UNIVERSITI SAINS MALAYSIA

Semester II Examination
Academic Session 2004/2005

March 2005

EEE 503 – STOCHASTIC PROCESS

Time : 3 hours

INSTRUCTION TO CANDIDATE:

Please ensure that this examination paper contains **SIX (6)** printed pages including Appendices (12 page) and **FIVE (5)** questions before answering.

Answer **ALL** questions.

Distribution of marks for each question is given accordingly.

All questions must be answered in English.

...2/-

1. (a) The voltage V at the output of a microphone is a uniform random variable with limits -1 volt and 1 volt. The microphone voltage is processed by a hard limiter with cut-off points -0.5 volt and 0.5 volt. The magnitude of the limiter output *L* is a random variable such that

$$L = \begin{cases} V & |V| \le 0.5 \\ 0.5 & otherwise \end{cases}$$

(i) What is $P[L = 0.5]$?

(ii) What is $F_L(l)$?

(iii) What is $E[L]$?

(35%)

(b) The amplitude V (as measured in volts) of a sinusoidal signal is a random variable with PMF:

$$P_V(v) = \begin{cases} 1/7 & v = -3,-2,\ldots, 2, 3 \\ 0 & otherwise \end{cases}$$

Let $Y = V^2/2$ watts denote the average power of the transmitted signal. Find $P_Y(y)$.

(25%)

(c) The joint CDF of a bivariate random variable (X, Y) is given by;

$$F_{XY}(x,y) = \begin{cases} (1-e^{-\alpha x})(1-e^{-\beta y}) & x \ge 0, y \ge 0, \alpha, \beta > 0 \\ 0 & otherwise \end{cases}$$

...3/-

(i) Find the marginal cdf's of X and Y

(ii) Show that X and Y are independent

(iii) Find $P(X > x, Y > y)$

(40%)

2. (a) The input to a digital filter is a sequence of random variables ...., $X_{-1}, X_0, X_1, \ldots$. The output is also a sequence of random variables ..., $W_{-1}, W_0, W_1, \ldots$. The relationship between the input and output is

$$W_n = \frac{1}{2}(X_n + X_{n-1})$$

Let the input be a sequence of an identically independent random variables with $E[X_i] = 0$ and $Var[X_i] = 1$. Find the following properties of the output sequence:

(i) $E[W_i]$

(ii) $Var[W_i]$

(iii) $Cov[W_{i+1}, W_i]$

(iv) $\rho_{Wi+1, Wi}$

(50%)

(b) The wide sense stationary process $X(t)$ with autocorrelation function $R_X(\tau)$ and power spectral density $S_X(f)$ is the input to a tapped delay line filter

$$H(f) = a_1 e^{-j2\pi f t_1} + a_2 e^{-j2\pi f t_2}$$

Find the output power spectral density $S_Y(f)$ and the output autocorrelation $R_Y(\tau)$.

(50%)

...4/-

3. (a) Let $X(t)$ be a stationary continuous time random process. By sampling $X(t)$ every Δ seconds, we obtain the discrete time random sequence $Y_n = X(n\Delta)$. Is Y_n a stationary random sequence ?

(25%)

(b) $X(t)$ is a wide sense stationary random processes with average power equal to 1. Let Θ denote a random variable with uniform distribution over $[0, 2\pi]$ such that $X(t)$ and Θ are independent.

(i) What is $E[X^2(t)]$?

(ii) What is $E[\cos(2\pi f_c t + \Theta)]$?

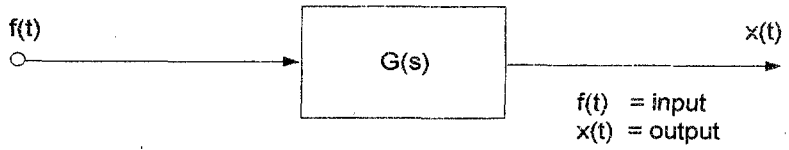
(iii) Let $Y(t) = X(t)\cos(2\pi f_c t + \Theta)$. What is $E[Y(t)]$?

(iv) What is the average power of $Y(t)$?

