Massachusetts Institute of Technology
Department of Electrical Engineering and Computer Science

6.432 Stochastic Processes, Detection and Estimation

Problem Set 11

Spring 2004

Issued: Thursday, May 6, 2004 **Due:** Next time the Red Sox win the World Series

---

Final Exam: Our final will take place on **May 19, 2004, from 9:00am to 12:00 noon.**
You are allowed to bring three $8\frac{1}{2}'' \times 11''$ sheets of notes (both sides).

Problem 11.1
Let

$$\begin{aligned} y[n] &= x[n] + w[n] \\ x[n] &= \sum_k h[k] v[n-k] \end{aligned}$$

be a discrete-time random processes, where $v[n]$ and $w[n]$ are uncorrelated, zero-mean, white Gaussian noise processes with variances of $\sigma_v^2 = 2/3$ and $\sigma_w^2 = 1$, respectively and

$$h[n] = \left(\sqrt{\frac{2}{3}}\right)^n u[n].$$

(a) Determine the non-causal Wiener filter for estimating $x[n]$ from $y[n]$ and the associated mean-square estimation error.

(b) Determine the causal Wiener filter for estimating $x[n]$ from $y[n]$ and the associated mean-square estimation error.

(c) Determine the Kalman filter equations for generating $\hat{x}[n|n]$, the linear least-squares estimate of $x[n]$ at time n based on the data $y[0], y[1], \ldots, y[n]$. Also, specify a recursion for the mean- square error in the estimate $\hat{x}[n|n]$.

(d) Determine $\lim_{n\to\infty} \lambda_e[n|n]$, the steady-state estimation error variance for your estimator in part (c). Explain the similarity or differences between your answer and the error variances you calculated in parts (a) and (b).

(e) The sequence $v[n]$ could represent a message being transmitted through a noisy, dispersive communication channel, in which case this is the signal of interest. Give an efficient algorithm for generating $\hat{v}[n|n]$, the estimate of $v[n]$ based on $y[0], \ldots, y[n]$.

Problem 11.2
Suppose we are transmitting a zero-mean, WSS, white sequence $x[n]$ with variance σ_x^2 through the channel

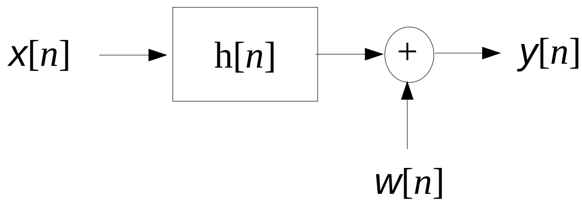


Figure 2-1

where $w[n]$ is zero-mean, WSS white noise with variance σ_w^2 that is uncorrelated with $x[n]$ and where $h[n]$ is the FIR filter

$$h[n] = \delta[n] - \frac{1}{2}\delta[n-1].$$

(a) Write down a state-space model for the problem using

$$s[n] = \begin{bmatrix} x[n] \\ x[n-1] \end{bmatrix}$$

as the state vector.

(b) Determine a recursive algorithm for computing $\hat{x}[n|n]$, the linear least-squares estimate of $x[n]$ based on $y[0], \ldots, y[n]$.

(c) Determine $\lim_{n\to\infty} \lambda_e[n|n]$, the steady-state filtering error variance.

(d) Compare your answer to part (c) to the error variances of the causal and non-causal Wiener filters.

Problem 11.3
Consider a system with the following state space description

$$\begin{aligned} x[n+1] &= a\,x[n] + v[n] \\ y[n] &= x[n] + w[n] \end{aligned}$$

where $v[n]$ and $w[n]$ are uncorrelated, wide-sense stationary, zero-mean, white random processes with variances 1 and σ_w^2 respectively.

(a) Suppose the process $x[n]$ starts at time $n = -\infty$. Determine the unit-sample response $h[n]$ such that the system can be expressed in terms of the following block diagram.

