5. OPTIMAL LINEAR SMOOTHING

Smoothing is a non-real-time data processing scheme that uses all measurements between 0 and T to estimate the state of a system at a certain time t, where $0 \leqslant t \leqslant T$. The smoothed estimate of $\underline{x}(t)$ based on all the measurements between 0 and T is denoted by $\underline{\hat{x}}(t|T)$. An optimal smoother can be thought of as a suitable combination of two optimal filters. One of the filters, called a "forward filter," operates on all the data before time t and produces the estimate $\underline{\hat{x}}(t)$; the other filter, called a "backward filter," operates on all the data after time t and produces the estimate $\underline{\hat{x}}_b(t)$. Together these two filters utilize *all* the available information; see Fig. 5.0-1. The two estimates they provide have

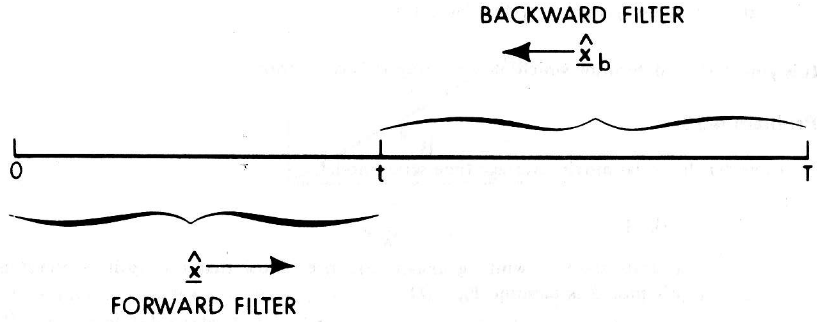

Figure 5.0-1 Relationship of Forward and Backward Filters

uncorrelated errors, since process and measurement noises are assumed white. This suggests that the optimal combination of $\underline{\hat{x}}(t)$ and $\underline{\hat{x}}_b(t)$ will, indeed, yield the optimal smoother; proof of this assertion can be found in Ref. 1.

Three types of smoothing are of interest. In *fixed-interval smoothing*, the initial and final times 0 and T are fixed and the estimate $\underline{\hat{x}}(t|T)$ is sought, where t varies from 0 to T. In *fixed-point smoothing*, t is fixed and $\underline{\hat{x}}(t|T)$ is sought as T increases. In *fixed-lag smoothing*, $\underline{\hat{x}}(T-\Delta|T)$ is sought as T increases, with Δ held fixed.

In this chapter the two-filter form of optimal smoother is used as a point of departure. Fixed-interval, fixed-point and fixed-lag smoothers are derived for the continuous-time case, with corresponding results presented for the discrete-time case, and several examples are discussed.

5.1 FORM OF THE OPTIMAL SMOOTHER

Following the lead of the previous chapter, we seek the optimal smoother in the form

$$\underline{\hat{x}}(t|T) = A\underline{\hat{x}}(t) + A'\underline{\hat{x}}_b(t) \tag{5.1-1}$$

where A and A′ are weighting matrices to be determined. Replacing each of the estimates in this expression by the corresponding true value plus an estimation error, we obtain

$$\underline{\tilde{x}}(t|T) = [A + A' - I]\ \underline{x}(t) + A\underline{\tilde{x}}(t) + A'\underline{\tilde{x}}_b(t) \tag{5.1-2}$$

For unbiased filtering errors, $\underline{\tilde{x}}(t)$ and $\underline{\tilde{x}}_b(t)$, we wish to obtain an unbiased smoothing error, $\underline{\tilde{x}}(t|T)$; thus, we set the expression in brackets to zero. This yields

$$A' = I - A \tag{5.1-3}$$

and, consequently,

$$\underline{\tilde{x}}(t|T) = A\underline{\tilde{x}}(t) + (I - A)\ \underline{\tilde{x}}_b(t) \tag{5.1-4}$$

Computing the smoother error covariance, we find

$$\begin{aligned} P(t|T) &= E[\underline{\tilde{x}}(t|T)\ \underline{\tilde{x}}^T(t|T)] \\ &= AP(t)\ A^T + (I - A)\ P_b(t)(I - A)^T \end{aligned} \tag{5.1-5}$$

where product terms involving $\underline{\tilde{x}}(t)$ and $\underline{\tilde{x}}_b(t)$ do not appear. P(t|T) denotes the smoother error covariance matrix, while P(t) and $P_b(t)$ denote forward and backward optimal filter error covariance matrices, respectively.

