1(b) Conditionals of Gaussian Random Variables are Gaussian Random Variables. If (x, y) are jointly distributed as:

$$p(x,y) = \mathcal{N}\left(\begin{bmatrix} x \\ y \end{bmatrix}; \begin{bmatrix} \mu_x \\ \mu_y \end{bmatrix}, \begin{bmatrix} \Sigma_{xx} & \Sigma_{xy} \\ \Sigma_{yx} & \Sigma_{yy} \end{bmatrix}\right)$$

then the conditional $p(x|y) = \dfrac{p(x,y)}{p(y)}$

is given by:

$$p(x|y) = \mathcal{N}\left(x;\ \mu_x + \Sigma_{xy}\Sigma_{yy}^{-1}(y-\mu_y),\ \Sigma_{xx} - \Sigma_{xy}\Sigma_{yy}^{-1}\Sigma_{yx}\right)$$

Answer: without loss of generalility we consider the case of $p(\tilde{x}, \tilde{y})$ such that $\tilde{x} = x - \mu_x$
$\tilde{y} = y - \mu_y$

★ For the simplicity of notation, $\tilde{x}$ and $\tilde{y}$ are referred as x and y in the calcutions

① We start by computing two equivalent forms of the inverse using different Schur decompositions, to compute the precision matrix of the joint $p(\tilde{x}, \tilde{y})$

$$\Lambda_{xy} = \begin{pmatrix} \Sigma_{xx} & \Sigma_{xy} \\ \Sigma_{yx} & \Sigma_{yy} \end{pmatrix}^{-1}$$

For a matrix $M = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$

The first Schur inverse form is given by.

$$M_1^{-1} = \begin{pmatrix} A & B \\ C & D \end{pmatrix}^{-1} = \begin{pmatrix} A^{-1} + A^{-1}B(D - CA^{-1}B)^{-1}CA^{-1} & -A^{-1}B(D-CA^{-1}B)^{-1} \\ -(D-CA^{-1}B)^{-1}CA^{-1} & (D-CA^{-1}B)^{-1} \end{pmatrix}$$

The second Schur inverse form is given by.

$$M_2^{-1} = \begin{pmatrix} A & B \\ C & D \end{pmatrix}^{-1} = \begin{pmatrix} (A - BD^{-1}C)^{-1} & -(A-BD^{-1}C)^{-1}BD^{-1} \\ -D^{-1}C(A-BD^{-1}C)^{-1} & D^{-1} + D^{-1}C(A-BD^{-1}C)^{-1}BD^{-1} \end{pmatrix}$$

Thus the precision matrix Λ_{xy} of the joint can be written as.

$A = \Sigma_{xx} \quad B = \Sigma_{xy} \quad C = \Sigma_{yx} \quad D = \Sigma_{yy}.$

$$M_1^{-1} = \begin{pmatrix} \Sigma_{xx}^{-1} + \Sigma_{xx}^{-1}\Sigma_{xy}(\Sigma_{yy} - \Sigma_{yx}\Sigma_{xx}^{-1}\Sigma_{xy})^{-1}\Sigma_{yx}\Sigma_{xx}^{-1} & -\Sigma_{xx}^{-1}\Sigma_{xy}(\Sigma_{yy} - \Sigma_{yx}\Sigma_{xx}^{-1}\Sigma_{xy})^{-1} \\ -(\Sigma_{yy} - \Sigma_{yx}\Sigma_{xx}^{-1}\Sigma_{xy})^{-1}\Sigma_{yx}\Sigma_{xx}^{-1} & (\Sigma_{yy} - \Sigma_{yx}\Sigma_{xx}^{-1}\Sigma_{xy})^{-1} \end{pmatrix}$$

We use notation $\Lambda_{y|x} = (\Sigma_{yy} - \Sigma_{yx}\Sigma_{xx}^{-1}\Sigma_{xy})^{-1}$

and $\Lambda_{x|y} = (\Sigma_{xx} - \Sigma_{xy}\Sigma_{yy}^{-1}\Sigma_{yx})^{-1}$

as covariance matrices are symmetric

$\Sigma_{yx} = \Sigma_{xy}$ ..

