2) a)

(i) $h(x,y) = -\cos(x^2) + e^{xy} - 2y^2$

Gradient :

$$\nabla h(x,y) = \begin{pmatrix} \nabla h_x(x,y) \\ \nabla h_y(x,y) \end{pmatrix}$$

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

$$= 2x\sin(x^2) + ye^{xy}$$

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

$$= xe^{xy} - 4y$$

$$\nabla h(x,y) = \begin{pmatrix} 2x\sin(x^2) + ye^{xy} \\ xe^{xy} - 4y \end{pmatrix}$$

Hessian

$$H(x,y) = \begin{pmatrix} \frac{\partial^2}{\partial x^2} h(x,y) & \frac{\partial^2}{\partial x \partial y} h(x,y) \\ \frac{\partial^2}{\partial y \partial x} h(x,y) & \frac{\partial^2}{\partial y^2} h(x,y) \end{pmatrix}$$

$$\frac{\partial^2}{\partial x^2} h(x,y) = 2(\sin(x^2) + 2x^2\cos(x^2)) + y^2 e^{xy}$$

$$\frac{\partial^2}{\partial x \partial y} h(x,y) = \frac{\partial^2}{\partial y \partial x} h(x,y) = xye^{xy} + e^{xy}$$

$$\frac{\partial^2}{\partial y^2} h(x,y) = x^2 e^{xy} - 4$$

$$H(x,y) = \begin{pmatrix} 2(\sin(x^2) + 2x^2\cos(x^2)) + y^2 e^{xy} & xye^{xy} + e^{xy} \\ xy\,e^{xy} + \cancel{xy}\, e^{xy} & x^2 e^{xy} - 4 \end{pmatrix}$$

2) a)

(ii) Second order Taylor expansion at point $(x = x_0, y = y_0)$

$$g(x,y) = f(x_0,y_0) + f'(x_0,y_0)\begin{bmatrix} x - x_0 \\ y - y_0 \end{bmatrix} + \frac{1}{2}\begin{bmatrix} x - x_0 & y - y_0 \end{bmatrix} f''(x_0,y_0)\begin{bmatrix} x - x_0 \\ y - y_0 \end{bmatrix}$$

$$= -\cos(x_0^2) + e^{x_0y_0} - 2y_0^2 + \begin{pmatrix} 2x_0\sin(x_0^2) + y_0 e^{x_0y_0} \\ x_0 e^{x_0y_0} - 4y_0 \end{pmatrix}\begin{pmatrix} x - x_0 \\ y - y_0 \end{pmatrix}$$

$$+ \frac{1}{2}(x - x_0 \quad y - y_0)\begin{pmatrix} 2(\sin(x_0^2) + 2x_0^2\cos(x_0^2)) + y_0^2 e^{x_0y_0} & x_0y_0 e^{x_0y_0} + e^{x_0y_0} \\ x_0y_0 e^{x_0y_0} + e^{x_0y_0} & x_0^2 e^{x_0y_0} - 4 \end{pmatrix}$$

1) a)
iii) At point $(x_0=0, y_0=0)$

$$f(0,0) = -\cos(0) + e^0$$
$$= 0$$

$$\frac{\partial}{\partial x}(0,0) = 0 \qquad \frac{\partial}{\partial y}(0,0) = 0 \qquad \frac{\partial^2}{\partial x^2}(0,0) = 2(\sin(0)+0)$$
$$= 0$$

$$\frac{\partial^2}{\partial y^2}(0,0) = -4 \qquad \frac{\partial^2}{\partial x \partial y}(0,0) = e^0 = 1$$

$$g(x,y) = f(0,0) + f'(0,0)\begin{pmatrix} x-0 \\ y-0 \end{pmatrix} + \frac{1}{2}(x-0 \;\; y-0) f''(0,0)\begin{pmatrix} x-0 \\ y-0 \end{pmatrix}$$

$$= 0 + (0 \;\; 0)\begin{pmatrix} x \\ y \end{pmatrix} + \frac{1}{2}(x \;\; y)\begin{pmatrix} 0 & 1 \\ 1 & -4 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}$$

$$= \frac{1}{2}(y \;\; x-4y)\begin{pmatrix} x \\ y \end{pmatrix}$$

$$= \frac{1}{2}(xy + y(x-4y)) = \frac{1}{2}(2xy - 4y^2)$$

$$= \underline{\underline{xy - 2y^2}}$$

2) a)
(iv) $H(0,0) = \begin{pmatrix} 0 & 1 \\ 1 & -4 \end{pmatrix}$

Eigen values are

$\lambda_1 = -2 + \sqrt{5}$ $\qquad \lambda_2 = -2 - \sqrt{5}$

$\lambda_1 \simeq 0.23$ $\qquad \lambda_2 \simeq -4.23$

Some eigen values are positive and some are negative so the hessian of h around the point $(x_0=0, y_0=0)$ is

<u>indefinite</u>

2) b) (i) $\vec{x}, \vec{b} \in \mathbb{R}^n$, $M \in \mathbb{R}^{n\times n}$

To prove

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

Expanding LHS

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

$$\vec{x}^T M \vec{x} - \vec{x}^T \vec{b} - M^{-1} \vec{b}^T M \vec{x} + \vec{b}^T M^{-1} \vec{b} \quad (\because M M^{-1} = I)$$

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

according to inner product notation)

Hence proved..

