Stats 310B / Math 230B Midterm Solutions (7 February 2014)

Problem 1. Write $X \equiv (X_1, \ldots, X_n)$ and $\langle a, X\rangle \equiv \sum_{i=1}^n a_i X_i$.

(a) By Thm. 4.3.10 it suffices to show that $\mathcal{W} \equiv \mathcal{W}(\{X_i\}_{i\in[n]})$ is a Hilbert subspace (Defn. 4.3.5) of $L^2(\Omega, \mathcal{F}, \mathbb{P})$. Clearly $\mathcal{W}$ is closed under addition and scalar multiplication, so it remains to show that every Cauchy sequence in $\mathcal{W}$ has a limit in $\mathcal{W}$. Since $\mathcal{W}$ is the span of the X_i we assume without loss that the X_i are linearly independent; in particular, each X_i has L^2-norm $\|X_i\| > 0$. We can also apply Gram–Schmidt orthogonalization to ensure $\mathbb{E}[X_i X_j] = 0$ for every $i \neq j$. If $X^{(t)} = \langle a^{(t)}, X\rangle$ ($t \geqslant 1$) defines a Cauchy sequence, then

$$0 = \lim_{s\to\infty} \Big(\sup_{t\geqslant s} \|X^t - X^s\|^2 \Big) = \lim_{s\to\infty} \Big(\sup_{t\geqslant s} \|a^t - a^s\|^2 \Big)$$

where the first equality is by definition of a Cauchy sequence, and the second is by orthogonality. Therefore $(a^t)_{t\geqslant 1}$ is a Cauchy sequence in the Hilbert space $\mathbb{R}^n$, so converges to a limit $a \in \mathbb{R}^n$. It follows that X^t converges to $X \equiv \langle a, X\rangle \in \mathcal{W}$. Thm. 4.3.10 then implies the existence of $\widehat{Y} \in \mathcal{W}$ satisfying the conditions stated in the problem.

(b) From part (a) we have $\mathbb{E}[(Y - \widehat{Y})W] = 0$ for all $W \in \mathcal{W}$. If $(X, Y) \equiv (X_1, \ldots, X_n, Y)$ is a zero-mean jointly-Gaussian vector, then so is $(X, Y - \widehat{Y})$, with

$$\mathrm{Cov}(Y - \widehat{Y}, X_i) = \mathbb{E}[(Y - \widehat{Y})X_i] = 0 \quad \text{for all } 1 \leqslant i \leqslant n$$

by part (a). It follows that $Y - \widehat{Y}$ is independent of X, so for any $B \in \mathcal{F} \equiv \sigma(X)$ we have

$$\mathbb{E}[(Y - \widehat{Y})\mathbf{1}_B] = \mathbb{E}[Y - \widehat{Y}]\mathbb{E}[\mathbf{1}_B] = 0,$$

proving that $\widehat{Y}$ is a version of $\mathbb{E}[Y|\mathcal{F}] \equiv \mathbb{E}[Y|X]$.

(c) We wish to show that $\widehat{\mathbb{P}}_{Y|\mathcal{F}}(A, \omega) \equiv \gamma_{\widehat{Y}(\omega),\sigma^2}(A)$ where $\sigma^2 \equiv \mathbb{E}[(Y - \widehat{Y})^2]$ is an RCPD for Y given $\mathcal{F} \equiv \sigma(X)$. We check the conditions of Defn. 4.4.2 with $(\mathbb{S}, \mathcal{S}) = (\mathbb{R}, \mathcal{L})$. It is clear that for each fixed $\omega \in \Omega$, the set function $A \mapsto \widehat{\mathbb{P}}_{Y|\mathcal{F}}(A, \omega)$ defines a probability measure on $(\mathbb{R}, \mathcal{L})$. It remains to show that for each fixed $A \in \mathcal{L}$, $\widehat{\mathbb{P}}_{Y|\mathcal{F}}(A, \omega)$ is a version of $\mathbb{E}[\mathbf{1}\{Y \in A\}|\mathcal{F}]$ — equivalently, for all $A \in \mathcal{L}$ and $B \in \mathcal{F}$,

$$\mathbb{E}[\mathbf{1}_B(\mathbf{1}\{Y \in A\} - \widehat{\mathbb{P}}_{Y|\mathcal{F}}(A, \omega))] = 0.$$