$\Sigma_{yx}^{T} = \Sigma_{yx}$

We note $M_1^{-1}[1,1] = \Sigma_{xx}^{-1} + \Sigma_{xx}^{-1}\Sigma_{xy}\left(\Sigma_{yy} - \Sigma_{yx}\Sigma_{xx}^{-1}\Sigma_{xy}\right)^{-1}\Sigma_{yx}\Sigma_{xx}^{-1}$

the expression is of the form

$$(Z + UWV^{T})^{-1} = Z^{-1} - Z^{-1}U(W^{-1} + V^{T}Z^{-1}U)^{-1}V^{T}Z^{-1}$$

where $Z^{-1} = -\Sigma_{xx}^{-1}$ $\quad U = \Sigma_{xy}$ $\quad V = \Sigma_{yx}$

$W^{-1} = \Sigma_{yy}$ $\quad W = \Sigma_{yy}^{-1}$

$$-\left[(-\Sigma_{xx})^{-1} - (-\Sigma_{xx})(\Sigma_{xy})\left((\Sigma_{yy}) + (\Sigma_{yx})^{T}(-\Sigma_{xx})(\Sigma_{xy})\right)^{-1}(\Sigma_{yx})^{T}(-\Sigma_{xx})\right]$$

$$\Rightarrow -\left(-\Sigma_{xx}^{-1} + \Sigma_{xy}\Sigma_{yy}^{-1}\Sigma_{yx}\right)^{-1} \Rightarrow \Lambda_{x|y}$$

$$M_1^{-1} = \begin{bmatrix} \Lambda_{x|y} & -\Sigma_{xx}^{-1}\Sigma_{xy}\Lambda_{y|x} \\ -\Lambda_{y|x}\Sigma_{yx}\Sigma_{xx}^{-1} & \Lambda_{y|x} \end{bmatrix}$$

We now compute the alternate form of

$$M_2^{-1} = \begin{pmatrix} A & B \\ C & D \end{pmatrix}^{-1} = \begin{pmatrix} (A - BD^{-1}C)^{-1} & -(A-BD^{-1}C)^{-1}BD^{-1} \\ -(D^{-1}C(A-BD^{-1}C)^{-1} & D^{-1} + D^{-1}C(A-BD^{-1}C^{-1})BD^{-1} \end{pmatrix}$$

$A = \Sigma_{xx}$ $\quad B = \Sigma_{xy}$ $\quad C = \Sigma_{yx}$ $\quad D = \Sigma_{yy}$.

$$M_2^{-1} = \begin{pmatrix} \Sigma_{xx} & \Sigma_{xy} \\ \Sigma_{yx} & \Sigma_{yy} \end{pmatrix}^{-1} = \begin{pmatrix} (\Sigma_{xx} - \Sigma_{xy}\Sigma_{yy}^{-1}\Sigma_{yx})^{-1} & -(\Sigma_{xx} - \Sigma_{xy}\Sigma_{y}^{-1}\Sigma_{yx})\Sigma_{xy}\Sigma_{yy}^{-1} \\ -(\Sigma_{yy}^{-1}\Sigma_{yx}(\Sigma_{xx} - \Sigma_{xy}\Sigma_{yy}^{-1}\Sigma_{yx})^{-1}) & \Sigma_{yy}^{-1} + \Sigma_{yy}^{-1}\Sigma_{yx}(\Sigma_{xx} - \Sigma_{xy}\Sigma_{y}^{-1}\Sigma_{yx})\Sigma_{xy}\Sigma_{yy}^{-1} \end{pmatrix}$$

$$M_2^{-1} = \begin{pmatrix} \Lambda_{x|y} & -\Lambda_{x|y}\Sigma_{xy}\Sigma_{yy}^{-1} \\ -\Sigma_{yy}^{-1}\Sigma_{yx}\Lambda_{x|y} & \Sigma_{yy}^{-1} + \Sigma_{yy}^{-1}\Sigma_{yx}(\Sigma_{xx} - \Sigma_{xy}\Sigma_{yy}^{-1}\Sigma_{yx})^{-1}\Sigma_{xy}\Sigma_{yy}^{-1} \end{pmatrix}$$

again using

$$(Z + UWV^T)^{-1} = Z^{-1} - Z^{-1}U(W^{-1} + V^TZ^{-1}U)^{-1}V^TZ^{-1}$$

using $\quad Z^{-1} = -\Sigma_{yy}^{-1}$ $\quad U = \Sigma_{yx}$ $\quad W^{-1} = \Sigma_{xx}$ $\quad V^T = \Sigma_{xy}$ $\quad U = \Sigma_{yx}$

