SOLUTION TO THE EXAM 2004 (SUMMER COURSE)

P. CHIGANSKY

Problem 1.

(a) The optimal receiver is given by the Kalman formulae

$$\widehat{X}_n = \widehat{X}_{n-1} + \frac{b_n P_{n-1}}{1+b_n^2 P_{n-1}}\big(Y_n - a_n - b_n\widehat{X}_{n-1}\big) \tag{1}$$

$$P_n = P_{n-1} - \frac{b_n^2 P_{n-1}^2}{1+b_n^2 P_{n-1}} \tag{2}$$

subject to $\widehat{X}_0 = 0$ and $P_0 = 1$.

(b) Note that

$$P_n = \frac{P_{n-1}}{1+b_n^2 P_{n-1}}$$

and

$$Q_n := P_n^{-1} = Q_{n-1} + b_n^2 = 1 + \sum_{m=1}^n b_m^2$$

so that

$$P_n = \frac{1}{1+\sum_{m=1}^n b_m^2}.$$

Since $\gamma_n = E(a_n + b_n X)^2 = a_n^2 + b_n^2 \le \gamma$ the choice $a_n = 0$ and $b_n = \sqrt{\gamma}$ minimizes P_n for any $n \ge 1$ and keeps the power within the required limit $\gamma_n \le \gamma$. The minimal error is then

$$P_n = 1/(1+n\gamma).$$

(c) If ξ_1 were not Gaussian, smaller estimation error may be attained for any chosen transmitter (why?) - in particular for the transmitter from **(b)**. Thus the error may only decrease.

(d) The generalized Kalman filter implements the optimal receiver: the optimal filter and the conditional mean square error are given by the same equations (1)-(2) with a_n depending on the past of Y.

(e) Y is not necessarily Gaussian, since a_n is allowed to depend nonlinearly on Y - e.g. if $a_2 = \mathrm{sign}(Y_1)$, Y_2 is non Gaussian (why?)

(f) Note that

$$\gamma_n = E\big(a_n(Y_1^{n-1}) + b_n X\big)^2 = E\big(a_n(Y_1^{n-1}) + b_n\widehat{X}_{n-1} + b_n(X - \widehat{X}_{n-1})\big)^2 = E\big(a_n(Y_1^{n-1}) + b_n\widehat{X}_{n-1}\big)^2 + b_n^2 P_{n-1} \le \gamma. \tag{3}$$

On the other hand, the equation of P_n depends only on b_n:

$$P_n = P_{n-1}\frac{1}{b_n^2 P_{n-1}+1}, \quad P_0 = 1.$$

By (3) $b_n^2 P_{n-1} \le \gamma$ and so the latter implies

$$P_n = \prod_{k=1}^n \frac{1}{b_k^2 P_{k-1}+1} \ge \prod_{k=1}^n \frac{1}{\gamma+1} = \left(\frac{1}{\gamma+1}\right)^n.$$

This bound is attained if $b_n^2 P_{n-1} = \gamma$, which requires that $a_n(Y_1^{n-1}) = -\widehat{X}_{n-1}$ is chosen. Then

$$b_n = \sqrt{\gamma/P_{n-1}} = \sqrt{\gamma(1+\gamma)^{n-1}}$$

(g) As before

$$P_n = P_{n-1}\frac{1}{1+b_n^2(Y_1^{n-1})P_{n-1}}$$

and so

$$P_n = \prod_{k=1}^n \frac{1}{1+b_k^2(Y_1^{k-1})P_{k-1}} = \exp\left\{\sum_{k=1}^n - \log\left(1+b_k^2(Y_1^{k-1})P_{k-1}\right)\right\}$$

Both $\exp(\cdot)$ and $-\log(\cdot)$ are convex functions and so

$$EP_n \ge \exp\left\{\sum_{k=1}^n - \log\left(1+Eb_k^2(Y_1^{k-1})P_{k-1}\right)\right\}$$

which in turn implies

$$EP_n \ge \prod_{k=1}^n \frac{1}{1+Eb_k^2(Y_1^{k-1})P_{k-1}}. \tag{4}$$

Now the power constraint gives

$$\begin{aligned}\gamma_n =& E\big(a_n(Y_1^{n-1}) + b_n(Y_1^{n-1})X\big)^2 = \\ & E\big((a_n(Y_1^{n-1}) + b_n(Y_1^{n-1})\widehat{X}_{n-1}) + b_n(Y_1^{n-1})(X-\widehat{X}_{n-1})\big)^2 = \\ & E\big(a_n(Y_1^{n-1}) + b_n(Y_1^{n-1})\widehat{X}_{n-1}\big)^2 + Eb_n^2(Y_1^{n-1})(X-\widehat{X}_{n-1})^2 = \\ & E\big(a_n(Y_1^{n-1}) + b_n(Y_1^{n-1})\widehat{X}_{n-1}\big)^2 + Eb_n^2(Y_1^{n-1})P_{n-1} \le \gamma\end{aligned}$$

