Homework 1: answers

Solution: Exercise 1 We first show that we may find $A \in \mathbb{R}^{p\times q}$ such that

$$X = AY + U, \quad \text{where } U \perp\!\!\!\perp Y, \text{ is a random vector of dimension } p \tag{1}$$

1. For $A \in \mathbb{R}^{p+q}$, and $U = X - AY$, we have that (U, Y) is a Gaussian vector, because it is a linear tranform of the Gaussian vector (X, Y). Thus, if $\mathbb{C}\mathrm{ov}(X - AY, Y) = \mathbf{0}_{\mathbb{R}^{p\times q}}$, then $U \perp\!\!\!\perp Y$. The converse is immediate: the covariance matrix of two independent vectors is null.

2. If A and U satisfy (1), then $U = X - AY$ is a linear transform of (X, Y), thus U is a Gaussian vector as well.

3. If A and U satisfy (1), then for all bounded, continuous function h, for P_Y-almost all y,

$$\mathbb{E}\left(h(X) \mid Y = y\right) = \mathbb{E}\left(h(AY + U) \mid Y = y\right) = \mathbb{E}\left(h(Ay + U)\right),$$

with $Ay + U \sim \mathcal{N}(Ay + m_U, \Sigma_U)$. Thus

$$\mathcal{L}(X|Y = y) = \mathcal{N}(Ay + m_U, \Sigma_U).$$

4. We now solve (1) w.r.t. A, U. According to question 1., an equivalent condition is that $\mathbb{C}\mathrm{ov}(X - AY, Y) = 0$.

$$\begin{aligned}\mathbb{C}\mathrm{ov}(X - AY, Y) = 0 &\iff \mathbb{C}\mathrm{ov}(X, Y) - A\mathbb{C}\mathrm{ov}(Y, Y) = 0\\ &\iff \Sigma_{X,Y} - A\Sigma_{YY} = 0\\ &\iff A = \Sigma_{YY}^{-1}\Sigma_{X,Y}.\end{aligned}$$

Now, according to question 3. we only need to compute Σ_U and m_U. The above display implies that

$$m_U = \mathbb{E}\left(X - \Sigma_{YY}^{-1}\Sigma_{X,Y}Y\right) = m_X - \Sigma_{YY}^{-1}\Sigma_{X,Y}m_Y \tag{2}$$

and letting $\tilde{X} = X - m_X$, $\tilde{Y} = Y - m_Y$,

$$\begin{aligned}\Sigma_U &= \mathbb{E}\left(\left(\tilde{X} - \Sigma_{XY}\Sigma_{YY}^{-1}\tilde{Y}\right)\left(\tilde{X} - \Sigma_{XY}\Sigma_{YY}^{-1}\tilde{Y}\right)^\top\right)\\ &= \Sigma_{XX} - \Sigma_{XY}\Sigma_{YY}^{-1}\Sigma_{YX} - \Sigma_{XY}\Sigma_{YY}^{-1}\Sigma_{YX} + \Sigma_{XY}\Sigma_{YY}^{-1}\Sigma_{YY}\Sigma_{YY}^{-1}\Sigma_{YX}\\ &= \Sigma_{XX} - \Sigma_{XY}\Sigma_{YY}^{-1}\Sigma_{YX}\end{aligned} \tag{3}$$

Combining question 3. and equations (2) and (3) yields

$$\mathcal{L}(X|Y = y) = \mathcal{N}\Big(m_X + \Sigma_{XY}\Sigma_{YY}^{-1}(y - m_X), \Sigma_{XX} - \Sigma_{XY}\Sigma_{YY}^{-1}\Sigma_{YX}\Big).$$

Solution: Exercise 2

- First the vector $\Big(X, AX + Y\Big)$ is Gaussian, as a linear transformation of the Gaussian vector (X, Y).

- Since $\mathbb{E}\left(AX+Y\right)=m_X+Am_Y$, $\mathbb{C}\mathrm{ov}(AX+Y)=A\Sigma_X A^\top+\Sigma_Y$ (because $\Sigma_{XY}=0$) and $\mathbb{C}\mathrm{ov}(X,AX+Y)=\Sigma_X A^\top$, we obtain

$$\begin{pmatrix} X \\ AX+Y \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} m_X \\ m_X+AY \end{pmatrix}, \begin{pmatrix} \Sigma_X & \Sigma_X A^\top \\ A\Sigma_X & \Sigma_Y + A\Sigma_X A^\top \end{pmatrix}\right).$$

Solution: Exercise 3

1. We check that $(\Lambda+A^\top LA)(\Lambda^{-1}-\Lambda^{-1}A^\top K^{-1}A\Lambda^{-1})=I_p$.

$$\begin{aligned} &(\Lambda+A^\top LA)(\Lambda^{-1}-\Lambda^{-1}A^\top K^{-1}A\Lambda^{-1}) \\ &= I_p - A^\top K^{-1}A\Lambda^{-1} + A^\top LA\Lambda^{-1} - A^\top LA\Lambda^{-1}A^\top K^{-1}A\Lambda^{-1} \\ &= I_p - A^\top(L-K^{-1}-LA\Lambda^{-1}A^\top K^{-1})A\Lambda^{-1} \end{aligned}$$

Alos, since $(A\Lambda^{-1}A^\top+L^{-1})K^{-1}=I_q$; we have

$$A\Lambda^{-1}A^\top K^{-1} = I_q - L^{-1}K^{-1}$$

The two latter displays yield

$$\begin{aligned} &(\Lambda+A^\top LA)(\Lambda^{-1}-\Lambda^{-1}A^\top K^{-1}A\Lambda^{-1}) \\ &= I_p + A^\top\Big(L-K^{-1}-L(I_q-L^{-1}K^{-1}\Big)A\Lambda^{-1} \\ &= I_p \end{aligned}$$

2. The expression for $\Sigma_{X|y}$ is a direct consequence of Exercise 2 and of the latter question. On the other hand, using Exercise 2 again,

$$\begin{aligned} \mu_{X|y} &= \mu + \Lambda^{-1}A^\top K^{-1}(y-A\mu-b) \\ &= (I_p-\Lambda^{-1}A^\top K^{-1}A)\mu + \Lambda^{-1}A^\top K^{-1}(y-b) \quad (4) \\ nonumber &= \big(I_p-(\Lambda^{-1}-S)\Lambda\big)\mu + \Lambda^{-1}A^\top K^{-1}(y-b) \quad \text{using the latter question.} \quad (5) \end{aligned}$$

Also

$$\begin{aligned} S\Lambda^{-1}A^\top K^{-1} &= (A^\top LA+\Lambda)\Lambda^{-1}A^\top(A\Lambda^{-1}A^\top+L^{-1})^{-1} \\ &= (A^\top LA\Lambda^{-1}A^\top + A^\top)(A\Lambda^{-1}A^\top+L^{-1})^{-1} \\ &= A^\top L(A\Lambda^{-1}A^\top+L^{-1})(A\Lambda^{-1}A^\top+L^{-1})^{-1} \\ &= I_q \end{aligned}$$

Thus $\Lambda^{-1}A^\top K^{-1}=SA^\top L$, which combined with (5) yields

$$\mu_{X|y} = S\Lambda\mu + SA^\top L(Y-b).$$