\partial x^2} & \frac{\partial^2(k(x,y))}{\partial x \partial y} \\ \frac{\partial^2(k(x,y))}{\partial x \partial y} & \frac{\partial^2(k(x,y))}{\partial y^2} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & -4 \end{bmatrix}$$

since A, is positive and 2 is negative, we can swaly say about definitenes. So its termed as indefinite.

2b) given 
$$\vec{x}$$
,  $\vec{b} \in \mathbb{R}^{n}$  and  $M \in \mathbb{R}^{N \times n}$  are symmetric and invertible matrix

(i)  $(\vec{x} - M^{-1}b)^{T} M (\vec{x} - M^{-1}\vec{b})$ 

$$= (\vec{x} - M^{-1}b)^{T} (M\vec{x} - \vec{b})$$

$$= \vec{x}^{T}M\vec{x} - \vec{x}^{T}M\vec{x} + \vec{b}^{T}M^{-1}\vec{b}$$

$$= \vec{x}^{T}M\vec{x} - 2\vec{b}^{T}\vec{x}^{T} + \vec{b}^{T}M^{-1}\vec{b}$$

[since inner product of symmetric matrix are same. we can say
$$\vec{b}^{T}\vec{x} = \vec{x}^{T}\vec{b}$$

$$= \vec{\chi}^{T} M \vec{x} - 2\vec{b}^{T} \vec{\chi}^{T} + \vec{b}^{T} M^{-1} \vec{b}$$
[Since inner product of symmetric matrix are same we can say 
$$\vec{b}^{T} \vec{x}^{T} = \vec{\chi}^{T} \vec{b}$$
]

(ii) Given  $\vec{\chi}^{T} , \vec{k} \in \mathbb{R}^{n}$ ,  $\theta \in \mathbb{R}^{n}$  and  $A + B$  is invertible and  $A, B \in \mathbb{R}^{M + n}$ 

$$f(x) = (\vec{\chi} - \vec{\mu})^{T} A (\vec{\chi} - \vec{\mu}) + (\vec{\chi} - \vec{R})^{T} B (\vec{\chi} - \vec{R})$$

$$f(x) = (\vec{x} - \vec{\mu})^T A (\vec{x} - \vec{\mu}) + (\vec{x} - \vec{\theta})^T B (\vec{x} - \vec{\theta})$$

$$= \vec{x}^T A \vec{x} - 2 \vec{\mu} A \vec{x} + \vec{\mu} A \mu + \vec{x}^T B \vec{x} - 2 \vec{\theta}^T B \vec{x} + \vec{\theta}^T B \vec{\theta}$$

Let 
$$M$$
 be  $(A+B)$  constant term  $(c)$  and  $b = A^T \mu + B^T \theta$ 

xt (A+B) x - 2 (ptA+0tB) x + ptAp + 0B0

Combing terms to make (A+B) Quadratic form,

Using the hint from (i), we can say. 

 $= (\vec{x} - M^{-1}b)^{T} M (\vec{x} - M^{-1}\vec{b}) - \vec{b}^{T} M^{-1}\vec{b} + C$ single quadratic form.

3 a) 
$$A = \begin{bmatrix} 4 & 1 \\ 1 & 4 \end{bmatrix}$$

$$A^{T}A = \begin{bmatrix} 4 & 1 \\ 1 & 4 \end{bmatrix} \begin{bmatrix} 4 & 1 \\ 1 & 4 \end{bmatrix} = \begin{bmatrix} 17 & 8 \\ 8 & 17 \end{bmatrix}$$

$$det(A^{T}A - \lambda I) = \begin{bmatrix} 17 - \lambda & 8 \\ 8 & 17 - \lambda \end{bmatrix}$$

$$\lambda_{1} = \sqrt{25}, \lambda_{2} = \sqrt{9}$$
lo finding vector corresponding to  $\lambda_{1}$ , i.e.  $v_{1} = \begin{bmatrix} -1 & 1 \\ 1-1 \end{bmatrix} \begin{bmatrix} x_{1} \\ x_{2} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ 

Eigen vector corresponding to 
$$A_2$$
,  $\Rightarrow xy = x$ 

$$\lambda_2 = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \chi_1 \\ \chi_2 \end{bmatrix} \Rightarrow \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\lambda_2 = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \chi_1 \\ \chi_2 \end{bmatrix} \Rightarrow \begin{bmatrix} \chi_1 \\ \chi_2 \end{bmatrix} \Rightarrow \begin{bmatrix} \chi_2 \\ \chi_3 \end{bmatrix}$$