(75%)

4. Consider the input process to the system in Figure 1 to have an exponential autocorrelation function of the form $R_f(\tau) = \sigma^2 e^{-\beta|\tau|}$. The transfer function of the system filter is given by

$$G(s) = \frac{Ts}{(1+Ts)^2}$$



f(t) = input
x(t) = output

$$G(s) = \frac{X(s)}{F(s)} = \text{transfer function}$$

Figure 1

...5/-

(a) Find the output spectral function, Sx. (50%)

(b) Find the mean-squared value of the output x. (50%)

5. Consider a Wiener filter problem as shown in Figure 2, where input f(t) is a combination of Gauss-Markov signal s(t) with exponential auto correlation function $R_s(\tau) = \sigma^2 e^{-\beta/\tau/}$ and white-noise n(t) with $S_n(j\omega) = \frac{No}{2}$.

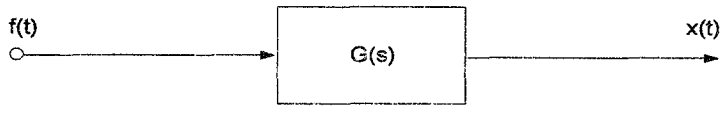


Figure 2

If the filter transfer function is chosen to be $G(s) = \frac{2}{Ks+3}$, where K is a constant for the filter.

(a) Find the mean-squared error for the output x(t). (25%)

(b) Determine the x for the filter so that the mean-squared error is minimum. (25%)

(c) For the same Markov signal and white-noise combination input to system in Figure 2, find the optimal noncausal filter (i.e. α = 0). For your calculations use No = 2, σ^2 = 4, β = 1. (50%)

...6/-

6. Consider a stationary Gauss-Markov process having an autocorrelation function of the form $R_x(\tau) = 1.e^{-|\tau|}$. The shaping filter that shapes white-noise into the process is shown in Figure 3.

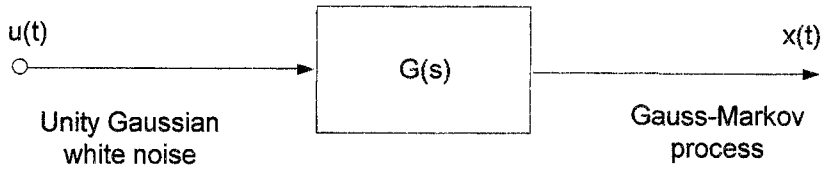


Figure 3

Suppose we have a sequence of noisy measurements of this process taken .02 sec apart starting at t = 0. The measurement error will be assumed to have a variance of unity. We need to process these via a Kalman filter and obtain an optimal estimate of x(t). Determine the followings.

(a) Spectral function for the process (25%)

(b) State equation (25%)

(c) The filter parameter ϕ_k, H_k, Q_k and R_k. (25%)

(d) Find the optimal estimate of x(t) for only 3 steps with intial conditions for $\hat{x}_0^- = 0$ and $P_0^- = 1$. (25%)

ooo0ooo

THE STATIONARY OPTIMIZATION PROBLEM—WEIGHTING FUNCTION APPROACH*

We now consider the filter optimization problem that Wiener first solved in the 1940s (1). Referring to Fig. 4.3, we assume the following:

1. The filter input is an additive combination of signal and noise, both of which are covariance stationary with known auto- and crosscorrelation functions (or corresponding spectral functions).
2. The filter is linear and not time-varying. No further assumption is made as to its form.
3. The output is covariance stationary. (A long time has elapsed since any switching operation.)
4. The performance criterion is minimum mean-square error, where the error is defined as $e(t) = s(t + \alpha) - x(t)$.

In addition to the generalization relative to the form of the filter transfer function, we are also generalizing by saying the ideal filter output is to be $s(t + \alpha)$ rather than just $s(t)$. The following terminology has evolved relative to the choice of the α parameter:

1. α ***positive:*** This is called the *prediction* problem. (The filter is trying to predict the signal value α units ahead of the present time t.)
2. $\alpha = 0$**:** This is called the *filter* problem. (The usual problem we have considered before.)

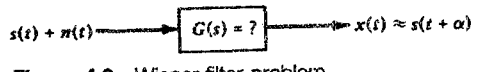


Figure 4.3 Wiener filter problem.

3. α ***negative:*** This is called the *smoothing* problem. (The filter is trying to estimate the signal value α units in the past.)

This is an important generalization and there are numerous physical applications corresponding to all three cases. The α parameter is chosen to fit the particular application at hand, and it is fixed in the optimization process.

