RANDOM PROCESSES - SOLUTION OF THE FINAL EXAM 2001

PAVEL CHIGANSKY

Problem 1.

a. $\mathbf{E}\big(\widehat{\mathbf{E}}(\xi|\eta)\big|\eta\big) = \mathbf{E}(\xi|\eta)$ is generally **FALSE**, since $\mathbf{E}\big(\widehat{\mathbf{E}}(\xi|\eta)\big|\eta\big) = \widehat{\mathbf{E}}(\xi|\eta)$, $\mathbf{P}$-a.s. (see **b.**)

b. $\mathbf{E}\big(\widehat{\mathbf{E}}(\xi|\eta)\big|\eta\big) = \widehat{\mathbf{E}}(\xi|\eta)$ is **TRUE**, since $\widehat{\mathbf{E}}(\xi|\eta)$ is a (linear) function of η.

c. $\widehat{\mathbf{E}}\big(\mathbf{E}(\xi|\eta)\big|\eta\big) = \mathbf{E}(\xi|\eta)$ is generally **FALSE**, since the left side expression is a linear function of η, and the conditional expectation on the right side is nonlinear function of η generally.

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

$$\widehat{\mathbf{E}}\big(\mathbf{E}(\xi|\eta)\big|\eta\big) = \mathbf{E}\mathbf{E}(\xi|\eta) + \frac{\mathbf{E}\mathbf{E}(\xi|\eta)\eta}{\mathbf{E}\eta^2}\big(\eta - \mathbf{E}\eta\big) \overset{\dagger}{=} \mathbf{E}\xi + \frac{\mathbf{E}\xi\eta}{\mathbf{E}\eta^2}(\eta - \mathbf{E}\eta) = \widehat{\mathbf{E}}(\xi|\eta) \tag{1.1}$$

where † follows from the definition of cond. exp.

Problem 2. To show convergence in $\mathbb{L}^2$ sense (and hence also in prob. and in law) it suffices to verify the Cauchy property:

$$\mathbf{E}\big(I(X_n = i) - I(X_m = i)\big)^2 = p_n(i) + p_m(i) - 2\mathbf{P}(X_n = i|X_m = i)p_m(i) \xrightarrow{n,m\to\infty} 0$$

where $p_n(i) = \mathbf{P}(X_n = i)$ for convenience.

Calculate the probabilities $p_n(i)$, $i = -1, 0, 1$:

$$p_n(0) = \mathbf{P}(X_n = 0) = \mathbf{P}(X_n = 0|X_0 = 0)\beta = (1/4)^n\beta, \quad n \ge 0$$

$$p_n(-1) = \mathbf{P}(X_n = -1) = 1 \cdot \mathbf{P}(X_{n-1} = -1) + 1/4\mathbf{P}(X_{n-1} = 0)$$
$$= p_{n-1}(-1) + 1/4p_{n-1}(0)$$

so

$$\begin{aligned} p_n(-1) &= \alpha + 1/4\beta\sum_{i=0}^{n-1}(1/4)^i = \alpha + 1/4\beta\frac{1-(1/4)^n}{1-1/4} = \\ &= \alpha + 1/3\beta - 1/3\beta(1/4)^n \end{aligned}$$

and similarly

$$\begin{aligned} p_n(1) &= p_{n-1}(1) + 1/2p_{n-1}(0) = \gamma + 1/2\beta\sum_{i=0}^{n-1}(1/4)^i = \\ &= \gamma + 1/2\beta\frac{1-(1/4)^n}{1-1/4} = \gamma + 2/3\beta - 2/3\beta(1/4)^n \end{aligned}$$