OPTIMIZATION OF THE SMOOTHER

Once again, following the previous chapter, we choose that value of A which minimizes the trace of P(t|T). Forming this quantity, differentiating with respect to A and setting the result to zero, we find

$$0 = 2AP + 2(I - A)P_b(-I) \tag{5.1-6}$$

or

$$A = P_b(P + P_b)^{-1} \tag{5.1-7}$$

and, correspondingly

$$I - A = P(P + P_b)^{-1} \tag{5.1-8}$$

Inserting these results into Eq. (5.1-5), we obtain

$$P(t|T) = P_b(P + P_b)^{-1}P(P + P_b)^{-1}P_b + P(P + P_b)^{-1}P_b(P + P_b)^{-1}P \tag{5.1-9}$$

By systematically combining factors in each of the two right-side terms of this equation, we arrive at a far more compact result. The algebraic steps are sketched below,

$$\begin{aligned} P(t|T) &= P_b(P + P_b)^{-1}P\left(I + P_b^{-1}P\right)^{-1} + P(P + P_b)^{-1}P_b\left(P^{-1}P_b + I\right)^{-1} \\ &= P_b(P + P_b)^{-1}\left(P^{-1} + P_b^{-1}\right)^{-1} + P(P + P_b)^{-1}\left(P^{-1} + P_b^{-1}\right) \\ &= \left(P^{-1} + P_b^{-1}\right)^{-1} \end{aligned} \tag{5.1-10}$$

or

$$P^{-1}(t|T) = P^{-1}(t) + P_b^{-1}(t) \tag{5.1-11}$$

From Eq. (5.1-11), P(t|T) ≤ P(t), which means that the smoothed estimate of $\underline{x}(t)$ is always better than or equal to its filtered estimate. This is shown graphically in Fig. 5.1-1. Performing similar manipulations on Eq. (5.1-1), we find

$$\begin{aligned} \hat{\underline{x}}(t|T) &= A\hat{\underline{x}}(t) + (I - A)\,\hat{\underline{x}}_b(t) \\ &= P_b(P + P_b)^{-1}\,\hat{\underline{x}}(t) + P(P + P_b)^{-1}\,\hat{\underline{x}}_b(t) \\ &= \left(P^{-1} + P_b^{-1}\right)^{-1} P^{-1}\hat{\underline{x}}(t) + \left(P^{-1} + P_b^{-1}\right)^{-1} P_b^{-1}\hat{\underline{x}}_b(t) \\ &= P(t|T)\left[P^{-1}(t)\,\hat{\underline{x}}(t) + P_b^{-1}(t)\,\hat{\underline{x}}_b(t)\right] \end{aligned} \tag{5.1-12}$$

Equations (5.1-11) and (5.1-12) are the results of interest.