We begin by defining the filter error as

$$e(t) = s(t + \alpha) - x(t) \tag{4.3.1}$$

The squared error is then

$$e^2(t) = s^2(t + \alpha) - 2s(t + \alpha)x(t) + x^2(t) \tag{4.3.2}$$

We next write $x(t)$ as a convolution integral:

$$\int_{-\infty}^{\infty} g(u)[s(t - u) + n(t - u)]\, du \tag{4.3.3}$$

This can be substituted into Eq. (4.3.2) and both sides averaged to yield*

$$E(e^2) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(u)g(v)R_{s+n}(u - v)\, du\, dv - 2\int_{-\infty}^{\infty} g(u)R_{s+n,s}(\alpha + u)\, du + R_s(0) \tag{4.3.4}$$

where

R_s = autocorrelation function of $s(t)$

R_{s+n} = autocorrelation function of $s(t) + n(t)$

$R_{s+n,s}$ = crosscorrelation between $s(t) + n(t)$ and $s(t)$

Note that if signal and noise have zero crosscorrelation,

$$\left.\begin{aligned} R_{s+n} &= R_s + R_n \\ R_{s+n,s} &= R_s \end{aligned}\right\} \quad \text{(for zero correlation)} \tag{4.3.5}$$

We wish to find the function $g(u)$ in Eq. (4.3.4) that minimizes $E(e^2)$. This will be recognized as a problem in calculus of variations (5). Following the usual procedure, we replace $g(u)$ with a perturbed weighting function $g(u) + \varepsilon\eta(u)$

* We have chosen here to write the mean-square error in terms of the filter weighting function and input autocorrelation functions. In the stationary problem, one can also write $E(e^2)$ in terms of the filter transfer function and input spectral functions, and then proceed with the optimization on that basis (6, 7). We have chosen the time-domain approach because it is easily generalized to the non-stationary problem that is considered in Section 4.4. The frequency-domain approach is not readily generalized.

where

$g(u)$ = optimum weighting function [*Note*: From this point on in the solution, $g(u)$ will denote the *optimal* weighting function.]
$\eta(u)$ = an arbitrary perturbing function
ε = small perturbation factor such that the perturbed function approaches the optimum one as ε goes to zero

The optimum and perturbed weighting functions are sketched in Fig. 4.4. Replacing $g(u)$ with $g(u) + \varepsilon\eta(u)$ in Eq. (4.3.4) then leads to

$$E(e^2) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} [g(u) + \varepsilon\eta(u)][g(v) + \varepsilon\eta(v)]R_{s+n}(u - v)\,du\,dv - 2\int_{-\infty}^{\infty} [g(u) + \varepsilon\eta(u)]R_{s+n,s}(\alpha + u)\,du + R_s(0) \tag{4.3.6}$$

Note that $E(e^2)$ is now a function of ε, and it is to be a minimum when $\varepsilon = 0$. Now, using differential calculus methods, we differentiate $E(e^2)$ with respect to ε and set the result equal to zero for $\varepsilon = 0$. After interchanging dummy variables of integration freely, the result is

$$\int_{-\infty}^{\infty} \eta(\tau)\left[-R_{s+n,s}(\alpha + \tau) + \int_{-\infty}^{\infty} g(u)R_{s+n}(u - \tau)\,du\right] d\tau = 0 \tag{4.3.7}$$

A subtlety in the solution arises at this point; therefore, it is convenient to look at the causal and noncausal cases separately.

Noncausal Solution

If we put no constraint on the filter weighting function, we will very likely obtain a $g(u)$ that is nontrivial for negative as well as positive u. This weighting function is noncausal because it requires the filter to "look ahead" of real time and use data that are not yet available. This is, of course, not possible if the filter is operating on-line. However, in off-line applications, such as postflight analysis of recorded data, the noncausal solution is possible and very much of interest. Thus, it should not be ignored.

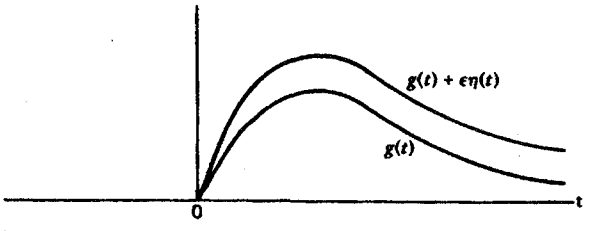


Figure 4.4 Optimal and perturbed weighting functions.