Indeed, writing Z for a centered Gaussian random variable with variance σ^2 independent of everything else, it follows from the independence noted in part (b) that $Y = (Y - \widehat{Y}) + \widehat{Y}$ where $\widehat{Y} \in \mathcal{F}$, while $Y - \widehat{Y}$ independent of $\mathcal{F}$ and has the same law as Z. Therefore

$$\mathbb{E}[\mathbf{1}_B\mathbf{1}\{Y \in A\}] = \mathbb{E}[\mathbf{1}_B\mathbf{1}\{Z + \widehat{Y} \in A\}] = \mathbb{E}[\mathbf{1}_B\widehat{\mathbb{P}}_{Y|\mathcal{F}}(A, \omega)],$$

where the last step follows since, conditioned on $\mathcal{F}$, $Z + \widehat{Y}$ is distributed as a Gaussian random variable with variance σ^2 centered at $\widehat{Y}$. This concludes the proof.

Problem 2. First note that since X and Y are absolutely continuous, $\mathbb{P}(X = Y = M) = 0$. The symmetry between X and Y then gives, for any bounded measurable function g,

$$\mathbb{E}[g(M)] = \mathbb{E}[g(M)\mathbf{1}\{M = X\}] + \mathbb{E}[g(M)\mathbf{1}\{M = Y\}] = 2\mathbb{E}[g(X)\mathbf{1}\{M = X\}]. \tag{1}$$

Write $F(x) \equiv \int_{-\infty}^x f(x)\,dx$. We will show that the RCPD of X given M is

$$\widehat{\mathbb{P}}_{X|M}(A, \omega) = \begin{cases} (1/2)\mathbf{1}\{M(\omega) \in A\} + (1/2)\dfrac{\int_{-\infty}^{M(\omega)} \mathbf{1}_A(x) f(x)\,dx}{F(M(\omega))} & \text{if } F(M(\omega)) > 0, \\ \int_A f(x)\,dx & \text{otherwise.} \end{cases}$$

We verify the conditions of Defn. 4.4.2 with $(\mathbb{S}, \mathcal{S}) = (\mathbb{R}, \mathcal{L})$. It is clear that for each fixed $\omega \in \Omega$, the set function $A \mapsto \widehat{\mathbb{P}}_{X|M}(A, \omega)$ defines a probability measure on $(\mathbb{R}, \mathcal{L})$. It remains to show that for each fixed $A \in \mathcal{L}$, $\widehat{\mathbb{P}}_{X|M}(A,\cdot)$ is a version of $\mathbb{E}[\mathbf{1}\{X \in A\}|M]$ — equivalently, for all $A, B \in \mathcal{L}$,

$$\mathbb{E}[\mathbf{1}\{M \in B\}\mathbf{1}\{X \in A\}] = \mathbb{E}[\mathbf{1}\{M \in B\}\widehat{\mathbb{P}}_{X|M}(A,\omega)]. \tag{2}$$

Using the π-λ theorem it suffices to check (2) for $B = (-\infty, b]$ and $A = (-\infty, a]$. Using (1),

$$\begin{aligned} \text{LHS of (2)} &= \mathbb{P}(M \leqslant b, X \leqslant a, M = X) + \mathbb{P}(M \leqslant b, X \leqslant a, M = Y) \\ &= (1/2)\mathbb{P}(M \leqslant a \wedge b) + \mathbb{P}(Y \leqslant b, X \leqslant Y \wedge a). \end{aligned}$$

Meanwhile, since $\mathbb{P}(F(M) = 0) = 0$, we have

$$\text{RHS of (2)} = (1/2)\mathbb{P}(M \leqslant a \wedge b) + (1/2)\mathbb{E}\Big[\mathbf{1}\{M \leqslant b\}\frac{F(M \wedge a)}{F(M)}\Big].$$

Using (1) again we can re-express the second term as

$$\begin{aligned} &\mathbb{E}\Big[\mathbf{1}\{M = Y\}\mathbf{1}\{Y \leqslant b\}\frac{F(Y \wedge a)}{F(Y)}\Big] = \mathbb{E}\Big[\mathbf{1}\{Y \leqslant b\}\frac{F(Y \wedge a)}{F(Y)}\mathbf{1}\{X < Y\}\Big] \\ &= \mathbb{E}\Big[\mathbf{1}\{Y \leqslant b\}\frac{F(Y \wedge a)}{F(Y)}\mathbb{P}(X < Y|Y)\Big] = \mathbb{E}[\mathbf{1}\{Y \leqslant b\}F(Y \wedge a)] = \mathbb{P}(Y \leqslant b, X \leqslant Y \wedge a) \end{aligned}$$

where we repeatedly used the tower property together with the independence of X from Y. This completes the verification of (2) and proves that the RCPD is as written above.

