RANDOM PROCESSES - SOLUTION OF THE FINAL EXAM 2001 (B)

PAVEL CHIGANSKY

Problem 1.

(a) $\mathbf{E}\big(\eta|\widehat{\mathbf{E}}(\eta|\xi)\big) = \mathbf{E}(\eta|\xi)$ is (generally) **TRUE** if and only if $\alpha := \mathbf{E}\eta\xi/\mathbf{E}\xi^2 \neq 0$, since in this case ξ and $\widehat{\mathbf{E}}(\eta|\xi)$ are related by one to one correspondence, i.e. $\widehat{\mathbf{E}}(\eta|\xi) = \alpha\xi$.

(b) $\widehat{\mathbf{E}}\big(\eta|\mathbf{E}(\eta|\xi)\big) = \mathbf{E}(\eta|\xi)$ is **TRUE**:

$$\widehat{\mathbf{E}}\big(\eta|\mathbf{E}(\eta|\xi)\big) = \frac{\mathbf{E}\eta\mathbf{E}(\eta|\xi)}{\mathbf{E}\big(\mathbf{E}(\eta|\xi)\big)^2}\mathbf{E}(\eta|\xi) = \mathbf{E}(\eta|\xi)$$

(if $\mathbf{E}\big(\mathbf{E}(\eta|\xi)\big)^2 = 0$ the statement holds by definition).

Another argument is that $\eta - \mathbf{E}(\eta|\xi)$ is orthogonal to any (integrable) function of ξ and thus in particular to random variables of the form $\alpha + \beta\mathbf{E}(\eta|\xi)$ for any α, β. The result follows by uniqueness (**P**-a.s) of the orthogonal projection.

(c) in view of **(b)** $\widehat{\mathbf{E}}\big(\eta|\mathbf{E}(\eta|\xi)\big) = \widehat{\mathbf{E}}(\eta|\xi)$ is obviously **FALSE**, whenever $\mathbf{E}(\eta|\xi) \neq \widehat{\mathbf{E}}(\eta|\xi)$.

(d) $\widehat{\mathbf{E}}\big(\eta|\widehat{\mathbf{E}}(\eta|\xi)\big) = \widehat{\mathbf{E}}(\eta|\xi)$ is **TRUE**:

$$\widehat{\mathbf{E}}\big(\eta|\widehat{\mathbf{E}}(\eta|\xi)\big) = \frac{\mathbf{E}\eta\widehat{\mathbf{E}}(\eta|\xi)}{\mathbf{E}\big(\widehat{\mathbf{E}}(\eta|\xi)\big)^2}\widehat{\mathbf{E}}(\eta|\xi) = \frac{\alpha\mathbf{E}\eta\xi}{\alpha^2\mathbf{E}\xi^2}\alpha\xi = \widehat{\mathbf{E}}(\eta|\xi)$$

Alternatively, the argument similar to (**b**) can be used.

(e) the statement is **TRUE**:

$$\widehat{\mathbf{E}}(\eta|\xi) = \xi \quad \Longrightarrow \quad \mathbf{E}(\eta - \xi)\xi = 0$$
$$\widehat{\mathbf{E}}(\xi|\eta) = \eta \quad \Longrightarrow \quad \mathbf{E}(\eta - \xi)\eta = 0$$

By subtracting the two equations one obtains:

$$\mathbf{E}(\eta - \xi)^2 = 0$$

that is $\eta = \xi$, $\mathbf{P} - a.s.$

(f) the statement is **TRUE**:

$$\mathbf{E}\big(\mathbf{E}(\eta|\xi) - \eta\big)^2 = \mathbf{E}\mathbf{E}(\eta|\xi)\big(\mathbf{E}(\eta|\xi) - \eta\big) - \mathbf{E}\eta\big(\mathbf{E}(\eta|\xi) - \eta\big) = 0$$

where the first term vanishes by virtue of orthogonality property of cond. expectation and the second term equals zero, since $\mathbf{E}\big(\mathbf{E}(\eta|\xi)|\eta\big) = \eta$ and $\mathbf{E}\eta^2 < \infty$ implies $\mathbf{E}\eta(\mathbf{E}(\eta|\xi) - \eta) = 0$.

Problem 2.

