5. Empirical Mean and Covariance Matrix of a Vector Data Set II

Formula for the Empirical Covariance Matrix

1/1 point (ungraded)

Let $\mathbf{X}_1, \ldots, \mathbf{X}_n \in \mathbb{R}^d$ denote a data set, and let

$$\mathbb{X} = \begin{pmatrix} \longleftarrow & \mathbf{X}_1^T & \longrightarrow \\ \longleftarrow & \mathbf{X}_2^T & \longrightarrow \\ \vdots & \vdots & \vdots \\ \longleftarrow & \mathbf{X}_n^T & \longrightarrow \end{pmatrix}.$$

Let $\mathbf{S}$ denote its empirical covariance matrix.

Which of the following is the correct formula for $\mathbf{S}$?

In the choices below, I_n is the $n \times n$ identity matrix, and $\mathbf{1} \in \mathbb{R}^n$ is the vector with all 1 entries.

○ $\frac{1}{n}\mathbb{X}^T\left(I_n - \mathbf{1}^T\mathbf{1}\right)\mathbb{X}$

● $\frac{1}{n}\mathbb{X}^T\left(I_n - \frac{1}{n}\mathbf{1}\mathbf{1}^T\right)\mathbb{X}$

○ $\frac{1}{n}\mathbb{X}\,(I_d - \frac{1}{d}\mathbf{1}\mathbf{1}^T)\,\mathbb{X}^T$

○ $\frac{1}{n}\mathbb{X}\,(I_n - \frac{1}{n}\mathbf{1}\mathbf{1}^T)\,\mathbb{X}^T$

✓

Solution:

The second choice is correct:

$$\mathbf{S} = \frac{1}{n}\mathbb{X}^T\,(I_n - \frac{1}{n}\mathbf{1}\mathbf{1}^T)\,\mathbb{X}.$$

We make some notes about the incorrect choices.

- In the first choice, we have $\mathbf{1}^T\mathbf{1}$, which is a number, instead of $\mathbf{1}\mathbf{1}^T$, which is a matrix. The former is the vector **inner product** whereas the latter is the vector **outer product**.
- In the third choice, one can check that the formula has incompatible matrix operations. Specifically, I_d is $d \times d$ and $\mathbf{1}\mathbf{1}^T$ is $n \times n$, which makes $I_d - \frac{1}{d}\mathbf{1}\mathbf{1}^T$ an incompatible operation.
- In the fourth choice, the matrix product written is undefined because $\mathbb{X} \in \mathbb{R}^{n\times d}$ and $I_n - \frac{1}{n}\mathbf{1}\mathbf{1}^T \in \mathbb{R}^{n\times n}$.

Submit

You have used 1 of 3 attempts

ℹ Answers are displayed within the problem

Matrix Products Involving Outer Products

1/1 point (ungraded)

Let $\mathbf{x}$ denote a 3×1 column vector and let $\mathbf{1}$ denote the 3×1 column vector with all entries equal to 1. Suppose that the entries of $\mathbf{x}$ sum to 0. That is,

$$\mathbf{x}^1 + \mathbf{x}^2 + \mathbf{x}^3 = 0.$$

What is $(\mathbf{1}\mathbf{1}^T)\,\mathbf{x}$?

● $(0 \quad 0 \quad 0)^T$

○ It depends on the value of $\mathbf{x}$.

○ It is not defined.

✓

Solution:

Matrix multiplication is associative, so we may write

$$(\mathbf{1}\mathbf{1}^T)\,\mathbf{x} = \mathbf{1}\,(\mathbf{1}^T\mathbf{x})\,.$$

Note that

$$\mathbf{1}^T\mathbf{x} = \mathbf{x}^1 + \mathbf{x}^2 + \mathbf{x}^3 = 0$$

by assumption. Hence,

$$(\mathbf{1}\mathbf{1}^T)\,\mathbf{x} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}(0) = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

The first response is correct.

Submit You have used 1 of 3 attempts

Answers are displayed within the problem

An Orthogonal Projection Matrix I

6/6 points (ungraded)

The matrix

$$H = I_n - \frac{1}{n}\mathbf{1}\mathbf{1}^T$$

arises in the formula for the empirical covariance matrix of a data set. For simplicity, let's set $n = 3$. Then we have

Let $\mathbf{x} = (2, -1, -2)^T$.

