GAUSSIAN PROCESSES
EXERCISE SHEET 8: LINEAR STRUCTURE OF GAUSSIANS II

Exercise 1.

It is enough to show that $C_{YY} - C_{YX}C_{XX}^{-1}C_{XY}$ is a covariance matrix and that

$$\begin{pmatrix} C_{XX} & C_{X(C_{YX}C_{XX}^{-1}X-Z)} \\ C_{(C_{YX}C_{XX}^{-1}X-Z)X} & C_{(C_{YX}C_{XX}^{-1}X-Z)(C_{YX}C_{XX}^{-1}X-Z)} \end{pmatrix} = \begin{pmatrix} C_{XX} & C_{XY} \\ C_{YX} & C_{YY} \end{pmatrix}.$$

To check that $C_{YY} - C_{YX}C_{XX}^{-1}C_{XY}$ is a covariance, we may just compute that it is the covariance of the Gaussian vector $C_{YX}C_{XX}^{-1}X - Y$:

$$\begin{aligned}
&\mathbb{E}(C_{YX}C_{XX}^{-1}X - Y)(C_{YX}C_{XX}^{-1}X - Y)^T \\
&= \mathbb{E}(C_{YX}C_{XX}^{-1}XX^TC_{XX}^{-1}C_{YX}^T - C_{YX}C_{XX}^{-1}XY^T - YX^TC_{XX}^{-1}C_{YX}^T + YY^T) \\
&= C_{YX}C_{XX}^{-1}C_{XY} - C_{YX}C_{XX}^{-1}C_{XY} - C_{YX}C_{XX}^{-1}C_{XY} + C_{YY} = C_{YY} - C_{YX}C_{XX}^{-1}C_{XY}.
\end{aligned}$$

Next we check that $C_{X(C_{YX}C_{XX}^{-1}X-Z)} = C_{XY}$:

$$\mathbb{E}X(C_{YX}C_{XX}^{-1}X - Z)^T = \mathbb{E}(C_{YX}C_{XX}^{-1}XX^T - XZ^T) = C_{YX}.$$

Similarly $C_{(C_{YX}C_{XX}^{-1}X-Z)X} = C_{YX}$, and finally

$$\begin{aligned}
\mathbb{E}(C_{YX}C_{XX}^{-1}X - Z)(C_{YX}C_{XX}^{-1}X - Z)^T &= \mathbb{E}(C_{YX}C_{XX}^{-1}XX^TC_{XX}^{-1}C_{XY} + ZZ^T) \\
&= C_{YX}C_{XX}^{-1}C_{XY} + C_{ZZ} = C_{YY}.
\end{aligned}$$

□

Exercise 2.

We recall that $\Gamma(k)$ was given by $\sum_{i=1}^k W(i)$, where $W(i)$ are i.i.d. standard normal random variables. Let us apply Exercise 1 in the case where $X = \Gamma(n)$ and $Y = (\Gamma(1), \ldots, \Gamma(n-1))$. We then see that $(X, Y) = (\Gamma(n), (\Gamma(1), \ldots, \Gamma(n-1))^T)$ has the same distribution as

$$(\Gamma(n), \frac{1}{\mathbb{E}\Gamma(n)^2}\begin{pmatrix} \mathbb{E}\Gamma(1)\Gamma(n) \\ \vdots \\ \mathbb{E}\Gamma(n-1)\Gamma(n) \end{pmatrix}\Gamma(n) - \overline{Z}) = (\Gamma(n), \frac{1}{n}\begin{pmatrix} 1 \\ \vdots \\ n-1 \end{pmatrix}\Gamma(n) - \overline{Z}),$$

where $\overline{Z}$ is independent of $\Gamma(n)$ and has covariance

$$C_{YY} - \frac{1}{n}C_{YX}C_{XY} = \Big(\min(i,j) - \frac{ij}{n}\Big)_{i,j=1}^n = \Big(\frac{\min(i,j)(n - \max(i,j))}{n}\Big)_{i,j=1}^n.$$

