1 a)

$$f(\alpha x+(1-\alpha)y) \leq \alpha f(x) + (1-\alpha)f(y)$$

$$\alpha^2 \sum_i^d x_i^2 + \sum_i^d 2\alpha(1-\alpha)x_i y_i + \sum_i^d (1-\alpha)^2 y_i^2 \leq \alpha x^2 + (1-\alpha)y^2$$

$$\alpha(\alpha-1)\sum_i^d x_i^2 + 2\alpha(1-\alpha)\sum_i^d x_i y_i + \alpha(\alpha-1)\sum_i^d y_i^2 \leq 0$$

$$\alpha(\alpha-1)\left(\sum_i^d x_i^2 + 2x_i y_i + y_i^2\right) \leq 0$$

$$\underbrace{\alpha(\alpha-1)}\left(\sum_i^d (x_i+y_i)^2\right) \leq 0$$

$\alpha(\alpha-1)$ is greater or equal to zero since $\alpha \in [0,1]$

$\sum_i^d (x_i+y_i)^2$ is also greater or equal to zero since something to the power to is always greater or equal to zero.

1 b): $f(x) = |x|$, $f: \mathbb{R} \mapsto \mathbb{R}$

def: $g: \Omega \mapsto \mathbb{R}$ is convex iff $\forall x, y \in \Omega$ and $\forall \alpha \in [0,1]$ we have $f(\alpha x + (1-\alpha)y) \leq \alpha f(x) + (1-\alpha) f(y)$ and $\Omega$ is a convex set

$$|\alpha x + (1-\alpha) y| \leq \alpha |x| + (1-\alpha)|y|$$

$\boxed{x=0, y>0}$ $(1-\alpha) y \leq (1-\alpha) y$ (Equal)

$\boxed{x>0, y>0}$ $\alpha x + (1-\alpha) y \leq (1-\alpha) y + \alpha x$ (Equal)

$\boxed{x<0, y>0}$ $-\alpha x + \cancel{(1-\alpha) y} \leq -\alpha x + \cancel{(1-\alpha) y}$

$x \leq -x$, true since $x<0$

$\boxed{x<0, y<0}$ $\alpha x + (1-\alpha) y \leq -(\alpha x + (1-\alpha) y)$

$\boxed{\alpha x + (1-\alpha) y = a}$

$a \leq -a$ True since $a<0$

c) $f(x)$ can't be strictly convex because $f(\alpha x + (1-\alpha) y) = \alpha f(x) + (1-\alpha) f(y)$ when $[x=0, y>0]$ and $[x>0, y>0]$

2 a)

A and B are 2 convex sets

$$\boxed{Z = A \cap B}$$

$\forall x, y \in Z \Rightarrow x, y \in A \wedge x, y \in B$

we know by definition of convex sets:

$c = \alpha x + (1-\alpha) y, \quad \alpha \in [0,1]$

$c \in A \wedge c \in B \Rightarrow c \in Z$

---

b) $Z = \bigcap_{i=0}^{n} a_i$ $\left\{ \begin{array}{l} Z \text{ is an intersection} \\ \text{between } n \text{ sets.} \end{array} \right.$

$A = \{a_i \mid \forall i \leq n\}$ A is a set of sets

$\forall x, y \in Z \Rightarrow \forall a \in A$ we have $x, y \in a$

we know by definition of convex sets

$c = \alpha x + (1-\alpha) y, \quad \alpha \in [0,1]$

$\forall a \in A$ we have $c \in a$

$\Rightarrow c \in Z$

d) $\sqrt{\alpha x + (1-\alpha) y} \leq \alpha \sqrt{x} + (1-\alpha)\sqrt{y}$ ①

$\boxed{x = 18, \quad y = 0, \quad \alpha = \frac{1}{2}}$

Inequality ① must hold for all x and y in the domain of f.

$$\sqrt{9} \leq \frac{1}{2}\sqrt{18}$$

$$3 \not\leq \sqrt{\frac{18}{4}} = \frac{3}{\sqrt{2}}$$

It doesn't hold and f(x) can't be convex

3 a)

