Q1 Conside the following function

$$f(x) = \frac{1}{2}x_1^2 + x_1x_2 + \frac{1}{2}x_2^2, \quad g(x) = \frac{1}{2}x_1^3 + x_1x_2 + \frac{1}{2}x_2^2$$

a) Is f, g coercive?

b) Is $x = (10, -10)$ a critical point of f? Is it a global minimizer?

A1: a) $f(x) = \frac{1}{2}(x_1+x_2)^2$, whenever $x_1 = -x_2$ $f(x)=0$

So when $\|x\| = \sqrt{x_1^2+x_2^2} \to \infty$, $f(x)$ doesn't go to $+\infty$

So no

$g(x) = \frac{1}{2}x_1^3 + x_1x_2 + \frac{1}{2}x_2^2$. to see this we set $x_2 = 0$

$\Rightarrow g(x_1, 0) = \frac{1}{2}x_1^3$ when $x_1 < 0$ and $\|x\| = \sqrt{x_1^2 + 0} = |x_1| \to \infty$

$g(x) = \frac{1}{2}x_1^3 \to -\infty$, So no

b). We compute the gradient

$$\nabla f(x) = \begin{bmatrix} x_1 + x_2 \\ x_1 + x_2 \end{bmatrix}$$

$\nabla f((10,-10)) = \begin{bmatrix} 0 \\ 0 \end{bmatrix} = 0 \Rightarrow (10, -10)$ is a critical point.

$Hf(x) = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \Rightarrow \Delta_1 > 0 \ \Delta_2 = 0 \Rightarrow Hf(x) \succeq 0$ for $\forall x \in \mathbb{R}^n$

$\Rightarrow x = (10, -10)$ is a global minimizer of f.

(To see this we can also refer to the property of function $f = \frac{1}{2}(x_1^2 + x_2^2) \geq 0$)

Q: let $A \in \mathbb{R}^{n\times n}$ be a symmetric matrix

a) show that if A is positive definate, then A^{-1} is symmetric and positive definite

b) Classify the following matrices according to whether they are positive or negative definite, semidefinite, or indefinite.

(i) $\begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0.1 \end{bmatrix}$ (ii) $\begin{bmatrix} 3 & 1 & 2 \\ 1 & 5 & 3 \\ 2 & 3 & 7 \end{bmatrix}$ (iii) $\begin{bmatrix} 2 & -4 & 0 \\ -4 & 8 & 0 \\ 0 & 0 & -3 \end{bmatrix}$

(iiii) $\begin{bmatrix} 2 & -4 & 3 \\ -4 & 0 & 6 \\ 3 & 6 & 5 \end{bmatrix}$

A2: (a).

① symmetric

A is symmetric $\Rightarrow A = A^T$

we want to prove $(A^{-1})^T = A^{-1}$ (*)

that is to prove $(A^{-1})^T \cdot A = (A^{-1})^T A^T = (A A^{-1})^T = I$

So (*) is proved, which means A^{-1} is symmetric

② $A \succ 0$

we want to prove $x^T A^{-1} x > 0$, for $\forall x$

we have that $y^T A y > 0$ for $\forall y$

$\Rightarrow \forall x$, x can be represented as $x = Ay$

$x^T A^{-1} x = (Ay)^T A^{-1} A y = y^T A^T A^{-1} A y = y^T A^T y = y^T A y > 0$

$\therefore A^{-1} \succ 0$

(Note this can also be seen from the eigenvalue view:

$\because A \succ 0$, $\Rightarrow \lambda_i > 0$ for all $i \in [1, \cdots n]$,

$\therefore$ all the eigenvalues of A^{-1} are $1/\lambda_i > 0$, $\Rightarrow A^{-1} \succ 0$)

(b) (i) all eigenvalues $\lambda_1 = 2, \lambda_2 = 1, \lambda_3 = 0.1 > 0 \Rightarrow A \succ 0$

(also $\Delta_1 = 2 > 0$, $\Delta_2 = 1 > 0$, $\Delta_3 = 0.1 > 0$)

(ii) $\Delta_1 = 3 > 0$ $\quad \Delta_2 = \det \begin{bmatrix} 3 & 1 \\ 1 & 5 \end{bmatrix} = 15 - 1 = 14 > 0$

