2. A Simple Singular Covariance Matrix

Suppose $\mathbf{X}$ is a random vector, where $\mathbf{X} = (X^{(1)}, \ldots, X^{(d)})^T$, with mean $\mathbf{0}$ and covariance matrix $\mathbf{v}\mathbf{v}^T$, for some vector $\mathbf{v} \in \mathbb{R}^d$.

(a)

0/1 point (graded)

If $d > 1$, is the covariance matrix $\mathbf{v}\mathbf{v}^T$ invertible?

Hint: Compute the determinant for the case $d = 2$. That result will generalize to higher dimension.

- (selected) $\mathbf{v}\mathbf{v}^T$ is invertible.
- $\mathbf{v}\mathbf{v}^T$ is **not** invertible.

Solution:

For $d > 1$, the matrix $\mathbf{v}\mathbf{v}^T$ where $\mathbf{v}$ is a vector in $\mathbb{R}^d$ is not invertible. To get an intuition, we start with an example in 2 dimensions:

$$\mathbf{v} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \implies \mathbf{v}\mathbf{v}^T = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

The matrix $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ is not invertible. One way to see this is that its determinant is $1(0) - (0)(0) = 0$. Another way to see it is that for any 2×2 matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$:

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a & b \\ 0 & 0 \end{pmatrix} \neq \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

In fact, the above argument works in general after a change of variables. Given $\mathbf{v} \in \mathbb{R}^d$, change coordinates of $\mathbb{R}^d$ so that the first axis points in the direction of $\mathbf{v}$ (and so that $\mathbf{v}$ has unit length). In this new coordinate system, $\mathbf{v}$ can be rewritten as $\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$, and the matrix

$$\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}\begin{pmatrix} 1 & 0 & \ldots & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & \ldots & 0 \\ 0 & & & \\ \vdots & & \ddots & \vdots \\ 0 & & \ldots & 0 \end{pmatrix}$$

is not invertible because no $d \times d$ matrix when multiplied by it will give the identity matrix.

提交 你已经尝试了1次（总共可以尝试1次）

Answers are displayed within the problem

(b)

1.0/1 point (graded)

Let $\mathbf{u}$ be a vector in $\mathbb{R}^d$ such that $\mathbf{u} \perp \mathbf{v}$, i.e. $\mathbf{u} \cdot \mathbf{v} = \mathbf{u}^T \mathbf{v} = \mathbf{v}^T \mathbf{u} = 0$.

Find the variance of $\mathbf{u}^T \mathbf{X}$.

(If applicable, enter **trans(v)** for the transpose $\mathbf{v}^T$ of a vector $\mathbf{v}$, and **norm(v)** for the norm $||\mathbf{v}||$ of a vector $\mathbf{v}$.)

$\text{Var}\left(\mathbf{u}^T \mathbf{X}\right) =$ 0 Answer: 0

STANDARD NOTATION

Solution:

Given two vectors $\mathbf{u}, \mathbf{X} \in \mathbb{R}^d$, the inner product $\mathbf{u}^T \mathbf{X}$ is a scalar, and its variance is also a scalar. Using the covariance matrix formula, we get

$$\begin{aligned} \text{Var}\left(\mathbf{u}^T \mathbf{X}\right) = \text{Cov}\left(\mathbf{u}^T \mathbf{X}\right) &= \mathbf{u}^T \text{Cov}\left(\mathbf{X}\right) \mathbf{u} \\ &= \mathbf{u}^T \left(\mathbf{v}\mathbf{v}^T\right) \mathbf{u} \\ &= \left(\mathbf{u}^T \mathbf{v}\right)\left(\mathbf{v}^T \mathbf{u}\right) = 0. \end{aligned}$$

提交 你已经尝试了1次（总共可以尝试3次）

Answers are displayed within the problem

(c)

1.0/1 point (graded)

Let $\overline{\mathbf{v}} = \frac{\mathbf{v}}{\|\mathbf{v}\|}$ (i.e., $\overline{\mathbf{v}}$ is the normalized version of $\mathbf{v}$). What is the variance of $\overline{\mathbf{v}}^T \mathbf{X}$?

(If applicable, enter **trans(v)** for the transpose $\mathbf{v}^T$ of $\mathbf{v}$, and **norm(v)** for the norm $||\mathbf{v}||$ of a vector $\mathbf{v}$.)