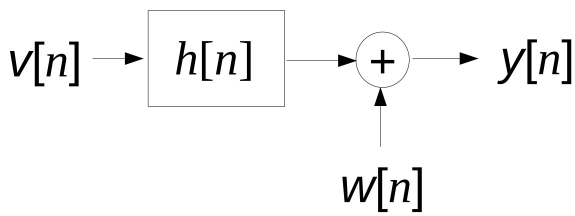


Figure 3-1

(b) Suppose the process $x[n]$ starts at time $n = 0$ and that the initial condition $x[0]$ has mean zero, variance σ_0^2 and is uncorrelated with both $v[n]$ and $w[n]$ for all n. For which values of a and σ_0^2 is the process $y[n]$ wide-sense stationary (for $n \geq 0$)?

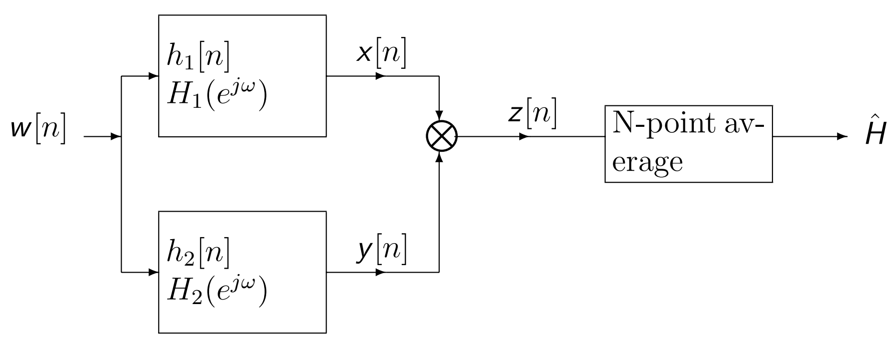
(c) Suppose that

$$\begin{aligned}\hat{x}[2|2] &= \frac{16}{51}\hat{x}[1|1] + \frac{19}{51}y[2], \\ \lambda_e[1|1] &= \frac{3}{4}\end{aligned}$$

where $\hat{x}[n|k]$ is the linear least-squares estimate of $x[n]$ based on $y[0], y[1], \ldots, y[k]$, and $\lambda_e[n|k]$ is the associated mean-square estimation error. Determine a and σ_w^2.

Problem 11.4

The figure shown below is a noise cross-correlation technique for measuring part of the frequency response of a discrete-time, linear, time- invariant system.



In this figure:

1. $w[n]$ is a zero-mean stationary white Gaussian discrete-time noise process with $S_{ww}(e^{j\omega}) = q$.

2. System 1 has unknown real-valued impulse response $h_1(t)$ and associated frequency response $H_1(e^{j\omega})$.

3. System 2 is a known passband reference filter with frequency response

$$H_2(e^{j\omega}) = \begin{cases} q^{-1} & |\omega - \omega_0| \le \Delta \quad \text{or} \quad |\omega + \omega_0| \le \Delta \\ 0 & \text{elsewhere in} \quad |\omega| \le \pi \end{cases}$$

where ω_0 is the center frequency of the passband of this filter.

4. The N-point averager produces as its output

$$\hat{H} = \frac{1}{N} \sum_{n=-(n-1)/2}^{(N-1)/2} z[n]$$

where N is an odd integer.

(a) Find the mean function $m_z[n]$ and the covariance function $K_{zz}[n,m]$ of the process $z[n]$. Express your mean function answer as an explicit function of the unknown frequency response $H_1(e^{j\omega})$. You may leave your covariance function answer in terms of $h_1[n]$ and $h_2[n]$, or $H_1(e^{j\omega})$ and $H_2(e^{j\omega})$.

(b) Find the mean and variance of $\hat{H}$ in terms of $m_z[n]$ and $K_{zz}[n,m]$.