If there are no restrictions on $g(u)$, then, similarly, there are no constraints on the perturbation function $\eta(\tau)$. It is arbitrary for all values of its argument. Thus, if the integral with respect to τ in Eq. (4.3.7) is to be zero, the bracketed term must be zero for *all* τ. This leads to

$$\int_{-\infty}^{\infty} g(u)R_{s+n}(u - \tau)\, du = R_{s+n,s}(\alpha + \tau), \qquad -\infty < \tau < \infty \tag{4.3.8}$$

This is an integral equation of the first kind, and in this case it can be solved readily using Fourier transform methods. Since R_{s+n} is symmetric, the term on the left side of Eq. (4.3.8) has the exact form of a convolution integral. Therefore, transforming both sides yields

$$G(s)S_{s+n}(s) = S_{s+n,s}(s)e^{\alpha s} \tag{4.3.9}$$

or

$$G(s) = \frac{S_{s+n,s}(s)e^{\alpha s}}{S_{s+n}(s)} \tag{4.3.10}$$

Remember that the transforms indicated in Eq. (4.3.10) are *two-sided* transforms rather than the usual single-sided transforms. Of course, if we wish to find the weighting function $g(u)$, we simply take the inverse transform of the expression given by Eq. (4.3.10).

The filter mean-square error is given by Eq. (4.3.4). If $g(u)$ is the optimal weighting function satisfying Eq. (4.3.8), the mean-square error equation may be simplified as follows. First write the second term of Eq. (4.3.4) as the sum of two equal terms and combine one of these with the double integral term. After we rearrange terms, this leads to

$$E(e^2) = R_s(0) - \int_{-\infty}^{\infty} g(u)R_{s+n,s}(\alpha + u)\, du + \int_{-\infty}^{\infty} g(u)\left[-R_{s+n,s}(\alpha + u) + \int_{-\infty}^{\infty} g(v)R_{s+n}(v - u)\, dv\right] du \tag{4.3.11}$$

The bracketed quantity in Eq. (4.3.11) is zero for optimal $g(v)$ for all u. Therefore, the mean-square error is

$$E(e^2) = R_s(0) - \int_{-\infty}^{\infty} g(u)R_{s+n,s}(\alpha + u)\, du \tag{4.3.12}$$

THE DISCRETE KALMAN FILTER

R. E. Kalman's paper describing a recursive solution of the discrete-data linear filtering problem was published in 1960 (1). About this same time, advances in digital computer technology made it possible to consider implementing his recursive solution in a number of real-time applications. This was a fortuitous circumstance, and Kalman filtering caught hold almost immediately. We will consider some examples shortly, but we must first develop the Kalman filter recursive equations, which are, in effect, the "filter."

We begin by assuming the random process to be estimated can be modeled in the form

$$\mathbf{x}_{k+1} = \boldsymbol{\phi}_k \mathbf{x}_k + \mathbf{w}_k \tag{5.5.1}$$

The observation (measurement) of the process is assumed to occur at discrete points in time in accordance with the linear relationship

$$\mathbf{z}_k = \mathbf{H}_k \mathbf{x}_k + \mathbf{v}_k \tag{5.5.2}$$

Some elaboration on notation and the various terms of Eqs. (5.5.1) and (5.5.2) is in order:

$\mathbf{x}_k$ = $(n \times 1)$ process state vector at time t_k

$\boldsymbol{\phi}_k$ = $(n \times n)$ matrix relating $\mathbf{x}_k$ to $\mathbf{x}_{k+1}$ in the absence of a forcing function (if $\mathbf{x}_k$ is a sample of continuous process, $\boldsymbol{\phi}_k$ is the usual state transition matrix)

$\mathbf{w}_k$ = $(n \times 1)$ vector—assumed to be a white sequence with known covariance structure

$\mathbf{z}_k$ = $(m \times 1)$ vector measurement at time t_k

$\mathbf{H}_k$ = $(m \times n)$ matrix giving the ideal (noiseless) connection between the measurement and the state vector at time t_k

