UC Berkeley
Department of Electrical Engineering and Computer Sciences

EECS 126: Probability and Random Processes

Discussion 14
Spring 2023

1. **MMSE for Jointly Gaussian Random Variables**

   Provide justification for each of the following steps to prove that the LLSE $g(X) \coloneqq \mathbb{L}(Y \mid X)$ is equal to the MMSE estimator for jointly Gaussian random variables X and Y.

$$
\begin{align}
& \mathbb{E}((Y - g(X)) \cdot X) = 0 \tag{1} \\
\implies & \operatorname{cov}(Y - g(X), X) = 0 \tag{2} \\
\implies & Y - g(X) \text{ is independent of } X \tag{3} \\
\implies & \mathbb{E}((Y - g(X)) \cdot f(X)) = 0 \ \forall f \tag{4} \\
\implies & g(X) = \mathbb{E}(Y \mid X). \tag{5}
\end{align}
$$

Solution:

a. As $g(X)$ is the LLSE, $Y - g(X)$ is orthogonal to all linear functions of X.
b. As $Y - g(X)$ is zero-mean, $\operatorname{cov}(Y - g(X), X) = \mathbb{E}((Y - g(X)) \cdot X) - \cancel{\mathbb{E}(Y - g(X))\,\mathbb{E}(X)}$.
c. Linear combinations of jointly Gaussian X, Y are also jointly Gaussian, in particular X and $Y - g(Y)$. For JG random variables, uncorrelatedness implies independence.
d. Functions of independent random variables are also independent. As $Y - g(X)$ and X are independent, $\mathbb{E}((Y - g(X)) \cdot f(X)) = \mathbb{E}(Y - g(X))\,\mathbb{E}(f(X)) = 0$.
e. The orthogonality principle for the MMSE states that $\mathbb{E}(Y \mid X)$ is the unique $g(X)$ that satisfies $\mathbb{E}((Y - g(X)) \cdot f(X)) = 0$ for any $f(X)$.

2. **Joint Gaussian Probability**

Let $X \sim \mathcal{N}(1,1)$ and $Y \sim \mathcal{N}(0,1)$ be jointly Gaussian with covariance ρ. What is $\mathbb{P}(X > Y)$?

Solution: Any linear combination of jointly Gaussian random variables is itself Gaussian. We observe that $\mathbb{P}(X > Y) = \mathbb{P}(X - Y > 0)$, and find

$$\begin{aligned} \mathbb{E}(X - Y) &= 1 \\ \operatorname{var}(X - Y) &= \operatorname{var}(X) - 2\operatorname{cov}(X, Y) + \operatorname{var}(Y) = 2 - 2\rho, \end{aligned}$$

so the distribution of $X - Y$ must be $\mathcal{N}(1, 2 - 2\rho)$. Then

$$\begin{aligned} \mathbb{P}(X > Y) &= \mathbb{P}(X - Y > 0) \\ &= \mathbb{P}\left(\mathcal{N}(0,1) > \frac{-1}{\sqrt{2 - 2\rho}}\right) \\ &= 1 - \Phi\left(\frac{-1}{\sqrt{2 - 2\rho}}\right) \\ &= \Phi\left(\frac{1}{\sqrt{2 - 2\rho}}\right). \end{aligned}$$

3. **Joint Gaussians As Linear Transformations of IID Gaussians**

Let $X = \begin{bmatrix} X_1 & X_2 \end{bmatrix}^\mathsf{T}$ be a jointly Gaussian random vector with mean $\begin{bmatrix} 0 & 0 \end{bmatrix}^\mathsf{T}$ and covariance

$$\Sigma_X = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} \frac{\sqrt{3}+1}{2} & \frac{\sqrt{3}-1}{2} \\ \frac{\sqrt{3}-1}{2} & \frac{\sqrt{3}+1}{2} \end{bmatrix} \begin{bmatrix} \frac{\sqrt{3}+1}{2} & \frac{\sqrt{3}-1}{2} \\ \frac{\sqrt{3}-1}{2} & \frac{\sqrt{3}+1}{2} \end{bmatrix} := AA^\mathsf{T}.$$

a. Express X_1, X_2 as linear combinations of i.i.d. standard Gaussian random variables.

b. Find a matrix B such that the components of BX are i.i.d. $\mathcal{N}(0,1)$. You do not have to simplify your answer.

Solution:

a. We claim that $X = AZ$ for some $Z = \begin{bmatrix} Z_1 & Z_2 \end{bmatrix}^\mathsf{T}$ with components $Z_1, Z_2 \sim_{\text{i.i.d.}} \mathcal{N}(0,1)$. AZ is a jointly Gaussian random vector as a linear transformation of Z, so it is uniquely characterized by its mean and covariance matrix:

$$\begin{aligned} \mathbb{E}(AZ) &= A\,\mathbb{E}(Z) \\ &= \vec{0} = \mathbb{E}(X) \\ \Sigma_{AZ} &= \mathbb{E}((AZ)(AZ)^\mathsf{T}) \\ &= \mathbb{E}(A(ZZ^\mathsf{T})A^\mathsf{T}) \\ &= A\,\mathbb{E}(ZZ^\mathsf{T})A^\mathsf{T} \\ &= AA^\mathsf{T} = \Sigma_X. \end{aligned}$$

Written out more explicitly, we have that

$$\begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = \begin{bmatrix} \frac{\sqrt{3}+1}{2} Z_1 + \frac{\sqrt{3}-1}{2} Z_2 \\ \frac{\sqrt{3}-1}{2} Z_1 + \frac{\sqrt{3}+1}{2} Z_2 \end{bmatrix}.$$

b. If $B = A^{-1}$, then $BX = (BA)Z = Z$ has i.i.d. $\mathcal{N}(0,1)$ components, as desired.