continued from the previous page

The remaining statement, $\Sigma_{xx}^{-1}\,\mu_x = \bar{\xi}_x$, is obtained analogously, exploiting the fact that $\mu = \Omega^{-1}\xi$ (see Equation(3.73)) and the equality of the two expressions marked "$(*)$" above:

$$\begin{aligned}
\begin{pmatrix} \Sigma_{xx}^{-1}\,\mu_x \\ 0 \end{pmatrix} &= \begin{pmatrix} \Sigma_{xx}^{-1} & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \mu_x \\ \mu_y \end{pmatrix} = \begin{pmatrix} \Sigma_{xx}^{-1} & 0 \\ 0 & 0 \end{pmatrix} \Omega^{-1} \begin{pmatrix} \xi_x \\ \xi_y \end{pmatrix} \\
&\overset{(*)}{=} \left[\Omega - \begin{pmatrix} 0 & \Omega_{xy}\,\Omega_{yy}^{-1} \\ 0 & 1 \end{pmatrix} \Omega \right] \Omega^{-1} \begin{pmatrix} \xi_x \\ \xi_y \end{pmatrix} \\
&= \begin{pmatrix} \xi_x \\ \xi_y \end{pmatrix} - \begin{pmatrix} 0 & \Omega_{xy}\,\Omega_{yy}^{-1} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \xi_x \\ \xi_y \end{pmatrix} = \begin{pmatrix} \bar{\xi}_x \\ 0 \end{pmatrix}
\end{aligned}$$

Table 11.6 Lemma for marginalizing Gaussians in information form. The form of the covariance $\bar{\Omega}_{xx}$ in this lemma is also known as *Schur complement*.

Conditionals of a multivariate Gaussian. Let the probability distribution $p(x, y)$ over the random vectors x and y be a Gaussian represented in the information form

$$\Omega = \begin{pmatrix} \Omega_{xx} & \Omega_{xy} \\ \Omega_{yx} & \Omega_{yy} \end{pmatrix} \qquad \text{and} \qquad \xi = \begin{pmatrix} \xi_x \\ \xi_y \end{pmatrix}$$

The conditional $p(x \mid y)$ is a Gaussian with information matrix Ω_{xx} and information vector $\xi_x - \Omega_{xy}\,y$.

Proof. The result follows trivially from the definition of a Gaussian in information form:

$$\begin{aligned}
&p(x \mid y) \\
&= \eta \exp\left\{ -\tfrac{1}{2} \begin{pmatrix} x \\ y \end{pmatrix}^T \begin{pmatrix} \Omega_{xx} & \Omega_{xy} \\ \Omega_{yx} & \Omega_{yy} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} x \\ y \end{pmatrix}^T \begin{pmatrix} \xi_x \\ \xi_y \end{pmatrix} \right\} \\
&= \eta \exp\left\{ -\tfrac{1}{2} x^T \Omega_{xx} x - x^T \Omega_{xy} y - \tfrac{1}{2} y^T \Omega_{yy} y + x^T \xi_x + y^T \xi_y \right\} \\
&= \eta \exp\{ -\tfrac{1}{2} x^T \Omega_{xx} x + x^T (\xi_x - \Omega_{xy} y) \underbrace{- \tfrac{1}{2} y^T \Omega_{yy} y + y^T \xi_y}_{\text{const.}} \} \\
&= \eta \exp\{ -\tfrac{1}{2} x^T \Omega_{xx} x + x^T (\xi_x - \Omega_{xy} y) \}
\end{aligned}$$

Table 11.7 Lemma for conditioning Gaussians in information form.