$\text{Var}\left(\overline{\mathbf{v}}^T \mathbf{X}\right) =$ trans(v)*(v*(trans(v)*v))/(Answer: norm(v)^2

STANDARD NOTATION

Solution:

Similarly

$$\begin{aligned} \text{Var}\left(\overline{\mathbf{v}}^T \mathbf{X}\right) = \text{Cov}\left(\overline{\mathbf{v}}^T \mathbf{X}\right) &= \overline{\mathbf{v}}^T \text{Cov}\left(X\right) \overline{\mathbf{v}} \\ &= \left(\frac{\mathbf{v}}{\|\mathbf{v}\|}\right)^T \left(\mathbf{v}\mathbf{v}^T\right) \left(\frac{\mathbf{v}}{\|\mathbf{v}\|}\right) \\ &= \frac{\left(\mathbf{v}^T \mathbf{v}\right)\left(\mathbf{v}^T \mathbf{v}\right)}{\|\mathbf{v}\|^2} = \|\mathbf{v}\|^2. \end{aligned}$$

提交 你已经尝试了1次（总共可以尝试3次）

Answers are displayed within the problem

(d)

1/1 point (graded)

Suppose we observe n independent copies of $\mathbf{X}$ and call them $\mathbf{X}_1, \ldots, \mathbf{X}_n$. What is the asymptotic distribution of $\overline{\mathbf{X}}_n = \frac{\sum_{i=1}^n \mathbf{X}_i}{n}$? (Select all that apply.)

- [ ] $\sqrt{n}\left(\overline{\mathbf{X}}_n - \mathbf{0}\right) \xrightarrow[n\to\infty]{(d)} \mathcal{N}\left(\mathbf{0}, \mathbf{I}_d\right)$ where $\mathbf{I}_d$ is the identity matrix in R^d
- [x] $\sqrt{n}\left(\overline{\mathbf{X}}_n - \mathbf{0}\right) \xrightarrow[n\to\infty]{(d)} \mathcal{N}\left(\mathbf{0}, \mathbf{v}\mathbf{v}^T\right)$
- [ ] $\sqrt{n}\left(\overline{\mathbf{X}}_n - \mathbf{0}\right) \xrightarrow[n\to\infty]{(d)} \mathcal{N}\left(\mathbf{0}, \|\mathbf{v}\|^2\right)$

Note on notation: In the choices above, $\mathcal{N}$ denotes a multivariate Gaussian distribution. In lecture and elsewhere, a multivariate Gaussian distribution in d dimension is also sometimes denoted with an extra subscript by $\mathcal{N}_d$. To be accurate, read the dimension from the arguments, i.e. the mean and the covariance matrix.

Solution:

By multivariate CLT,

$$\sqrt{n}\left(\overline{\mathbf{X}}_n - \mathbf{0}\right) \xrightarrow[n\to\infty]{(d)} \mathcal{N}\left(\mathbf{0}, \mathbf{v}\mathbf{v}^T\right)$$

However, $\mathbf{v}\mathbf{v}^T$ is not invertible, so the pdf of $\mathcal{N}\left(\mathbf{0}, \mathbf{v}\mathbf{v}^T\right)$ is not given by the usual formula that involves the inverse of the determinant of the covariance matrix of the multivariate Gaussian variable.

提交 你已经尝试了2次（总共可以尝试2次）

Answers are displayed within the problem

(e)

2.0/2 points (graded)

Let $\mathbf{Y}_i = \overline{\mathbf{v}}\left(\overline{\mathbf{v}}^T\mathbf{X}_i\right)$, or equivalently $\overline{\mathbf{v}}\left(\overline{\mathbf{v}}\cdot\mathbf{X}_i\right) = \left(\overline{\mathbf{v}}\cdot\mathbf{X}_i\right)\overline{\mathbf{v}}$, where $\overline{\mathbf{v}} = \dfrac{\mathbf{v}}{\|\mathbf{v}\|}$ is the same as in part (c).

We will compare the asymptotic distribution of $\overline{\mathbf{X}}_n$ you obtain in part (d) to the asymptotic distribution of $\overline{\mathbf{Y}}_n$ where $\overline{\mathbf{Y}}_n = \dfrac{\sum_i^n \mathbf{Y}_i}{n}$.