2) b) (ii) Expanding the quadratic form (i)

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

$$= \vec{x}^T A \vec{x} - 2\vec{\mu}^T A \vec{x} + \vec{\mu}^T A \vec{\mu} \rightarrow ①$$

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

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

Combining result of ① + ②

$$\Rightarrow \vec{x}^T (A+B) \vec{x} - 2(\vec{\mu}^T A + \vec{\theta}^T B)\vec{x} + \vec{\mu}^T A \vec{\mu} + \vec{\theta}^T B \vec{\theta}$$

consider $C = A+B$, $\vec{b} = A\vec{\mu} + B\vec{\theta}$, $R = \vec{\mu}^T A \vec{\mu} + \vec{\theta}^T B \vec{\theta}$

$\Rightarrow \vec{x}^T C \vec{x} - 2\vec{b}^T \vec{x} + R$

$\Rightarrow (\vec{x} - C^{-1}\vec{b})^T C (\vec{x} - C^{-1}\vec{b}) - \vec{b}^T C^{-1} \vec{b} + R$

$\therefore$ Single Quadriatic term + a constant term

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

Eigen vectors = $\begin{pmatrix} 1 \\ 1 \end{pmatrix} \begin{pmatrix} -1 \\ 1 \end{pmatrix}$

Eigen values = 5, 3

$A = U \Lambda U^{-1} \Rightarrow A^n = (U \Lambda U^{-1})^n$

$A^n = U \Lambda U^{-1} \cdot U \Lambda U^{-1} \cdot U \Lambda U^{-1} \ldots \ldots U \Lambda U^{-1}$

$A^n = U \Lambda^n U^{-1}$ $\quad (\because U U^{-1} = I)$

$A^n = U \begin{pmatrix} 3 & 0 \\ 0 & 5 \end{pmatrix}^n U^{-1}$

Deriving from eigen vectors

$U = \begin{pmatrix} -1 & 1 \\ 1 & 1 \end{pmatrix} \quad U^{-1} = \begin{pmatrix} -1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix}$

$A^n = \begin{pmatrix} -1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 3 & 0 \\ 0 & 5 \end{pmatrix}^n \begin{pmatrix} -1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix}$

$A^n = \begin{pmatrix} -1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 3^n & 0 \\ 0 & 5^n \end{pmatrix} \begin{pmatrix} -1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix}$

$= \begin{pmatrix} \frac{3^n + 5^n}{2} & \frac{-3^n + 5^n}{2} \\ \frac{-3^n + 5^n}{2} & \frac{3^n + 5^n}{2} \end{pmatrix}$

3) b) (i)

$$A = \begin{pmatrix} \frac{1+4\sqrt{3}}{4\sqrt{2}} & \frac{4-\sqrt{3}}{4\sqrt{2}} \\ \frac{4\sqrt{3}-1}{4\sqrt{2}} & \frac{4+\sqrt{3}}{4\sqrt{2}} \end{pmatrix}$$

$$U = \begin{pmatrix} \sqrt{2}/2 & -\sqrt{2}/2 \\ \sqrt{2}/2 & \sqrt{2}/2 \end{pmatrix} \quad \Sigma = \begin{pmatrix} 2 & 0 \\ 0 & 1/2 \end{pmatrix} \quad V^T = \begin{pmatrix} \sqrt{3}/2 & -1/2 \\ 1/2 & \sqrt{3}/2 \end{pmatrix}$$

For SVD, U, V are orthogonal matrices,

(matrix will have orthogonal columns)

i Consider first column of U as $\vec{u_1}$,

second column as $\vec{u_2}$

Two vectors are said to be orthogonal if their inner products are 0.

$$\langle \vec{u_1}, \vec{u_2} \rangle = 0$$

$$-2/4 + 2/4 = 0$$

$\therefore$ U is orthogonal matrix

Consider first column of V as $\vec{v_1}$,

second column of as $\vec{v_2}$.

$$\langle \vec{v_1}, \vec{v_2} \rangle = 0$$

$$-\sqrt{3}/4 + \sqrt{3}/4 = 0$$

$\therefore$ V is orthogonal matrix

Then the diagonal entries of Σ are non-negative.

Also, the rank of $A = 2$

number of non-zero diagonal entries in $\Sigma = 2$

$\therefore$ The given $U \, \Sigma \, V^T$ matrices are the SVD of A.

