

and of the covariance matrix Σ given by

$$\Sigma = \begin{pmatrix} \Sigma_{aa} & \Sigma_{ab} \\ \Sigma_{ba} & \Sigma_{bb} \end{pmatrix}. \quad (3.51)$$

Note that the symmetry $\Sigma^T = \Sigma$ of the covariance matrix implies that Σ_{aa} and Σ_{bb} are symmetric and that $\Sigma_{ba} = \Sigma_{ab}^T$.

In many situations, it will be convenient to work with the inverse of the covariance matrix:

$$\Lambda \equiv \Sigma^{-1}, \quad (3.52)$$

which is known as the *precision matrix*. In fact, we will see that some properties of Gaussian distributions are most naturally expressed in terms of the covariance, whereas others take a simpler form when viewed in terms of the precision. We therefore also introduce the partitioned form of the precision matrix:

$$\Lambda = \begin{pmatrix} \Lambda_{aa} & \Lambda_{ab} \\ \Lambda_{ba} & \Lambda_{bb} \end{pmatrix} \quad (3.53)$$

corresponding to the partitioning (3.49) of the vector \mathbf{x} . Because the inverse of a symmetric matrix is also symmetric, we see that Λ_{aa} and Λ_{bb} are symmetric and that $\Lambda_{ba} = \Lambda_{ab}^T$. It should be stressed at this point that, for instance, Λ_{aa} is not simply given by the inverse of Σ_{aa} . In fact, we will shortly examine the relation between the inverse of a partitioned matrix and the inverses of its partitions.

Exercise 3.16

We begin by finding an expression for the conditional distribution $p(\mathbf{x}_a|\mathbf{x}_b)$. From the product rule of probability, we see that this conditional distribution can be evaluated from the joint distribution $p(\mathbf{x}) = p(\mathbf{x}_a, \mathbf{x}_b)$ simply by fixing \mathbf{x}_b to the observed value and normalizing the resulting expression to obtain a valid probability distribution over \mathbf{x}_a . Instead of performing this normalization explicitly, we can obtain the solution more efficiently by considering the quadratic form in the exponent of the Gaussian distribution given by (3.27) and then reinstating the normalization coefficient at the end of the calculation. If we make use of the partitioning (3.49), (3.50), and (3.53), we obtain

$$\begin{aligned} -\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})^T \Sigma^{-1}(\mathbf{x} - \boldsymbol{\mu}) = & \\ & -\frac{1}{2}(\mathbf{x}_a - \boldsymbol{\mu}_a)^T \Lambda_{aa}(\mathbf{x}_a - \boldsymbol{\mu}_a) - \frac{1}{2}(\mathbf{x}_a - \boldsymbol{\mu}_a)^T \Lambda_{ab}(\mathbf{x}_b - \boldsymbol{\mu}_b) \\ & -\frac{1}{2}(\mathbf{x}_b - \boldsymbol{\mu}_b)^T \Lambda_{ba}(\mathbf{x}_a - \boldsymbol{\mu}_a) - \frac{1}{2}(\mathbf{x}_b - \boldsymbol{\mu}_b)^T \Lambda_{bb}(\mathbf{x}_b - \boldsymbol{\mu}_b). \end{aligned} \quad (3.54)$$

We see that as a function of \mathbf{x}_a , this is again a quadratic form, and hence, the corresponding conditional distribution $p(\mathbf{x}_a|\mathbf{x}_b)$ will be Gaussian. Because this distribution is completely characterized by its mean and its covariance, our goal will be to identify expressions for the mean and covariance of $p(\mathbf{x}_a|\mathbf{x}_b)$ by inspection of (3.54).

This is an example of a rather common operation associated with Gaussian distributions, sometimes called ‘completing the square’, in which we are given a