What is $H\mathbf{x}$?

$(H\mathbf{x})^{(1)} =$ 7/3 ✓ **Answer:** 7/3

$(H\mathbf{x})^{(2)} =$ -2/3 ✓ **Answer:** -2/3

$(H\mathbf{x})^{(3)} =$ -5/3 ✓ **Answer:** -5/3

What is $H^2\mathbf{x}$?

$(H^2\mathbf{x})^{(1)} =$ 7/3 ✓ **Answer:** 7/3

$(H^2\mathbf{x})^{(2)} =$ -2/3 ✓ **Answer:** -2/3

$(H^2\mathbf{x})^{(3)} =$ -5/3 ✓ **Answer:** -5/3

Solution:

Using the result of the previous problem, we know that

$$\mathbf{1}\mathbf{1}^T\mathbf{x} = \mathbf{1} * (\mathbf{1} \cdot (2,-1,-2)^T) = -\mathbf{1}.$$

Therefore,

$$H\mathbf{x} = (I - \frac{1}{3}\mathbf{1}\mathbf{1}^T)\,\mathbf{x} = \mathbf{x} + \frac{1}{3}\mathbf{1} = \begin{pmatrix} 7/3 \\ -2/3 \\ -5/3 \end{pmatrix}.$$

Observe the entries of $H\mathbf{x}$ sum to 0. Therefore,

$$\mathbf{1}\mathbf{1}^T H\mathbf{x} = \mathbf{1}\,(\mathbf{1}^T \cdot \mathbf{x}) = (0,0,0)^T.$$

Hence,

$$H^2\mathbf{x} = (I - \frac{1}{3}\mathbf{1}\mathbf{1}^T)\,H\mathbf{x} = H\mathbf{x} = \begin{pmatrix} 7/3 \\ -2/3 \\ -5/3 \end{pmatrix}.$$

Submit You have used 2 of 3 attempts

ℹ Answers are displayed within the problem

An Orthogonal Projection Matrix II

3/3 points (ungraded)

As in the previous problem, we consider the matrix

$$H = I_n - \frac{1}{n}\mathbf{1}\mathbf{1}^T$$

and for simplicity let $n = 3$. Let $\mathbf{x} = (2, -1, -2)^T$.

What is $H^{100}\mathbf{x}$?

(Enter the components of $H^{100}\mathbf{x}$ below. Below, $(H^{100}\mathbf{x})^{(i)}$ denotes the i^{th} component of $H^{100}\mathbf{x}$.)

$(H^{100}\mathbf{x})^{(1)} =$ 7/3 ✓ **Answer:** 7/3

$(H^{100}\mathbf{x})^{(2)} =$ -2/3 ✓ **Answer:** -2/3

$(H^{100}\mathbf{x})^{(3)} =$ -5/3 ✓ **Answer:** -5/3

Solution:

The matrix H is an **orthogonal projection matrix** , which means that

0. H is symmetric, and

0. $H^2 = H$.

Therefore,

$$H^{100}\mathbf{x} \; = H^2 H^{98}\mathbf{x}$$

$$
\begin{aligned}
&= H^{99}\mathbf{x} \\
&\vdots \\
&= H\mathbf{x} \\
&= \begin{pmatrix} 7/3 \\ -2/3 \\ -5/3 \end{pmatrix}
\end{aligned}
$$

Remark: One could start with computing $H^k\mathbf{x}$ for a few small values of k and observe that the output is always the same. However, it is important to keep in mind the conceptual point that the matrix H is an orthogonal projection.

Submit You have used 1 of 3 attempts

ℹ Answers are displayed within the problem

Concept Check: Orthogonal Projections

1/1 point (ungraded)

Let $\mathbf{X}_1, \ldots, \mathbf{X}_n \in \mathbb{R}^d$ denote a data set and let

$$
\mathbb{X} = \begin{pmatrix} \longleftarrow & \mathbf{X}_1^T & \longrightarrow \\ \longleftarrow & \mathbf{X}_2^T & \longrightarrow \\ \vdots & \vdots & \vdots \\ \longleftarrow & \mathbf{X}_n^T & \longrightarrow \end{pmatrix}.
$$

Recall that the empirical covariance matrix S of this data set can be expressed as

$$
S = \frac{1}{n}\mathbb{X}^T H \mathbb{X}
$$

where

$$H = I_n - \frac{1}{n}\mathbf{1}\mathbf{1}^T.$$

The matrix $H \in \mathbb{R}^n$ is an **orthogonal projection** .