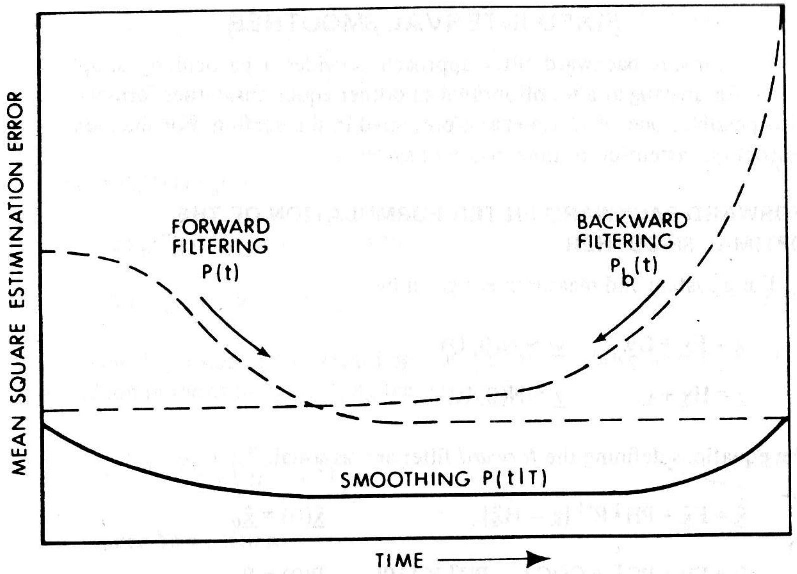

Figure 5.1-1 Advantage of Performing Optimal Smoothing

REINTERPRETATION OF PREVIOUS RESULTS

It is interesting to note that we could have arrived at these expressions by interpretation of optimal filter relationships. In the subsequent analogy, estimates $\hat{\underline{x}}_b$ from the backward filter will be thought of as providing "measurements" with which to update the forward filter. In the corresponding "measurement equation," H = I, as the total state vector is estimated by the backward filter. Clearly, the "measurement error" covariance matrix is then represented by P_b. From Eq (4.2-19), in which $P_k^{-1}(-)$ and $P_k^{-1}(+)$ are now interpreted as $P^{-1}(t)$ and $P^{-1}(t|T)$, respectively, we obtain

$$P^{-1}(t|T) = P^{-1}(t) + P_b^{-1}(t) \tag{5.1-13}$$

Equations (4.2-16b) and (4.2-20) provide the relationships

$$K_k = P_k(+)\,H_k^T R_k^{-1} \rightarrow P(t|T)\,P_b^{-1}(t)$$

$$(I - K_k H_k) = P_k(+)\,P_k^{-1}(-) \rightarrow P(t|T)\,P^{-1}(t)$$

which, when inserted in Eq. (4.2-5) yield the result

$$\hat{\underline{x}}(t|T) = P(t|T)\left[P^{-1}(t)\,\hat{\underline{x}}(t) + P_b^{-1}(t)\,\hat{\underline{x}}_b(t)\right], \tag{5.1-14}$$

where $\hat{\underline{x}}_k(-)$ and $\hat{\underline{x}}_k(+)$ have been interpreted as $\hat{\underline{x}}(t)$ and $\hat{\underline{x}}(t|T)$, respectively. Thus, we arrive at the same expressions for the optimal smoother and its error covariance matrix as obtained previously.

5.2 OPTIMAL FIXED-INTERVAL SMOOTHER

The forward-backward filter approach provides a particularly simple mechanism for arriving at a set of optimal smoother equations. Other formulations are also possible, one of which is also presented in this section. For the moment, we restrict our attention to time-invariant systems.

FORWARD-BACKWARD FILTER FORMULATION OF THE OPTIMAL SMOOTHER

For a system and measurement given by

$$\dot{\underline{x}} = F\underline{x} + G\underline{w}, \qquad \underline{w} \sim N(\underline{0}, Q)$$

$$\underline{z} = H\underline{x} + \underline{v}, \qquad \underline{v} \sim N(\underline{0}, R) \tag{5.2-1}$$

the equations defining the *forward* filter are, as usual,

$$\dot{\hat{\underline{x}}} = F\hat{\underline{x}} + PH^TR^{-1}[\underline{z} - H\hat{\underline{x}}], \qquad \hat{\underline{x}}(0) = \hat{\underline{x}}_0 \tag{5.2-2}$$

$$\dot{P} = FP + PF^T + GQG^T - PH^TR^{-1}HP, \qquad P(0) = P_0 \tag{5.2-3}$$

The equations defining the *backward* filter are quite similar. Since this filter runs backward in time, it is convenient to set $\tau = T - t$. Writing Eq. (5.2-1) in terms of τ, gives*

$$\frac{d}{d\tau}\underline{x} = -\frac{d}{dt}\underline{x}$$

$$= -F\underline{x} - G\underline{w} \tag{5.2-4}$$

$$\underline{z}(\tau) = H\underline{x} + \underline{v} \tag{5.2-5}$$

for $0 \leqslant \tau \leqslant T$. By analogy with the forward filter, the equations for the backward filter can be written changing F to –F and G to –G. This results in

