Covariance

Dávid Iván

December 4, 2021

Contents

1 Cholesky decomposition of the covariance **2**
1.1 Numerical example 2

2 Standardization **3**

3 Mahalanobis distance and Chi square **4**

4 How to visualize Covariance **5**
4.1 Points c^2 from the mean 5
4.2 How to find c^2? 5
4.3 Example 6

1 Cholesky decomposition of the covariance

Since the covariance matrix $P \in \mathbb{R}^{n\times n}$ is symmetric positive-definite, we can apply the Cholesky decomposition:

$$P = LL^T \tag{1}$$

where $L \in \mathbb{R}^{n\times n}$ is a lower triangular matrix. Let's apply SVD to it.

$$L = USV^T = \sum_{i=1}^{n} \sigma_i u_i v_i^T \tag{2}$$

$\sigma_i > 0$ $(i = 1, 2, \dots n)$ are the singular values of L, $u_i \in \mathbb{R}^n$ $(i = 1, 2, \dots n)$ are the left singular vectors, and $v_i \in \mathbb{R}^n$ $(i = 1, 2, \dots n)$ are the right singular vectors of L.

With this we can express both L^T and L^{-1}:

$$L^T = VS^TU^T = VSU^T = \sum_{i=1}^{n} \sigma_i v_i u_i^T \tag{3}$$

$$L^{-1} = VS^{-1}U^T = \sum_{i=1}^{n} \frac{1}{\sigma_i} v_i u_i^T \tag{4}$$

Let's express P:

$$P = LL^T = (USV^T)(VSU^T) = US^2U^T = \sum_{i=1}^{n} \sigma_i^2 u_i u_i^T \tag{5}$$

This is the eigendecomposition of P, so its eigenvalues are σ_i^2, its eigenvectors are u_i.

1.1 Numerical example

Consider the folloing covariance matrix:

$$P = \begin{bmatrix} 2 & 0.5 \\ 0.5 & 1 \end{bmatrix} \tag{6}$$

The Cholesky decomposition is:

$$P = LL^T \tag{7}$$

where

$$L = \begin{bmatrix} 1.4142 & 0 \\ 0.3536 & 0.9354 \end{bmatrix} \tag{8}$$

The SVD of L:

$$\begin{aligned} L &= \sigma_1 u_1 v_1^T + \sigma_2 u_2 v_2^T \\ &= 1.4856 \cdot \begin{bmatrix} -0.9239 \\ -0.3827 \end{bmatrix} \cdot \begin{bmatrix} -0.9705 & -0.2410 \end{bmatrix} \\ &+ 0.8904 \cdot \begin{bmatrix} -0.3827 \\ 0.9239 \end{bmatrix} \cdot \begin{bmatrix} -0.2410 & 0.9705 \end{bmatrix} \end{aligned} \tag{9}$$

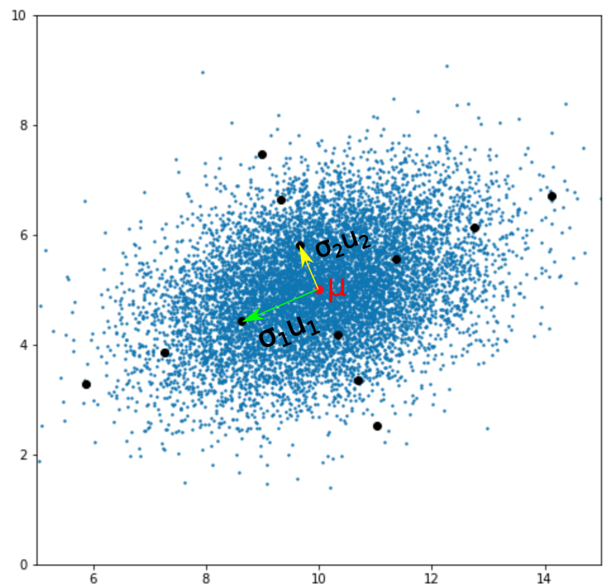

Figure 1: Illustration of the eigenvectors and singular values.

2 Standardization

How to standardize a multivariate normal distribution? $X \sim N(\mu, P) \in \mathbb{R}^n$ is normally distributed with mean μ, covariance P.

$$\boxed{Z = L^{-1}(X - \mu)} \tag{10}$$

Where L comes from equation (1). It is clear that Z is normally distributed, since it is a linear transformation of the normal X. The mean:

$$E(Z) = L^{-1}(E(X) - \mu) = L^{-1}(\mu - \mu) = 0 \tag{11}$$

The covariance:

$$\begin{aligned} P_Z &= Var(Z) = Var(L^{-1}(X - \mu)) \\ &= Var(L^{-1}X) = L^{-1}Var(X)L^{-T} \\ &= L^{-1}LL^TL^{-T} = I \end{aligned} \tag{12}$$

So we see that $Z \sim N(0, I)$, i.e., it is standard normally distributed.