Now (say for $n \ge m$)

$$\begin{aligned} & p_n(0)+p_m(0)-2\mathbf{P}(X_n=0|X_m=0)p_m(0)= \\ &=\beta(1/4)^m+\beta(1/4)^n-2(1/4)^{n-m}\beta(1/4)^m= \\ &=\beta(1/4)^m-\beta(1/4)^n \xrightarrow{n,m\to\infty} 0 \end{aligned}$$

$$\begin{aligned} & p_n(-1)+p_m(-1)-2\mathbf{P}(X_n=-1|X_m=-1)p_m(-1)=p_n(-1)-p_m(-1)= \\ &=1/3\beta(1/4)^n-1/3\beta(1/4)^m \xrightarrow{n,m\to\infty} 0 \end{aligned}$$

and

$$\begin{aligned} & p_n(1)+p_m(1)-2\mathbf{P}(X_n=1|X_m=1)p_m(1)=p_n(1)-p_m(1)= \\ &=2/3\beta(1/4)^n-2/3\beta(1/4)^m \xrightarrow{n,m\to\infty} 0 \end{aligned}$$

which means that X_n is a $\mathbb{L}^2$ Cauchy sequence and thus converges to a limit, which is a random variable, say X.

The sequence converges also with probability one. To show this it suffices (why ?) to verify that

$$\mathbf{E}\|I_n-I\|^q \le C\rho^n$$

for some C, $q>0$ and $0<\rho<1$, where

$$I_n=\begin{bmatrix} I(X_n=-1) \\ I(X_n=0) \\ I(X_n=1) \end{bmatrix}$$

and I is its limit. Since I_n converges in $\mathbb{L}^2$,

$$I=I_0+\sum_{m=1}^{\infty}(I_m-I_{m-1})$$

so that we have to verify (e.g. for $q=1$)

$$\sum_{m=n+1}^{\infty}\sqrt{\mathbf{E}\big(I(X_m=i)-I(X_{m-1}=i)\big)^2} \le C(i)\rho(i)^n$$

for $i=-1,0,1$. Obviously

$$\begin{aligned} \mathbf{E}\big(I(X_m=0)-I(X_{m-1}=0)\big)^2 &= \beta(1/4)^{m-1}(1-1/4) \\ \mathbf{E}\big(I(X_m=1)-I(X_{m-1}=1)\big)^2 &= \beta 1/3(1/4)^{m-1}(1-1/4) \\ \mathbf{E}\big(I(X_m=-1)-I(X_{m-1}=-1)\big)^2 &= \beta 2/3(1/4)^{m-1}(1-1/4) \end{aligned}$$

so that e.g.

$$\begin{aligned} \sum_{m=n+1}^{\infty}\sqrt{\mathbf{E}\big(I(X_m=0)-I(X_{m-1}=0)\big)^2} &\le \text{const.}\sum_{m=n+1}^{\infty}(1/2)^{m-1} \\ &\le \text{const.}(1/2)^n \end{aligned}$$

b.

$$F_n(x) := \mathbf{P}(X_n\le x)=\begin{cases} 0 & x\in(-\infty,-1) \\ \alpha+1/3\beta-1/3\beta(1/4)^n & x\in[-1,0) \\ \alpha+1/3\beta+2/3\beta(1/4)^n & x\in[0,1) \\ 1 & x\in[1,\infty) \end{cases}$$

Clearly $\lim_{n\to\infty} \sup_{x\in\mathbb{R}} |F_n(x) - F(x)| = 0$, where

$$F(x) \quad := \quad \begin{cases} 0 & x \in (-\infty, -1) \\ \alpha + 1/3\beta & x \in [-1, 1) \\ 1 & x \in [1, \infty) \end{cases}$$

which means that $X = \lim_{n\to\infty} X_n$ is a random variable with values $\{-1, 1\}$ and $\mathbf{P}(X = -1) = \alpha + 1/3\beta$ and $\mathbf{P}(X = 1) = \gamma + 2/3\beta$.

c. Clearly X is deterministic only if $\alpha = 1$ or $\gamma = 1$.