$\mathbf{v}_k$ = $(m \times 1)$ measurement error—assumed to be a white sequence with known covariance structure and having zero crosscorrelation with the $\mathbf{w}_k$ sequence

The covariance matrices for the $\mathbf{w}_k$ and $\mathbf{v}_k$ vectors are given by

$$E[\mathbf{w}_k\mathbf{w}_i^T] = \begin{cases} \mathbf{Q}_k, & i = k \\ 0, & i \neq k \end{cases} \tag{5.5.3}$$

$$E[\mathbf{v}_k\mathbf{v}_i^T] = \begin{cases} \mathbf{R}_k, & i = k \\ 0, & i \neq k \end{cases} \tag{5.5.4}$$

$$E[\mathbf{w}_k\mathbf{v}_i^T] = 0, \qquad \text{for all } k \text{ and } i \tag{5.5.5}$$

We assume at this point that we have an initial estimate of the process at some point in time t_k, and that this estimate is based on all our knowledge about the process prior to t_k. This prior (or *a priori*) estimate will be denoted as $\hat{\mathbf{x}}_k^-$ where the "hat" denotes estimate, and the "super minus" is a reminder that this is our best estimate prior to assimilating the measurement at t_k. (Note that super minus as used here is not related in any way to the super minus notation used in spectral factorization.) We also assume that we know the error covariance matrix associated with $\hat{\mathbf{x}}_k^-$. That is, we define the estimation error to be

$$\mathbf{e}_k^- = \mathbf{x}_k - \hat{\mathbf{x}}_k^- \tag{5.5.6}$$

and the associated error covariance matrix is*

$$\mathbf{P}_k^- = E[\mathbf{e}_k^-\mathbf{e}_k^{-T}] = E[(\mathbf{x}_k - \hat{\mathbf{x}}_k^-)(\mathbf{x}_k - \hat{\mathbf{x}}_k^-)^T] \tag{5.5.7}$$

* We tacitly assume here that the estimation error has zero mean, and thus it is proper to refer to $E[\mathbf{e}_k^-\mathbf{e}_k^{-T}]$ as a covariance matrix. It is also, of course, a moment matrix, but it is usually not referred to as such.

In many cases, we begin the estimation problem with no prior measurements. Thus, in this case, if the process mean is zero, the initial estimate is zero, and the associated error covariance matrix is just the covariance matrix of $\mathbf{x}$ itself.

With the assumption of a prior estimate $\hat{\mathbf{x}}_k^-$, we now seek to use the measurement $\mathbf{z}_k$ to improve the prior estimate. We choose a linear blending of the noisy measurement and the prior estimate in accordance with the equation

$$\hat{\mathbf{x}}_k = \hat{\mathbf{x}}_k^- + \mathbf{K}_k(\mathbf{z}_k - \mathbf{H}_k\hat{\mathbf{x}}_k^-) \tag{5.5.8}$$

where

$\hat{\mathbf{x}}_k$ = updated estimate

$\mathbf{K}_k$ = blending factor (yet to be determined)

The justification of the special form of Eq. (5.5.8) will be deferred until Section 5.8. The problem now is to find the particular blending factor $\mathbf{K}_k$ that yields an updated estimate that is optimal in some sense. Just as in the Wiener solution, we use minimum mean-square error as the performance criterion. Toward this end, we first form the expression for the error covariance matrix associated with the updated (*a posteriori*) estimate.

$$\mathbf{P}_k = E[\mathbf{e}_k\mathbf{e}_k^T] = E[(\mathbf{x}_k - \hat{\mathbf{x}}_k)(\mathbf{x}_k - \hat{\mathbf{x}}_k)^T] \tag{5.5.9}$$

Next, we substitute Eq. (5.5.2) into Eq. (5.5.8) and then substitute the resulting expression for $\hat{\mathbf{x}}_k$ into Eq. (5.5.9). The result is

$$\begin{aligned}\mathbf{P}_k = E\{&[(\mathbf{x}_k - \hat{\mathbf{x}}_k^-) - \mathbf{K}_k(\mathbf{H}_k\mathbf{x}_k + \mathbf{v}_k - \mathbf{H}_k\hat{\mathbf{x}}_k^-)] \\ &[(\mathbf{x}_k - \hat{\mathbf{x}}_k^-) - \mathbf{K}_k(\mathbf{H}_k\mathbf{x}_k + \mathbf{v}_k - \mathbf{H}_k\hat{\mathbf{x}}_k^-)]^T\}\end{aligned} \tag{5.5.10}$$

