1. (10 points) For the following matrix

\[ A = \begin{bmatrix}
1 & 1 & 1 & 1 \\
1 & -1 & 1 & -1 \\
1 & 1 & -1 & -1 \\
1 & -1 & -1 & 1
\end{bmatrix}, \]

Compute the matrix norms \( \|A\|_1, \|A\|_2, \|A\|_\infty \).

**Solution:** For \( p = 1, 2, \infty \), the matrix norm \( \|B\|_p \) is defined as

\[ \|B\|_p = \max_{x \neq 0} \frac{\|Bx\|_p}{\|x\|_p}. \]

The formula is given as (the dimension is \( m \times n \)):

\[ \|B\|_1 = \max_{i=1}^m \sum_{j=1}^n |B_{ij}|, \quad \|B\|_\infty = \max_{j=1}^n \sum_{i=1}^m |B_{ij}|, \]

\[ \|B\|_2 = \max_{x \neq 0} \frac{\|Bx\|_2}{\|x\|_2} = \sqrt{\lambda_{\text{max}}(B^T B)}. \]

For the above given matrix \( A \), we have \( A^T A = 4I_4 \) (\( I_n \) denotes the \( n \times n \) identity matrix), so

\[ \|A\|_1 = \|A\|_\infty = 4, \quad \|A\|_2 = 2. \]

2. (10 points) Determine if the set

\[ \mathcal{E} = \{(x_1, x_2, x_3) : 3x_1^2 + 12x_2^2 - 12x_1x_2 + x_3^2 \leq 1\} \]

is convex or not. Give reasons to justify the answer.

**Solution:** Write the quadratic form in the matrix form

\[ 3x_1^2 + 12x_2^2 - 12x_1x_2 + x_3^2 = x^T C x \]
where \( C \) is the matrix
\[
C = \begin{bmatrix} 3 & -6 & 0 \\ -6 & 12 & 0 \\ 0 & 0 & 1 \end{bmatrix} \succeq 0.
\]

It is a psd matrix. Note that \( C = RR^T \) for the matrix
\[
R = \begin{bmatrix} \sqrt{3} & -2\sqrt{3} & 0 \\ 0 & 0 & 1 \end{bmatrix}.
\]

Then the set \( \mathcal{E} \) is the same as
\[
\mathcal{E} = \{(x_1, x_2, x_3) : \|Rx\|_2 \leq 1\}.
\]

It is convex. This is because, for all \( x, y \in \mathcal{E} \) and for all \( \theta \in [0, 1] \),
\[
\|R(\theta x + (1 - \theta)y)\|_2 = \|\theta Rx + (1 - \theta)Ry\|_2 \\
\leq \theta \|Rx\|_2 + (1 - \theta)\|Ry\|_2 \leq \theta \cdot 1 + (1 - \theta) \cdot 1 = 1,
\]
which means that \( \theta x + (1 - \theta)y \in \mathcal{E} \).

3. (10 points) Determine the convergence order of the sequence
\[
\left\{ k \ln \left(1 - \frac{1}{k}\right) \right\}_{k=1}^{\infty}.
\]

Solution: First note that the limit is \(-1\). The convergence order is one. For \( x_k = k \ln \left(1 - \frac{1}{k}\right) \), we can see that
\[
\frac{|x_{k+1} + 1|}{|x_k + 1|} = \frac{|(k + 1) \ln \left(1 - \frac{1}{k+1}\right) + 1|}{|k \ln \left(1 - \frac{1}{k}\right) + 1|}.
\]

Note the Taylor expansion
\[
\frac{1}{x} \ln(1 - x) + 1 = -\frac{x}{2} - \frac{x^2}{3} - \frac{x^3}{4} - \cdots
\]

