8. Covariance Matrices

Note: Now is a good time to review the matrix exercises in Homework 0.

Note on Notation: In this course, we assume all vectors to be column vectors. Therefore, while

$$\mathbf{X} = \begin{bmatrix} X^{(1)} \\ X^{(2)} \\ \vdots \\ X^{(d)} \end{bmatrix},$$

we sometimes write it as $\mathbf{X} = \left(X^{(1)}, \ldots, X^{(d)}\right)^T$ to be more compact in representation.

Example of Covariance II

4/4 points (graded)

Let X, Y be random variables such that

- X takes the values ± 1 each with probability 0.5
- (Conditioned on X) Y is chosen uniformly from the set $\{-3X-1, -3X, -3X+1\}$.

(Round all answers to 2 decimal places.)

What is Cov (X, X) (equivalent to Var (X))?

Cov $(X, X) =$ 1 Answer: 1.0

What is Cov (Y, Y) (equivalent to Var (Y))?

Cov $(Y, Y) =$ 29/3 Answer: 9.67

What is Cov (X, Y)?

Cov $(X, Y) =$ -3 Answer: -3.00

What is Cov (Y, X)?

Cov $(Y, X) =$ -3 Answer: -3.00

Solution:

Observe that $\mathbb{E}[X]$ and $\mathbb{E}[Y]$ are both zero, since X is uniformly distributed over $\{\pm 1\}$ and Y is uniformly distributed over the set $\{-4, -3, -2, 2, 3, 4\}$.

- Cov (X, X) is the variance of X, which equals $\mathbb{E}[X^2] - \mathbb{E}[X]^2 = p + (1-p) = 1.$
- Cov (Y, Y) is the variance of Y, which equals $\mathbb{E}[Y^2] - \mathbb{E}[Y]^2 = \frac{16+9+4+4+9+16}{6} = \frac{29}{3} \approx 9.67.$
- Cov (X, Y) and Cov (Y, X) are always equal, by the symmetry of the definition. Observe that the joint density of (X, Y) is uniform over the pairs $(1,-4), (1,-3), (1,-2), (-1,2), (-1,3), (-1,4)$. Thus, either covariance can be computed as $\mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] = \frac{-4-3-2-2-3-4}{6} = -3.$

提交 你已经尝试了2次（总共可以尝试3次）

Answers are displayed within the problem

Covariance Matrix

4/4 points (graded)

Given random variables $X^{(1)}, X^{(2)}, \ldots, X^{(d)}$, one can write down the **covariance matrix** Σ, where $\Sigma_{i,j} = \mathrm{Cov}\left(X^{(i)}, X^{(j)}\right)$.

Let $X^{(1)}$, $X^{(2)}$ be random variables such that

- $X^{(1)}$ takes the values ± 1 each with probability 0.5
- (Conditioned on $X^{(1)}$) $X^{(2)}$ is chosen uniformly from the set $\{-3X^{(1)} - 1, -3X^{(1)}, -3X^{(1)} + 1\}$.

What is the covariance matrix Σ?

$\Sigma_{1,1} =$ 1 Answer: 1.0 $\Sigma_{1,2} =$ -3 Answer: -3.00

$\Sigma_{2,1} =$ -3 Answer: -3.00 $\Sigma_{2,2} =$ 29/3 Answer: 9.67

Solution:

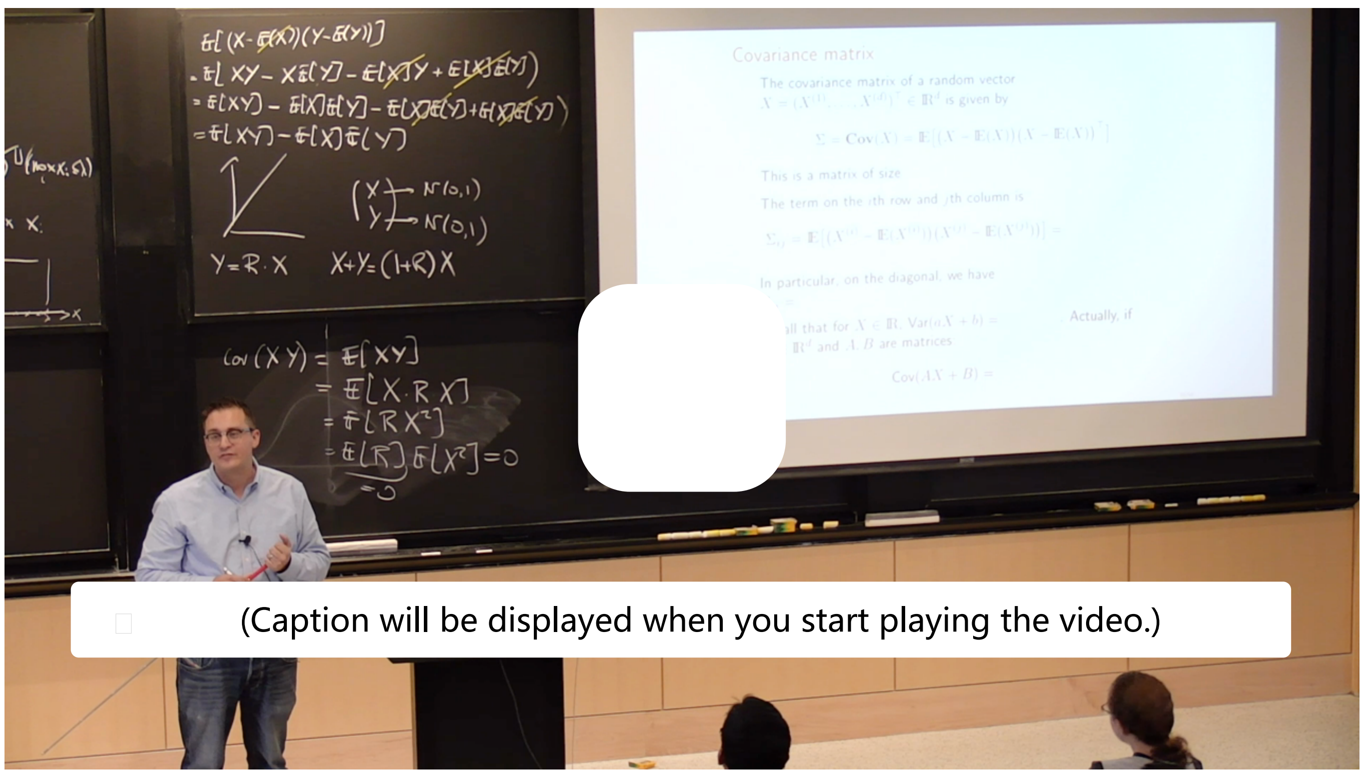
Using the answer to the previous question, the 2×2 covariance matrix Σ evaluates to

$$\Sigma = \begin{pmatrix} \Sigma_{1,1} & \Sigma_{1,2} \\ \Sigma_{2,1} & \Sigma_{2,2} \end{pmatrix} = \begin{pmatrix} 1 & -3 \\ -3 & \frac{29}{3} \end{pmatrix}$$

提交 你已经尝试了1次（总共可以尝试3次）

Answers are displayed within the problem

Covariance Matrix: Definitions



in a simple way, which is by informing the matrix that's

covariance of xx, covariance of yy, covariance of xy,

and covariance of yx.

So clearly, it's a symmetric matrix,

because these are symmetric numbers.

And so all I'm saying is that I can gather

all the information about covariance and variance

in one matrix.

In this case, it's 2 by 2.

You can do that more generally, when you have, say, d by d.

If you have a vector of size d, you

can talk about its covariance matrix.