$x \in \mathbb{R}^n$, $f : \mathbb{R}^n \mapsto \mathbb{R} \Rightarrow f(x) \in \mathbb{R}$

$\omega = f(x) \Rightarrow (x, \omega) \in E_f^n$

b) show $\boxed{\alpha(x_1, \omega_1) + (1-\alpha)(x_2, \omega_2) \in E_f}$

where $(x_1, f(x_1)), (x_2, f(x_2)) \in E_f$

$f(x_1) \leq \omega_1$ , $f(x_2) \leq \omega_2$

$x_\alpha = \alpha(x_1, \omega_1) + (1-\alpha)(x_2, \omega_2) \Rightarrow x_\alpha \in \mathbb{R}^n$

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

$\leq \alpha \omega_1 + (1-\alpha)\omega_2 = \omega \in \mathbb{R}$

$\Rightarrow (x_\alpha, \omega) \in E_f$

4.

a) $f: \mathbb{R}^2 \to \mathbb{R}$, $f(\vec{x}) = \frac{1}{2}x_1^4 + x_1 x_2 - e^{x_2}$

$f'_{x_1} = 2x_1^3 + x_2$, $f'_{x_2} = x_1 - e^{x_2}$

$$\boxed{\nabla f(\vec{x}) = \begin{bmatrix} 2x_1^3 + x_2 \\ x_1 - e^{x_2} \end{bmatrix}}$$ Gradient

$f''_{xx} = 6x_1$ $f''_{x_1 x_2} = 1$ $f''_{x_2 x_2} = -e^{x_2}$

$$\boxed{\nabla^2 f(\vec{x}) = \begin{bmatrix} 6x_1 & 1 \\ 1 & -e^{x_2} \end{bmatrix}}$$ Hessian

$$\vec{v}^T \nabla^2 f(\vec{x}) \vec{v} = [v_1, v_2] \begin{bmatrix} 6x_1 & 1 \\ 1 & -e^{x_2} \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$$

$$= 6x_1 v_1^2 + 2v_1 v_2 - e^{x_2} v_2^2$$

$\boxed{x_2 = v_1 = x_1 = 0 \quad v_2 = 1} \to -1 < 0$

It is not PSD $\Rightarrow$ f is not convex

4b) $f: \mathbb{R}^d \mapsto \mathbb{R}$, $f(x) = \langle a, x \rangle^2 + \langle b, x \rangle$

$$\nabla f = \begin{bmatrix} 2(a, x) \cdot a_1 + b_1 \\ \vdots \\ 2(a, x) a_d + b_n \end{bmatrix} (a x)^2 \quad \begin{pmatrix} a \\ \vdots \end{pmatrix} \begin{pmatrix} x \\ \vdots \end{pmatrix} \begin{bmatrix} (a_i x_i)^2 \\ \vdots \\ (a_n x_n)^2 \end{bmatrix}$$

$$\nabla^2 f = \begin{bmatrix} 2a_1^2 & \cdots & 0 \\ \vdots & \ddots & \\ 2a_n^2 & & 0 \end{bmatrix} \quad \cancel{2(a_i x_i)}$$

$$\boxed{f(x) = \left(\sum_{i=1}^{d} (a_i x_i)\right)^2 + \left(\sum_{i=1}^{d} b_i x_i\right)}$$

$$\nabla f = \begin{bmatrix} 2\left(\sum_{i=1}^{d} a_i x_i\right) a_1 + b_1 \\ \vdots \\ 2\left(\sum_{i=1}^{d} a_i x_i\right) a_n + b_n \end{bmatrix}$$

$$\nabla^2 f = \begin{bmatrix} 2a_1 \cdot a_1 & 2a_1 a_2 & \cdots & 2a_1 a_n \\ \vdots & \ddots & & \\ \vdots & & \ddots & \\ 2a_n a_1 & \cdots & \cdots & 2a_n^2 \end{bmatrix}$$