Then we get
\[
\frac{|x_{k+1} + 1|}{|x_k + 1|} = \frac{\left| -\frac{k}{2(k+1)} - \frac{k}{3(k+1)^2} - \cdots \right|}{\left| -\frac{1}{2} - \frac{1}{3k^2} - \cdots \right|} = \frac{\left| -\frac{k}{2(k+1)} - \frac{k}{3(k+1)^2} - \cdots \right|}{\left| -\frac{1}{2} - \frac{k}{3k^2} - \cdots \right|}
\]

As \( k \to \infty \), the limit is \( 1 > 0 \).
4. (10 points) Find the biggest integer \( r > 0 \) such that

\[
\exp\left(-\frac{1}{k^2}\right) - 1 + \frac{1}{k^2} = O\left(\frac{1}{k^r}\right)
\]
as \( k \) goes to infinity.

**Solution:** Note the Taylor expansion

\[
\exp(-x) = 1 - x + \frac{x^2}{2} - \frac{x^3}{6} + \cdots
\]

Then we can get

\[
\exp\left(-\frac{1}{k^2}\right) - 1 + \frac{1}{k^2} = \frac{1}{2k^4} - \frac{1}{6k^6} + \cdots = \frac{1}{k^4} \left(\frac{1}{2} - \frac{k^4}{6k^6} + \cdots\right)
\]

So, the biggest \( r \) for this question is 4.

5. (10 points) For a symmetric matrix \( A = (a_{ij}) \in \mathbb{R}^{n \times n} \), let \( f(x) \) be the function in \( x = (x_1, \ldots, x_n) \) such that

\[
f(x) = \sum_{i,j=1}^{n} a_{ij} x_i^2 x_j^2.
\]

Express the gradient and Hessian of \( f(x) \) in terms of \( A \) and \( x \).

**Solution:** For a vector \( x \), we use \( x^2 \) to denote the vector of squaring all its entries. Then

\[
f(x) = (x^2)^T A x^2.
\]

Let \( h(x) = x^T A x \), \( g(x) = x^2 \), then \( f(x) \) is the composite function

\[
f(x) = h(g(x)).
\]

Note that

\[
h'(x) = 2x^T A, \quad \nabla h(x) = 2Ax,
\]

\[
g'(x) = 2 \cdot \text{diag}(x)^2 \begin{bmatrix} x_1 & 0 & \cdots & 0 \\ 0 & x_2 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & x_n \end{bmatrix}.
\]

By the chain rule, the derivative of \( f(x) \) is

\[
f'(x) = h'(g(x))g'(x) = 2(x^2)^T A \cdot 2\text{diag}(x) = 4(x^2)^T A \cdot \text{diag}(x).
\]
So its gradient is 
\[ \nabla f(x) = f'(x)^T = 4 \cdot \text{diag}(x)Ax^2. \]
The Hessian of \( f(x) \) is the derivative of \( \nabla f(x) \). Note that 
\[ \nabla f(x + h) = 4 \cdot \text{diag}(x + h)A(x + h)^2 = 4 \cdot \text{diag}(x + h)A(x^2 + 2xh + h^2) = \]
\[ \nabla f(x) + 4 \cdot \text{diag}(h)Ax^2 + 8 \cdot \text{diag}(x)A(xh) + O(h^2) \]
Because \( \text{diag}(h)Ax^2 = \text{diag}(Ax^2)h \), \( \text{diag}(x)A(xh) = \text{diag}(x)A \text{diag}(x)h \),
we can get the expression 
\[ \nabla f(x + h) - \nabla f(x) = 4 \cdot \text{diag}(Ax^2)h + 8 \cdot \text{diag}(x)A \cdot \text{diag}(x)h + O(h^2). \]
Therefore, we get the Hessian 
\[ \nabla^2 f(x) = \left( \nabla f(x) \right)' = 4 \cdot \text{diag}(Ax^2) + 8 \cdot \text{diag}(x)A \cdot \text{diag}(x). \]

6. (10 points) Let \( A \in \mathbb{R}^{n \times n} \) be a symmetric positive definite matrix. Express the gradient and Hessian of \( f(x) := \sqrt{1 + x^TAx} \) in terms of \( A \) and \( x \).