Now, performing the indicated expectation and noting the $(\mathbf{x}_k - \hat{\mathbf{x}}_k^-)$ is the a priori estimation error that is uncorrelated with the measurement error $\mathbf{v}_k$, we have

$$\mathbf{P}_k = (\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)\mathbf{P}_k^-(\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)^T + \mathbf{K}_k\mathbf{R}_k\mathbf{K}_k^T \tag{5.5.11}$$

Notice here that Eq. (5.5.11) is a perfectly general expression for the updated error covariance matrix, and it applies for any gain $\mathbf{K}_k$, suboptimal or otherwise.

Returning to the optimization problem, we wish to find the particular $\mathbf{K}_k$ that minimizes the individual terms along the major diagonal of $\mathbf{P}_k$, because these terms represent the estimation error variances for the elements of the state vector being estimated. The optimization can be done in a number of ways. We will do this using a straightforward differential calculus approach, and to do so we need two matrix differentiation formulas. They are

$$\frac{d[\text{trace}(\mathbf{AB})]}{d\mathbf{A}} = \mathbf{B}^T \quad (\mathbf{AB} \text{ must be square}) \tag{5.5.12}$$

$$\frac{d[\text{trace}(\mathbf{ACA}^T)]}{d\mathbf{A}} = 2\mathbf{AC} \quad (\mathbf{C} \text{ must be symmetric}) \tag{5.5.13}$$

where the derivative of a scalar with respect to a matrix is defined as

$$\frac{ds}{d\mathbf{A}} = \begin{bmatrix} \frac{ds}{da_{11}} & \frac{ds}{da_{12}} & \cdots \\ \frac{ds}{da_{21}} & \frac{ds}{da_{22}} & \cdots \\ \vdots & & \end{bmatrix} \tag{5.5.14}$$

Proof of these two differentiation formulas will be left as an exercise (see Problem 5.16). We will now expand the general form for $\mathbf{P}_k$, Eq. (5.5.11), and rewrite it in the form:

$$\mathbf{P}_k = \mathbf{P}_k^- - \mathbf{K}_k\mathbf{H}_k\mathbf{P}_k^- - \mathbf{P}_k^-\mathbf{H}_k^T\mathbf{K}_k^T + \mathbf{K}_k(\mathbf{H}_k\mathbf{P}_k^-\mathbf{H}_k^T + \mathbf{R}_k)\mathbf{K}_k^T \tag{5.5.15}$$

Notice that the second and third terms are linear in $\mathbf{K}_k$ and that the fourth term is quadratic in $\mathbf{K}$. The two matrix differentiation formulas may now be applied to Eq. (5.5.15). We wish to minimize the trace of $\mathbf{P}$ because it is the sum of the mean-square errors in the estimates of all the elements of the state vector. We can use the argument here that the individual mean-square errors are also minimized when the total is minimized, provided that we have enough degrees of freedom in the variation of $\mathbf{K}_k$, which we do in this case. We proceed now to differentiate the trace of $\mathbf{P}_k$ with respect to $\mathbf{K}_k$, and we note that the trace of $\mathbf{P}_k^-\mathbf{H}_k^T\mathbf{K}_k^T$ is equal to the trace of its transpose $\mathbf{K}_k\mathbf{H}_k\mathbf{P}_k^-$. The result is

$$\frac{d(\text{trace } \mathbf{P}_k)}{d\mathbf{K}_k} = -2(\mathbf{H}_k\mathbf{P}_k^-)^T + 2\mathbf{K}_k(\mathbf{H}_k\mathbf{P}_k^-\mathbf{H}_k^T + \mathbf{R}_k) \tag{5.5.16}$$

We now set the derivative equal to zero and solve for the optimal gain. The result is

$$\mathbf{K}_k = \mathbf{P}_k^-\mathbf{H}_k^T(\mathbf{H}_k\mathbf{P}_k^-\mathbf{H}_k^T + \mathbf{R}_k)^{-1} \tag{5.5.17}$$

This particular $\mathbf{K}_k$, namely, the one that minimizes the mean-square estimation error, is called the *Kalman gain.*

