3. Expectation and Covariance of a Random Vector

Review: Vector Outer Product I

3/3 points (ungraded)

Let $\mathbf{x}, \mathbf{y} \in \mathbb{R}^3$ denote column vectors. Consider the product $\mathbf{x}\mathbf{y}^T$. This is referred to as the **outer product** of the vectors $\mathbf{x}$ and $\mathbf{y}$.

How many rows are in $\mathbf{x}\mathbf{y}^T$?

3

✓ **Answer:** 3

How many columns are in $\mathbf{x}\mathbf{y}^T$?

3

✓ **Answer:** 3

Is the matrix $\mathbf{x}\mathbf{y}^T$ always symmetric?

○ Yes

● No ✓

Solution:

The vector $\mathbf{x} \in \mathbb{R}$ is a column vector, so it can alternatively be thought of as a 3×1 matrix. Similarly, $\mathbf{y}^T$ is a 1×3 matrix, so the product $\mathbf{x}\mathbf{y}^T$ is a 3×3 matrix.

Moreover, by the rule for matrix multiplication,

$$(\mathbf{x}\mathbf{y}^T)_{ij} = \mathbf{x}^i \mathbf{y}^j.$$

Therefore, if $\mathbf{x}^i \mathbf{y}^j \neq \mathbf{x}^j \mathbf{y}^i$ for some i, j, then the matrix $\mathbf{x}\mathbf{y}^T$ is not symmetric. For example, if we let

$$\mathbf{x} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \quad \mathbf{y} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix},$$

then

$$\mathbf{x}\mathbf{y}^T = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 \\ 2 & 2 & 2 \\ 3 & 3 & 3 \end{pmatrix}$$

which is **not** symmetric.

Remark: In this chapter, we will usually have $\mathbf{y} = \mathbf{x}$, so we will be looking at the outer product of $\mathbf{x}$ with itself, which is $\mathbf{x}\mathbf{x}^T$. This is symmetric in general because

$$(\mathbf{x}\mathbf{x}^T)_{ij} = \mathbf{x}^i \mathbf{x}^j = (\mathbf{x}\mathbf{x}^T)_{ji}.$$

Submit You have used 2 of 3 attempts

Answers are displayed within the problem

Review: Vector Outer Product II

3/3 points (ungraded)
Consider the vector

$$\mathbf{x} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}.$$

Consider the matrix product $\mathbf{x}\mathbf{x}^T$.

What is $(\mathbf{x}\mathbf{x}^T)_{11}$?

1 ✓ **Answer:** 1

What is $(\mathbf{x}\mathbf{x}^T)_{21}$?

2 ✓ **Answer:** 2

What is $(\mathbf{x}\mathbf{x}^T)_{23}$?

6 ✓ **Answer:** 6

Solution:

The outer product of $\mathbf{x}$ with itself is given by

$$\mathbf{x}\mathbf{x}^T = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 3 & 6 & 9 \end{pmatrix}$$

so $(\mathbf{x}\mathbf{x}^T)_{11} = 1$, $(\mathbf{x}\mathbf{x}^T)_{21} = 2$, and $(\mathbf{x}\mathbf{x}^T)_{23} = 6$.

Submit You have used 1 of 3 attempts

Answers are displayed within the problem

Review: Expectation of a Random Vector

2.0/2 points (ungraded)
Let $\mathbf{X} \in \mathbb{R}^3$ denote a random vector.

Then $\mathbb{E}[\mathbf{X}]$ is...

○ A number in $\mathbb{R}$.

● A vector in $\mathbb{R}^3$. ✓

○ A matrix in $\mathbb{R}^{3\times 3}$

○ None of the above.

Suppose that

$$\mathbf{X} \sim N\left(\begin{pmatrix} -10 \\ 0 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 0 \\ 2 & 2 & 1 \\ 0 & 1 & 1 \end{pmatrix}\right).$$

What is $\mathbb{E}[\mathbf{X}]$?

(Enter your answer as a vector, e.g., type **[3,2]** for the vector $\begin{pmatrix} 3 \\ 2 \end{pmatrix}$).

$\mathbb{E}[\mathbf{X}] =$ [-10,0,2] ✓ **Answer:** [-10,0,2]

Solution:

It is important to remember the definition

$$\mathbf{E}[\mathbf{X}]_i = \mathbf{E}[\mathbf{X}^i].$$

Note that the diagonal entries of the given covariance matrix denote the variances of $\mathbf{X}^1, \mathbf{X}^2$, and $\mathbf{X}^3$. Therefore,

$$\mathbf{X}^1 \sim N(-10, 1), \ \mathbf{X}^2 \sim N(0, 2), \ \mathbf{X}^3 \sim N(2, 1).$$

It follows that

$$\begin{aligned} \mathbb{E}[\mathbf{X}]_1 &= \mathbb{E}[\mathbf{X}^1] = -10 \\ \mathbb{E}[\mathbf{X}]_2 &= \mathbb{E}[\mathbf{X}^2] = 0 \\ \mathbb{E}[\mathbf{X}]_3 &= \mathbb{E}[\mathbf{X}^3] = 2. \end{aligned}$$

Remark: Observe that the mean of $\mathbf{X}$ does not depend on the covariance structure.

