Problem 1.

Let $\boldsymbol{x} = (x_1, \cdots, x_m)^\top \in \mathbb{R}^m$ be the outcome of m repeated trials, and we denote the mean by μ. An empirical estimate of the second-order raw moment is

$$\frac{1}{m}\sum_{i=1}^{m}(x_i^2)$$

and a (biased) empirical estimate of the second-order central moment is

$$\frac{1}{m}\sum_{i=1}^{m}[(x_i - \mu)^2]$$

Express them in vector representations. In other words, you may use $\boldsymbol{x}$, μ, and vector/scalar operators. Summations are not allowed.

Here, nth order raw moment is $\mathbb{E}[X^n]$ and nth order central moment is $\mathbb{E}[(X-\mu)^n]$. The definitions per se are not needed for solving the problem.

Hint: The question is to write math equations, not code.

Solution:

Second-order raw moment in vector representation: $\frac{1}{m}\boldsymbol{x}^\top\boldsymbol{x}$
Second-order central moment: $\frac{1}{m}(\boldsymbol{x} - \mu\mathbb{1})^\top(\boldsymbol{x} - \mu\mathbb{1}))$.
The superscript T denotes transpose. The $\mathbb{1}$ is a vector of all ones of dimension m.

Problem 2.

Consider a two dimensional space $\mathbb{R}^2$. Determine whether the following sets are convex or not. Prove or disprove.

- $\{(x_1, x_2)\colon x_1^{\ 2} + x_2^{\ 2} = 1\}$
- $\{(x_1, x_2)\colon |x_1| + |x_2| \le 1\}$

Solution:

Part 1: Let $A_1 \doteq \{(x_1, x_2)|x_1^2 + x_2^2 = 1\}$. We show that A_1 is not convex by the following counterexample. Pick $(1,0) \in A_1$ and $(0,1) \in A_1$, and let $c = \alpha(1,0)+(1-\alpha)(0,1) = (\alpha,(1-\alpha))$, where $\alpha \in [0,1]$ such that $\alpha+(1-\alpha) = 1$. Just pick $\alpha = 0.5$, $c \notin A_1$. We can verify this as follows,

$$\alpha^2 + (1-\alpha)^2 = 0.5^2 + 0.5^2 = 0.50 \neq 1. \tag{1}$$

In summary, we cannot just pick any two arbitrary points from A_1 and have the convex combination of the two points also be in A_1, and thus A_1 is not convex.

Part 2: Let $A_2 \doteq \{(x_1, x_2)||x_1| + |x_2| \le 1\}$. Pick any $a, b \in A_2$, let $c = \alpha a + (1-\alpha)b$ for an arbitrary $\alpha \in [0,1]$ s.t. $\alpha + (1-\alpha) = 1$. We now show that $c \in A_2$. Note that $c = ((\alpha a_1 + (1-\alpha)b_1), (\alpha a_2 + (1-\alpha)b_2))$, then

$$\begin{aligned}
&|(\alpha a_1 + (1-\alpha)b_1)| + |(\alpha a_2 + (1-\alpha)b_2)| && (2)\\
&\le \alpha|a_1| + (1-\alpha)|b_1| + \alpha|a_2| + (1-\alpha)|b_2| \quad \text{by triangles's inequality} && (3)\\
&= \alpha(|a_1| + |a_2|) + (1-\alpha)(|b_1| + |b_2|) && (4)\\
&\le \alpha + (1-\alpha) && (5)\\
&= 1. && (6)
\end{aligned}$$

To get from eq. (4) to eq. (5) is because $a, b \in A_2$ by setup and this means that a, b both satify the condition $|a_1|+|a_2| \le 1$ and $|b_1|+|b_2| \le 1$ respectively. We've just shown that $c \in A_1$. Since a, b, α are arbitrary, and any convex combination of two points from A_2 would also be in A_2. In conclusion, A_2 is convex.

Problem 3.

Consider the function $f(x_1, x_2) = x_1^2 + x_2^2 - 4x_1x_2$

a) View x_1 as a variable and x_2 as a constant. Determine whether f is convex in x_1 and prove it.
b) View x_2 as a variable and x_1 as a constant. Determine whether f is convex in x_2 and prove it.
c) View $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ as a function of the input vector (x_1, x_2). Determine whether f is convex in (x_1, x_2) and prove it.

Hints: For a) and b), treat one variable as a constant, and calculate the second-order derivative of a single-variable function.

For c), calculate the Hessian matrix H first and choose a point, say, (0,0). You may use numpy in Python to calculate the eigenvalue

```
import numpy as np
from numpy import linalg as LA
H = np.array([ [11, 12], [21, 22]])  # your values here
eigenval, eigenvec = LA.eig(H)
```

Print `eigenval`. If any number is less than 0, then the function is not convex. Otherwise, it is convex. Eigenvalues may also be calculated manually.

The example shows that an element-wise convex function may not be jointly convex.

Solution:

Part a), fix x_2, and take the the derivative of $f(x_1, x_2)$ w.r.t. x_1, we get

$$2x_1 - 4x_2. \tag{7}$$

Taking the derivative of eq. (7), we get 2, which means that the function is convex in x_1 since $f''(x_1, x_2) \geq 0$.

Part b), fix x_1 and $f''(x_1, x_2) = 2$, which also means that the function is convex in x_2.

Part c), calculate the Hessian

$$\frac{\partial^2 f}{\partial x_1 \partial x_1} = 2 \tag{8}$$

$$\frac{\partial^2 f}{\partial x_2 \partial x_2} = 2 \tag{9}$$

$$\frac{\partial^2 f}{\partial x_1 \partial x_2} = \frac{\partial^2 f}{\partial x_2 \partial x_1} = -4. \tag{10}$$

Check if the Hessian matrix is positive semi-definite by entering the matrix $\begin{bmatrix} 2 & -4 \\ -4 & 2 \end{bmatrix}$ into Python and compute the eigenvalues, and we get $[6, -2]$, which indicate the Hessian is not positive semi-definite. The function is not jointly convex in x_1 and x_2.

Problem 4.

Prove that a differentiable convex function satisfies the first-order condition.

Hint:
http://www.princeton.edu/~aaa/Public/Teaching/ORF523/S16/ORF523_S16_Lec7_gh.pdf

Solution:

By the definition of convex function that for all $x, y \in dom(f)$, for all $\lambda \in (0, 1)$,

$$\lambda f(x) + (1 - \lambda) f(y) \geq f(\lambda x + (1 - \lambda) y). \tag{11}$$

We rearrange eq. (11), and get

$$f(x) \geq f(y) + \frac{f(\lambda(x - y) + y) - f(y)}{\lambda}. \tag{12}$$

Take the limit of both sides of eq. (12) as $\lambda \to 0$, we get $f(x) \geq f(y) + \nabla f(y)^\top (x - y)$. Hence, any convex differentiable function satisfies first-order condition.

END