**Solution:** Let \( h(t) = \sqrt{1 + t} \) and \( g(x) = x^TAx \), then 
\[ f(x) = h(g(x)), \quad f'(x) = h'(g(x))g'(x). \]
Since \( h'(t) = \frac{1}{2\sqrt{1+t}}, \quad g'(x) = 2x^TA \), we can get 
\[ \nabla f(x) = \left( f'(x) \right)^T = \frac{Ax}{\sqrt{1 + x^TAx}} \]
Similarly, we can get that
\[ \left( \frac{1}{\sqrt{1 + x^TAx}} \right)' = -\frac{x^TA}{\sqrt{1 + x^TAx}}. \]
The Hessian \( \nabla^2 f(x) \) is the derivative of the gradient \( \nabla f(x) \). By the product rule, we have 
\[ \nabla^2 f(x) = \left( \nabla f(x) \right)' = (Ax)' \frac{1}{\sqrt{1 + x^TAx}} + Ax \left( \frac{1}{\sqrt{1 + x^TAx}} \right)' \]
\[ = \frac{A}{\sqrt{1 + x^TAx}} - \frac{Axx^TA}{\sqrt{1 + x^TAx}}. \]
We can also do this by doing the expansion
\[ \nabla f(x+h) = \frac{A(x+h)}{2\sqrt{1 + (x+h)^T A(x+h)}} = \frac{Ax + Ah}{2\sqrt{1 + x^T Ax + 2x^T Ah + h^T Ah}} \]

Note the Taylor expansion
\[ \frac{1}{\sqrt{1 + x^T Ax + \alpha}} = \frac{1}{\sqrt{1 + x^T Ax}} - \frac{\alpha}{2\sqrt{1 + x^T Ax}} + O(\alpha^2) \]

If we let \( \alpha = 2x^T Ah + h^T Ah \), then we get that
\[ \nabla f(x+h) - \nabla f(x) = \frac{Ah}{\sqrt{1 + x^T Ax}} - \frac{Axx^T Ah}{\sqrt{1 + x^T Ax}} + \text{square terms in } h. \]

By the definition, the Hessian is
\[ \nabla^2 f(x) = \frac{A}{\sqrt{1 + x^T Ax}} - \frac{Axx^T A}{\sqrt{1 + x^T Ax}}. \]

7. (10 points) Let \( \| \cdot \| \) be a norm on \( \mathbb{R}^n \). Define the matrix function
\[ \| A \| := \max_{x \neq 0} \frac{\| Ax \|}{\| x \|}. \]

Show that \( \| A \| \) is a norm function in \( \mathbb{R}^{n \times n} \) and is an operator norm, i.e., show that
\[ \| AB \| \leq \| A \| \cdot \| B \| \]
for all \( A, B \in \mathbb{R}^{n \times n} \). In particular, show that \( \| A \| \geq |\lambda| \) for every real eigenvalue \( \lambda \) of \( A \).

**Solution:** We can see that \( \| A \| \geq 0 \) for all \( A \). If \( \| A \| = 0 \), then \( \| Ax \| = 0 \) for all vectors \( x \). Since \( \| \cdot \| \) is a vector norm, \( Ax = 0 \) for all \( x \), so \( A \) must be the zero matrix. Also note that
\[ \| \alpha A \| := \max_{x \neq 0} \frac{\| \alpha Ax \|}{\| x \|} = |\alpha| \max_{x \neq 0} \frac{\| Ax \|}{\| x \|} = |\alpha| \cdot \| A \|. \]

