UC Berkeley
Department of Electrical Engineering and Computer Sciences

EECS 126: Probability and Random Processes

Discussion 13
Spring 2023

1. **Orthogonal LLSE**

   a. Consider zero-mean random variables X, Y, Z with Y, Z orthogonal. Show that

   $$\mathbb{L}(X \mid Y, Z) = \mathbb{L}(X \mid Y) + \mathbb{L}(X \mid Z).$$

   b. Now, for *any* zero-mean random variables X, Y, Z, explain why it holds that

   $$\mathbb{L}(X \mid Y, Z) = \mathbb{L}(X \mid Y) + \mathbb{L}[X \mid (Z - \mathbb{L}(Z \mid Y))].$$

Solution:

a. Let $f(Y) := \mathbb{L}(X \mid Y)$ and $g(Z) := \mathbb{L}(X \mid Z)$. We observe that X, $f(Y)$, and $g(Z)$ are all zero-mean, with Y orthogonal to $g(Z) \in \operatorname{span}\{1, Z\}$. ($Y$ is orthogonal to 1 because it is zero-mean.) A similar argument establishes that $f(Y)$ and Z are orthogonal. Now,

$$\begin{aligned}
\mathbb{E}(X - f(Y) - g(Z)) &= 0 \\
\mathbb{E}((X - f(Y) - g(Z))Y) &= \mathbb{E}(g(Z)Y) = 0 \\
\mathbb{E}((X - f(Y) - g(Z))Z) &= \mathbb{E}(f(Y)Z) = 0,
\end{aligned}$$

where $X - f(Y) \perp Y$ and $X - g(Z) \perp Z$. As $X - (f(Y) + g(Z))$ is orthogonal to any linear function of $1, Y, Z$, by the orthogonality principle,

$$\begin{aligned}
\mathbb{L}(X \mid Y, Z) &= f(Y) + g(Z) \\
&= \mathbb{L}(X \mid Y) + \mathbb{L}(X \mid Z).
\end{aligned}$$

b. $W := Z - \mathbb{L}(Z \mid Y)$ is orthogonal to Y, so $\mathbb{L}(X \mid Y, W) = \mathbb{L}(X \mid Y) + \mathbb{L}(X \mid W)$ by part a. Then, $\mathbb{L}(X \mid Y, W) = \mathbb{L}(X \mid Y, Z)$ as $\operatorname{span}\{1, Y, W\} = \operatorname{span}\{1, Y, Z\}$.

2. **Improving Estimation Error**

Show that "more information yields better estimation error":

$$\mathbb{E}[(X - \mathbb{E}[X \mid Y, Z])^2] \le \mathbb{E}[(X - \mathbb{E}[X \mid Y])^2].$$

Hint: $\mathbb{E}[X \mid Y]$ can be interpreted as the orthogonal projection of X onto the subspace of all functions of Y.

Solution: $\mathbb{E}[X \mid Y, Z]$ is the orthogonal projection of X onto the subspace of all functions of Y, Z, so it is also the MMSE estimator, the minimizer of squared distance to X:

$$\mathbb{E}[X \mid Y, Z] = \operatorname*{argmin}_{h(Y,Z)} \mathbb{E}[(X - h(Y, Z))^2].$$

$\mathbb{E}[X \mid Y]$ is also a function of Y, Z, so its mean squared error $\mathbb{E}[(X - \mathbb{E}[X \mid Y])^2]$ is at least the minimum $\mathbb{E}[(X - \mathbb{E}[X \mid Y, Z])^2]$.

Alternatively, let $f(Y) \coloneqq \mathbb{E}[X \mid Y]$ and $g(Y, Z) \coloneqq \mathbb{E}[X \mid Y, Z]$. Then we have

$$\begin{aligned}
\mathbb{E}[(X - f(Y))^2] &= \mathbb{E}[(X - g(Y, Z) + g(Y, Z) - f(Y))^2] \\
&= \mathbb{E}[(X - g(Y, Z))^2] + \mathbb{E}[(g(Y, Z) - f(Y))^2] \\
&\ge \mathbb{E}[(X - g(Y, Z))^2],
\end{aligned}$$

where $X - g(Y, Z)$ is orthogonal to $g(Y, Z) - f(Y)$, an affine function of Y, Z.

3. **Photodetector LLSE**

Consider a photodetector in an optical communications system that counts the number of photons arriving during a certain interval. A user can convey information to the system using a photon transmitter. Suppose the transmitter is constantly on with probability p and constantly off otherwise.

- If the transmitter is on, the number of photons transmitted over the interval of interest is a Poisson random variable Θ with mean λ.
- If the transmitter is off, the number of photons transmitted is 0.

Unfortunately, regardless of whether the transmitter is on or off, photons may be detected due to "shot noise." The number N of detected shot noise photons is a Poisson(μ) random variable, independent of the number T of transmitted photons. If D is the total number of photons detected, find $\mathbb{L}(T \mid D)$.

(You do not have to simplify your final expression.)

Solution: We can compute the LLSE as

$$\mathbb{L}(T \mid D) = \mathbb{E}(T) + \frac{\operatorname{cov}(T, D)}{\operatorname{var}(D)}(D - \mathbb{E}(D)).$$

$\mathbb{E}(T) = p\lambda$ by the law of total expectation. Noting that $D = T + N$,

$$\begin{aligned} \operatorname{cov}(T, D) &= \operatorname{var}(T) + \operatorname{cov}(T, N) \\ &= \mathbb{E}(T^2) - \mathbb{E}(T)^2 \\ &= p(\lambda^2 + \lambda) - (p\lambda)^2. \end{aligned}$$

Now, again by the independence of T and N,

$$\begin{aligned} \operatorname{var}(D) &= \operatorname{var}(T) + \operatorname{var}(N) \\ &= p(\lambda^2 + \lambda) - (p\lambda)^2 + \mu. \end{aligned}$$

Lastly, we have $\mathbb{E}(D) = \mathbb{E}(T) + \mathbb{E}(N) = p\lambda + \mu$. Putting these together, we have the LLSE.