SOLUTION TO THE FINAL TEST
JULY 24, 1997

Problem 1

Non linear filter.
(a) $\pi_{0|0} = \mathbf{E}(X_0|Y_0) \equiv 1$

(b) Put $G_n(Y_0^{n-1}, Y_n) = \pi_{n|n} = \Pr\{X_n = 1|Y_0^n\} \equiv \mathbf{E}(X_n|Y_0^n)$. Then following Lecture Note # 9:

$$\mathbf{E}(g(Y_n)X_n|Y_0^{n-1}) = \mathbf{E}(g(Y_n)G_n(Y_0^{n-1}, Y_n)|Y_0^{n-1})$$

for any bounded function $g(x)$. Further:

$$\begin{aligned}
\mathbf{E}(g(Y_n)X_n|Y_0^{n-1}) &= \mathbf{E}(g(X_n + \xi_n)X_n|Y_0^{n-1}) = \\
&= \mathbf{E}\Big[\mathbf{E}\Big(X_n g(X_n + \xi_n)|Y_0^{n-1}, X_n\Big)|Y_0^{n-1}\Big] = \\
&= \mathbf{E}\left(X_n \int g(y) f_\xi(y - X_n) dy |Y_0^{n-1}\right) = \\
&= \pi_{n|n-1} \int g(y) f_\xi(y-1) dy
\end{aligned}$$

Similarly,

$$\begin{aligned}
\mathbf{E}(g(Y_n)G_n(Y_0^{n-1}, Y_n)) &= \pi_{n|n-1} \int g(y) G_n(Y_0^{n-1}, y) f_\xi(y-1) dy + \\
&\quad (1 - \pi_{n|n-1}) \int g(y) G_n(Y_0^{n-1}, y) f_\xi(y) dy
\end{aligned}$$

So that due to arbitrariness of $g(y)$:

$$\pi_{n|n} = G(Y_0^{n-1}, Y_n) = \frac{\pi_{n|n-1} f_\xi(Y_n - 1)}{\pi_{n|n-1} f_\xi(Y_n - 1) + (1 - \pi_{n|n-1}) f_\xi(Y_n)} \tag{1.1}$$

(c) Note that $\pi_{n|n-1} = \mathbf{E}(X_n|Y_0^{n-1})$ and hence:

$$\pi_{n|n-1} = \mathbf{E}(X_n|Y_0^{n-1}) = \mathbf{E}(X_{n-1}\varepsilon_n|Y_0^{n-1}) = \pi_{n-1|n-1}\mathbf{E}\varepsilon_n = \pi_{n-1|n-1} p_n \tag{1.2}$$

So the optimal filter is completely specified by (1.1) and (1.2) subject to initial condition $\pi_{0|0} = 1$.

(d) *(bonus+10)*

$$\tau \equiv \sum_{k=1}^{\tau} 1 = \sum_{k=1}^{\infty} I(\tau \geq k)$$

where $I(\cdot)$ is an indicator function. Take conditional expectation of both sides:

$$\mathbf{E}(\tau|Y_0^n) = \sum_{k=1}^{\infty} \Pr(\tau \ge k|Y_0^n) = \sum_{k=1}^{\infty} \Pr(X_k = 1|Y_0^n)$$

Kalman filter

(a)

$$X_n = X_{n-1}\varepsilon_n = p_n X_{n-1} + (\varepsilon_n - p_n)X_{n-1}$$

Define $\widetilde{\varepsilon}_n = (\varepsilon_n - p_n)X_{n-1}$ Then:

$$\begin{aligned}
\mathbf{E}\widetilde{\varepsilon}_n X_m &= \mathbf{E}(\varepsilon_n - p_n)\mathbf{E}X_{n-1}X_m = 0 \quad \text{for } m \le n-1 \\
\mathbf{E}\widetilde{\varepsilon}_n &= \mathbf{E}(\varepsilon_n - p_n)\mathbf{E}X_{n-1} \equiv 0 \\
\mathbf{E}\widetilde{\varepsilon}_n^2 &= \mathbf{E}(\varepsilon_n - p_n)^2\mathbf{E}X_{n-1}^2 = p_n(1-p_n)\prod_{k=1}^{n-1} p_k = (1-p_n)\prod_{k=1}^{n} p_k
\end{aligned}$$