For all \( A, B \), we have that
\[ \| (A + B)x \| = \| Ax + Bx \| \leq \| Ax \| + \| Bx \|. \]
Therefore, we can see
\[ \|A + B\|_\infty = \max_{x \neq 0} \frac{\|(A + B)x\|_\infty}{\|x\|_\infty} \leq \max_{x \neq 0} \frac{\|Ax\|_\infty + \|Bx\|_\infty}{\|x\|_\infty} \]
\[ \leq \max_{x \neq 0} \frac{\|Ax\|_\infty}{\|x\|_\infty} + \max_{x \neq 0} \frac{\|Bx\|_\infty}{\|x\|_\infty} = \|A\|_\infty + \|B\|_\infty. \]

So \(\|A\|_\infty\) is a matrix norm. It is also an operator norm, because
\[ \|AB\|_\infty = \max_{x \neq 0} \frac{\|ABx\|_\infty}{\|x\|_\infty} = \max_{x \neq 0} \frac{\|ABx\|_\infty}{\|x\|_\infty} \]
\[ \leq \left( \max_{x \neq 0} \frac{\|ABx\|_\infty}{\|x\|_\infty} \right) \left( \max_{x \neq 0} \frac{\|Bx\|_\infty}{\|x\|_\infty} \right) \leq \|A\|_\infty \cdot \|B\|_\infty. \]

If \(\lambda\) is a real eigenvalue, then there is a real eigenvector \(u \neq 0\) such that
\[ Au = \lambda u, \]
so
\[ \|Au\|_\infty = \|\lambda u\|_\infty = |\lambda|. \]

By the definition, we know \(\|A\|_\infty \geq |\lambda|\).

8. (10 points) Let \(\{x_k\} \subseteq \mathbb{R}\) be a sequence such that \(x_k \to x^* > 0\), with convergence order \(r \geq 1\). Determine the convergence order of the sequence \(\{x_k^2\}\).

**Solution:** By the definition, we know
\[ \lim_{k \to \infty} \frac{|x_{k+1} - x^*|}{|x_k - x^*|^r} = a > 0. \]

Note that \(x_k^2 \to (x^*)^2\), so we have
\[ \lim_{k \to \infty} \frac{|x_{k+1}^2 - (x^*)^2|}{|x_k^2 - (x^*)^2|^r} = \lim_{k \to \infty} \frac{|(x_{k+1} + x^*)(x_{k+1} - x^*)|}{|(x_k + x^*)(x_k - x^*)|^r} = a \frac{|2x^*|}{|2x^*|^r} > 0. \]

So the convergence order is still \(r\).

9. (10 points) Let \(f(x) \in C^2[a, b]\) be a function such that \(f'(a) \geq 0\) and \(f''(x) \geq 0\) for all \(x \in (a, b)\). show that \(f(x) \geq f(a)\) for all \(x \in [a, b]\).

**Solution:** By the Taylor expansion, for all \(x \in [a, b]\), there exists \(\theta \in (0, 1)\) such that
\[ f(x) = f(a) + f'(a)(x - a) + \frac{f''(a + \theta(x-a))}{2}(x-a)^2. \]
For the given condition, we know \( f''(a + \theta(x - a)) \geq 0 \), so
\[
f(x) \geq f(a) + f'(a)(x - a) \geq f(a).
\]

10. (10 points) Suppose \( f(x) \in C[a, b] \). Show that there exists a number \( \xi \in [a, b] \) such that
\[
f(\xi) = \frac{f(a) + f(\frac{a+b}{2}) + f(b)}{3}.
\]

**Solution:** Let \( f_{\text{max}}, f_{\text{min}} \) be the maximum, minimum value of \( f(x) \) in the interval \([a, b]\). Then note that
\[
f_{\text{min}} \leq \frac{f(a) + f(\frac{a+b}{2}) + f(b)}{3} \leq f_{\text{max}}.
\]

By the intermediate value theorem, there must exist \( \xi \in [a, b] \) such that
\[
f(\xi) = \frac{f(a) + f(\frac{a+b}{2}) + f(b)}{3}.
\]