3 Mahalanobis distance and Chi square

The Mahalanobis (squared) distance is defined as

$$d_M^2(x; \mu, P) = (x - \mu)^T P^{-1}(x - \mu) \tag{13}$$

where $\mu, x \in \mathbb{R}^n$, $P \in \mathbb{R}^{n \times n}$ is a symmetric positive-definite matrix (the covariance).

Assuming that X is normally distributed with μ mean, and P (nonsingular) covariance, how is $d^2(X; \mu, P)$ distributed?

We know that $Z^TZ \sim \chi_n^2$, where Z comes from the standardization of X, equation (10).

$$\begin{aligned} Z^TZ &= (X - \mu)^T L^{-T} L^{-1}(X - \mu) \\ &= (X - \mu)^T (LL^T)^{-1}(X - \mu) \\ &= (X - \mu)^T P^{-1}(X - \mu) \\ &= d_M^2(X; \mu, P) \\ &\sim \chi_n^2 \end{aligned} \tag{14}$$

So the Mahalanobis squared distance $d_M^2(X; \mu, P)$ is Chi-squared distributed with n degrees of freedom.

4 How to visualize Covariance

4.1 Points c^2 from the mean

How to find those points that are in a specific Mahalanobis distance from the mean (μ)?

Construct the following vector ($\alpha_1, \alpha_2 \in \mathbb{R}$):

$$x = \mu + \sum_{i=1}^{n} \alpha_i \sigma_i u_i \tag{15}$$

The Mahalanobis squared distance is:

$$\begin{aligned} d_M^2(x; \mu, P) &= (x - \mu)^T P^{-1} (x - \mu) \\ &= (x - \mu)^T L^{-T} L^{-1} (x - \mu) \\ &= \left\| L^{-1}(x - \mu) \right\|^2 \\ &= \left\| \sum_{i=1}^{n} \alpha_i \sigma_i L^{-1} u_i \right\|^2 \\ &= \left\| \sum_{i=1}^{n} \alpha_i \sigma_i \frac{1}{\sigma_i} v_i \right\|^2 \\ &= \left\| \sum_{i=1}^{n} \alpha_i v_i \right\|^2 \\ &= \sum_{i=1}^{n} \alpha_i^2 \end{aligned} \tag{16}$$

We used equation (4) to compute $L^{-1}u_i$. So if we seek vectors x for which the Mahalanobis squared distance is c^2, then we can construct it by picking a point on the n-dimensional sphere with radius c, i.e.,

$$\sum_{i=1}^{n} \alpha_i^2 = c^2 \tag{17}$$

Once we have a specific α, we can construct x. For this $x(\alpha)$, $d_M^2(x(\alpha); \mu, P) = c^2$

4.2 How to find c^2?

Say I want to find a c^2 for which it is 95% probable that $d_M^2(X; \mu, P) < c^2$, where $X \sim N(\mu, P)$. This is easy, since we know that $d_M^2(X; \mu, P) \sim \chi_n^2$.

We need the quantile function (or percent point function) of the Chi-squared distribution.

$$c^2 = Q_{\chi^2_n}(0.95) \tag{18}$$

In 2 dimensions, $Q_{\chi^2_n}(0.95) = 5.99$.

4.3 Example

In 2 dimensions, we have α_1 and α_2. With $t \in (0, 2\pi)$

$$\alpha_1 = c \cdot \sin(t) \tag{19}$$

$$\alpha_2 = c \cdot \cos(t) \tag{20}$$

The following figure shows the ellipse for this scenario.

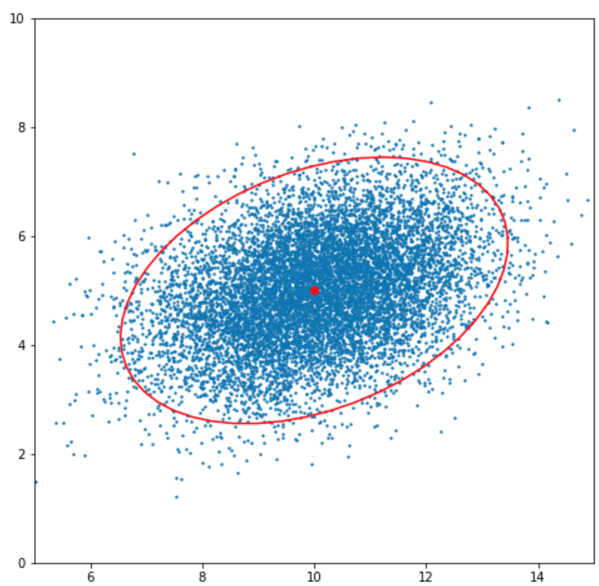

Figure 2: ellipse of the 95% confidence region.