Problem 3.

a. A standard derivation of the optimal filter: put $\pi_{n|n-1}(i) = \mathbf{P}(X_n = a_i|Y_0^{n-1})$ and $\pi_n(i) = \mathbf{P}(X_n = a_i|Y_0^n) := G(Y_n; Y_0^{n-1})$. Fix an arbitrary function $h(x) : \mathbb{R} \to \mathbb{R}$. The cond. exp. $G(Y_n; Y_0^{n-1})$ should satisfy $\mathbf{P}$-a.s.

$$\mathbf{E}\big(I(X_n = a_i)h(Y_n)|Y_0^{n-1}\big) = \mathbf{E}\big(h(Y_n)G(Y_n; Y_0^{n-1})|Y_0^{n-1}\big)$$

The left hand side gives:

$$\begin{aligned}
&\mathbf{E}\big(I(X_n = a_i)h(Y_n)|Y_0^{n-1}\big) = \\
&= \mathbf{E}\big(I(X_n = a_i)\big[I(a_i \in \mathcal{J})h(1) + I(a_i \notin \mathcal{J})h(0)\big]|Y_0^{n-1}\big) = \\
&= \pi_{n|n-1}(i)\big[I(a_i \in \mathcal{J})h(1) + I(a_i \notin \mathcal{J})h(0)\big]
\end{aligned}$$

Similarly the right hand side gives:

$$\begin{aligned}
&\mathbf{E}\big(h(Y_n)G(Y_n; Y_0^{n-1})|Y_0^{n-1}\big) = \\
&= \mathbf{E}\big[I(X_n \in \mathcal{J})h(1)G(1; Y_0^{n-1}) + I(X_n \notin \mathcal{J})h(0)G(1; Y_0^{n-1})|Y_0^{n-1}\big] = \\
&= \mathbf{P}(X_n \in \mathcal{J}|Y_0^{n-1})h(1)G(1; Y_0^{n-1}) + \mathbf{P}(X_n \notin \mathcal{J}|Y_0^{n-1})h(0)G(0; Y_0^{n-1}) = \\
&= h(1)G(1; Y_0^{n-1}) \sum_{i\in\mathcal{J}} \pi_{n|n-1}(i) + h(0)G(0; Y_0^{n-1}) \sum_{i\notin\mathcal{J}} \pi_{n|n-1}(i)
\end{aligned}$$

comparing the above expressions we arrive at

$$\pi_n(i) = \frac{\pi_{n|n-1}(i)I(a_i \in \mathcal{J})}{\sum_{i\in\mathcal{J}} \pi_{n|n-1}(i)} I(Y_n = 1) + \frac{\pi_{n|n-1}(i)I(a_i \notin \mathcal{J})}{\sum_{i\notin\mathcal{J}} \pi_{n|n-1}(i)} I(Y_n = 0)$$

so **(ii)** is correct.

Remark: the correct answer can be found also by excluding answers, which do not satisfy obvious requirements, e.g. $\sum_i \pi_n(i) \equiv 1$, or $\pi_n(i) \equiv 0$ if $Y_n = 1$ and $i \notin \mathcal{J}$, etc.

b. Use familiar state-space representation for Markov chains:

$$I_n = \Lambda^* I_{n-1} + \varepsilon_n$$

where ε_n is a sequence of zero mean vector random variables such that

$$\mathbf{E}\varepsilon_n\varepsilon_m^* = 0, \quad n \neq m$$

and

$$\mathbf{E}\varepsilon_n\varepsilon_n^* = \mathsf{diag}(p_n) - \Lambda\mathsf{diag}(p_{n-1})\Lambda^* := D_n$$

where $p_n = \mathbf{E}I_n$. Note also that $Y_n = u^*I_n = u^*\Lambda^*I_{n-1} + u^*\varepsilon_n$, where u is a column vector with ones at indices corresponding to $\mathcal{J}$ and zeros otherwise. So the Kalman filter recursion is

$$\begin{aligned}
\widehat{\pi}_n &= \Lambda^*\widehat{\pi}_{n-1} + \big(\Lambda^*P_{n-1}\Lambda + D_n\big)u\big(u^*\Lambda^*P_{n-1}\Lambda u + u^*D_nu\big)^+\big(Y_n - u^*\Lambda^*\widehat{\pi}_{n-1}\big)\\
P_n &= \Lambda^*P_{n-1}\Lambda + D_n -\\
&\quad -\big(\Lambda^*P_{n-1}\Lambda + D_n\big)u(u^*\Lambda^*P_{n-1}\Lambda u + u^*D_nu\big)^+u^*\big(\Lambda^*P_{n-1}\Lambda + D_n\big)
\end{aligned}$$

subject to $\widehat{\pi}_0 = p_0$ and $P_0 = \mathsf{diag}(p_0) - p_0p_0^*$.