$\Delta_3 = 3 \times (35 - 9) - 1 \times (7 - 6) + 2(3 - 10) = 63 > 0$

$\therefore A \succ 0$

(iii) $\Delta_1 = 2 > 0$ $\quad \Delta_2 = 16 + 16 = 32 > 0$ $\quad \Delta_3 = -3 \times 32 = -96 < 0$

$\Rightarrow A \nsucc 0 \quad A \nsucceq 0 \quad A \nprec 0 \quad A \npreceq 0$

we check by one more step with eigenvalues:

$Ax = \lambda x$

$\Rightarrow (A-\lambda I)x = 0$

$$A-\lambda I = \begin{bmatrix} 2 & -4 & 0 \\ -4 & 8 & 0 \\ 0 & 0 & -3 \end{bmatrix} - \begin{bmatrix} \lambda & & \\ & \lambda & \\ & & \lambda \end{bmatrix} = \begin{bmatrix} 2-\lambda & -4 & 0 \\ -4 & 8-\lambda & 0 \\ 0 & 0 & -3-\lambda \end{bmatrix}$$

$$\det(A-\lambda I) = (-3-\lambda)\left[(8-\lambda)(2-\lambda) - 16\right]$$
$$= -(3+\lambda)(\lambda^2 - 10\lambda)$$
$$= -(3+\lambda)(\lambda-10)\lambda$$

by taking $\det(A-\lambda I) = 0 \Rightarrow \lambda_1 = 10,\ \lambda_2 = 0,\ \lambda_3 = -3$.

Since there is a positive and a negative eigenvalue, the matrix is indefinite.

(iiii) It is not symmetric.

Q3: Compute the gradient of the following functions.

① $f(x) = y^T x$

② $f(x) = x^T y$ $\qquad x, y \in \mathbb{R}^n$

③ $f(x) = x^T A x$ $\qquad A \in \mathbb{R}^{n\times n}$

④ $f(x) = \frac{1}{2}\|Ax - y\|^2$

A: ① $f(x) = y^T x = \sum_{i=1}^{n} y_i x_i$

$$\Rightarrow \nabla f(x) = \begin{bmatrix} \frac{\partial f(x)}{\partial x_1} \\ \vdots \\ \frac{\partial f(x)}{\partial x_i} \\ \vdots \\ \frac{\partial f(x)}{\partial x_n} \end{bmatrix} = \left[\frac{\partial f(x)}{\partial x_1}, \cdots, \frac{\partial f(x)}{\partial x_i}, \cdots \frac{\partial f(x)}{\partial x_n} \right]^T = [y_1, \cdots y_i \cdots y_n]^T$$

$= y$

② $f(x) = x^T y = \sum_{i=1}^{n} y_i x_i$

Similarily. $\nabla f(x) = y$

③ $f(x) = x^T A x$

$$= [x_1 \cdots x_i \cdots x_n] \cdot \begin{bmatrix} \vec{a}_1 \\ \vec{a}_2 \\ \vdots \\ \vec{a}_n \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

$$= \left(\sum_{i=1}^{n} x_i \vec{a}_i\right) \cdot x = \sum_{i=1}^{n} (x_i \vec{a}_i \cdot x) = \sum_{i=1}^{n} x_i (\vec{a}_i x)$$

$$= \sum_{i=1}^{n} x_i \sum_{j=1}^{n} a_{ij} x_j = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j$$

$$\Rightarrow \frac{\partial f(x)}{\partial x_k} = \sum_{j=1}^{n} a_{kj} x_j + \sum_{i=1}^{n} a_{ik} x_i = \vec{a}_k \cdot x + \vec{\beta}_k^T \cdot x$$

$\Rightarrow \nabla f(x) = Ax + A^T x = (A + A^T)x$ if A is symmetric $\Rightarrow \nabla f(x) = 2Ax$

④. $f(x) = \frac{1}{2}\|Ax - y\|^2$, $x, y \in \mathbb{R}^n$

$$f(x) = \frac{1}{2}\|Ax - y\|^2$$
$$= \frac{1}{2}(Ax - y)^T (Ax - y)$$
$$= \frac{1}{2}(x^T A^T - y^T)(Ax - y)$$
$$= \frac{1}{2}\left[x^T A^T A x - x^T A^T y - y^T A x + y^T y\right]$$

$$\nabla f(x) = \frac{1}{2}\left[2A^T A \cdot x - A^T y - (y^T A)^T\right]$$
$$= A^T A x - A^T y$$
$$= A^T (Ax - y)$$