Solution to Exercise 4.4

★4.4 Consider the random variables x_1 and x_2, which are bivariate normal with $x_1 \sim \mathrm{N}(0, \sigma_1^2)$, $x_2 \sim \mathrm{N}(0, \sigma_2^2)$, and correlation ρ. Show that the expectation of x_1 conditional on x_2 is $\rho(\sigma_1/\sigma_2)x_2$ and that the variance of x_1 conditional on x_2 is $\sigma_1^2(1-\rho^2)$. How are these results modified if the means of x_1 and x_2 are μ_1 and μ_2, respectively?

By the definition of the correlation coefficient,

$$\rho = \frac{\mathrm{Cov}(x_1, x_2)}{\big(\mathrm{Var}(x_1)\,\mathrm{Var}(x_2)\big)^{1/2}},$$

from which we deduce that $\mathrm{Cov}(x_1, x_2) = \rho\sigma_1\sigma_2$. Thus x_1 and x_2 are bivariate normal with mean zero and covariance matrix

$$\boldsymbol{\Omega} \equiv \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix}.$$

At this point, the fastest way to proceed is to note that $\boldsymbol{\Omega} = \boldsymbol{A}\boldsymbol{A}^\top$, with

$$\boldsymbol{A} = \begin{bmatrix} \sigma_1(1-\rho^2)^{1/2} & \rho\sigma_1 \\ 0 & \sigma_2 \end{bmatrix}.$$

It follows that x_1 and x_2 can be expressed in terms of two independent standard normal variables, z_1 and z_2, as follows:

$$x_1 = \sigma_1(1-\rho^2)^{1/2}z_1 + \rho\sigma_1 z_2, \quad \text{and} \quad x_2 = \sigma_2 z_2,$$

from which we find that $x_1 = \sigma_1(1-\rho^2)^{1/2}z_1 + (\rho\sigma_1/\sigma_2)x_2$, where z_1 and x_2 are independent, since x_2 depends only on z_2, which is independent of z_1 by construction. Thus $\mathrm{E}(z_1 \,|\, x_2) = 0$, and it is then immediate that $\mathrm{E}(x_1 \,|\, x_2) = \rho(\sigma_1/\sigma_2)x_2$. For the conditional variance, note that

$$x_1 - \mathrm{E}(x_1 \,|\, x_2) = \sigma_1(1-\rho^2)^{1/2}z_1,$$

of which the variance is $\sigma_1^2(1-\rho^2)$, as required.

If the means of x_1 and x_2 are μ_1 and μ_2, respectively, then we have

$$x_1 = \mu_1 + \sigma_1(1-\rho^2)^{1/2}z_1 + \rho\sigma_1 z_2, \quad \text{and} \quad x_2 = \mu_2 + \sigma_2 z_2, \tag{S4.01}$$

where z_1 and z_2 are still independent standard normal variables. Since $z_2 = (x_2 - \mu_2)/\sigma_2$, we find that

$$\mathrm{E}(x_1 \,|\, x_2) = \mu_1 + \rho(\sigma_1/\sigma_2)(x_2 - \mu_2). \tag{S4.02}$$

It remains true that $x_1 - \mathrm{E}(x_1 \,|\, x_2) = \sigma_1(1-\rho^2)^{1/2} z_1$, and so the conditional variance is unchanged.

The factorization of $\boldsymbol{\Omega}$ as $\boldsymbol{A}\boldsymbol{A}^\top$ is somewhat tricky unless one applies Crout's algorithm directly. A clumsier, but more direct, way to proceed is to note that $\boldsymbol{x}^\top\boldsymbol{\Omega}^{-1}\boldsymbol{x}$ is distributed as $\chi^2(2)$, and can thus be expressed as the sum of the squares of two independent standard normal variables. By $\boldsymbol{x}$ we mean the vector with x_1 and x_2 as the only two components. By brute force, one can compute that

$$\boldsymbol{\Omega}^{-1} = \frac{1}{\sigma_1^2\sigma_2^2(1-\rho^2)}\begin{bmatrix} \sigma_2^2 & -\rho\sigma_1\sigma_2 \\ -\rho\sigma_1\sigma_2 & \sigma_1^2 \end{bmatrix},$$

so that

$$\boldsymbol{x}^\top\boldsymbol{\Omega}^{-1}\boldsymbol{x} = \frac{1}{\sigma_1^2\sigma_2^2(1-\rho^2)}\big(\sigma_2^2x_1^2 - 2\rho\sigma_1\sigma_2x_1x_2 + \sigma_1^2x_2^2\big).$$

By the operation of "completing the square," the right-hand side of this equation becomes

$$\begin{aligned}&\frac{1}{\sigma_1^2(1-\rho^2)}\Big(\big(x_1 - (\rho\sigma_1/\sigma_2)x_2\big)^2 + x_2^2\big(\sigma_1^2/\sigma_2^2 - \rho^2\sigma_1^2/\sigma_2^2\big)\Big)\\ &\qquad = \left(\frac{x_1 - (\rho\sigma_1/\sigma_2)x_2}{\big(\sigma_1^2(1-\rho^2)\big)^{1/2}}\right)^2 + \left(\frac{x_2}{\sigma_2}\right)^2.\end{aligned}$$

The two variables of which the squares are summed here can now be identified with z_1 and z_2 of (S4.01), and the rest of the argument is as above.

Since the above calculations are all either tricky or heavy, it is convenient to have a simpler way to rederive them when needed. This can be done by making use of the fact that, from (S4.02), the conditional expectation of x_1 is *linear* (more strictly, affine) with respect to x_2. Therefore, consider the linear regression

$$x_1 = \alpha + \beta x_2 + u. \tag{S4.03}$$

If both x_1 and x_2 have mean 0, it is obvious that $\alpha = 0$. The population analog of the standard formula for the OLS estimator of β tells us that

$$\beta = \frac{\mathrm{Cov}(x_1, x_2)}{\mathrm{Var}(x_2)} = \frac{\rho\sigma_1\sigma_2}{\sigma_2^2} = \frac{\rho\sigma_1}{\sigma_2}.$$

Thus we see that

$$x_1 = \frac{\rho\sigma_1}{\sigma_2}x_2 + u. \tag{S4.04}$$

It is then immediate that $\mathrm{E}(x_1 \,|\, x_2) = \rho(\sigma_1/\sigma_2)x_2$. The conditional variance follows by the same argument as above, and it can also be seen to be simply the variance of u in (S4.04).