Problem 4.

a. The pair (θ, Y_n) is Gaussian and obeys the model $(\theta_n \equiv \theta)$

$$\begin{aligned}
\theta_n &= \theta_{n-1}\\
Y_n &= \theta_{n-1} + \xi_n, \quad n \ge 1
\end{aligned}$$

subject to $\theta_0 = \theta$. The optimal estimate is given by Kalman filter

$$\begin{aligned}
m_n &= m_{n-1} + \frac{P^m_{n-1}}{P^m_{n-1}+1}\big(Y_n - m_{n-1}\big)\\
P^m_n &= P^m_{n-1} - \frac{(P^m_{n-1})^2}{P^m_{n-1}+1}
\end{aligned}$$

or

$$\begin{aligned}
m_n &= m_{n-1} + P^m_n\big(Y_n - m_{n-1}\big)\\
P^m_n &= \frac{P^m_{n-1}}{P^m_{n-1}+1}
\end{aligned}$$

subject to $m_0 = 0$ and $P^m_0 = 1$.

b. From the mouse point of view the signal (cat's position) is m_n and the observation is θ, i.e. it sees the following model

$$\begin{aligned}
m_n &= (1 - P^m_n)m_{n-1} + P^m_n\big(\theta_{n-1} + \xi_n\big)\\
\theta_n &= \theta_{n-1}
\end{aligned}$$

Let $c_n = \mathbf{E}(m_n|\theta) \equiv \mathbf{E}(m_n|\theta^n_0)$ and $P^c_n = \mathbf{E}(m_n - c_n)^2$. The pair (θ, m_n) is Gaussian so the optimal estimate is given by Kalman filter:

$$\begin{aligned}
c_n &= (1 - P^m_n)c_{n-1} + P^m_n\theta_{n-1} \equiv (1 - P^m_n)c_{n-1} + P^m_n\theta\\
P^c_n &= (1 - P^m_n)^2P^c_{n-1} + (P^m_n)^2
\end{aligned}$$

subject to $c_0 = 0$ and $P^c_0 = 0$ (why?)

c. Consider a simple average estimate of $\breve{m}_n = \frac{1}{n}\sum_{k=1}^n Y_n$. Clearly $\mathbf{E}(\theta - m_n)^2 \le \mathbf{E}(\theta - \breve{m}_n)^2 \xrightarrow{n\to\infty} 0$, so $\lim_{n\to\infty} P^m_n = 0$. Now consider a simple constant estimate $\breve{c}_n \equiv \theta$. Clearly

$$P^c_n = \mathbf{E}(c_n - m_n)^2 \le \mathbf{E}(\breve{c}_n - m_n)^2 = \mathbf{E}(\theta - m_n)^2 = P^m_n \xrightarrow{n\to\infty} 0 \tag{1.2}$$

d. The correct answer is $P^c_n \le P^m_n$ as follows from (1.2).

e. The relation of **(d)** holds also generally by the very same argument as in (1.2): let $(\theta_n, Y_n)_{n\geq 0}$ be a pair of random sequences, then

$$P_n^c = \mathbf{E}\big[\mathbf{E}(\theta_n|Y_0^n) - \mathbf{E}\big(\mathbf{E}(\theta_n|Y_0^n)|\theta_0^n\big)\big]^2 \leq \mathbf{E}\big[\mathbf{E}(\theta_n|Y_0^n) - \theta_n\big]^2 = P_n^m$$