(c) Combine your results of (a) and (b) to obtain a simple approximate expression for $E(\hat{H})$ which applies when

$$H_1(e^{j\omega}) \approx H_1(e^{j\omega_0})$$

over the passband of $H_2(e^{j\omega})$.

(d) Under the assumption that $|H_1(e^{j\omega})| \le M < \infty$ for all ω where M is a known constant, show that $\mathrm{var}(\hat{H}) \to 0$ as $N \to \infty$.

(e) Suppose $w[n]$ is applied only for $|n| \le K - 1$, where K is an odd integer, and zero input is applied to the two systems for $|n| \ge K$. For what range of (Δ, N) values will $\hat{H}$ be about the same for this $2K$-point input as it would be for the case when $w[n]$ is applied for $-\infty < n < \infty$. Briefly discuss how the resolution and variance of $\hat{H}$ behave for different (Δ, N) choices within this range.

Problem 11.5

We are interested in detecting the presence of a random signal $x[n]$ based on observations $y[n]$ corrupted by random noise $w[n]$. Our hypotheses, therefore, are:

$$\begin{aligned} H_0 : y[n] &= w[n] \\ H_1 : y[n] &= x[n] + w[n] \end{aligned}$$

where $n = 0, 1, \ldots, N-1$. The processes $x[n]$ and $w[n]$ are independent, zero-mean, stationary, Gaussian random processes with power spectral densities $S_{xx}(e^{j\omega})$ and σ_w^2 respectively. The hypotheses H_0 and H_1 are equally likely.

Consider a detector of the form

$$\ell = \frac{1}{2\pi} \int_{-\pi}^{\pi} \hat{S}_N(e^{j\omega})\, F(e^{j\omega})\, d\omega \underset{\hat{H}=H_0}{\overset{\hat{H}=H_1}{\gtrless}} \gamma \tag{1}$$

where $\hat{S}_N(e^{j\omega})$ denotes the periodogram of the data, i.e.,

$$\hat{S}_N(e^{j\omega}) = \left|Y_N(e^{j\omega})\right|^2 = \frac{1}{N}\left|\sum_{n=0}^{N-1} y[n]\, e^{-j\omega n}\right|^2.$$

We restrict our attention throughout this problem to the large N scenario. In this regime, you may exploit, among other approximations,

$$E\left[\hat{S}_N(e^{j\omega})\right] \approx S_{yy}(e^{j\omega})$$
$$\operatorname{cov}\left(\hat{S}_N(e^{j\omega}), \hat{S}_N(e^{j\nu})\right) \approx \frac{1}{N} S_{yy}^2(e^{j\omega})\, \delta(\omega - \nu)$$

(a) Determine the function $F(e^{j\omega})$ in (1) that maximizes the performance metric

$$D = \frac{\left(E\left[\ell \mid H = H_1\right] - E\left[\ell \mid H = H_0\right]\right)^2}{\operatorname{var}\left[\ell \mid H = H_0\right]},$$

which is referred to as the "deflection" metric. Also determine the corresponding value of D.

(b) Suppose that instead of maximizing the deflection performance metric, we wish to minimize the probability of error. Determine $F(e^{j\omega})$ and γ in (1) so that the resulting detector asymptotically minimizes the probability of error in this Gaussian, "stationary process, long observation time" scenario.

(c) Under what conditions on $S_{xx}(e^{j\omega})$ can the detector you determine in (a) achieve a probability of error close to that achieved by the detector you determine in (b).

Problem 11.6

(a) A discrete-time Gaussian stochastic process $y[n]$ is generated by a first-order system driven by white Gaussian noise $w[n]$:

$$y[n+1] = \alpha y[n] + w[n], \qquad n = 0, 1, 2, \ldots$$

where $y[0] \sim N(0, P)$ is independent of the noise sequence. Also, $w[n] \sim N(0, \sigma^2)$, and α is unknown. Find the ML estimate of α based on observation of $y[0], \ldots, y[N+1]$. (*Hint:* $y[n]$ is a Gauss-Markov process.)