and thus the lower bound in (4) implies

$$EP_n \ge \prod_{k=1}^n \frac{1}{1+\gamma}.$$

The power constraint is clearly satisfied if $b_n^2(Y_1^{n-1})P_{n-1} = \gamma$ and $a_n(Y_1^{n-1}) = -b_n\widehat{X}_{n-1}$ are set. Moreover in this case

$$P_n = \prod_{k=1}^n \frac{1}{1+\gamma} = \left(\frac{1}{1+\gamma}\right)^n$$

so that the lower bound for EP_n is attained. So the optimal transmitter is given by

$$\sqrt{\frac{\gamma}{P_{n-1}}}\big(Y_n - \widehat{X}_{n-1}\big) = \sqrt{\gamma(1+\gamma)^{n-1}}\big(Y_n - \widehat{X}_{n-1}\big).$$

Surprisingly no further improvement is gained by letting b_n depend on $\{Y_1, ..., Y_{n-1}\}$

Problem 2.

(a) Since $EX_i^4 < \infty$, the strong (e.g. Cantelli) law of large numbers implies

$$S_n = \frac{1}{n}\sum_{i=1}^n X_i^2 \xrightarrow[P-a.s.]{n\to\infty} EX_1^2.$$

The function $1/\sqrt{x}$ is continuous at $x = 1$ and so $1/\sqrt{S_n}$ converges P-a.s. and thus also in probability to 1. Then

$$\sqrt{n}Y^n(1) = \frac{\sqrt{n}X_1}{\sqrt{\sum_{i=1}^n X_i}} = \frac{X_1}{\sqrt{S_n}} \xrightarrow[P-a.s.]{n\to\infty} X_1$$

and also in probability (see problem 1.8 in the exercises file #1). Hence $Y^n(1)$ converges weakly as well. Since $Y^n(1)$ is bounded it also converges in $\mathbb{L}^2$.

(b) Suppose X_1 is a standard Gaussian r.v. (without loss of generality $EX_1^2 = 1$ is assumed) and let U be an orthogonal matrix. Denote by X^n the vector with entries $X_1, ..., X_n$, so that $Y^n = X^n/\|X^n\|$. Then the rotated vector satisfies

$$\widetilde{Y}^n := UY^n = UX^n/\|X^n\| = UX^n/\|UX^n\|$$

where the latter is due to $UU^* = I$. The vector UX^n is Gaussian with zero mean and unit covariance (why?), so it is distributed exactly as X^n. So $\widetilde{Y}^n$ is distributed as Y^n for any rotation U.

Clearly atomic X_1 cannot lead to the uniform distribution, since rotations would translate the atoms all over the sphere.

(c) Since the uniform distribution is unique and can be realized by Gaussian X_i's, $Y^n(1)$ should converge to a Gaussian r.v. weakly (in distribution). Obviously it may not converge in probability - imagine that for each n independent X_i's are used!

Problem 3.

(a) Since

$$P(Z_n = 1) = P\big(X_n = 1, ..., X_0 = 1\big) = \\ P(X_n = 1|X_{n=1} = 1)...P(X_1 = 1|X_0 = 1)p(1) = \big(\lambda_{11}\big)^n p(1) \xrightarrow{n\to\infty} 0,$$

Z_n converges to zero in probability and hence in law. Since $|Z_n| \le 1$, it also converges in $\mathbb{L}^2$. Moreover since $\sum_{n=0}^\infty P(Z_n = 1) < \infty$, $P(Z_n = 1, i.o.) = 0$ and so Z_n converges to zero P-a.s. by the Borel-Cantelli lemma.

(b) Clearly

$$EP(Z_n = 1|Y_1^n) = P(Z_n = 1)$$

and thus $P(Z_n = 1|Y_1^n)$ also converges to zero in prob., weakly and $\mathbb{L}^2$.

(c) Z is a Markov process:

$$P(Z_n=1|Z_0^{n-1})=E(Z_n|Z_0^{n-1})=Z_{n-1}E(X_n|Z_0^{n-1})=Z_{n-1}E(X_n|Z_{n-1}=1)=$$
$$Z_{n-1}E(X_n|X_0=1,...,X_{n-1}=1)=Z_{n-1}P(X_n=1|X_{n-1}=1)=Z_{n-1}\lambda_1.$$

The transition probabilities are

$$P(Z_n=0|Z_{n-1}=0)=1,\quad P(Z_n=1|Z_{n-1}=1)=\lambda_1.$$

(d) Apply the formulae, developed in class

$$\pi_n=\left(1+\frac{f(Y_n)\big((1-\lambda_1)\pi_{n-1}+\lambda_0(1-\pi_{n-1})\big)}{f(Y_n-1)\big(\lambda_1\pi_{n-1}+(1-\lambda_0)(1-\pi_{n-1})\big)}\right)^{-1}.$$

subject to $\pi_0=p(1)=1/2$.

