1. [**5 marks**] *Determine whether the following function h on the region Ω_0 is convex, strictly convex, concave, strictly concave, or neither.*

$$h(x_1, x_2) = \frac{x_1^2}{x_2}, \qquad \Omega_0 = \{(x_1, x_2) \in \mathbb{R}^2 \mid x_2 > 0\}.$$

Solution. The gradient and the Hessian of h are given by

$$\nabla h(x_1, x_2) = \begin{pmatrix} \frac{2x_1}{x_2} \\ -\frac{x_1^2}{x_2^2} \end{pmatrix}, \text{ and } \nabla^2 h(x_1, x_2) = \begin{pmatrix} \frac{2}{x_2} & -\frac{2x_1}{x_2^2} \\ -\frac{2x_1}{x_2^2} & \frac{2x_1^2}{x_2^3} \end{pmatrix}.$$

Then, $\operatorname{tr}(\nabla^2 h(x_1, x_2)) = \frac{2}{x_2} + \frac{2x_1^2}{x_2^3} > 0$ on Ω_0 and $\det(\nabla^2 h(x_1, x_2)) = \frac{4x_1^2}{x_2^4} - \frac{4x_1^2}{x_2^4} = 0$, implying that one of the eigenvalues is zero and the other is positive. Hence, $\nabla^2 h(x_1, x_2)$ is positive semi-definite on Ω_0, and h is a convex function on Ω_0.

2. [**15 marks**] *Consider the problem of optimizing the function*

$$f(\boldsymbol{x}) = x_1^4 + 2x_2^4 + 2x_1^2x_2^2 - 2x_1^2 - 4x_2^2 + 2$$

on $\mathbb{R}^2$. The gradient and Hessian of f are **given** *respectively by*

$$\nabla f(\boldsymbol{x}) = 4\begin{pmatrix} x_1(x_1^2 + x_2^2 - 1) \\ x_2(x_1^2 + 2x_2^2 - 2) \end{pmatrix}, \text{ and } \nabla^2 f(\boldsymbol{x}) = 4\begin{pmatrix} 3x_1^2 + x_2^2 - 1 & 2x_1x_2 \\ 2x_1x_2 & x_1^2 + 6x_2^2 - 2 \end{pmatrix}.$$

i) *Find the five stationary points of f on $\mathbb{R}^2$.*

Solution. Solving $\nabla f(\boldsymbol{x}) = 0$ requires that

$$x_1(x_1^2 + x_2^2 - 1) = 0, \tag{1}$$

and

$$x_2(x_1^2 + 2x_2^2 - 2) = 0. \tag{2}$$

It follows from Eqn. (1) that $x_1 = 0$ or $x_1^2 + x_2^2 - 1 = 0$. When $x_1 = 0$, Eqn. (2) gives $x_2 = 0$ or $x_2 = \pm 1$, so $(0, 0)$, $(0, 1)$, $(0, -1)$ are stationary points. When $x_1^2 + x_2^2 - 1 = 0$, we require $x_2 = 0$ or $x_1^2 + 2x_2^2 - 2 = 0$ from Eqn. (2). Solving $x_1^2 + x_2^2 - 1 = 0$ and $x_2 = 0$ gives the stationary points $(1, 0)$ and $(-1, 0)$. Finally, solving $x_1^2 + x_2^2 - 1 = 0$ and $x_1^2 + 2x_2^2 - 2 = 0$ gives the stationary points $(0, 1)$ and $(0, -1)$. In conclusion, the stationary points are

$$\boldsymbol{a} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \boldsymbol{b} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \boldsymbol{c} = \begin{pmatrix} -1 \\ 0 \end{pmatrix}, \boldsymbol{d} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \boldsymbol{e} = \begin{pmatrix} 0 \\ -1 \end{pmatrix}.$$

ii) *Identify, as far as possible using Hessian information, the five stationary points of f as (strict) local minimizers, (strict) local maximizers or saddle points, etc.*

Solution.

- For $\boldsymbol{a} = (0, 0)^T$, the Hessian matrix is $\nabla^2 f(\boldsymbol{a}) = 4\begin{pmatrix} -1 & 0 \\ 0 & -2 \end{pmatrix}$. The eigenvalues of $\nabla^2 f(\boldsymbol{a})$ are -4 and -8, so $\nabla^2 f(\boldsymbol{a})$ is negative definite. By the second-order sufficient optimality conditions, $\boldsymbol{a}$ is a strict local maximiser of f.

- For $\boldsymbol{b} = (1,0)^T$, $\nabla^2 f(b) = 4\begin{pmatrix} 2 & 0 \\ 0 & -1 \end{pmatrix}$. The eigenvalues are $\lambda_1 = 8 > 0$ and $\lambda_2 = -4 < 0$, so $\nabla^2 f(\boldsymbol{b})$ is indefinite, and $\boldsymbol{b}$ is a saddle point of f.
- For $\boldsymbol{c} = (-1,0)^T$, $\nabla^2 f(\boldsymbol{c}) = 4\begin{pmatrix} 2 & 0 \\ 0 & -1 \end{pmatrix}$, which is indefinite, and $\boldsymbol{c}$ is a saddle point of f.
- For $\boldsymbol{d} = (0,1)^T$, $\nabla^2 f(\boldsymbol{d}) = 4\begin{pmatrix} 0 & 0 \\ 0 & 4 \end{pmatrix}$, which is positive semi-definite as the eigenvalues are 0 and 16. It is not enough to conclude that $\boldsymbol{d}$ is a local minimiser, but we can conclude that $\boldsymbol{d}$ is not a local maximiser by the second-order necessary conditions for a local maximizer.
- For $\boldsymbol{e} = (0,-1)^T$, $\nabla^2 f(\boldsymbol{e}) = 4\begin{pmatrix} 0 & 0 \\ 0 & 4 \end{pmatrix}$, which is positive semi-definite, so $\boldsymbol{e}$ cannot be a local maximiser of f by the second-order necessary conditions for a local maximizer.

iii) *Does the function f have a global maximizer on $\mathbb{R}^2$? Give reasons for your answer.*

Solution. The function f does not have a global maximiser because

$$f(x_1, 0) = x_1^4 - 2x_1^2 + 2 = (x_1^2 - 1)^2 + 1 \to +\infty$$

as $x_1 \to +\infty$, and the function is not bounded above on $\mathbb{R}^2$.