(b) Now suppose $y[n]$ is a Gaussian process generated by an $(M+1)$th-order system driven by white Gaussian noise $w[n]$:

$$y[n+1] = a_0 y[n] + a_1 y[n-1] + \cdots + a_M y[n-M] + w[n], \qquad n = 0, 1, 2, \ldots,$$

where $[y[-M] \cdots y[0]]^T \sim N(0, \mathbf{R})$ is independent of the white noise sequence. Also, $w[n] \sim N(0, \sigma^2)$, and $\mathbf{a} = [a_0 \cdots a_M]^T$ is unknown. Find the ML estimate of $\mathbf{a}$ based on observation of $y[-M], \ldots, y[N+1]$.

(*Comment*: In reality, we may not know the density for the initial conditions, or this density may depend on $\mathbf{a}$. In this case, we can use the same solution as an approximate ML estimate. Since the approximation effectively does not attempt to model the first $M+1$ points, it is good only for large N.)

(c) Sometimes it is useful to take a discrete-time sequence $y[-M], \ldots, y[N+1]$ and try to model each sample as a linear combination of its past samples. That is, we seek to find the best $\mathbf{a} = [a_0 \cdots a_M]^T$ so that we minimize

$$\sum_{k=1}^{N+1} \left(y[k] - \sum_{i=0}^{M} a_i y[k-1-i] \right)^2 .$$

Find the optimal $\mathbf{a}$, and compare with part (b).

(*Comment:* This technique is referred to as the "autocovariance method of linear prediction" and is widely used. One successful application is speech compression: A deterministic speech waveform is represented in terms of linear combinations of its past samples, and the error signal is transmitted. Because the error signal has smaller energy than the speech signal, for a given quantization error level, it takes fewer bits to transmit the error signal.)

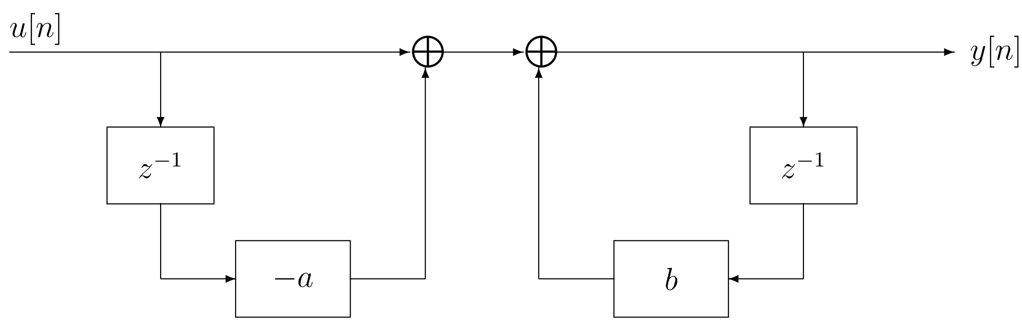
Problem 11.7

(a) Consider a causal LTI system, with input $u[n]$ and output $y[n]$, described by the system function

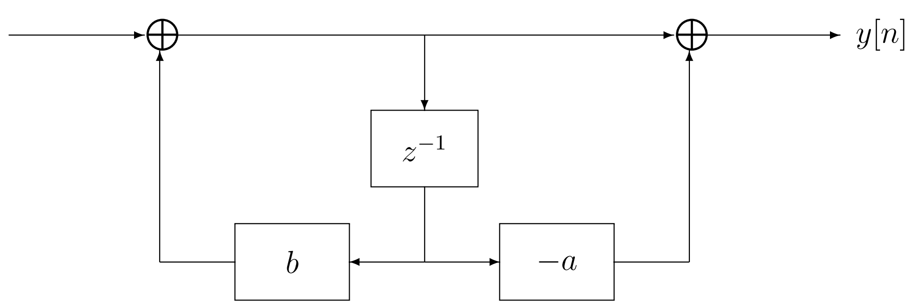
$$H(z) = \frac{1 - az^{-1}}{1 - bz^{-1}}$$