(a) The solution is quite standard - fix an arbitrary function $h:\{0,1\}\mapsto\mathbb{R}$, then

$$\begin{aligned}&\mathbf{E}\big(I(X_n=i)h(\nu_n)|\nu_1^{n-1}\big)=\mathbf{E}\big(I(X_n=i)\sum_j I(X_{n-1}=j)h(\nu_n)|\nu_1^{n-1}\big)\\&=h(0)\mathbf{E}\big(I(X_n=i)I(X_{n-1}=i)|\nu_1^{n-1}\big)+h(1)\mathbf{E}\big(\sum_{j\neq i}I(X_n=i)I(X_{n-1}=j)|\nu_1^{n-1}\big)\\&=h(0)\lambda_{ii}\pi_{n-1}(i)+h(1)\sum_{j\neq i}\lambda_{ji}\pi_{n-1}(j)\end{aligned}$$

So that if $\pi_n(i)=G(\nu_n;\nu_1^{n-1})$, then

$$\begin{aligned}&\mathbf{E}\big(h(\nu_n)G(\nu_n;\nu_1^{n-1})|\nu_1^{n-1}\big)=...\\&\qquad=h(0)G(0;\nu_1^{n-1})\sum_\ell\lambda_{\ell\ell}\pi_{n-1}(\ell)+h(1)G(1;\nu_1^{n-1})\sum_i\sum_{j\neq i}\lambda_{ji}\pi_{n-1}(j)\end{aligned}$$

Comparing $h(0)$ and $h(1)$ terms in the above equation obtain:

$$\pi_n(i)=\frac{\lambda_{ii}\pi_{n-1}(i)}{\sum_j\lambda_{jj}\pi_{n-1}(j)}I(\nu_n=0)+\frac{\sum_{\ell\neq i}\lambda_{\ell i}\pi_{n-1}(\ell)}{\sum_k\sum_{\ell\neq k}\lambda_{\ell k}\pi_{n-1}(\ell)}I(\nu_n=1)$$

(b) Similarly, fix a pair of bounded functions $h:\{0,1\}\mapsto\mathbb{R}$ and $g:\mathbb{R}\to\mathbb{R}$, then

$$\begin{aligned}&\mathbf{E}\big(I(X_n=i)h(\nu_n)g(Y_n)\big|\nu_1^{n-1},Y_1^{n-1}\big)=\\&=h(0)\mathbf{E}\big(I(X_n=i)I(X_{n-1}=i)g(a_i+\xi_n)\big|\nu_1^{n-1},Y_1^{n-1}\big)+\\&+h(1)\mathbf{E}\big(\sum_{j\neq i}I(X_n=i)I(X_{n-1}=j)g(a_i+\xi_n)\big|\nu_1^{n-1},Y_1^{n-1}\big)=\\&=h(0)\lambda_{ii}\zeta_{n-1}(i)\int g(a_i+x)f(x)dx+h(1)\sum_{j\neq i}\lambda_{ji}\zeta_{n-1}(j)\int g(a_i+x)f(x)dx\end{aligned}$$

from where it is not difficult to guess the correct answer:

$$\zeta_n(i)=\frac{\lambda_{ii}\zeta_{n-1}(i)f(Y_n-a_i)}{\sum_j\lambda_{jj}\zeta_{n-1}(j)f(Y_n-a_j)}I(\nu_n=0)+\frac{f(Y_n-a_i)\sum_{\ell\neq i}\lambda_{\ell i}\zeta_{n-1}(\ell)}{\sum_k f(Y_n-a_k)\sum_{\ell\neq k}\lambda_{\ell k}\zeta_{n-1}(\ell)}I(\nu_n=1)$$

Problem 3.

(a) *(1)* (X_n) converges in all mentioned senses, since $Q_n:=\mathbf{E}X_n^2\to 0$:

$$Q_n=\left(\frac{3}{4}\right)^2Q_{n-1}+\left(\frac{3}{4}\right)^2Q_{n-1}\mathbf{E}\varepsilon_1^2=\left(\frac{9}{16}+\frac{9}{16}\cdot\frac{1}{3}\right)Q_{n-1}$$

(2) (X_n) does not converge in $\mathbb{L}^2$:

$$Q_n=\left(\frac{9}{16}+\frac{9}{16}\cdot 1\right)Q_{n-1}\implies Q_n\nearrow\infty$$

Still we have other types of convergence, since $\mu_n=\mathbf{E}|X_n|$ obeys

$$\mu_n=\mu_{n-1}\mathbf{E}\left|\frac{3}{4}+\frac{3}{4}\varepsilon_1\right|=\mu_{n-1}\left(0+\frac{1}{2}\left|\frac{3}{4}+\frac{3}{4}\right|\right)=\frac{3}{4}\mu_{n-1}$$

and hence $\mathbf{E}|X_n| \to 0$, as $n \to \infty$.

(3) Obviously X_n is a Markov chain with two possible values: ± 1 - moreover , since $\mathbf{P}(\varepsilon_1 = 1) = 1/2$, $(X_n)_{n\ge 1}$ is i.i.d. sequence, and hence it converges in distribution to an equiprobable binary random variable. Clearly it does not converge in any strong sense (e.g. it is not fundamental in Cauchy sense) !