So victor 
$$V = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$$
, Let S be invertible matrix i.e  $\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$ 

$$A = 5\begin{bmatrix} 3 & 0 \\ 0 & 5 \end{bmatrix}$$

$$A = 5\begin{bmatrix} 3 & 0 \\ 0 & 5 \end{bmatrix}$$
Therefore for n

$$A^{n} = \left(5\begin{bmatrix}3 & 0\\0 & 5\end{bmatrix}s^{-1}\right)^{n} = 5\begin{bmatrix}3 & 0\\0 & 5\end{bmatrix}^{n}s^{-1}$$

$$= \frac{1}{2} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 3^n & 0 \\ 0 & 5^n \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$$

$$A^{n} = \underbrace{1}_{2} \begin{bmatrix} 3^{n} + 5^{n} & -3^{n} + 5^{n} \\ -3^{n} + 5^{n} & 3^{n} + 5^{n} \end{bmatrix}$$

Prop 1 
$$\Rightarrow$$
 U is matrix with orthonormal columns. mean  $V^TV = I$ .

$$\begin{bmatrix} \sqrt{2} & \sqrt{2} \\ 2 & \sqrt{2} \\ -\sqrt{2} & \sqrt{2} \end{bmatrix} = \begin{bmatrix} \sqrt{2} & -\sqrt{2} \\ \sqrt{2} & \sqrt{2} \end{bmatrix}^2 + (\sqrt{2})^2 & 0 \\ 0 & (-\frac{12}{2})^2 + (\frac{72}{2})^2 \end{bmatrix}$$
Prop 2  $\Rightarrow$  Similarly V should a normal vector. So  $V^T = V^{-1}$ 

$$V^{-1} = \begin{bmatrix} \sqrt{3} & +\frac{1}{2} \\ -\frac{1}{2} & \frac{3}{2} \end{bmatrix}$$
and  $dit(V) = \frac{3}{4} + \frac{1}{4} \Rightarrow 1$ .

Prop 3  $\Rightarrow$   $\geq$  Should a non-negative diagonal matrix. which stating the condition since  $\begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$ 

9,00 (7) Finally, UEVT by multiply it perovides real valued matrix A.

Hence U, & and VT are SVD of matrix A. Proved using

JU, E, V<sup>T</sup> are singular value decomposition of A, then

 $\begin{pmatrix} \sqrt{2} & -\sqrt{2} \\ \sqrt{2} & \sqrt{2} \\ \sqrt{2} & \sqrt{2} \end{pmatrix} \begin{pmatrix} 2 & 0 \\ 0 & \frac{1}{2} \end{pmatrix} \begin{pmatrix} \sqrt{3} & \frac{1}{2} \\ \frac{1}{2} & \sqrt{3} \\ \frac{1}{2} & \sqrt{3} \end{pmatrix}$ 

3 b i) given SVD of A as

properties.

→ dot product with any of the now with itself should 1 → dot product with any of the column/now other than itself should 0.

$$V = \begin{pmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{pmatrix}$$

$$\det(V) = \begin{pmatrix} \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{pmatrix}^2 - \begin{pmatrix} -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{pmatrix}$$

$$= 1$$
since  $\det(U) = 1$ .

Moreover a notation matrix satisfy  $\begin{bmatrix} \cos \theta - \sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}$ 

We can say for 
$$\theta = \overline{14}$$
.  $V = \sqrt{\frac{1}{4}} = \sin\left(\frac{\pi}{4}\right) - \sin\left(\frac{\pi}{4}\right)$ 

$$\frac{1}{14} = \cos\left(\frac{\pi}{4}\right) - \cos\left(\frac{\pi}{4}\right) = \cos\left(\frac{\pi}{4}\right)$$

In similar way, we can also, write V in Ro terms as

$$V^{T} = \begin{bmatrix} \cos\left(\frac{\pi}{6}\right) - \sin\left(\frac{\pi}{6}\right) \end{bmatrix}^{T} \qquad \det\left(V\right) = \begin{bmatrix} \left(\frac{\pi}{2}\right)^{2} - \left(-\frac{1}{2}\right)\left(\frac{1}{2}\right) \end{bmatrix} = 1$$
Hence we can U and V both one notation matrixes

(sh) To find the angle  $\theta u$  . we can use [trace(V) = trace(Ro)]

$$\cos \theta_{k} = \frac{1}{\sqrt{2}} \implies \theta_{k} \implies \frac{\pi}{4}$$
.