2)
(ii) $U = \begin{pmatrix} \sqrt{2}/2 & -\sqrt{2}/2 \\ \sqrt{2}/2 & \sqrt{2}/2 \end{pmatrix}$ $\quad V^T = \begin{pmatrix} \sqrt{3}/2 & -1/2 \\ 1/2 & \sqrt{3}/2 \end{pmatrix}$

Determinant of $U = \left(\frac{\sqrt{2}}{2} \times \frac{\sqrt{2}}{2}\right) - \left(-\frac{\sqrt{2}}{2} \times \frac{\sqrt{2}}{2}\right)$

$= 1$

$\therefore$ U is rotational matrix

Determinant of $V = \left(\sqrt{3}/2 \times \sqrt{3}/2\right) - \left(-1/2 \times 1/2\right)$

$= 1$

$\therefore$ V is rotational matrix

$$U = \begin{pmatrix} \cos^{-1}(\sqrt{2}/2) & -\sin^{-1}(\sqrt{2}/2) \\ \sin^{-1}(\sqrt{2}/2) & \cos^{-1}(\sqrt{2}/2) \end{pmatrix} = \begin{pmatrix} \pi/4 & -\pi/4 \\ \pi/4 & \pi/4 \end{pmatrix}$$

~~If~~ $\theta_U = \pi/4$

$$V^T = \begin{pmatrix} \cos^{-1}(\sqrt{3}/2) & -\sin^{-1}(1/2) \\ \sin^{-1}(1/2) & \cos^{-1}(\sqrt{3}/2) \end{pmatrix}^T = \begin{pmatrix} \pi/6 & -\pi/6 \\ \pi/6 & \pi/6 \end{pmatrix}^T$$

$\theta_{V^T} = \pi/6$

3) b)
(iii) * Transformations are performed in the order of

(i) U (orthogonal matrix) (rotation)

(ii) Σ (diagonal matrix) (scaling)

(iii) V^T (orthogonal matrix) (rotation)

* For example. in the given matrix, V^T will be performed.

first for (rotation/reflection), then with Σ (scaling along axes) (shrinking/expand), then again with U gives (rotation/reflection)

3) b)
(iv)
When the original A matrix is considered,

SVD of $A = U \Sigma V^T$

when unit circle and plane basis $\vec{x}, \vec{y}$ is transformed by A,

(i) The unit circle is rotated counter clockwise by $\pi/6$ (θ_{V})

(ii) Then it is scaled along the x-axis (by twice) and y-axis by $1/2$ (Σ). This forms an ellipse

(iii) Then it is rotated again counter clockwise by $\pi/4$ by θ_U

Thus the answer will be A_1 (option b)

(This can also be confirmed from the values of A which are all non-negative) $A = \begin{pmatrix} 1.401 & 0.400 \\ 1.047 & 1.013 \end{pmatrix}$

1) a) $ReLU(x) = \begin{cases} x, & \forall x > 0 \\ 0, & \forall x \leq 0 \end{cases}$

$f(x) = x$	$f(x) = 0$
$f'(x) = 1$	$f'(x) = 0$
$f''(x) = 0$	$f''(x) = 0$

∴ Twice differentiable $Hf(x) \geq 0$

∴ The function $ReLU(x)$ is convex.

1) b) $\|\vec{v} + \vec{w}\| \leq \|\vec{v}\| + \|\vec{w}\|$ - (by triangular inequality of norms)

$\|\vec{x} + \vec{b}\| \leq \|\vec{x}\| + \|\vec{b}\|$

From the property of convex functions

$f(\lambda\vec{x} + (1-\lambda)\vec{b}) \leq \lambda f(\vec{x}) + (1-\lambda) f(\vec{y})$

$f(A\vec{x} + \vec{b}) \leq Af(\vec{x}) + f(\vec{y})$

$\|A\vec{x} + \vec{b}\|_2 \leq \|A\vec{x}\|_2 + \|\vec{b}\|_2$

∴ $\|A\vec{x} + \vec{b}\|_2$ is convex.

L_1 norm : $\|\vec{x}\|_1 = \sum_{i=1}^{n} |x_i| \Rightarrow$ value is always positive

$\therefore \|\vec{x}\|_1$ is always convex.

Using the property $\sum_i w_i f_i(\vec{x})$ is convex if f_i are convex and $w_i \geq 0$

$1 \, \|A\vec{x}+\vec{b}\|_2 + \lambda \|\vec{x}\|_1$

Since $\lambda \geq 0$

$\therefore \|A\vec{x}+\vec{b}\|_2 + \lambda\|\vec{x}\|_1$ is convex

1) c)

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

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

$$f''(1) = \frac{e^{-2}(-e+1)}{(1+e^{-1})^3} = -0.09085$$

$$-f''(1) = \frac{e^{-2}(-e+1)}{(1+e^{-1})^3} = 0.09085$$

$$f''(2) = \frac{e^{-4}(-e^{2}+1)}{(1+e^{-2})^3} = -0.07996$$

$$-f''(2) = -\frac{e^{-4}(-e^{2}+1)}{(1+e^{-2})^{3}} = 0.07996$$

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