(4) It is easy to see that X_n does not converge (to zero) neither in $\mathbb{L}^2$:

$$Q_n = Q_{n-1}\left(1 + \frac{1}{4}\cdot 1\right) \implies Q_n \nearrow \infty$$

nor in $\mathbb{L}^1$:

$$\mu_n = \mu_{n-1}\left(\frac{1}{2}\cdot\frac{3}{2} + \frac{1}{2}\cdot\frac{1}{2}\right) = \mu_{n-1} \implies \mu_n \equiv 1$$

It is logical to use Chebyshev inequality for lower moments to discover convergence in probability: let ξ be a positive r.v. with $\mathbf{E}\xi^q < \infty$ for some $q > 0$, then with $\varepsilon > 0$

$$\mathbf{E}\xi^q = \mathbf{E}\xi^q I(\xi \ge \varepsilon) + \mathbf{E}\xi^q I(\xi < \varepsilon) \ge \mathbf{E}\xi^q I(\xi \ge \varepsilon) \ge \varepsilon^q \mathbf{P}(\xi \ge \varepsilon) \implies \mathbf{P}(\xi \ge \varepsilon) \le \mathbf{E}\xi^q/\varepsilon^q$$

Try $q = 1/2$:

$$\begin{aligned}\mathbf{E}\sqrt{|X_n|} &= \mathbf{E}\sqrt{|X_{n-1}|}\mathbf{E}\sqrt{|1 + 1/2\varepsilon_1|} = \\ &= \mathbf{E}\sqrt{|X_{n-1}|}\underbrace{\left(\frac{1}{2}\sqrt{\frac{3}{2}} + \frac{1}{2}\sqrt{\frac{1}{2}}\right)}_{\le 0.97} \implies \mathbf{E}|X_n|^{1/2} \to 0\end{aligned}$$

so X_n converges to zero in probability and hence also in distribution.

Note: it is possible to use the *strong law of large numbers* to show the convergence (how ?)

(b) Define $m_n = \mathbf{E}X_n$, $\Delta_n = X_n - m_n$ and $V_n = \mathbf{E}\Delta_n^2$. Then

$$X_n = \alpha X_{n-1} + \beta X_{n-1}\varepsilon_n = \alpha X_{n-1} + \beta m_{n-1}\varepsilon_n + \beta(X_{n-1} - m_{n-1})\varepsilon_n$$

Set $\widetilde{\varepsilon}_n = (X_{n-1} - m_{n-1})\varepsilon_n$, so that the signal/observation model is obtained:

$$\begin{aligned}X_n &= \alpha X_{n-1} + \beta m_{n-1}\varepsilon_n + \beta\widetilde{\varepsilon}_n \\ Y_n &= X_{n-1} + \varepsilon_n\end{aligned}$$

Note that $(\widetilde{\varepsilon}_n)$ is an orthogonal sequence, and it is uncorrelated with (ε_n):

$$\begin{aligned}\mathbf{E}\widetilde{\varepsilon}_n\varepsilon_k &= \mathbf{E}(X_{n-1} - m_{n-1})\varepsilon_k\mathbf{E}\varepsilon_n = 0, \quad k < n \\ \mathbf{E}\widetilde{\varepsilon}_n\varepsilon_n &= \mathbf{E}(X_{n-1} - m_{n-1})\mathbf{E}\varepsilon_n^2 = 0 \\ \mathbf{E}\widetilde{\varepsilon}_n\varepsilon_k &= \mathbf{E}(X_{n-1} - m_{n-1})\varepsilon_n\mathbf{E}\varepsilon_k = 0, \quad k > n\end{aligned}$$

So we can apply the Kalman filter:

$$\begin{aligned}\widehat{X}_n &= \alpha\widehat{X}_{n-1} + \frac{\alpha P_{n-1} + \beta m_{n-1}}{P_{n-1} + 1}\left(Y_n - \widehat{X}_{n-1}\right) \\ P_n &= \alpha^2 P_{n-1} + \beta^2 m_{n-1}^2 + \beta^2 V_{n-1} - \frac{\left(\alpha P_{n-1} + \beta m_{n-1}\right)^2}{P_{n-1} + 1}\end{aligned} \tag{1.1}$$

where m_n and $Q_n = V_n + m_n^2 = \mathbf{E}X_n^2$ are generated by

$$\begin{aligned}m_n &= \alpha m_{n-1}, \quad m_0 = 1 \\ Q_n &= (\alpha^2 + \beta^2)Q_{n-1}, \quad Q_0 = 1\end{aligned}$$

(c) For the special case under consideration X_n can be precisely estimated out of Y_1^n, e.g.

$$\mathbf{E}(X_n|Y_1^n) = Y_n^2/2 - 1 = (1 + 2X_{n-1}\varepsilon_n + 1)/2 - 1 = X_{n-1}\varepsilon_n = X_n$$

It is not difficult (e.g. from the previous question), however, to verify that the optimal linear filter gives a trivial estimate $\widehat{X}_n \equiv 0, \quad n \geq 1$