UC Berkeley
Department of Electrical Engineering and Computer Sciences

EECS 126: Probability and Random Processes

Problem Set 13

Fall 2021

1. Projections

The following exercises are from the note on the Hilbert space of random variables. See the notes for some hints.

(a) Let $\mathcal{H} := \{X : X \text{ is a real-valued random variable with } \mathbb{E}[X^2] < \infty\}$. Prove that $\langle X, Y\rangle := \mathbb{E}[XY]$ makes $\mathcal{H}$ into a real inner product space. [1]

(b) Let U be a subspace of a real inner product space V and let P be the projection map onto U. Prove that P is a linear transformation.

(c) Suppose that U is finite-dimensional, $n := \dim U$, with basis $\{v_i\}_{i=1}^n$. Suppose that the basis is orthonormal. Show that $Py = \sum_{i=1}^n \langle y, v_i\rangle v_i$. (Note: If we take $U = \mathbb{R}^n$ with the standard inner product, then P can be represented as a matrix in the form $P = \sum_{i=1}^n v_i v_i^{\mathsf{T}}$.)

Solution:

(a) To satisfy the definition of a vector space, we must check that $\mathcal{H}$ is closed under vector addition and scalar multiplication, that is, if $X, Y \in \mathcal{H}$ and $c \in \mathbb{R}$, then $X + Y \in \mathcal{H}$ and $cX \in \mathcal{H}$. Since

$$\mathbb{E}[(X+Y)^2] = \mathbb{E}[X^2] + 2\mathbb{E}[XY] + \mathbb{E}[Y^2]$$

and $\mathbb{E}[X^2], \mathbb{E}[Y^2] < \infty$, we must show that $\mathbb{E}[XY] < \infty$, but by the Cauchy-Schwarz Inequality, $\mathbb{E}[XY] \le \sqrt{\mathbb{E}[X^2]\mathbb{E}[Y^2]} < \infty$. Also,

$$\mathbb{E}[(cX)^2] = c^2\mathbb{E}[X^2] < \infty.$$

We must also check the following properties of a vector space: for all $X, Y, Z \in \mathcal{H}$ and $c, d \in \mathbb{R}$,

- $X + (Y + Z) = (X + Y) + Z$;
- $X + Y = Y + X$;
- $0 + X = X$;
- $1 \cdot X = X$;
- $c(dX) = (cd)X$;
- $c(X + Y) = cX + cY$;
- $(c + d)X = cX + dX$.

All of the above are familiar properties of functions, so we do not bother to check them. Finally, we have to check the properties of the inner product:

[1] To be perfectly correct, it is possible for $X \neq 0$ but $\mathbb{E}[X^2] = 0$; this occurs if $X = 0$ with probability 1. To fix this, we need to define two random variables X and Y to be *equal* if $P(X = Y) = 1$. In other words, we consider *equivalence classes* of random variables, defined by the relation $\overset{a.s.}{=}$. With this definition, then if $X \neq 0$ we do indeed have $\mathbb{E}[X^2] > 0$.

- $\mathbb{E}[XY] = \mathbb{E}[YX]$;
- $\mathbb{E}[(X + cY)Z] = \mathbb{E}[XZ] + c\mathbb{E}[XZ]$;
- $\mathbb{E}[X^2] > 0$ if $X \neq 0$.

The first property is clear; the second property is linearity of expectation. Perhaps surprisingly, the third property is technically challenging, but do not worry about the details.

(b) Let $u, v \in V$ and $c \in \mathbb{R}$. Then, we claim that $P(u + cv) = Pu + cPv$. It suffices to check that $Pu + cPv \in U$ and $u + cv - Pu - cPv \in U^\perp$. Since $Pu \in U$ and $Pv \in U$, then $Pu + cPv \in U$ since U is a subspace. Also, for any $w \in U$, we get

$$\langle w, u + cv - Pu - cPv\rangle = \langle w, u - Pu\rangle + c\langle w, v - Pv\rangle = 0$$

since $u - Pu \in U^\perp$ and $v - Pv \in U^\perp$. Thus, $u + cv - Pu - cPv \in U^\perp$. We therefore have $P(u + cv) = Pu + cPv$ and P is linear.

(c) Since each $v_i \in U$, then $\sum_{i=1}^n \langle y, v_i\rangle v_i \in U$ since U is a subspace. Also, for any v_j, we get

$$\left\langle v_j, y - \sum_{i=1}^n \langle y, v_i\rangle v_i \right\rangle = \langle v_j, y\rangle - \sum_{i=1}^n \langle y, v_i\rangle\langle v_j, v_i\rangle = \langle v_j, y\rangle - \langle y, v_j\rangle = 0,$$

where we have used the fact that the basis is orthonormal (so $\langle v_i, v_j\rangle = 0$ if $i \neq j$ and $\langle v_j, v_j\rangle = 1$). Since we have $\langle v_j, y - \sum_{i=1}^n \langle y, v_i\rangle v_i\rangle$ for any v_j and $v_1, \ldots, v_n$ is a basis for U, then we get $\langle u, y - \sum_{i=1}^n \langle y, v_i\rangle v_i\rangle = 0$ for any $u \in U$, i.e., $y - \sum_{i=1}^n \langle y, v_i\rangle v_i \in U^\perp$. Hence, $Py = \sum_{i=1}^n \langle y, v_i\rangle v_i$.

2. Exam Difficulties

The difficulty of an EECS 126 exam, Θ, is uniformly distributed on $[0, 100]$ (i.e. continuous distribution, not discrete), and Alice gets a score X that is uniformly distributed on $[0, \Theta]$. Alice gets her score back and wants to estimate the difficulty of the exam.

