

given by

$$\mathbf{R} = \begin{pmatrix} \mathbf{\Lambda} + \mathbf{A}^T \mathbf{L} \mathbf{A} & -\mathbf{A}^T \mathbf{L} \\ -\mathbf{L} \mathbf{A} & \mathbf{L} \end{pmatrix}. \quad (3.88)$$

The covariance matrix is found by taking the inverse of the precision, which can be done using the matrix inversion formula (3.60) to give

Exercise 3.23

$$\text{cov}[\mathbf{z}] = \mathbf{R}^{-1} = \begin{pmatrix} \mathbf{\Lambda}^{-1} & \mathbf{\Lambda}^{-1} \mathbf{A}^T \\ \mathbf{A} \mathbf{\Lambda}^{-1} & \mathbf{L}^{-1} + \mathbf{A} \mathbf{\Lambda}^{-1} \mathbf{A}^T \end{pmatrix}. \quad (3.89)$$

Similarly, we can find the mean of the Gaussian distribution over \mathbf{z} by identifying the linear terms in (3.86), which are given by

$$\mathbf{x}^T \mathbf{\Lambda} \boldsymbol{\mu} - \mathbf{x}^T \mathbf{A}^T \mathbf{L} \mathbf{b} + \mathbf{y}^T \mathbf{L} \mathbf{b} = \begin{pmatrix} \mathbf{x} \\ \mathbf{y} \end{pmatrix}^T \begin{pmatrix} \mathbf{\Lambda} \boldsymbol{\mu} - \mathbf{A}^T \mathbf{L} \mathbf{b} \\ \mathbf{L} \mathbf{b} \end{pmatrix}. \quad (3.90)$$

Using our earlier result (3.55) obtained by completing the square over the quadratic form of a multivariate Gaussian, we find that the mean of \mathbf{z} is given by

$$\mathbb{E}[\mathbf{z}] = \mathbf{R}^{-1} \begin{pmatrix} \mathbf{\Lambda} \boldsymbol{\mu} - \mathbf{A}^T \mathbf{L} \mathbf{b} \\ \mathbf{L} \mathbf{b} \end{pmatrix}. \quad (3.91)$$

Exercise 3.24

Making use of (3.89), we then obtain

$$\mathbb{E}[\mathbf{z}] = \begin{pmatrix} \boldsymbol{\mu} \\ \mathbf{A} \boldsymbol{\mu} + \mathbf{b} \end{pmatrix}. \quad (3.92)$$

Next we find an expression for the marginal distribution $p(\mathbf{y})$ in which we have marginalized over \mathbf{x} . Recall that the marginal distribution over a subset of the components of a Gaussian random vector takes a particularly simple form when expressed in terms of the partitioned covariance matrix. Specifically, its mean and covariance are given by (3.76) and (3.77), respectively. Making use of (3.89) and (3.92), we see that the mean and covariance of the marginal distribution $p(\mathbf{y})$ are given by

Section 3.2

$$\mathbb{E}[\mathbf{y}] = \mathbf{A} \boldsymbol{\mu} + \mathbf{b} \quad (3.93)$$

$$\text{cov}[\mathbf{y}] = \mathbf{L}^{-1} + \mathbf{A} \mathbf{\Lambda}^{-1} \mathbf{A}^T. \quad (3.94)$$

A special case of this result is when $\mathbf{A} = \mathbf{I}$, in which case the marginal distribution reduces to the convolution of two Gaussians, for which we see that the mean of the convolution is the sum of the means of the two Gaussians and the covariance of the convolution is the sum of their covariances.

Finally, we seek an expression for the conditional $p(\mathbf{x}|\mathbf{y})$. Recall that the results for the conditional distribution are most easily expressed in terms of the partitioned precision matrix, using (3.57) and (3.59). Applying these results to (3.89) and (3.92), we see that the conditional distribution $p(\mathbf{x}|\mathbf{y})$ has mean and covariance given by

Section 3.2

$$\mathbb{E}[\mathbf{x}|\mathbf{y}] = (\mathbf{\Lambda} + \mathbf{A}^T \mathbf{L} \mathbf{A})^{-1} \{ \mathbf{A}^T \mathbf{L} (\mathbf{y} - \mathbf{b}) + \mathbf{\Lambda} \boldsymbol{\mu} \} \quad (3.95)$$

$$\text{cov}[\mathbf{x}|\mathbf{y}] = (\mathbf{\Lambda} + \mathbf{A}^T \mathbf{L} \mathbf{A})^{-1}. \quad (3.96)$$