

Thus, we obtain the intuitively satisfying result that the marginal distribution  $p(\mathbf{x}_a)$  has mean and covariance given by

$$\mathbb{E}[\mathbf{x}_a] = \boldsymbol{\mu}_a \quad (3.76)$$

$$\text{cov}[\mathbf{x}_a] = \boldsymbol{\Sigma}_{aa}. \quad (3.77)$$

We see that for a marginal distribution, the mean and covariance are most simply expressed in terms of the partitioned covariance matrix, in contrast to the conditional distribution for which the partitioned precision matrix gives rise to simpler expressions.

Our results for the marginal and conditional distributions of a partitioned Gaussian can be summarized as follows. Given a joint Gaussian distribution  $\mathcal{N}(\mathbf{x}|\boldsymbol{\mu}, \boldsymbol{\Sigma})$  with  $\boldsymbol{\Lambda} \equiv \boldsymbol{\Sigma}^{-1}$  and the following partitions

$$\mathbf{x} = \begin{pmatrix} \mathbf{x}_a \\ \mathbf{x}_b \end{pmatrix}, \quad \boldsymbol{\mu} = \begin{pmatrix} \boldsymbol{\mu}_a \\ \boldsymbol{\mu}_b \end{pmatrix} \quad (3.78)$$

$$\boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_{aa} & \boldsymbol{\Sigma}_{ab} \\ \boldsymbol{\Sigma}_{ba} & \boldsymbol{\Sigma}_{bb} \end{pmatrix}, \quad \boldsymbol{\Lambda} = \begin{pmatrix} \boldsymbol{\Lambda}_{aa} & \boldsymbol{\Lambda}_{ab} \\ \boldsymbol{\Lambda}_{ba} & \boldsymbol{\Lambda}_{bb} \end{pmatrix} \quad (3.79)$$

then the conditional distribution is given by

$$p(\mathbf{x}_a|\mathbf{x}_b) = \mathcal{N}(\mathbf{x}|\boldsymbol{\mu}_{a|b}, \boldsymbol{\Lambda}_{aa}^{-1}) \quad (3.80)$$

$$\boldsymbol{\mu}_{a|b} = \boldsymbol{\mu}_a - \boldsymbol{\Lambda}_{aa}^{-1} \boldsymbol{\Lambda}_{ab}(\mathbf{x}_b - \boldsymbol{\mu}_b) \quad (3.81)$$

and the marginal distribution is given by

$$p(\mathbf{x}_a) = \mathcal{N}(\mathbf{x}_a|\boldsymbol{\mu}_a, \boldsymbol{\Sigma}_{aa}). \quad (3.82)$$

We illustrate the idea of conditional and marginal distributions associated with a multivariate Gaussian using an example involving two variables in [Figure 3.5](#).

### 3.2.6 Bayes' theorem

In Sections 3.2.4 and 3.2.5 we considered a Gaussian  $p(\mathbf{x})$  in which we partitioned the vector  $\mathbf{x}$  into two subvectors  $\mathbf{x} = (\mathbf{x}_a, \mathbf{x}_b)$  and then found expressions for the conditional distribution  $p(\mathbf{x}_a|\mathbf{x}_b)$  and the marginal distribution  $p(\mathbf{x}_a)$ . We noted that the mean of the conditional distribution  $p(\mathbf{x}_a|\mathbf{x}_b)$  was a linear function of  $\mathbf{x}_b$ . Here we will suppose that we are given a Gaussian marginal distribution  $p(\mathbf{x})$  and a Gaussian conditional distribution  $p(\mathbf{y}|\mathbf{x})$  in which  $p(\mathbf{y}|\mathbf{x})$  has a mean that is a linear function of  $\mathbf{x}$  and a covariance that is independent of  $\mathbf{x}$ . This is an example of a *linear-Gaussian model* (Roweis and Ghahramani, 1999). We wish to find the marginal distribution  $p(\mathbf{y})$  and the conditional distribution  $p(\mathbf{x}|\mathbf{y})$ . This is a structure that arises in several types of generative model and it will prove convenient to derive the general results here.

We will take the marginal and conditional distributions to be

$$p(\mathbf{x}) = \mathcal{N}(\mathbf{x}|\boldsymbol{\mu}, \boldsymbol{\Lambda}^{-1}) \quad (3.83)$$

$$p(\mathbf{y}|\mathbf{x}) = \mathcal{N}(\mathbf{y}|\mathbf{A}\mathbf{x} + \mathbf{b}, \mathbf{L}^{-1}) \quad (3.84)$$