What is the expectation $\mathbb{E}\left[\mathbf{Y}_i\right]$ of $\mathbf{Y}_i$?
(Choose all that apply.)

- [x] $\overline{\mathbf{v}}\overline{\mathbf{v}}^T\mathbb{E}\left[\mathbf{X}_i\right]$
- [x] $\mathbf{0}$ (the zero vector in $\mathbb{R}^d$)
- [ ] 0 (the real number zero)
- [ ] $\overline{\mathbf{v}}^T\mathbf{v}$

Find the covariance matrix $\Sigma_{\mathbf{Y}_i}$ of $\mathbf{Y}_i$ in terms of the vector $\mathbf{v}$.

(If applicable, enter **trans(v)** for the transpose $\mathbf{v}^T$ of $\mathbf{v}$, and **norm(v)** for the norm $||\mathbf{v}||$ of a vector $\mathbf{v}$.)

$\Sigma_{\mathbf{Y}_i} =$ v*(trans(v)*(v*(trans(v)*(**Answer:** v*trans(v)

可以化简

(There is no answer box for the following question.)

Notice that $\mathbf{Y}_i$ is a scalar multiple of the vector $\mathbf{v}$ and hence lies on the same line as $\mathbf{v}$ no matter what value $\mathbf{X}_i$ takes. (Specifically, $\mathbf{Y}_i = (\overline{\mathbf{v}}^T \mathbf{X}_i)\,\overline{\mathbf{v}}$ is the projection of $\mathbf{X}_i$ onto the vector $\mathbf{v}$.) Use your answers for $\mathbb{E}[\mathbf{Y}_i]$ and $\Sigma_{\mathbf{Y}_i}$ to find the asymptotic distribution of $\overline{\mathbf{Y}}_n$. Compare this with the asymptotic distribution of $\overline{\mathbf{X}}_n$ from the previous part. What can you conclude about the asymptotic distribution of $\overline{\mathbf{X}}_n$?

STANDARD NOTATION

Solution:

Since $\mathbf{Y}_i = \overline{\mathbf{v}}\,(\overline{\mathbf{v}}^T \mathbf{X}_i)$, the covariance matrix of $\mathbf{Y}_i$ is

$$\begin{aligned}
\mathsf{Cov}(\mathbf{Y}_i) &= \overline{\mathbf{v}}\overline{\mathbf{v}}^T \mathsf{Cov}(\mathbf{X}_i)\,(\overline{\mathbf{v}}\overline{\mathbf{v}}^T)^T \\
&= \overline{\mathbf{v}}\overline{\mathbf{v}}^T \mathbf{v}\mathbf{v}^T (\overline{\mathbf{v}}\overline{\mathbf{v}}^T)^T \\
&= \frac{\mathbf{v}\boxed{(\mathbf{v}^T\mathbf{v}\mathbf{v}^T\mathbf{v})}\mathbf{v}^T}{\boxed{\|\mathbf{v}\|^4}} = \mathbf{v}\mathbf{v}^T.
\end{aligned}$$

This implies

$$\sqrt{n}\,(\overline{\mathbf{Y}}_n - \mathbf{0}) = \sqrt{n}\,(\overline{\mathbf{Y}}_n) \xrightarrow[n\to\infty]{(d)} \mathcal{N}(\mathbf{0}, \mathsf{Cov}(\mathbf{Y}_i)) = \mathcal{N}(\mathbf{0}, \mathbf{v}\mathbf{v}^T).$$

Observe that $\mathcal{N}(\mathbf{0}, \mathbf{v}\mathbf{v}^T)$ is also the asymptotic distribution of $\overline{\mathbf{X}}_n$. Since $\overline{\mathbf{Y}}_n$ is a random vector that lies along the line spanned by a single vector $\mathbf{v}$ for all n, the support of its asymptotic distribution, $\mathcal{N}(\mathbf{0}, \mathbf{v}, \mathbf{v}^T)$ also lies within the same line. Hence, geometrically, $\overline{\mathbf{X}}_n$, where $\mathbf{X}_i$ has covariance matrix $\mathbf{v}\mathbf{v}^T$, approaches a one-dimensional Gaussian random variable along the line spanned by $\mathbf{v}$ as $n \to \infty$.

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:Homework 5: Maximum Likelihood Estimation / 2. A Simple Singular Covariance Matrix