Results for normal distribution

Fanny Øverbø Næss

November 13, 2022

1 The bivariate normal distribution

We assume $\mathbf{X}$ is a bivariate normal random variable with

$$\mu = E\mathbf{X} = \begin{pmatrix} 0 \\ 2 \end{pmatrix} \qquad \text{and} \qquad \Sigma = \text{Cov}\mathbf{X} = \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix} \tag{1}$$

We let

$$\mathbf{Y} = \begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} = \begin{pmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix} \mathbf{X} \tag{2}$$

1.1 Mean and covariance matrix

We find the mean and covariance matrix of $\mathbf{Y}$. We denote

$$A = \begin{pmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix}$$

$$E(A\mathbf{X}) = AE\mathbf{X} = \begin{pmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix} \begin{pmatrix} 0 \\ 2 \end{pmatrix} = \begin{pmatrix} -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{pmatrix}$$

$$\text{Cov}(A\mathbf{X}) = A(\text{Cov}\mathbf{X})A^T = \begin{pmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix} \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 4 \end{pmatrix}$$

$$\boldsymbol{\mu}_Y = \begin{pmatrix} -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{pmatrix} \qquad \text{and} \qquad \Sigma_Y = \text{Cov}\mathbf{Y} = \begin{pmatrix} 2 & 0 \\ 0 & 4 \end{pmatrix}$$

We know that any linear combination of a the bivariate normal random variable $\mathbf{X}$ is also a bivariate random normal variable. $\mathbf{Y}$ is a linear combination of $\mathbf{X}$, hence $\mathbf{Y}$ is a bivariate normal with distribution

$$Y \sim N(\boldsymbol{\mu}_Y, \Sigma_Y)$$

For two normal random variables we have that zero covariance implies independence. Given that the covariance between Y_1 and Y_2

$$\sigma_{Y_1 Y_2} = \sigma_{Y_2 Y_1} = 0$$

$\Rightarrow$ Y_1 and Y_2 are independent random variables.

1.2 Geometry of the Gaussian

- We understand the geometry of the Gaussian through the eigendecomposition of the covariance matrix Σ. The parameters of the ellipsoid are determined by the mean vector and covariance matrix of $\mathbf{X}$

  According to Theorem 2.7 in HS the half lengths of the ellipse are $\sqrt{d^2\lambda_i}$, where the λ_i are the eigenvalues of the covariance matrix Σ. From the calculations in R we see that the eigenvalues

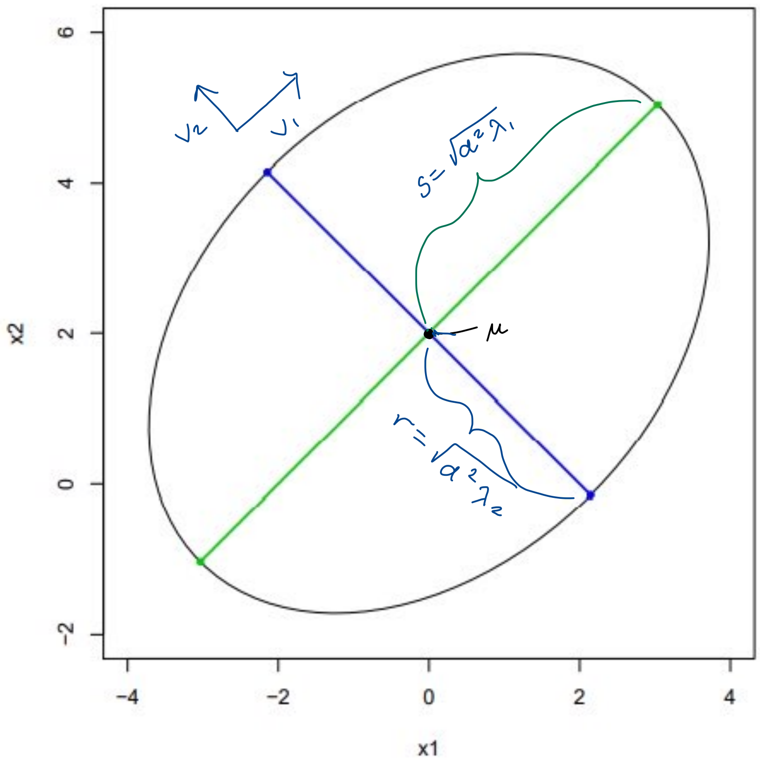

Figure 1: The contour ellipse with marked half lengths s and r

of Σ are $\lambda_1 = 4$ and $\lambda_2 = 2$ and from the definition of the ellipse we have $d^2 = b = 4.6$. The marked half lengths s and r are

$$s = 2\sqrt{4.6} \approx 4.29 \qquad \text{and} \qquad r = \sqrt{9.2} \approx 3.03$$

Each of the eigenvalues $\lambda_1 = 4$ and $\lambda_2 = 2$ have corresponding eigenvectors which are marked on the figure. The eigenvectors of Σ, which are shown in the R calculations in the problem description to be normalized eigenvectors gives us the direction of the principal axes.

- The mean vector μ gives us the point where the principal axes of the ellipse meet, which is marked on the figure in point (0,2)
- The probability that X falls within the ellipse. Theorem 4.7 from HS states that

$$\text{If} \qquad x \sim \mathcal{N}_p(\mu, \Sigma)$$

then the variable

$$U = (x-\mu)^T \Sigma^- 1(x-\mu) \sim \mathcal{X}_p^2$$

We see from the calculations in R that $P(\mathcal{X}_2^2 \leq 4.6) \approx 0.9$. This means that by choosing $b = 4.6$ then approximately 90% of the data set will be contained within the ellipse.