Similarly  $\theta_{V} = \cos^{-1}\left(\frac{\sqrt{3}}{2}\right) \implies \frac{\pi}{6}$ 

 $2\cos\theta_{1}=\sqrt{2}$ 

So as in quiet step matrix A, is notated by angle of  $\theta_V$  by nector  $V^T$ .

• Then using diagonal matrix  $\Xi$ , would seale the vector. So secondly there would be scaling operation or transformation.
• Finally, after scaling, matrix A is again related by  $\Theta u$  from vector U.

Trivulore matrix went through Rotation & Scaling - Rotation (Ou)

(iv) give unt circle with 
$$\vec{x} = (1,0)$$
 and  $\vec{y} = (0,1)$ . Performing the

(iv) give unit calle with 
$$x' = (1,0)$$
 and  $y' = (0,1)$ . Performing the transformation from (iii),

$$V^{T}x = \begin{bmatrix} \sqrt{3} & \frac{1}{2} \\ -\frac{1}{2} & \frac{\sqrt{3}}{2} \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} \sqrt{3} \\ \frac{2}{1} \end{bmatrix} \qquad x' = \begin{pmatrix} \sqrt{3} \\ \frac{2}{2} \end{pmatrix}, -\frac{1}{2}$$

$$V^{\mathsf{T}} y = \begin{bmatrix} \overline{3} & \underline{1} \\ \overline{2} & \overline{12} \\ -\underline{1} & \underline{3} \\ \overline{2} \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{pmatrix} \overline{3} & \underline{1} \\ \overline{12} & \underline{73} \end{pmatrix} \Rightarrow y = \begin{pmatrix} \underline{1} \\ \overline{12} & \underline{73} \end{pmatrix}$$

Then we scale the above mention new yester.

$$\Xi \left( \sqrt{1} \chi \right) = \begin{bmatrix} 2 & 0 \\ 0 & \frac{1}{2} \end{bmatrix} \begin{bmatrix} \sqrt{3} \\ -\frac{1}{2} \end{bmatrix} = \begin{bmatrix} \sqrt{3} \\ -\frac{1}{4} \end{bmatrix}$$

 $\Sigma (Y^T Y) = \begin{bmatrix} 2 & 0 \\ 0 & \frac{1}{2} \end{bmatrix} \begin{bmatrix} \frac{1}{2} \\ \frac{13}{4} \end{bmatrix} = \begin{bmatrix} 12 \\ \frac{13}{4} \end{bmatrix}$ 

Finally we apply 
$$3^{\text{ed}}$$
 transformation by multiplying  $U$  vectors:  $U\left(\mathbb{E}V^{T}\chi\right) = \begin{bmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix} \begin{bmatrix} \sqrt{3} \\ -\frac{1}{4} \end{bmatrix} = \begin{pmatrix} 4\sqrt{3}+1 \\ 4\sqrt{2} \end{pmatrix}, \frac{4\sqrt{3}-1}{4\sqrt{2}}$ 

$$U\left(\mathbb{E}V^{T}\chi\right) = \begin{bmatrix} \sqrt{2} \\ 2 \end{bmatrix} \begin{bmatrix} \sqrt{2} \\ \sqrt{2} \end{bmatrix} \begin{bmatrix} \sqrt{2} \\ \sqrt{2} \end{bmatrix} \begin{bmatrix} \sqrt{2} \\ \sqrt{2} \end{bmatrix} = \begin{pmatrix} 4-\sqrt{3} \\ 4 \end{pmatrix}, \frac{4+\sqrt{3}}{4} \end{pmatrix}$$

so after all the transformation  $\frac{1}{12}$  changed to  $\left(\frac{453+1}{412}, \frac{453-1}{4-12}\right) \approx \left(1.401, 1.04\right)$  $\frac{1}{4}$  changed to  $\left(\frac{4-13}{4}, \frac{4+13}{4}\right) \approx (0.566, 1.433)$ Fig (b) i.e A, x and A, y represents correct transformation on unit circle by A.