where the last implication is due to the fact, that:

$$\mathbf{E}X_n^2 = \mathbf{E}X_{n-1}^2\mathbf{E}\varepsilon_n^2 = p_n\mathbf{E}X_{n-1}^2 \quad \Longrightarrow \quad \mathbf{E}X_n^2 = \prod_{k=1}^{n} p_k$$

So this model suits the Kalman filter setting:

$$\begin{aligned}
X_n &= p_n X_{n-1} + \widetilde{\varepsilon}_n \\
Y_n &= X_n + \xi_n
\end{aligned}$$

where $(\widetilde{\varepsilon}_n)_{n\ge1}$ is a sequence of independent zero mean random variable with $\mathbf{E}\widetilde{\varepsilon}_n^2 = P_n = (1 - p_n)\prod_{k=1}^{n} p_k$ and $(\xi_n)_{n\ge1}$ is a zero mean i.i.d. sequence with $\mathbf{E}\xi_1^2 = \int x^2 f_\xi(x)dx \stackrel{\triangle}{=} \sigma_\xi^2$.

(b) The Kalman filter is given by:

$$\begin{aligned}
\widehat{X}_n &= p_n\widehat{X}_{n-1} + \frac{p_n^2 V_{n-1} + P_n}{p_n^2 V_{n-1} + P_n + \sigma_\xi^2}(Y_n - p_n\widehat{X}_{n-1}) \\
V_n &= p_n^2 V_{n-1} + P_n - \frac{(p_n^2 V_{n-1} + P_n)^2}{p_n^2 V_{n-1} + P_n + \sigma_\xi^2}
\end{aligned}$$

where $V_n = \mathbf{E}(X_n - \widehat{X}_n)^2$.

Degenerate case

Note that if $p_m = 0$ for some m, then $X_n \equiv 0$ for $n \ge m$ and thus $\pi_{n|n} = \mathbf{E}(X_n = 1|Y_0^n) \equiv 0$ for $n \ge m$. It is easy to see that Kalman filter also gives $\widehat{X}_n \equiv 0$ for $n \ge m$ in this case, since $(P_n)_{n\ge m} \equiv 0$.

Problem 2

1. Show that $\Delta_n = \mathbf{E}(\theta - \widehat{\theta})^2$ is an non increasing sequence:

$$\begin{aligned}\Delta_n &= \mathbf{E}\Big(\theta - \mathbf{E}(\theta|X_0^n)\Big)^2 \stackrel{\dagger}{=} \mathbf{E}\Big(\theta^2 - [\mathbf{E}(\theta|X_0^n)]^2\Big) \stackrel{\dagger\dagger}{=} \\ &= \mathbf{E}\Big(\theta^2 - \big[\mathbf{E}\big(\mathbf{E}(\theta|X_0^{n+1})|X_0^n\big)\big]^2\Big) \geq \mathbf{E}\Big(\theta^2 - \mathbf{E}\Big(\big[\mathbf{E}(\theta|X_0^{n+1})\big]^2 |X_0^n\Big)\Big) \stackrel{\dagger\dagger}{=} \\ &= \mathbf{E}\theta^2 - \mathbf{E}\left[\mathbf{E}(\theta|X_0^{n+1})\right]^2 \stackrel{\dagger}{=} \Delta_{n+1}\end{aligned}$$

where Jensen inequality for conditional expectations had been used. Equalities † is due to:

$$\begin{aligned}\mathbf{E}\Big(\theta - \mathbf{E}(\theta|X_0^n)\Big)^2 &= \mathbf{E}\theta^2 - 2\mathbf{E}\theta\mathbf{E}(\theta|X_0^n) + \mathbf{E}\,[\mathbf{E}(\theta|X_0^n)]^2 = \\ &= \mathbf{E}\theta^2 - 2\mathbf{E}\big[\mathbf{E}(\theta\mathbf{E}(\theta|X_0^n)|X_0^n)\big] + \mathbf{E}\,[\mathbf{E}(\theta|X_0^n)]^2 = \\ &= \mathbf{E}\Big(\theta^2 - [\mathbf{E}(\theta|X_0^n)]^2\Big)\end{aligned}$$

Equality †† is obtained by smoothing property of conditional expectation.