2 Distributional results for $\overline{X}$ and S^2 from a univariate normal sample

2.1 Showing that $\overline{X} = \frac{1}{n}\mathbf{1}^T x$

$$\frac{1}{n}\mathbf{1}^T x = \frac{1}{n}[11\ldots1]\begin{bmatrix} x_1 \\ x_2 \\ \ldots \\ x_n \end{bmatrix} = \frac{1}{n}\sum_{i=1}^{n} x_i = \overline{X}$$

Now we show that $S^2 = \frac{1}{n-1}\mathbf{X}^T C\mathbf{X}$.

$$C = I - \frac{1}{n}\mathbf{1}\mathbf{1}^T$$

Due to that C is idempotent ($C = C^2$) we have that

$$\frac{1}{n-1}\mathbf{X}^T C\mathbf{X} = \frac{1}{n-1}\mathbf{X}^T CC\mathbf{X}$$

And since C is symmetric we have that $\mathbf{X}^T C = (C\mathbf{X})^T$, so

$$\frac{1}{n-1}\mathbf{X}^T CC\mathbf{X} = \frac{1}{n-1}(C\mathbf{X})^T C\mathbf{X}$$

We expand $C\mathbf{X}$ to

$$C\mathbf{X} = (I - \frac{1}{n}\mathbf{1}\mathbf{1}^T)\mathbf{X} = \begin{pmatrix} x_1 - \overline{x} \\ \ldots \\ x_n - \overline{x} \end{pmatrix}$$

So that we can see more clearly that

$$\frac{1}{n-1}(C\mathbf{X})^T C\mathbf{X} = \frac{1}{n-1}\begin{pmatrix} x_1 - \overline{x} & \ldots & x_n - \overline{x} \end{pmatrix}\begin{pmatrix} x_1 - \overline{x} \\ \ldots \\ x_n - \overline{x} \end{pmatrix} = \frac{1}{n-1}\sum_{i=1}^{n}(x_n - \overline{x})^2 = S^2$$

2.2 Independence of $\overline{X}$ and S^2

First we show that $\frac{1}{n}\mathbf{1}^T C = \mathbf{0}^T$. We note that $\mathbf{1}^T\mathbf{1} = \sum_{i=1}^{n} 1 = n$, so

$$\frac{1}{n}\mathbf{1}^T C = \frac{1}{n}(\mathbf{1}^T - \frac{1}{n}\mathbf{1}^T\mathbf{1}\mathbf{1}^T) = \frac{1}{n}(\mathbf{1}^T - \mathbf{1}^T) = \mathbf{0}^T$$

This means that

$$\frac{1}{n}\mathbf{1}^T\sigma^2 IC^T = \frac{\sigma^2}{n}\mathbf{1}^T C = \mathbf{0}^T$$

hence, since $X \sim (\mu, \sigma^2 I)$, $\frac{1}{n}\mathbf{1}^T\mathbf{X}$ and $C\mathbf{X}$ must be independent of each other. Since $\frac{1}{n}\mathbf{1}^T x = \overline{X}$ and $\frac{1}{n-1}(C\mathbf{X})^T C\mathbf{X} = S^2$ then $\overline{\mathbf{X}}$ and S^2 must also be independent of each other.

2.3 Showing that $\mathbf{Z}^T C\mathbf{Z} \sim \mathbf{X}^2_{n-1}$

We have that $\mathbf{X} \sim \mathcal{N}(\mu\mathbf{1}, \sigma^2 I)$, then a transformation $\mathbf{X} = \sigma\mathbf{Y}$ would give that $\mathbf{Y}$ has the distribution $\mathbf{X} \sim \mathcal{N}(\mu\mathbf{1}, I)$, so we can write the following using the definition of S^2 from task 2b).

$$(n-1)\frac{S^2}{\sigma^2} = \frac{n-1}{n-1}\frac{\mathbf{X}^T C\mathbf{X}}{\sigma^2} = \mathbf{Y}^T C\mathbf{Y}$$

Now we let $\mathbf{Y} = \mathbf{Z} - \mu\mathbf{1}$, then we have the distribution $\mathbf{Z} \sim \mathcal{N}(0, I)$ and

$$\mathbf{Y}^T C\mathbf{Y} = (\mathbf{Z} - \mu\mathbf{1})^T C(\mathbf{Z} - \mu\mathbf{1})$$

We note that

$$\mu\mathbf{1}^T C = \mu\mathbf{1}^T(I - \frac{1}{n}\mathbf{1}\mathbf{1}^T) = (\mu\mathbf{1}^T - \frac{1}{n}\mu\mathbf{1}^T\mathbf{1}) = (\mu\mathbf{1}^T - \mu\mathbf{1}^T) = \mathbf{0}^T$$

Hence

$$(\mathbf{Z} - \mu\mathbf{1})^T C(\mathbf{Z} - \mu\mathbf{1}) = \mathbf{Z}^T C\mathbf{Z} \sim \mathbf{X}_r^2$$

where r denotes the range of C. Since C is idempodent,

$$\text{rank}(C) = tr(C) = \sum_{i=1}^{n} 1 - \frac{1}{n} = n - 1$$

so

$$\mathbf{Z}^T C\mathbf{Z} \sim \mathbf{X}_{n-1}^2$$

which is what we wanted to show.