And each entry on the diagonal will just be a variance term,

视频

下载视频文件

字幕

下载 SubRip (.srt) file

下载 Text (.txt) file

Here is a compact formula for the covariance matrix using vector notation.

Let $\mathbf{X} = \begin{pmatrix} X^{(1)} \\ \vdots \\ X^{(d)} \end{pmatrix}$ be a random vector of size $d \times 1$.

Let $\mu \triangleq \mathbb{E}[\mathbf{X}]$ denote the **entry-wise** mean, i.e.

$$\mathbb{E}[\mathbf{X}] = \begin{pmatrix} \mathbb{E}[X^{(1)}] \\ \vdots \\ \mathbb{E}[X^{(d)}] \end{pmatrix}.$$

Consider the vector outer product (refer to Homework 0) $(\mathbf{X} - \mu)(\mathbf{X} - \mu)^T$, which is a random $d \times d$ matrix. Then the **covariance matrix** Σ can be written as

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

Note: The following exercises will be discussed as properties of covariance in the video that follows, but we encourage you attempt these exercises before watching the video.

Covariance Matrix: Properties I

1/1 point (graded)

Let $\mathbf{X}$ be a random vector and let $\mathbf{Y} = \mathbf{X} + B$, where B is a constant vector. Let $\mu_{\mathbf{X}}$ be the mean vector of $\mathbf{X}$ and let $\Sigma_{\mathbf{X}}$ be the covariance matrix of $\mathbf{X}$. Select from the following all statements that are correct.

- [ ] The covariance matrix of $\mathbf{Y}$ could potentially be equal to $\Sigma_{\mathbf{X}}$ only under some conditions imposed on B
- [x] The covariance matrix of $\mathbf{Y}$ is the same as $\Sigma_{\mathbf{X}}$ for all vectors B
- [x] The covariance matrix of $\mathbf{Y}$ has the same size as the matrix $\Sigma_{\mathbf{X}}$
- [ ] The covariance matrix of $\mathbf{Y}$ is the same as $\Sigma_{\mathbf{X}}$ if and only if vector B is equal to 0

Solution:

Choices 2 and 3 are correct. Let the covariance matrix of $\mathbf{Y}$ be denoted $\Sigma_{\mathbf{Y}}$. Note that $\mathbb{E}[\mathbf{X} + B] = \mu_{\mathbf{X}} + B$ for any vector B.

$$\Sigma_{\mathbf{Y}} = \mathbb{E}\left[(\mathbf{X} + B - \mu_{\mathbf{X}} - B)(\mathbf{X} + B - \mu_{\mathbf{X}} - B)^T\right] = \Sigma_{\mathbf{X}}$$

Since choice 2 is correct, choices 1 and 4 that impose certain conditions on B are technically incorrect as we do not require that B satisfiy some conditions for $\Sigma_{\mathbf{Y}}$ to be the same as $\Sigma_{\mathbf{X}}$.

提交 你已经尝试了1次（总共可以尝试2次）

Answers are displayed within the problem

Covariance Matrix: Properties II

1/1 point (graded)

Let $\mathbf{X}$ be a random vector of size $d \times 1$ and let $\mathbf{Y} = A\mathbf{X} + B$, where A is a constant matrix of size $n \times d$ and B is a constant vector of size $n \times 1$. Let $\mu_{\mathbf{X}}$ be the mean vector of $\mathbf{X}$ and let $\Sigma_{\mathbf{X}}$ be the covariance matrix of $\mathbf{X}$. Let $\mu_{\mathbf{Y}}$ be the mean vector of $\mathbf{Y}$ and let $\Sigma_{\mathbf{Y}}$ be the covariance matrix of $\mathbf{Y}$.

Select from the following all statements that are correct.