Show that each of the following two block diagrams represents a realization of this system, i.e., the system function for the causal LTI systems described by each of these systems equals $H(z)$:

System A



System B



(b) Write the state equations

$$\mathbf{x}[n+1] = A\mathbf{x}[n] + Bu[n]$$
$$y[n] = C\mathbf{x}[n] + Du[n]$$

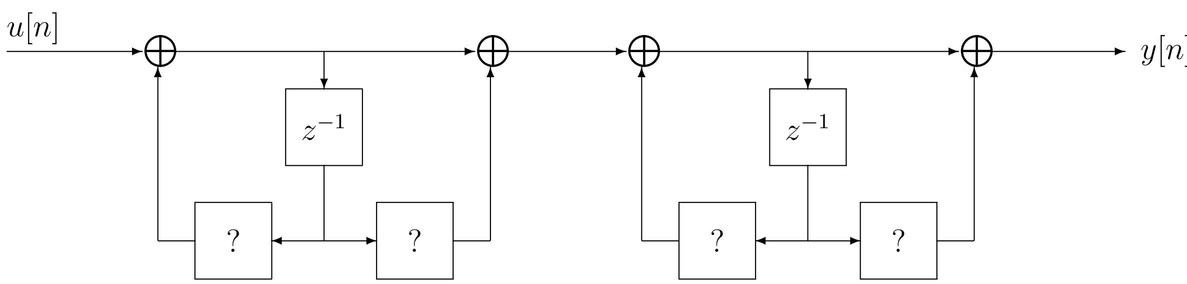
for each of the two systems in part (a), taking the outputs of the delay elements as state variables.

(c) Consider a causal LTI system, with input $u[n]$ and output $y[n]$ described by

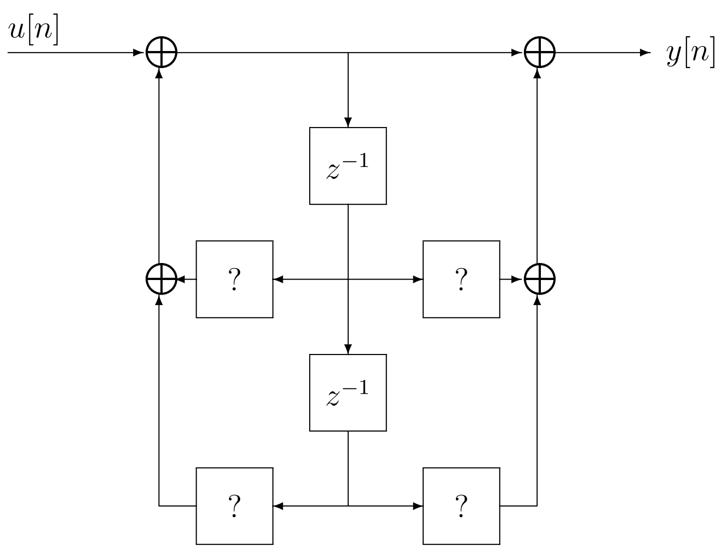
$$H(z) = \frac{1 + z^{-1} - 2z^{-2}}{1 - \frac{5}{2}z^{-1} - \frac{3}{2}z^{-2}} = \frac{(1 - z^{-1})(1 + 2z^{-1})}{(1 + \frac{1}{2}z^{-1})(1 - 3z^{-1})}$$

Determine two realizations of this system of the following two forms (i.e. find the coefficient values indicated by question marks).

System C



System D



(d) Write the state equations for the two systems of part (c), again taking the outputs of unit delay elements as state variables.

(e) (practice) Write the system function as

$$H(z) = a + \frac{b}{1 + \frac{1}{2}z^{-1}} + \frac{c}{1 - 3z^{-1}}$$

Find a, b, and c, construct a block diagram for a realization suggested by this form, and write the corresponding state equations.