2. To show that the limit $\lim_{n\to\infty} \mathbf{E}\widehat{\theta}_n^2$ exists it is sufficient to show that the sequence $(\mathbf{E}\widehat{\theta}_n^2)_{n\geq 0}$ is non decreasing and bounded from above. Indeed, it is bounded:

$$\mathbf{E}\widehat{\theta}_n^2 = \mathbf{E}\,[\mathbf{E}(\theta|X_0^n)]^2 \leq \mathbf{E}\mathbf{E}(\theta^2|X_0^n) = \mathbf{E}\theta^2$$

and non decreasing:

$$\begin{aligned}\mathbf{E}\widehat{\theta}_n^2 &= \mathbf{E}\,[\mathbf{E}(\theta|X_0^n)]^2 = \mathbf{E}\left[\mathbf{E}\Big(\mathbf{E}(\theta|X_0^{n+1})|X_0^n\Big)\right]^2 \leq \\ &\leq \mathbf{E}\mathbf{E}\Big(\big[\mathbf{E}(\theta|X_0^{n+1})\big]^2 |X_0^n\Big) = \mathbf{E}\left[\mathbf{E}(\theta|X_0^{n+1})\right]^2 = \mathbf{E}\widehat{\theta}_{n+1}^2\end{aligned}$$

by virtue of Jensen inequality.

Problem 3.

1.) Put $w_k = W(t_{k+1}) - W(t_k)$ and $v_k = V(t_{k+1}) - V(t_k)$. Then by the properties of Wiener process, the sequences $(v_k)_{k\geq 0}$ and $(w_k)_{k\geq 0}$ are independent Gaussian i.i.d sequences with zero mean and $\mathbf{E}w_k^2 = \mathbf{E}v_k^2 = \Delta$. Kalman filter is given by:

$$\begin{aligned}\widehat{X}(t_{k+1}) &= \widehat{X}(t_k) + a\Delta\widehat{X}(t_k) + \frac{(1+a\Delta)A\Delta V(t_k)}{A^2\Delta^2 V(t_k) + B^2\Delta}(Y(t_{k+1}) - Y(t_k) - A\widehat{X}(t_k)\Delta) \\ V(t_{k+1}) &= (1+a\Delta)^2 V(t_k) + b^2\Delta - \frac{[(1+a\Delta)A\Delta V(t_k)]^2}{A^2\Delta^2 V(t_k) + B^2\Delta}\end{aligned}$$

2.) Neglecting terms of order $\mathcal{O}(\Delta^2)$:

$$\begin{aligned}\widehat{X}(t_{k+1}) &= \widehat{X}(t_k) + a\Delta\widehat{X}(t_k) + \frac{AV(t_k)}{B^2}(Y(t_{k+1}) - Y(t_k) - A\widehat{X}(t_k)\Delta) \\ V(t_{k+1}) &= V(t_k) + 2a\Delta V(t_k) + b^2\Delta - \frac{A^2V^2(t_k)\Delta}{B^2}\end{aligned}$$

which is nothing but an approximation to differential equations:

$$\begin{aligned}\dot{\widehat{X}}(t) &= a\widehat{X}(t) + \frac{AV(t)}{B^2}(\dot{Y}(t) - A\widehat{X}(t)) \\ \dot{V}(t) &= 2aV(t) + b^2 - \frac{A^2V^2(t)}{B^2}\end{aligned}$$

These equations are known as Kalman-Bucy filter.