- [x] $\Sigma_{\mathbf{Y}}$ is the same as covariance matrix of $A\mathbf{X}$
- [x] $\Sigma_{\mathbf{Y}}$ is of size $n \times n$
- [ ] $\Sigma_{\mathbf{Y}} = A^2 \Sigma_{\mathbf{X}}$
- [x] $\Sigma_{\mathbf{Y}} = A \Sigma_{\mathbf{X}} A^T$
- [ ] $\Sigma_{\mathbf{Y}} = A^T \Sigma_{\mathbf{X}} A$

Solution:

As $\mathbf{Y}$ is an $n \times 1$ random vector, $\Sigma_{\mathbf{Y}}$ is of size $n \times n$.

From the previous problem we know that $\Sigma_{\mathbf{Y}}$ is the same as the covariance matrix of $A\mathbf{X}$. Therefore, it suffices to find this matrix, which we denote $\Sigma_{A\mathbf{X}}$.

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

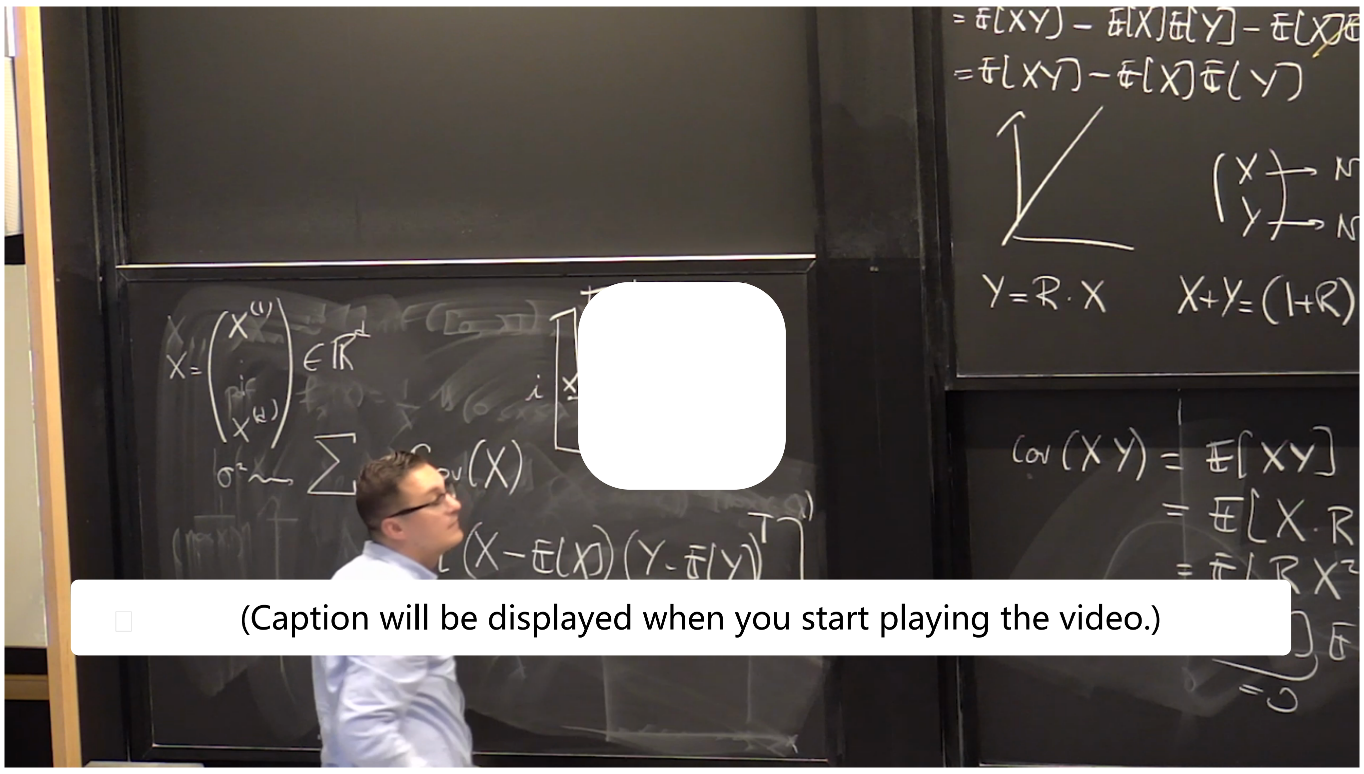
Therefore, choices 1, 2, and 4 are correct.