The covariance matrix associated with the optimal estimate may now be computed. Referring to Eq. (5.5.11), we have

$$\mathbf{P}_k = (\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)\mathbf{P}_k^-(\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)^T + \mathbf{K}_k\mathbf{R}_k\mathbf{K}_k^T \tag{5.5.18}$$

$$= \mathbf{P}_k^- - \mathbf{K}_k\mathbf{H}_k\mathbf{P}_k^- - \mathbf{P}_k^-\mathbf{H}_k^T\mathbf{K}_k^T + \mathbf{K}_k(\mathbf{H}_k\mathbf{P}_k^-\mathbf{H}_k^T + \mathbf{R}_k)\mathbf{K}_k^T \tag{5.5.19}$$

Routine substitution of the optimal gain expression, Eq. (5.5.17), into Eq. (5.5.19) leads to

$$\mathbf{P}_k = \mathbf{P}_k^- - \mathbf{P}_k^-\mathbf{H}_k^T(\mathbf{H}_k\mathbf{P}_k^-\mathbf{H}_k^T + \mathbf{R}_k)^{-1}\mathbf{H}_k\mathbf{P}_k^- \tag{5.5.20}$$

or

$$\mathbf{P}_k = \mathbf{P}_k^- - \mathbf{K}_k(\mathbf{H}_k\mathbf{P}_k^-\mathbf{H}_k^T + \mathbf{R}_k)\mathbf{K}_k^T \tag{5.5.21}$$

or

$$\mathbf{P}_k = (\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)\mathbf{P}_k^- \tag{5.5.22}$$

Note that we have four expressions for computing the updated $\mathbf{P}_k$ from the prior $\mathbf{P}_k^-$. Three of these, Eqs. (5.5.20), (5.5.21), and (5.5.22), are only valid for the optimal gain condition. However, Eq. (5.5.18) is valid for any gain, optimal or suboptimal. All four equations yield identical results for optimal gain with perfect arithmetic. We note, though, that in the real engineering world Kalman filtering is a numerical procedure, and some of the *P*-update equations may perform better numerically than others under unusual conditions. More will be said of this later in Chapter 6. For now, we will list the simplest update equation, that is, Eq. (5.5.22), as the usual way to update the error covariance. One should remember, though, that there are alternative equations for implementing the error covariance update.

We now have a means of assimilating the measurement at t_k by the use of Eq. (5.5.8) with $\mathbf{K}_k$ set equal to the Kalman gain as given by Eq. (5.5.17). Note that we need $\hat{\mathbf{x}}_k^-$ and $\mathbf{P}_k^-$ to accomplish this, and we can anticipate a similar need at the next step in order to make optimal use of the measurement $\mathbf{z}_{k+1}$. The updated estimated $\hat{\mathbf{x}}_k$ is easily projected ahead via the transition matrix. We are justified in ignoring the contribution of $\mathbf{w}_k$ in Eq. (5.5.1) because it has zero mean and is not correlated with any of the previous w's.* Thus, we have

$$\hat{\mathbf{x}}_{k+1}^- = \boldsymbol{\phi}_k\hat{\mathbf{x}}_k \tag{5.5.23}$$

* Recall that in our notation $\mathbf{w}_k$ is the process noise that accumulates during the step ahead from t_k to t_{k+1}. This is purely a matter of notation (but an important one), and in some books it is denoted as $\mathbf{w}_{k+1}$ rather than $\mathbf{w}_k$ (6, 9). Consistency in notation is the important thing here. Conceptually, we are thinking of doing real-time filtering in contrast to smoothing, which we usually think of doing offline (see Chapter 8). Therefore, if we begin with a discrete ARMA model, the white forcing sequence $w(k)$ must conform to this same step-ahead notation, and we need to restrict the order of the MA part of the model to be at least one less than the AR part (see Section 5.3). This then assures us that $w(k)$ contributes only to the state vector *after* time t_k and not before.