Thus conditioning on $\Gamma(n) = 0$ we get a process $-\overline{Z} =: \widehat{\Gamma}$ with the above covariance $\widehat{C} := \Big(\frac{\min(i,j)(n-\max(i,j))}{n}\Big)_{i,j=1}^n$ (the minus sign makes no difference because everything is mean zero). We recall from the model solution of Exercise 3 of Exercise sheet 7 that the inverse of $\widehat{C}$ is given by the following tridiagonal matrix:

$$\widehat{C}^{-1} = \begin{pmatrix} 2 & -1 & 0 & 0 & \dots & 0 & 0 & 0 \\ -1 & 2 & -1 & 0 & \dots & 0 & 0 & 0 \\ 0 & -1 & 2 & -1 & \dots & 0 & 0 & 0 \\ 0 & 0 & -1 & 2 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & 2 & -1 & 0 \\ 0 & 0 & 0 & 0 & \dots & -1 & 2 & -1 \\ 0 & 0 & 0 & 0 & \dots & 0 & -1 & 2 \end{pmatrix}.$$

Moreover, in the same exercise it was shown that the functions $\phi_i(j) = \sin(\pi \frac{ij}{n})$ are eigenfunctions of $\widehat{C}^{-1}$ with respective eigenvalues $\lambda_i = 4\sin(\frac{\pi i}{2n})^2$. Thus we may write the process $\widehat{\Gamma}$ as

$$\widehat{\Gamma}(j) = \sum_{i=1}^{n-1} \frac{1}{\sqrt{\lambda_i}} A_j \frac{\phi_i(j)}{\|\phi_i\|},$$

where A_j are i.i.d. $N(0,1)$ random variables and $\|\phi_i\| = \sqrt{\phi_i(1)^2 + \cdots + \phi_i(n-1)^2}$. It was computed in the Exercise 2 of Exercise sheet 7 that $\|\phi_i\| = \sqrt{n/2}$, so we get

$$\widehat{\Gamma}(j) = \sum_{i=1}^{n-1} \frac{1}{\sqrt{2n}\sin(\frac{\pi i}{2n})} A_j \sin(\pi \frac{ij}{n}).$$

$\square$

Exercise 3.

We may write

$$\begin{aligned} A &= \frac{\mathbb{E}[A(Ax_1+\epsilon_1)]}{\mathbb{E}[(Ax_1+\epsilon_1)^2]}(Ax_1+\epsilon_1) + B = \frac{x_1}{1+x_1^2}(Ax_1+\epsilon_1) + U \\ \epsilon_1 &= \frac{\mathbb{E}[\epsilon_1(Ax_1+\epsilon_1)]}{\mathbb{E}[(Ax_1+\epsilon_1)^2]}(Ax_1+\epsilon_1) + C = \frac{1}{1+x_1^2}(Ax_1+\epsilon_1) + V, \end{aligned}$$

where U and V are independent of $Ax_1 + \epsilon_1$. Moreover, solving for U and V we get

$$\begin{aligned} U &= \frac{1}{1+x_1^2}A - \frac{x_1}{1+x_1^2}\epsilon_1 \\ V &= \frac{-x_1}{1+x_1^2}A + \frac{x_1^2}{1+x_1^2}\epsilon_1. \end{aligned}$$

From here it follows that $V = -x_1 U$ and $U \sim N(0, \frac{1-x_1\sigma^2}{1+x_1^2})$, and (A, ϵ_1) has conditional distribution $(\frac{x_1 y_1}{1+x_1^2} + U, \frac{y_1}{1+x_1^2} + V)$ when conditioned on $Ax_1 + \epsilon_1 = y_1$.