(a) What is the MLE of Θ? What is the MAP of Θ?

(b) What is the LLSE for Θ?

Solution:

(a) Since our prior on Θ is the uniform prior, the MLE and MAP estimates will be the same. Given X, $X \le \Theta \le 100$. In order to maximize

$$f_{X\mid\Theta}(x \mid \theta) = \frac{1}{\theta},$$

one should choose $\hat{\Theta} = X$.

(b) We need $\mathbb{E}[\Theta]$, $\operatorname{var} X$, and $\operatorname{cov}(X, \Theta)$. First, $\mathbb{E}[\Theta] = 50$. By the **Law of Total Variance**,

$$\operatorname{var} X = \mathbb{E}[\operatorname{var}(X \mid \Theta)] + \operatorname{var} \mathbb{E}[X \mid \Theta].$$

The first term can be found as follows.

$$\operatorname{var}(X \mid \Theta) = \frac{\Theta^2}{12} \implies \mathbb{E}[\operatorname{var}(X \mid \Theta)] = \int_0^{100} \frac{\theta^2}{12} \cdot \frac{1}{100}\,\mathrm{d}\theta = \frac{10000}{36}$$

By noting $\mathbb{E}[X \mid \Theta] = \Theta/2$, the second term is

$$\frac{1}{4}\frac{10000}{12} = \frac{10000}{48}.$$

Thus,

$$\operatorname{var} X = \frac{70000}{144}.$$

Now, the covariance can be found as follows.

$$\operatorname{cov}(X, \Theta) = \mathbb{E}[\Theta X] - \mathbb{E}[\Theta]\mathbb{E}[X]$$

We found $\mathbb{E}[\Theta]$ above, and $\mathbb{E}[X] = \mathbb{E}[\mathbb{E}[X \mid \Theta]] = \mathbb{E}[\Theta/2] = 25$. Also,

$$\begin{aligned}\mathbb{E}[\Theta X] &= \mathbb{E}\big[\mathbb{E}[\Theta X \mid \Theta]\big] = \mathbb{E}\Big[\frac{\Theta^2}{2}\Big] = \frac{\operatorname{var}\Theta + \mathbb{E}[\Theta]^2}{2} = \frac{10000/12 + 2500}{2} \\ &= \frac{5000}{3}.\end{aligned}$$

Thus,

$$\operatorname{cov}(X, \Theta) = \frac{1250}{3}.$$

Then, the LLSE of Θ is

$$L[\Theta \mid X] = \mathbb{E}[\Theta] + \frac{\operatorname{cov}(X, \Theta)}{\sigma_X^2}(X - \mathbb{E}[X]) = 50 + \frac{6}{7}(X - 25).$$

3. **Jointly Gaussian Decomposition**

Let U and V be jointly Gaussian random variables with means $\mu_U = 1$, $\mu_V = 4$, respectively, with variances $\sigma_U^2 = 2.5$, $\sigma_V^2 = 2$, respectively, and with covariance $\rho = 1$. Can we write U as $U = aV + Z$, where a is a scalar and Z is independent of V? If you think we can, find the value of a and the distribution of Z; otherwise please explain the reason.

Solution:

The LLSE of U given V is

$$L[U \mid V] = \mu_U + \frac{\rho}{\sigma_V^2}(V - \mu_V).$$

For jointly Gaussian random variables, $Z' = U - L[U \mid V]$ is independent of V, so we have

$$U = \mu_U + \frac{\rho}{\sigma_V^2}(V - \mu_V) + Z'.$$

Let

$$\begin{aligned}a &= \frac{\rho}{\sigma_V^2}, \\ Z &= \mu_U - \frac{\rho}{\sigma_V^2}\mu_V + Z'.\end{aligned}$$

Then $U = aV + Z$ and Z is independent of V. We can find that

$$\mathbb{E}[Z] = \mu_U - \frac{\rho}{\sigma_V^2}\mu_V,$$

and

$$\operatorname{var} Z = \sigma_U^2 - \frac{\rho^2}{\sigma_V^2}.$$

Then we know that

$$Z \sim \mathcal{N}\Big(\mu_U - \frac{\rho}{\sigma_V^2}\mu_V, \sigma_U^2 - \frac{\rho^2}{\sigma_V^2}\Big).$$

Then we know that $a = 0.5$ and $Z \sim \mathcal{N}(-1, 2)$.

4. **Photodetector LLSE**

Consider a photodetector in an optical communications system that counts the number of photons arriving during a certain interval. A user conveys information by switching a photon transmitter on or off. Assume that the probability of the transmitter being on is p. If the transmitter is on, the number of photons transmitted over the interval of interest is a Poisson random variable Θ with mean λ, and if it is off, the number of photons transmitted is 0. Unfortunately, regardless of whether or not 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 random variable N with mean μ, independent of the transmitted photons. Let T be the number of transmitted photons and D be the number of detected photons. Find $L[T \mid D]$.

Solution:

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

We find each of these terms separately. We can see by the law of total expectation that $\mathbb{E}[T] = p\lambda$. Now, we have:

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

where the second to last equality follows since T and N are independent. We now find:

$$\operatorname{var} D = \operatorname{var}(T + N) = \operatorname{var} T + \operatorname{var} N = p(\lambda^2 + \lambda) - (p\lambda)^2 + \mu$$

Finally, we have $\mathbb{E}[D] = \mathbb{E}[T] + \mathbb{E}[N] = p\lambda + \mu$. Putting these together, we have the LLSE (no need to simplify the equation).