In general, we say that a matrix M is a **orthogonal projection onto a subspace** S if

0. M is symmetric,
0. $M^2 = M$, and
0. $S = \{\mathbf{y} : M\mathbf{x} = y \text{ for some } \mathbf{x} \in \mathbb{R}^n\}$

Which of the following are true about the matrix H? (Choose all that apply.)

☑ For any positive integer k and any vector $\mathbf{x} \in \mathbb{R}^n$, we have $H^k\mathbf{x} = H\mathbf{x}$.

☐ For any positive integer k and any vector $\mathbf{x} \in \mathbb{R}^n$, we have $H^k\mathbf{x} = \mathbf{x}$.

☑ The matrix H is a projection onto the subspace of vectors perpendicular to the vector $\mathbf{1} \in \mathbb{R}^n$, which has all of its entries equal to 1.

☑ The matrix H is a projections onto the subspace $\{\mathbf{x} : \frac{1}{n}\sum_{i=1}^n \mathbf{x}^i = 0\} \subset \mathbb{R}^n$. (In other words, this is the set of vectors having coordinate-wise average equal to 0.)

✓

Solution:

We examine the choices in order.

- The first choice is correct. Since H is an orthogonal projection, we know that $H^2 = H$. It follows by iterating that for any $k \geq 2$, we also have

$$H^k = H^{k-1} = H^{k-2} = \cdots = H^2 = H.$$

- The second choice is incorrect. Note that H is **not** equal to the identity matrix, and note that the identity matrix I is the only matrix satisfying

  $$I\mathbf{x} = \mathbf{x}$$

  for **all** $\mathbf{x} \in \mathbb{R}^n$. Hence, the given statement must be false when $k = 1$, for example.

  Remark: One could also use the example from the question "An Orthogonal Matrix Projection I" on this page to see that $H\mathbf{x} \neq \mathbf{x}$ when $n = 3$ and $\mathbf{x} = (2, -1, -2)^T$.

- The third choice is correct. If $\mathbf{x} \perp \mathbf{1}$, then we have

  $$H\mathbf{x} = \mathbf{x} - \frac{1}{n}\mathbf{1}\,(\mathbf{1} \cdot \mathbf{x}) = \mathbf{x}.$$

  Moreover, $H\mathbf{x} \perp \mathbf{1}$ because

  $$\begin{aligned} H\mathbf{x} \cdot \mathbf{1} &= (\mathbf{x} - \frac{1}{n}\mathbf{1}\,(\mathbf{1} \cdot \mathbf{x})) \cdot \mathbf{1} \\ &= \mathbf{x} \cdot \mathbf{1} - \frac{1}{n}(\mathbf{1} \cdot \mathbf{1})\,(\mathbf{1} \cdot \mathbf{x}) \\ &= 0. \end{aligned}$$

  These two facts imply that the outputs of H consist of all vectors that are perpendicular to $\mathbf{1}$.

- The fourth choice is correct. This follows from the explanation of the third choice because

$$\mathbf{x} \perp \mathbf{1} \Leftrightarrow \frac{1}{n}\sum_{i=1}^{n} \mathbf{x}^i = 0.$$

The above equivalence is true because $\mathbf{x} \cdot \mathbf{1} = \sum_{i=1}^{n} \mathbf{x}^i$.

Submit

You have used 1 of 3 attempts

Answers are displayed within the problem

Discussion

Hide Discussion

Topic: (Optional) Unit 8 Principal component analysis:(Optional) Preparation Exercises for Principal Component Analysis / 5. Empirical Mean and Covariance Matrix of a Vector Data Set II

Add a Post

Show all posts ▼

by recent activity ▼

There are no posts in this topic yet.

✖