For the second part, we use Exercise 1 with $X = (Ax_1 + \epsilon_1, \dots, Ax_n + \epsilon_n)$ and $Y = A$. We have

$$(C_{XX})_{ij} = \begin{cases} x_i x_j, & \text{if } i \neq j \\ 1 + x_i^2, & \text{if } i = j \end{cases}, \quad C_{XY} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \quad \text{and} \quad C_{YY} = 1.$$

With e.g. little guessing with small examples we see that

$$(C_{XX}^{-1})_{ij} = \frac{1}{1+x_1^2+x_2^2\cdots+x_n^2}\begin{cases}1+\sum_{k\neq i} x_k^2, & \text{if } i=j\\ -x_ix_j, & \text{if } i\neq j\end{cases}.$$

Indeed, if we D is a matrix with the elements as above, and $i \neq j$, then

$$(C_{XX}D)_{ij} = \sum_{k=1}^{n}(C_{XX})_{ik}D_{kj} = \frac{1}{1+\sum_{k=1}^n x_k^2}(-(1+x_i^2)x_ix_j + x_ix_j(1+\sum_{k\neq j} x_k^2) - \sum_{k\neq i,j} x_ix_jx_k^2) = 0.$$

Similarly, if $i=j$, then

$$(C_{XX}D)_{ii} = \sum_{k=1}^{n}(C_{XX})_{ik}D_{ki} = \frac{1}{1+\sum_{k=1}^n x_k^2}(\sum_{k\neq i} -x_i^2x_k^2 + (1+x_i^2)(1+\sum_{k\neq i} x_k^2)) = 1.$$

Hence

$$\begin{aligned}C_{YY} - C_{YX}C_{XX}^{-1}C_{XY} &= 1 - \sum_{i,j=1}^{n}(C_{YX})_{1i}(C_{XX}^{-1})_{ij}(C_{XY})_{j1} = 1 - \sum_{i,j=1}^{n} x_ix_j(C_{XX}^{-1})_{ij}\\ &= 1 - \frac{1}{1+\sum_{k=1}^n x_k^2}\Big(\sum_{i=1}^{n} x_i^2(1+\sum_{k\neq i} x_k^2) - \sum_{i\neq j} x_i^2x_j^2\Big)\\ &= \frac{1}{1+\sum_{k=1}^n x_k^2}.\end{aligned}$$

Thus when we condition by $X = (y_1,\dots,y_n)^T$, we get that A has variance $\frac{1}{1+\sum_{k=1}^n x_k^2}$ and mean

$$\begin{aligned}C_{YX}C_{XX}^{-1}(y_1,\dots,y_n)^T &= \sum_{i,j=1}^{n} x_i(C_{XX}^{-1})_{ij}y_j = \frac{1}{1+\sum_{k=1}^n x_k^2}\Big(\sum_{i=1}^{n} x_iy_i(1+\sum_{k\neq i} x_k^2) - \sum_{i\neq j} x_i^2x_jy_j\Big)\\ &= \frac{\sum_{i=1}^n x_iy_i}{1+\sum_{k=1}^n x_k^2}.\end{aligned}$$

Assuming next that $(x_i, y_i) \sim (X, 2X)$ are independent with $X \sim N(0,1)$, we see that the mean is

$$\frac{2\sum_{i=1}^n x_i^2}{1+\sum_{i=1}^n x_i^2},$$

which a.s. converges to 2, and the variance a.s. converges to 0. Thus the conditional law of A converges a.s. to the Dirac delta measure at the constant 2. $\square$

Exercise 4.

We consider the vector (Y, X) where $Y := MX$. Note that by Exercise 1, it has the same law as $(Y, C_{XY}C_{YY}^{-1}Y - Z)$ where Z is a zero mean Gaussian with covariance $C_{XX} - C_{XY}C_{YY}^{-1}C_{YX}$. To see this one simply check that C_{YY} is invertible, so the proof of Exercise 1 still works. Now we consider the conditioning $Y = MX = 0$, in this case, the covariance matrix of X would simply be the same as Z, which is

$$C_{XX} - C_{XY}C_{YY}^{-1}C_{YX} = C - CM^T(MCM^T)^{-1}MC. \tag{0.1}$$

$\square$