Choices 3 and 5 are not correct in general (even if A is a square matrix) because matrix multiplication is not commutative.

提交 你已经尝试了3次（总共可以尝试3次）

Answers are displayed within the problem

Covariance Matrix: Affine Transformation



non-negative, right?

So if I give you a covariance matrix

with a negative entry on the diagonal,

you can probably just--

you can skip and say you made a mistake.

OK, now recall that for x and y, If I

look at the variance of a linear and a fine transformation of x,

so something of the form AX plus B, what is this?

a squared variance of x.

And since the covariance matrix is just

a multivariate, a matrix generalization of the variance,

it's going to have similar properties.

In particular, it should, because I

know that if I do a linear transformation, then at least

视频
下载视频文件

字幕
下载 SubRip (.srt) file
下载 Text (.txt) file

Effect of Linear Transformations of Covariance Matrix

4/4 points (graded)

Let $\mathbf{X} = \begin{pmatrix} X^{(1)} \\ X^{(2)} \end{pmatrix}$ be a random vector with covariance Matrix $\Sigma_{\mathbf{X}} = \begin{pmatrix} 1 & 1/2 \\ 1/2 & 1 \end{pmatrix}$.

Let $\mathbf{Y} = M\mathbf{X}$, where $M = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$.

Observe that $Y^{(1)} = X^{(1)} - X^{(2)}$ and $Y^{(2)} = X^{(1)} + X^{(2)}$. What is the new covariance matrix Σ_Y?

$(\Sigma_{\mathbf{Y}})_{1,1} =$ 1 **Answer:** 1 $(\Sigma_{\mathbf{Y}})_{1,2} =$ 0 **Answer:** 0.0

$(\Sigma_{\mathbf{Y}})_{2,1} =$ 0 **Answer:** 0.0 $(\Sigma_{\mathbf{Y}})_{2,2} =$ 3 **Answer:** 3

Solution:

Recall from an earlier problem that for any pair of random variables A, B with the same variance $\textsf{Var}(A) = \textsf{Var}(B) = \sigma^2$, $\textsf{Cov}(A - B, A + B) = \textsf{Var}(A) - \textsf{Var}(B) = 0$.

Therefore, given the matrix M, $\Sigma_{\mathbf{Y}}$ must be a diagonal matrix.

We have
$\textsf{Cov}(Y^{(1)}, Y^{(1)}) = \textsf{Cov}(X^{(1)} - X^{(2)}, X^{(1)} - X^{(2)}) = \textsf{Cov}(X^{(1)}, X^{(1)}) - 2\textsf{Cov}(X^{(1)}, X^{(2)}) + \textsf{Cov}(X^{(2)}, X^{(2)}) = 1 - 1 + 1 = 1.$

Similarly,
$\textsf{Cov}(Y^{(2)}, Y^{(2)}) = \textsf{Cov}(X^{(1)} + X^{(2)}, X^{(1)} + X^{(2)}) = \textsf{Cov}(X^{(1)}, X^{(1)}) + 2\textsf{Cov}(X^{(1)}, X^{(2)}) + \textsf{Cov}(X^{(2)}, X^{(2)}) = 1 + 1 + 1 = 3.$

提交 你已经尝试了1次（总共可以尝试3次）

Answers are displayed within the problem

讨论

显示讨论

主题： Unit 3 Methods of Estimation:Lecture 10: Consistency of MLE, Covariance Matrices, and Multivariate Statistics / 8. Covariance Matrices