$$-\Big(\underbrace{(-\Sigma_{yy}^{-1})}_{Z^{-1}} - (-\Sigma_{yy})^{-1}(\Sigma_{yx})\big(\Sigma_{xx} + (\Sigma_{xy})^T(-\Sigma_{yy})(\Sigma_{yx})\big)^{-1}(\Sigma_{xy})(-\Sigma_{yy})$$

$$\Rightarrow \quad -(-\Sigma_{yy} + \Sigma_{yx}\Sigma_{xx}^{-1}\Sigma_{xy})$$

$$\Rightarrow \quad \Lambda_{y|x} \quad \text{also} = \Sigma_{yy}^{-1} + \Sigma_{yy}^{-1}\Sigma_{yx}\Lambda_{x}^{-1}\Sigma_{xy}\Sigma_{yy}^{-1}$$

Comparing entries in the inverses we have.

$$① \quad -\Lambda_{x|y}\Sigma_{xy}\Sigma_{yy}^{-1} = -\Sigma_{xx}^{-1}\Sigma_{xy}\Lambda_{y|x}$$

$$② \quad -\Lambda_{y|x}\Sigma_{yx}\Sigma_{xx}^{-1} = -\Sigma_{yy}\Sigma_{yx}\Lambda_{x|y}$$

$$\Lambda_{x|y}\Sigma_{xy}\Sigma_{yy}^{-1} = \Sigma_{xx}\Sigma_{xy}\Lambda_{y|x}$$

$$\Lambda_{x|y} = \Sigma_{xx}\Sigma_{xy}\Lambda_{y|x}\Sigma_{yy}\Sigma_{xy} \qquad \Lambda_{y|x} = \Sigma_{xy}^{-1}\Sigma_{xx}^{-1}\Lambda_{y|x}\Sigma_{xy}\Sigma_{yy}^{-1}$$

$$\Lambda_{x|y} = \Sigma_{yx}^{-1}\Sigma_{yy}^{-1}\Lambda_{y|x}\Sigma_{yx}\Sigma_{xx}^{-1} \qquad \Lambda_{y|x} = \Sigma_{yy}\Sigma_{yx}\Lambda_{x|y}\Sigma_{xx}\Sigma_{yx}^{-1}$$

Computing $p(x,y)/p(y)$
exponent

$$\begin{bmatrix} x & y \end{bmatrix}^T \begin{bmatrix} \Lambda_{x|y} & -\Lambda_{x|y}\Sigma_{xy}\Sigma_{yy}^{-1} \\ -\Sigma_{yy}^{-1}\Sigma_{yx}\Lambda_{x|y} & \Lambda_{y|x} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} - y^T\Sigma_{yy}^{-1}y$$

$$\Rightarrow \begin{bmatrix} x & y \end{bmatrix}^T \begin{bmatrix} \Lambda_{x|y} & -\Lambda_{x|y}\Sigma_{xy}\Sigma_{yy}^{-1} \\ -\Sigma_{yy}^{-1}\Sigma_{yx}\Lambda_{x|y} & \Lambda_{y|x} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} - y^T\Sigma_{yy}^{-1}y$$

$$\Rightarrow \quad x^T\Lambda_{x|y}x - x^T\Lambda_{x|y}\Sigma_{yy}^{-1}y - y^T\Sigma_{yy}^{-1}\Sigma_{yx}\Lambda_{x|y}x + y^T(\Lambda_{y|x} - \Sigma_{yy}^{-1})y$$

$$p(x|y) = \mathcal{N}(x; \Sigma_{xy}\Sigma_{yy}^{-1} y, \Sigma_{xx} - \Sigma_{xy}\Sigma_{yy}^{-1}\Sigma_{yx})$$

exponent : $(x - \Sigma_{xy}\Sigma_{yy}^{-1} y)^T \Lambda_{x|y} (x - \Sigma_{xy}\Sigma_{yy}^{-1} y)$

$$\Rightarrow (x^T - y^T \Sigma_{yy}^{-1}\Sigma_{xy}) \Lambda_{x|y} (x - \Sigma_{xy}\Sigma_{yy}^{-1} y)$$

$$\Rightarrow x^T \Lambda_{x|y} x - x^T \Lambda_{xy} \Sigma_{xy}\Sigma_{yy}^{-1} y - y^T \Sigma_{yy}^{-1}\Sigma_{xy}\Lambda_{x|y} x + y^T \Sigma_{yy}^{-1}\Sigma_{xy}\Lambda_{x|y}\Sigma_{xy}\Sigma_{yy}^{-1} y$$