Problem 11.8
Consider the following state variable model driven by zero-mean white noise

$$\mathsf{x}[n+1] = A[n]\mathsf{x}[n] + G[n]\mathsf{w}[n]$$

where $\mathsf{w}[n]$ is uncorrelated with the initial condition $\mathsf{x}[n_0]$, and

$$K_{\mathsf{ww}}[n, m] = Q[n]\delta[n - m]$$

(a) Let $P_{\mathsf{x}}[n] = \text{cov}(\mathsf{x}[n], \mathsf{x}[n])$. Show that $P_{\mathsf{x}}[n]$ satisfies the recursion

$$P_{\mathsf{x}}[n+1] = A[n]P_{\mathsf{x}}[n]A^T[n] + G[n]Q[n]G^T[n] \tag{2}$$

with initial condition $P_{\mathsf{x}}[n_0] = \text{cov}(\mathsf{x}[n_0], \mathsf{x}[n_0])$

(b) Suppose for the rest of this problem that $A[n] = A$, $G[n] = G$, and $Q[n] = Q$, i.e. that the coefficient matrices in the state model are constant and $\mathsf{w}[n]$ is WSS, with the covariance of $\mathsf{w}[n]$ equaling Q. Let

$$K_{\mathsf{xx}}[n, m] = \text{cov}(\mathsf{x}[n], \mathsf{x}[m])$$

(i) First, show that

$$K_{\mathsf{xx}}[n, m] = K_{\mathsf{xx}}^T[m, n]$$

(You don't need the state equations or anything since the second lecture to show this.)

(ii) Show that

$$K_{\mathsf{xx}}[n, m] = \begin{cases} A^{n-m}P_{\mathsf{x}}[m] & n \geq m \\ P_{\mathsf{x}}[n](A^T)^{m-n} & n \leq m \end{cases} \tag{3}$$

where for $k > 0$, A^k denotes the product of the matrix A with itself k times, and by convention $A^0 = I$.

(c) A result we will <u>not</u> show (but which is nevertheless true) in general is that if all of the eigenvalues of A have magnitude less that 1, then $P_{\mathsf{x}}[n]$, which is the constant parameter case satisfies

$$P_{\mathsf{x}}[n+1] = AP_{\mathsf{x}}[n]A^T + GQG^T$$

converges to a limit, which we denote by P_{x}.

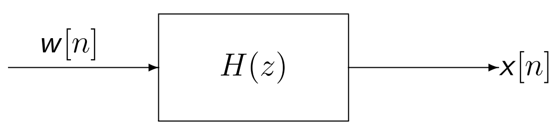
(i) Show that this is true in the scalar case, i.e. when $\mathsf{x}[n]$, A, G, W, Q are all numbers, and compute the limiting value of P_{x}.

(ii) Explain why, in the vector case, the limiting value of P_{x} satisfies the <u>discrete-time Algebraic Lyapunov Equation</u>

$$P_{\mathsf{x}} = AP_{\mathsf{x}}A^T + GQG^T \tag{4}$$

(d) Suppose that $P_{\mathsf{x}}[n_0] = P_{\mathsf{x}}$, the solution to (3). Show that in this case, $\mathsf{x}[n]$ is WSS for $n \geq n_0$, i.e. $K_{\mathsf{xx}}[n, m] = K_{\mathsf{xx}}[n-m]$.

(e) Consider again the scaler case, and suppose that $P_{\mathsf{x}}[n_0] = P_{\mathsf{x}}$, the value determined in part (c), so that $\mathsf{x}[n]$ is stationary. Determine <u>two</u> forms for $S_{\mathsf{xx}}(z)$, one from (2) and the other from the block diagram



where $K_{\mathsf{ww}}[m] = Q$ and $H(z)$ is the system functions from $\mathsf{w}[n]$ to $\mathsf{x}[n]$.