Inferring The Mean of A Multivariate Normal Distribution When Its Covariance Is Known

Saffet Gökçen Şen

May 27, 2020

1 Introduction

It is assumed that there is a set of vector data. These vectors are assumed to be from a multivariate Gaussian distribution:

$$\mathbf{x}_i \sim N\left(\boldsymbol{\mu}, \boldsymbol{\Sigma}\right) \;\; (i = 1, 2, 3, \ldots, N) \tag{1}$$

It is assumed that $\boldsymbol{\mu}$ is unknown while $\boldsymbol{\Sigma}$ is known. First, the likelihood of the data given the mean and the covariance will be obtained. Then, the prior for the mean will be set. Lastly, the posterior distribution of $\boldsymbol{\mu}$ is to be obtained.

2 The Likelihood

The likelihood is the probability of the occurrence of the data at hand. In the calculation of the likelihood, data samples are supposed to be independent. The likelihood is denoted by $p\left(\mathbf{X}|\boldsymbol{\mu}\right)$ and has the following mathematical expression:

$$p\left(\mathbf{X}|\boldsymbol{\mu}\right) = p\left(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \ldots, \mathbf{x}_N|\boldsymbol{\mu}\right) = \prod_{i=1}^{N} p\left(\mathbf{x}_i|\boldsymbol{\mu}\right) = \tag{2}$$

$$\prod_{i=1}^{N} \frac{1}{\left(2\pi\right)^{D/2} \left|\boldsymbol{\Sigma}\right|^{1/2}} \exp\left[-\frac{1}{2}\left(\mathbf{x}_i - \boldsymbol{\mu}\right)^T \boldsymbol{\Sigma}^{-1}\left(\mathbf{x}_i - \boldsymbol{\mu}\right)\right] \Rightarrow \tag{3}$$

$$p\left(\mathbf{X}|\boldsymbol{\mu}\right) = \frac{1}{\left(2\pi\right)^{ND/2} \left|\boldsymbol{\Sigma}\right|^{N/2}} \exp\left[\sum_{i=1}^{N} -\frac{1}{2}\left(\mathbf{x}_i - \boldsymbol{\mu}\right)^T \boldsymbol{\Sigma}^{-1}\left(\mathbf{x}_i - \boldsymbol{\mu}\right)\right] \tag{4}$$

The exponent of a multivariate Gaussian distribution is given as follows:

$$-\frac{1}{2}\left(\mathbf{x} - \boldsymbol{\mu}\right)^T \boldsymbol{\Sigma}^{-1}\left(\mathbf{x} - \boldsymbol{\mu}\right) = -\frac{1}{2}\mathbf{x}^T\boldsymbol{\Sigma}^{-1}\mathbf{x} + \mathbf{x}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu} - \frac{1}{2}\boldsymbol{\mu}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu} \Rightarrow \tag{5}$$

$$-\frac{1}{2}\left(\mathbf{x} - \boldsymbol{\mu}\right)^T \boldsymbol{\Sigma}^{-1}\left(\mathbf{x} - \boldsymbol{\mu}\right) = -\frac{1}{2}\mathbf{x}^T\boldsymbol{\Sigma}^{-1}\mathbf{x} + \mathbf{x}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu} + \text{constant} \tag{6}$$

The exponent of the likelihood is to be put into the form derived in the above. This is the usual operation in multivariate Gaussian distributions related operations. In the likelihood, N Gaussians are multiplied. The product of two Gaussians with different means and covariances will be used to derive the expression. The exponent of the product is

$$-\frac{1}{2}\left(\mathbf{x}-\boldsymbol{\mu}_1\right)^T\boldsymbol{\Sigma}_1^{-1}\left(\mathbf{x}-\boldsymbol{\mu}_1\right)-\frac{1}{2}\left(\mathbf{x}-\boldsymbol{\mu}_2\right)^T\boldsymbol{\Sigma}_2^{-1}\left(\mathbf{x}-\boldsymbol{\mu}_2\right). \tag{7}$$

This expression is to be equated to

$$-\frac{1}{2}\mathbf{x}^T\boldsymbol{\Sigma}_c^{-1}\mathbf{x}+\mathbf{x}^T\boldsymbol{\Sigma}_c^{-1}\boldsymbol{\mu}_c+\mathrm{c}. \tag{8}$$

where $\mathrm{c}=-\frac{1}{2}\boldsymbol{\mu}_c^T\boldsymbol{\Sigma}_c^{-1}\boldsymbol{\mu}_c$.