3. [**4 marks**] *Let A be a real $(n \times n)$ square matrix. Let a_{ii} and a_{jj} be the i^{th} and j^{th} diagonal elements of A respectively, where $i, j \in \{1, 2, \ldots, n\}$ and $i \neq j$. Using the definition of an indefinite matrix, show that if $a_{ii}a_{jj} < 0$, then A is indefinite.*

Proof. We take the standard basis vectors $\boldsymbol{e}_i \in \mathbb{R}^n$, which has entry 1 on the i-th component and zero otherwise, and $\boldsymbol{e}_j \in \mathbb{R}^n$, which has entry 1 on the j-th component and zero otherwise, $\boldsymbol{e}_i \neq \boldsymbol{e}_j$. Then, $\boldsymbol{e}_i^\top A\boldsymbol{e}_i = a_{ii}$ and $\boldsymbol{e}_j^\top A\boldsymbol{e}_j = a_{jj}$. We consider two scenarios when $a_{ii}a_{jj} < 0$, i.e., $a_{ii} < 0$, $a_{jj} > 0$ or $a_{ii} > 0$, $a_{jj} < 0$. For the former scenario, $\boldsymbol{e}_i^\top A\boldsymbol{e}_i < 0$ and $\boldsymbol{e}_j^\top A\boldsymbol{e}_j > 0$. By the definition, A is an indefinite matrix. For the latter scenario, $\boldsymbol{e}_i^\top A\boldsymbol{e}_i > 0$ and $\boldsymbol{e}_j^\top A\boldsymbol{e}_j < 0$, again implying that A is indefinite. □

4. [**6 marks**] *Let C be an $m \times n$ matrix, $\boldsymbol{d} \in \mathbb{R}^m$, $\boldsymbol{a} \in \mathbb{R}^n$ and $r \in \mathbb{R}$. Show that*

$$\Omega = \{\boldsymbol{x} \in \mathbb{R}^n \ : \ \|\mathrm{C}\boldsymbol{x} - \boldsymbol{d}\| \leq \mathbf{a}^\top \boldsymbol{x} + r\}$$

is a convex set for any norm $\|\cdot\|$ on $\mathbb{R}^m$. State any properties of a norm that you use.

Proof. Let $\boldsymbol{x}, \boldsymbol{y} \in \Omega$ and $\theta \in (0,1)$ be arbitrary. Then, $\|C\boldsymbol{x} - \boldsymbol{d}\| \leq \boldsymbol{a}^\top \boldsymbol{x} + r$ and $\|C\boldsymbol{y} - \boldsymbol{d}\| \leq \boldsymbol{a}^\top \boldsymbol{y} + r$. Define $\boldsymbol{z} = \theta\boldsymbol{x} + (1-\theta)\boldsymbol{y}$. We need to show that $\boldsymbol{z} \in \Omega$.

$$\begin{aligned}
\|C\boldsymbol{z} - \boldsymbol{d}\| &= \|C(\theta\boldsymbol{x} + (1-\theta)\boldsymbol{y}) - \boldsymbol{d}\| \\
&= \|\theta C\boldsymbol{x} + (1-\theta)C\boldsymbol{y} - \theta\boldsymbol{d} - (1-\theta)\boldsymbol{d}\| \\
&= \|\theta(C\boldsymbol{x} - \boldsymbol{d}) + (1-\theta)(C\boldsymbol{y} - \boldsymbol{d})\| \\
&\leq \|\theta(C\boldsymbol{x} - \boldsymbol{d})\| + \|(1-\theta)(C\boldsymbol{y} - \boldsymbol{d})\| && \text{(triangle inequality)} \\
&= |\theta|\|C\boldsymbol{x} - \boldsymbol{d}\| + |1-\theta|\|C\boldsymbol{y} - \boldsymbol{d}\| && \text{(positive homogenity)} \\
&= \theta\|C\boldsymbol{x} - \boldsymbol{d}\| + (1-\theta)\|C\boldsymbol{y} - \boldsymbol{d}\| && (\theta > 0 \text{ and } 1-\theta > 0) \\
&\leq \theta(\boldsymbol{a}^\top\boldsymbol{x} + r) + (1-\theta)(\boldsymbol{a}^\top\boldsymbol{y} + r) \\
&= \boldsymbol{a}^\top(\theta\boldsymbol{x} + (1-\theta)\boldsymbol{y}) + r && (\boldsymbol{a}^\top\boldsymbol{x} \text{ and } \boldsymbol{a}^\top\boldsymbol{y} \text{ are linear in } \boldsymbol{x} \text{ and } \boldsymbol{y} \text{ respectively)} \\
&= \boldsymbol{a}^\top\boldsymbol{z} + r.
\end{aligned}$$

Hence, $\boldsymbol{z} \in \Omega$, and by the definition of convexity, Ω is a convex set.