(e) Let $r_n:=G(Y_n;Y_1^{n-1})$, then

$$E\Big(\big(Z_n-G(Y_n;Y_1^{n-1})\big)\varphi(Y_n)\big|Y_1^{n-1}\Big)=0$$

for any bounded $\varphi(x)$. The left hand side gives

$$E\big(Z_n\varphi(Y_n)|Y_1^{n-1}\big)=E\big(Z_n\varphi(X_n+\varepsilon_n)|Y_1^{n-1}\big)=...=\widehat{Z}_{n|n-1}\int_{\mathbb{R}}\varphi(u)f(u-1)du$$

and so[1]

$$\widehat{Z}_n=\frac{\widehat{Z}_{n|n-1}f(Y_n-1)}{f(Y_n-1)\big(\lambda_1\pi_{n-1}+(1-\lambda_0)(1-\pi_{n-1})\big)+f(Y_n)\big((1-\lambda_1)\pi_{n-1}+\lambda_0(1-\pi_{n-1})\big)}$$

as usual. Further

$$\begin{aligned}\widehat{Z}_{n|n-1}=&E\big(Z_n|Y_1^{n-1}\big)=E\Big(Z_{n-1}E(X_n|X_{n-1})\big|Y_1^{n-1}\Big)=\\&E\Big(Z_{n-1}\big[\lambda_1X_{n-1}+(1-\lambda_0)(1-X_{n-1})\big]\big|Y_1^{n-1}\Big)=\\&\lambda_1\widehat{Z}_{n-1}+(1-\lambda_0)(1-\widehat{Z}_{n-1})\end{aligned}$$

where the latter follows since $Z_{n-1}X_{n-1}=Z_{n-1}$. So

$$\widehat{Z}_n=\frac{f(Y_n-1)\big(\lambda_1\widehat{Z}_{n-1}+(1-\lambda_0)(1-\widehat{Z}_{n-1})\big)}{f(Y_n-1)\big(\lambda_1\pi_{n-1}+(1-\lambda_0)(1-\pi_{n-1})\big)+f(Y_n)\big((1-\lambda_1)\pi_{n-1}+\lambda_0(1-\pi_{n-1})\big)}$$

or

$$\widehat{Z}_n=\frac{\lambda_1\widehat{Z}_{n-1}+(1-\lambda_0)(1-\widehat{Z}_{n-1})}{\lambda_1\pi_{n-1}+(1-\lambda_0)(1-\pi_{n-1})}\pi_n.$$

[1]the right hand side is calculated as in the case of π_n

Problem 4.

(a) Clearly

$$B_1 = \int_0^1 dB_s$$

(b) By the Itô formula $B_t^2 = \int_0^t 2B_u dB_u + t$ and so

$$B_1 = 1 + \int_0^1 2B_u dB_u$$

(c) Applying the Itô formula to $B_t t$ one gets the integration by parts rule

$$d(B_t t) = B_t dt + t dB_t \quad \Longrightarrow \quad B_1 = \int_0^1 B_t dt + \int_0^1 t dB_t.$$

So

$$\int_0^1 B_t dt = B_1 - \int_0^1 t B_t = \int_0^1 (1-t) B_t. \tag{5}$$

(d) Apply the Itô formula to B_t^3

$$B_1^3 = 3\int_0^1 B_t^2 dB_t + \frac{1}{2}\int_0^1 6B_t dt$$

Combining this with (5) one gets

$$B_1^3 = \int_0^1 3\big(B_t^2 + 1 - t\big) dB_t$$

(e) Apply the Itô formula to $e^{B_t - t/2}$:

$$d\big(e^{B_t - t/2}\big) = -\frac{1}{2} e^{B_t - t/2} dt + e^{B_t - t/2} dB_t + \frac{1}{2} e^{B_t - t/2} dt = e^{B_t - t/2} dB_t$$

which implies

$$e^{B_1 - 1/2} - 1 = \int_0^1 e^{B_t - t/2} dB_t$$

or

$$e^{B_1} = e^{1/2} + \int_0^1 e^{(B_t - t/2 + 1/2)} dB_t.$$

(f) Apply the Itô formula to $e^{t/2} \sin B_t$:

$$d\big(e^{t/2} \sin B_t\big) = \frac{1}{2} e^{t/2} \sin B_t dt + e^{t/2} d \sin B_t =$$

$$\frac{1}{2} e^{t/2} \sin B_t dt + e^{t/2}\big(\cos B_t dB_t - \frac{1}{2} \sin B_t dt\big) = e^{t/2} \cos B_t dB_t$$

and so

$$e^{1/2} \sin B_1 = \int_0^1 e^{t/2} \cos B_t dB_t \quad \Longrightarrow \quad \sin B_1 = \int_0^1 e^{(t-1)/2} \cos B_t dB_t$$