Let the exponent of the product of the two Gaussians be expanded:

$$-\frac{1}{2}\left(\mathbf{x}-\boldsymbol{\mu}_1\right)^T\boldsymbol{\Sigma}_1^{-1}\left(\mathbf{x}-\boldsymbol{\mu}_1\right)-\frac{1}{2}\left(\mathbf{x}-\boldsymbol{\mu}_2\right)^T\boldsymbol{\Sigma}_2^{-1}\left(\mathbf{x}-\boldsymbol{\mu}_2\right)= \tag{9}$$

$$-\frac{1}{2}\mathbf{x}^T\boldsymbol{\Sigma}_1^{-1}\mathbf{x}+\mathbf{x}^T\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\mu}_1+\mathrm{c}_1-\frac{1}{2}\mathbf{x}^T\boldsymbol{\Sigma}_2^{-1}\mathbf{x}+\mathbf{x}^T\boldsymbol{\Sigma}_2^{-1}\boldsymbol{\mu}_2+\mathrm{c}_2= \tag{10}$$

$$-\frac{1}{2}\mathbf{x}^T\left(\boldsymbol{\Sigma}_1^{-1}+\boldsymbol{\Sigma}_2^{-1}\right)\mathbf{x}+\mathbf{x}^T\left(\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\mu}_1+\boldsymbol{\Sigma}_2^{-1}\boldsymbol{\mu}_2\right)+\mathrm{c} \tag{11}$$

Let this expression be equated to the standard form given in equation (8):

$$-\frac{1}{2}\mathbf{x}^T\boldsymbol{\Sigma}_c^{-1}\mathbf{x}+\mathbf{x}^T\boldsymbol{\Sigma}_c^{-1}\boldsymbol{\mu}_c+\mathrm{c}=$$

$$-\frac{1}{2}\mathbf{x}^T\left(\boldsymbol{\Sigma}_1^{-1}+\boldsymbol{\Sigma}_2^{-1}\right)\mathbf{x}+\mathbf{x}^T\left(\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\mu}_1+\boldsymbol{\Sigma}_2^{-1}\boldsymbol{\mu}_2\right)+\mathrm{c}_1+\mathrm{c}_2\Rightarrow \tag{12}$$

$$\boldsymbol{\Sigma}_c^{-1}=\boldsymbol{\Sigma}_1^{-1}+\boldsymbol{\Sigma}_2^{-1} \tag{13}$$

$$\boldsymbol{\Sigma}_c^{-1}\boldsymbol{\mu}_c=\left(\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\mu}_1+\boldsymbol{\Sigma}_2^{-1}\boldsymbol{\mu}_2\right)\Rightarrow \tag{14}$$

$$\boldsymbol{\mu}_c = \left(\boldsymbol{\Sigma}_1^{-1} + \boldsymbol{\Sigma}_2^{-1}\right)^{-1} \left(\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\mu}_1 + \boldsymbol{\Sigma}_2^{-1}\boldsymbol{\mu}_2\right) \tag{15}$$

$$\mathrm{c} = -\frac{1}{2}\boldsymbol{\mu}_{c_1}^T \boldsymbol{\Sigma}_{c_1}^{-1}\boldsymbol{\mu}_{c_1} - \frac{1}{2}\boldsymbol{\mu}_{c_2}^T \boldsymbol{\Sigma}_{c_2}^{-1}\boldsymbol{\mu}_{c_2} \tag{16}$$

Hence, $\boldsymbol{\Sigma}_c^{-1}$ for the likelihood is $N\boldsymbol{\Sigma}^{-1}$. Then, $\boldsymbol{\Sigma}_c$ for the likelihood is $\frac{1}{N}\boldsymbol{\Sigma}$. $\boldsymbol{\mu}_c$ for the likelihood is $\left(N\boldsymbol{\Sigma}^{-1}\right)^{-1} N\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}$. Then, $\boldsymbol{\mu}_c$ for the likelihood is $\boldsymbol{\mu}$. The constant c for the likelihood is $-\frac{N}{2}\boldsymbol{\mu}_c^T\boldsymbol{\Sigma}_c^{-1}\boldsymbol{\mu}_c$.

The mathematical expression for the likelihood can now be written as follows:

$$p\left(\mathbf{X}|\boldsymbol{\mu}\right) = \frac{\exp\left(-\frac{N}{2}\boldsymbol{\mu}_c^T\boldsymbol{\Sigma}_c^{-1}\boldsymbol{\mu}_c\right)}{\left(2\pi\right)^{ND/2}\left|\boldsymbol{\Sigma}\right|^{N/2}} \exp\left(-\frac{1}{2}\mathbf{x}^T\boldsymbol{\Sigma}_c^{-1}\mathbf{x} + \mathbf{x}^T\boldsymbol{\Sigma}_c^{-1}\boldsymbol{\mu}_c\right) \tag{17}$$

The likelihood is not a multivariate Gaussian distribution.

3 Prior

4 Reference

Pattern Recognition and Machine Learning, Christopher M. Bishop.

Machine Learning A Probabilistic Perspective, Kevin P. Murphy.