4b)

$$x^T \nabla^2 f x = [x_1 \cdots x_n] \begin{bmatrix} 2a_1^2 & \cdots & 2a_1a_n \\ \vdots & \ddots & \vdots \\ 2a_na_1 & \cdots & 2a_n^2 \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

$$= [x_1, \cdots x_n] \begin{bmatrix} 2a_1^2 x_1 + \cdots + 2a_1a_n x_n \\ \vdots \\ 2a_na_1 x_1 + \cdots + 2a_n^2 x_n \end{bmatrix}$$

$$2x_1^2 a_1^2 + \cdots + 2a_1a_n x_1 x_n + \cdots + 2a_na_1x_1x_n + 2a_n^2 x_n^2$$

$$= x_1(2a_1^2 x_1 + \cdots + 2a_1a_nx_n) + \cdots + x_n(2a_na_1x_1 + 2a_n^2x_n)$$

$$= \sum_{i=1}^{n} x_n \left( \sum_{i=1}^{n} 2a_i \left( \sum_{i=1}^{n} a_i x_i \right) \right)$$

$$= 2(a_1^2x_1^2 + a_1a_2x_1x_2 + a_1a_3x_1x_3 + \cdots a_1a_nx_1x_n +$$

$$\cdots + a_1a_nx_1x_n + a_2a_nx_2x_n + \cdots + a_n^2x_n^2)$$

$$= 2\left( \sum_{i=1}^{n} \sum_{j=i}^{n} a_i a_j x_i x_j \right) = 2\langle a, x\rangle^2 \geq 0$$

5.

a) $f(x) = \|x\|_2^2 = \left(\sqrt{\sum x_i^2}\right)^2 = \sum x_i^2$

$$\nabla f(\vec{x}) = \begin{bmatrix} 2x_1 \\ \vdots \\ 2x_n \end{bmatrix}$$

b) $= \nabla(f \circ g)(x) = J(g(x))^T \cdot (\nabla f(y))$

$f(u) = \|u\|^2 = u^T u \longrightarrow \frac{dh}{du} = 2u$

$\frac{d\, g(x)}{dx} = A \longrightarrow f = h(g(x))$

$\longrightarrow \frac{df}{dx} = 2 A^T A x$

d)

$$f_1(x) = \|Ax-b\|_2^2 \quad , \quad f_2(x) = \delta \|x\|_2^2$$

$$f_3(x) = f_1(x) + \gamma f_2(x)$$

$$\frac{\partial f_3}{\partial x} = \frac{\partial f_1(x)}{\partial x} + \gamma \cdot \frac{\partial f_2(x)}{\partial x}$$

we have calculated $\frac{\partial f_1(x)}{\partial x}$ and $\frac{\partial f_2(x)}{\partial x}$ in part (a) and (c) of question 5.

$$\frac{\partial f_1}{\partial x} = 2(A\vec{x}-B)^T A \cdot \quad \frac{\partial f_2}{\partial x} = \begin{bmatrix} 2x_1 \\ 2x_2 \\ \vdots \\ 2x_n \end{bmatrix} = 2\vec{x}$$

$$\frac{\partial f_3}{\partial x} = 2(A\vec{x}-B)^T A \cdot 2\gamma \vec{x} = 4\gamma (Ax-B)^T Ax$$

answer: $4\gamma (Ax-B)^T Ax$

$g(u) = u^T u$

c) $f(x) = g(h(x))$, $g(z)$ $z =$

$$\frac{d\,g(u)}{du} = 2u^T \qquad \frac{d\,h(x)}{dx} = A$$

answer: $2A^T(Ax - B)$

6. a)

$$\nabla f(x) = 2A^T(Ax - B)$$

$$= 2A^TAx - 2A^TB$$

$$\nabla^2 f(x) = 2A^TA$$

$$\vec{z} \cdot 2A^TA\vec{z} = 2(\vec{z}^tA^TA\vec{z})$$

$$= 2(A\vec{z})^t(A\vec{z})$$

$$= 2\|Az\|_2^2 \geq 0$$

$\Rightarrow$ A is PSD

$\|A\|_2$ must be greater than zero for A f to be strict convex which means that A must have full rank.