Consider coefficient $y^T \cdot * y$

$$\Sigma_{yy}^{-1}\Sigma_{xy}\Lambda_{x|y}\Sigma_{xy}\Sigma_{yy}^{-1}$$

$$(Z + UWV^T)^{-1} = Z^{-1} - Z^{-1}U(W^{-1} + V^T Z^{-1} U)^{-1} V^T Z^{-1}$$

$$\Rightarrow \Lambda_{y|x} \quad \text{also} = \Sigma_{yy}^{-1} + \Sigma_{yy}^{-1}\Sigma_{yx}\Lambda_{x|y}\Sigma_{xy}\Sigma_{yy}^{-1}$$

$$(\Lambda_{y|x} - \Sigma_{yy}^{-1}) = \Sigma_{yy}^{-1}\Sigma_{yx}\Lambda_{x|y}\Sigma_{xy}\Sigma_{yy}^{-1}$$

Thus coefficient for $y^T \cdot * y$ $p(x|y)$ is same as $\frac{p(x,y)}{p(y)}$ thus by comparison all exponent coefficients are the same.

We now compute the determinat

$$\Sigma_{x|y} = \Sigma_{xx} - \Sigma_{xy} \Sigma_{yy}^{-1} \Sigma_{yx}$$

$$\det(|\Sigma_{x|y}|) = |\Sigma_{xx} - \Sigma_{xy} \Sigma_{yy}^{-1} \Sigma_{yx})$$

We can decompose the determinat of the block matrix $|\Sigma_{xy}|$ into a product of determinat of Σ_{yy} and remaining variance.

$$\det\left|\Sigma_{xy}\right| = \det\begin{vmatrix} \Sigma_{xx} & \Sigma_{xy} \\ \Sigma_{yx} & \Sigma_{yy} \end{vmatrix} = \det(\Sigma_{yy}) \det(\Sigma_{xx} - \Sigma_{xy} \Sigma_{yy}^{-1} \Sigma_{yx})$$

$$\det|\Sigma_{xy}| = \det(\Sigma_{x|y}) \det(\Sigma_{yy})$$

thus we have $$\frac{\det|\Sigma_{xy})}{\det|\Sigma_{yy})} = \det(\Sigma_{x|y})$$

Thus Finally we have -

where $\dim(x) + \dim(y) = D + D = 2D$

$$\frac{P(x,y)}{P(y)} = \frac{\dfrac{1}{(2\pi)^{2D/2} \det(\Sigma_{xy})^{\frac{1}{2}}} \exp\left\{-\frac{1}{2}\left|[x\ y]^T \Lambda_{xy} \begin{bmatrix} x \\ y \end{bmatrix}\right|\right\}}{\dfrac{1}{(2\pi)^{D/2}} \det(\Sigma_y)^{\frac{1}{2}} \exp\left\{-\frac{1}{2}\, y^T \Lambda_y y\right\}}$$

$$= \frac{1}{(2\pi)^{D/2}} \frac{1}{\det(\Sigma_{xy})^{\frac{1}{2}} / \det(\Sigma_y)^{\frac{1}{2}}} \exp\left\{-\frac{1}{2}\left([x\ y]^T \Lambda_{xy} \begin{bmatrix} x \\ y \end{bmatrix} - y^T \Lambda_y y\right)\right\}$$

$$= \frac{1}{(2\pi)^{p/2}} \frac{1}{\det(\Sigma_{x|y})^{\frac{1}{2}}} \exp\left\{ (x - \Sigma_{xy}\Sigma_{yy}^{-1}y)^T \Lambda_{x|y} (x - \Sigma_{xy}\Sigma_{yy}^{-1}y) \right\}$$

$$= \mathcal{N}(x \,; -\Sigma_{xy}\Sigma_{yy}^{-1}y, \Lambda_{x|y})$$

$= P(x|y)$ as per definition.

We know that here x and y actually refers to $\tilde{x}$ and $\tilde{y}$ by our initial substitution ☆

Resubstitute $x \triangleq x - \mu_x$, $y \triangleq y - \mu_y$, $\Lambda_{x|y} = (\Sigma_{xx} - \Sigma_{xy}\Sigma_{yy}^{-1}\Sigma_{yx})^{-1}$

$$\boxed{\mathcal{N}(x ; \mu_x - \Sigma_{xy}\Sigma_{yy}^{-1}(y - \mu_y), \Sigma_{xx} - \Sigma_{xy}\Sigma_{yy}^{-1}\Sigma_{yx})}$$

Q E D.