Problem 3.

(a) From the recursive definition of X_n and the bound $0 \leqslant f(x) \leqslant x \wedge (1-x)$ we have

$$0 \leqslant X_{n-1} - f(X_{n-1}) \leqslant X_n \leqslant X_{n-1} + f(X_{n-1}) \leqslant 1,$$

so X_n is well-defined and $[0,1]$-valued for all $n \geqslant 0$. Since ξ_n has zero mean and is independent of X_{n-1} we have $\mathbb{E}[X_n|\mathcal{F}_n] = X_{n-1} + (\mathbb{E}\xi_n)f(X_{n-1}) = X_{n-1}$ for all $n \geqslant 1$, proving that X_n is an $\mathcal{F}_n$-martingale. It follows from Thm. 5.3.12 that $X_n \to X_\infty$ both a.s. and in L^1 for some $[0,1]$-valued random variable X_∞.

(b) Since $X_n \to X_\infty$ a.s., $f(X_n) = |X_{n+1} - X_n| \to 0$ a.s. On the other hand continuity of f implies $f(X_n) \to f(X_\infty)$ a.s., so we conclude $f(X_\infty) = 0$ a.s.

(c) If $f(x) > 0$ for all $x \in (0,1)$ then part (b) implies $X_\infty \in \{0,1\}$ a.s., with mean $\mathbb{E}X_\infty = \mathbb{E}X_0 = \epsilon$ (since we saw in part (a) that X_n is a martingale, and $X_n \to X_\infty$ in L^1). Therefore $X_\infty \sim \text{Ber}(\epsilon)$.

Problem 4. Write $C \equiv \sup_n \mathbb{E}[X_n^2] < \infty$, and note that Cauchy–Schwarz gives

$$\mathbb{E}[(X_k - X_{k-1})^2] \leqslant 2C + 2|\mathbb{E}[X_k X_{k-1}]| \leqslant 2C + 2\mathbb{E}[X_k^2]^{1/2}\mathbb{E}[X_{k-1}^2]^{1/2} \leqslant 4C.$$

Using Cauchy–Schwarz again and recalling Exercise 5.1.8(a) we have, for all n,

$$(\mathbb{E}|Y_n|)^2 \leqslant \mathbb{E}[Y_n^2] = \sum_{k=1}^n a_k^2 \mathbb{E}[(X_k - X_{k-1})^2] \leqslant 4C\sum_{k=1}^\infty a_k^2 = 4C\|a\|^2 < \infty.$$

In particular, Y_n is integrable, and it is easy to see that it is a martingale. Since $\sup_n \mathbb{E}|Y_n| < \infty$, Doob's convergence theorem (Thm. 5.3.2) implies $Y_n \to Y_\infty$ a.s. with $\mathbb{E}|Y_\infty| < \infty$ and $\mathbb{E}[(Y_\infty)^2] < \infty$ (Fatou's lemma). Also, for $n < m$,

$$\mathbb{E}[(Y_m - Y_n)^2] = \mathbb{E}\Big[\Big(\sum_{k=n+1}^m a_k(X_k - X_{k-1})\Big)^2\Big] = \mathbb{E}\Big[\sum_{k=n+1}^m a_k^2(X_k - X_{k-1})^2\Big] \leqslant C\sum_{k>n} a_k^2.$$

Thus Y_n defines an L^2 Cauchy sequence, so it must converge in L^2 to a limit Z_∞. For any $\epsilon > 0$,

$$\begin{aligned} &\mathbb{P}(|Z_\infty - Y_\infty| > \epsilon) \leqslant \mathbb{P}(|Y_n - Z_\infty| > \epsilon/2) + \mathbb{P}(|Y_n - Y_\infty| > \epsilon/2) \\ &\leqslant (\epsilon/2)^{-2}\mathbb{E}[(Y_n - Z_\infty)^2] + \mathbb{P}(|Y_n - Y_\infty| > \epsilon/2) \to 0 \ \text{ as } n \to \infty, \end{aligned}$$

therefore $Z_\infty = Y_\infty$ a.s. which concludes the proof.