Submit You have used 1 of 3 attempts

ℹ Answers are displayed within the problem

Review: Variance and Covariance of Random Variables

2/2 points (ungraded)
Let $X \in [0, 1]$ denote a bounded random variable. The variance of X is defined to be

$$\mathsf{Var}(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2.$$

Equivalently, we may write

$$\mathsf{Var}(X) = \mathbb{E}[(X - A)^2]$$

for some constant A that depends on the distribution of X.

What is A?

● $\mathbb{E}[X]$ ✓

○ $\mathbb{E}[X^2]$

○ $(\mathbb{E}[X])^2$

○ None of the above.

Let $Y \in [0,1]$ denote another bounded random variable. Assume that X and Y have a joint distribution, but are not necessarily independent. The covariance of X and Y is defined to be

$$\text{Cov}(X,Y) = \mathbb{E}[XY] - \mathbb{E}[X]\,\mathbb{E}[Y]\,.$$

Equivalently, we may write

$$\text{Cov}(X,Y) = \mathbb{E}[(X-B)(Y-C)]$$

for some constants B and C that depend on the distribution of X and Y, respectively.

● $B = \mathbb{E}[X], C = \mathbb{E}[Y]$ ✓

○ $B = \mathbb{E}[Y], C = \mathbb{E}[X]$

○ $B = (\mathbb{E}[Y])^2, C = (\mathbb{E}[X])^2$

○ $B = \mathbb{E}[Y^2], C = \mathbb{E}[X^2]$

Solution:

We examine the questions in order. First we note that $\mathbb{E}[X^2]$, $\mathbb{E}[Y^2]$, and $\mathbb{E}[XY]$ are all finite because the random variables $X, Y \in [0,1]$ are finite.

For the first question, observe that

$$\mathbb{E}[(X-\mathbb{E}[X])^2] = \mathbb{E}[X^2 - 2X\mathbb{E}[X] + (\mathbb{E}X)^2] = \mathbb{E}[X]^2 - (\mathbb{E}[X])^2.$$

Hence, the correct response to the first question is $A = \mathbb{E}[X]$.

For the second question, observe that

$$\mathbb{E}[(X-\mathbb{E}[X])(Y-\mathbb{E}Y)] = \mathbb{E}[XY - X\mathbb{E}Y - Y\mathbb{E}X + \mathbb{E}[X]\,\mathbb{E}[Y]] = \mathbb{E}[XY] - \mathbb{E}[X]\,\mathbb{E}[Y]$$

Submit You have used 1 of 3 attempts

ℹ Answers are displayed within the problem

Review: Covariance of Random Vectors

2/2 points (ungraded)

Let $\mathbf{X} \in \mathbb{R}^d$ denote a random vector. Recall the **covariance matrix** of $\mathbf{X}$ is defined to be

$$\Sigma = \mathbb{E}\left[\mathbf{X}\mathbf{X}^T\right] - \mathbb{E}\left[\mathbf{X}\right]\mathbb{E}[\mathbf{X}]^T.$$

The covariance matrix can also be expressed as

$$\Sigma = \mathbb{E}\left[(\mathbf{X} - A)(\mathbf{X} - A)^T\right]$$

where A is a matrix that depends on the distribution of $\mathbf{X}$.

What is A?

● $\mathbf{E}[\mathbf{X}]$ ✓

○ $\mathbf{E}[\mathbf{X}\mathbf{X}^T]$

○ $\mathbf{E}[\mathbf{X}^T\mathbf{X}]$

○ None of the above.

What is Σ_{ij}?

○ $\mathbb{E}[\mathbf{X}^i\mathbf{X}^j]$

○ $\mathbb{E}[\mathbf{X}^i]\,\mathbb{E}[\mathbf{X}^j]$

○ $(\mathbb{E}[\mathbf{X}^i\mathbf{X}^j])^2$

● Cov $(\mathbf{X}^i, \mathbf{X}^j)$ ✓

Solution:

We examine the questions in order.

For the first question, observe that

$$\begin{aligned}\mathbb{E}\left[(\mathbf{X} - \mathbb{E}[\mathbf{X}])(\mathbf{X} - \mathbb{E}[\mathbf{X}])^T\right] &= \mathbb{E}\left[\mathbf{X}\mathbf{X}^T - \mathbf{X}\mathbb{E}[\mathbf{X}]^T - \mathbb{E}[\mathbf{X}]\mathbf{X}^T + \mathbb{E}[\mathbf{X}]\mathbb{E}[\mathbf{X}]^T\right] \\ &= \mathbb{E}[\mathbf{X}\mathbf{X}^T] - \mathbb{E}[\mathbf{X}]\mathbb{E}[\mathbf{X}]^T.\end{aligned}$$

The second line follows by the linearity of expectation. Therefore $A = \mathbb{E}[\mathbf{X}]$.

For the second question, observe that

$$\begin{aligned}\Sigma_{ij} &= (\mathbb{E}[\mathbf{X}\mathbf{X}^T]_{ij} - (\mathbb{E}[\mathbf{X}]\mathbb{E}[\mathbf{X}]^T)_{ij} \\ &= \mathbb{E}[\mathbf{X}^i\mathbf{X}^j] - \mathbb{E}[\mathbf{X}]^i\mathbb{E}[\mathbf{X}]^j \\ &= \text{Cov}(\mathbf{X}^i, \mathbf{X}^j).\end{aligned}$$

Submit

You have used 1 of 3 attempts