The error covariance matrix associated with $\hat{\mathbf{x}}_{k+1}^-$ is obtained by first forming the expression for the a priori error

$$\begin{aligned}\mathbf{e}_{k+1}^- &= \mathbf{x}_{k+1} - \hat{\mathbf{x}}_{k+1}^- \\ &= (\boldsymbol{\phi}_k \mathbf{x}_k + \mathbf{w}_k) - \boldsymbol{\phi}_k \hat{\mathbf{x}}_k \\ &= \boldsymbol{\phi}_k \mathbf{e}_k + \mathbf{w}_k \end{aligned} \tag{5.5.24}$$

We now note that $\mathbf{w}_k$ and $\mathbf{e}_k$ have zero crosscorrelation, because $\mathbf{w}_k$ is the process noise for the step ahead of t_k. Thus, we can write the expression for $\mathbf{P}_{k+1}^-$ as

$$\begin{aligned}\mathbf{P}_{k+1}^- &= E[\mathbf{e}_{k+1}^- \mathbf{e}_{k+1}^{-T}] = E[(\boldsymbol{\phi}_k \mathbf{e}_k + \mathbf{w}_k)(\boldsymbol{\phi}_k \mathbf{e}_k + \mathbf{w}_k)^T] \\ &= \boldsymbol{\phi}_k \mathbf{P}_k \boldsymbol{\phi}_k^T + \mathbf{Q}_k \end{aligned} \tag{5.5.25}$$

We now have the needed quantities at time t_{k+1}, and the measurement $\mathbf{z}_{k+1}$ can be assimilated just as in the previous step.

Equations (5.5.8), (5.5.17), (5.5.22), (5.5.23), and (5.5.25) comprise the Kalman filter recursive equations. It should be clear that once the loop is entered, it can be continued ad infinitum. The pertinent equations and the sequence of computational steps are shown pictorially in Fig. 5.8. This summarizes what is now known as the *Kalman filter*.

Before we proceed to some examples, it is interesting to reflect on the Kalman filter in perspective. If you were to stumble onto the recursive process of Fig. 5.8 without benefit of previous history, you might logically ask, "Why in the world did somebody call that a filter? It looks more like a computer algorithm." You would, of course, be quite right in your observation. The Kalman filter is just a computer algorithm for processing discrete measurements (the input) into optimal estimates (the output). Its roots, though, go back to the

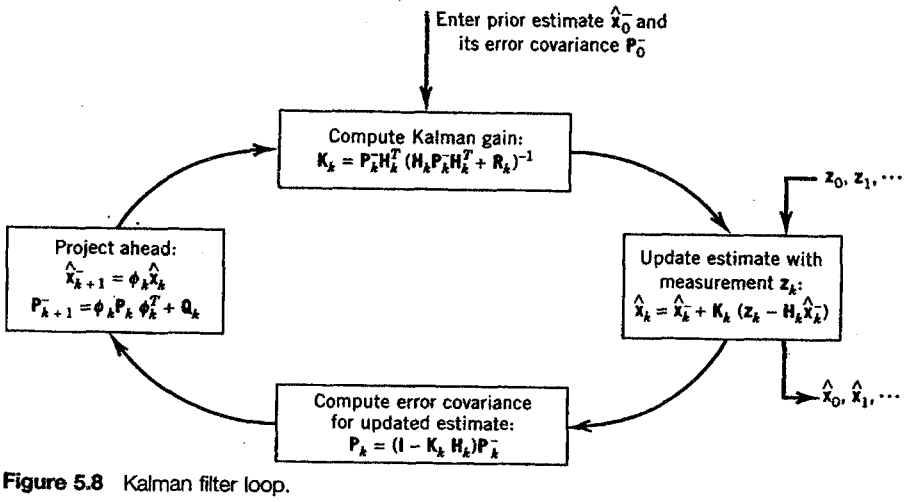


Figure 5.8 Kalman filter loop.

days when filters were made of electrical elements wired together in such a way as to yield the desired frequency response. The design was often heuristic. Wiener then came on the scene in the 1940s and added a more sophisticated type of filter problem. The end result of his solution was a filter weighting function or a corresponding transfer function in the complex domain. Implementation in terms of electrical elements was left as a further exercise for the designer. The discrete-time version of the Wiener problem remained unsolved (in a practical sense, at least) until Kalman's paper of 1960. Even though his presentation appeared to be quite abstract at first glance, engineers soon realized that this work provided a practical solution to a number of unsolved filtering problems, especially in the field of navigation. More than 35 years have elapsed since Kalman's original paper, and there are still numerous current papers dealing with new applications and variations on the basic Kalman filter. It has withstood the test of time!