$$\frac{d}{d\tau}\hat{\underline{x}}_b = -F\hat{\underline{x}}_b + P_bH^TR^{-1}[\underline{z} - H\hat{\underline{x}}_b] \tag{5.2-6}$$

$$\frac{d}{d\tau}P_b = -FP_b - P_bF^T + GQG^T - P_bH^TR^{-1}HP_b \tag{5.2-7}$$

At time t = T, the smoothed estimate must be the same as the forward filter estimate. Therefore, $\hat{\underline{x}}(T|T) = \hat{\underline{x}}(T)$ and $P(T|T) = P(T)$. The latter result, in combination with Eq. (5.1-11), yields the boundary condition on P_b^{-1},

*In this chapter, a dot denotes differentiation with respect to (forward) time t. Differentiation with respect to backward time is denoted by $d/d\tau$.

$$P_b^{-1}(t = T) = 0 \quad \text{or} \quad P_b^{-1}(\tau = 0) = 0 \tag{5.2-8}$$

but the boundary condition on $\hat{\underline{x}}_b(T)$ is yet unknown. One way of avoiding this problem is to transform Eq. (5.2-6) by defining the new variable

$$\underline{s}(t) = P_b^{-1}(t)\,\hat{\underline{x}}_b(t) \tag{5.2-9}$$

where, since $\hat{\underline{x}}_b(T)$ is finite, it follows that

$$\underline{s}(t = T) = \underline{0} \quad \text{or} \quad \underline{s}(\tau = 0) = \underline{0} \tag{5.2-10}$$

Computational considerations regarding the equations above lead us to their reformulation in terms of P_b^{-1}. Using the relationship

$$\frac{d}{d\tau}P_b^{-1} = -P_b^{-1}\left(\frac{d}{d\tau}P_b\right)P_b^{-1} \tag{5.2-11}$$

Eq. (5.2-7) can be written as

$$\frac{d}{d\tau}P_b^{-1} = P_b^{-1}F + F^TP_b^{-1} - P_b^{-1}GQG^TP_b^{-1} + H^TR^{-1}H \tag{5.2-12}$$

for which Eq. (5.2-8) is the appropriate boundary condition. Differentiating Eq. (5.2-9) with respect to τ and employing Eqs. (5.2-6) and (5.2-12) and manipulating, yields

$$\frac{d}{d\tau}\underline{s} = (F^T - P_b^{-1}GQG^T)\,\underline{s} + H^TR^{-1}\underline{z} \tag{5.2-13}$$

for which Eq. (5.2-10) is the appropriate boundary condition. Equations (5.1-11, 12) and (5.2-2, 3, 12, 13) define the optimal smoother. See Table 5.2-1, in which alternate expressions for $\hat{\underline{x}}(t|T)$ and $P(t|T)$, which obviate the need for unnecessary matrix inversions, are also presented (Ref. 1). These can be verified by algebraic manipulation. The results presented in Table 5.2-1 are for the general, time-varying case.

ANOTHER FORM OF THE EQUATIONS

Several other forms of the smoothing equations may also be derived. One is the Rauch-Tung-Striebel form (Ref. 3), which we utilize in the sequel. This form, which does not involve backward filtering *per se*, can be obtained by differentiating Eqs. (5.1-11) and (5.1-12) and using Eq. (5.2-12). It is given by Eqs. (5.2-2) and (5.2-3) and*

*From this point on, all discussion pertains to the general time-varying case unless stated otherwise. However, for notational convenience, explicit dependence of F, G, H, Q, R upon t may not be shown.

TABLE 5.2-1 SUMMARY OF CONTINUOUS, FIXED-INTERVAL OPTIMAL LINEAR SMOOTHER EQUATIONS, TWO-FILTER FORM

System Model	$\dot{\underline{x}}(t) = F(t)\underline{x}(t) + G(t)\underline{w}(t), \quad \underline{w}(t) \sim N[\underline{0}, Q(t)]$
Measurement Model	$\underline{z}(t) = H(t)\underline{x}(t) + \underline{v}(t), \quad \underline{v}(t) \sim N[\underline{0}, R(t)]$
Initial Conditions	$E[\underline{x}(0)] = \hat{\underline{x}}_0, \quad E[(\underline{x}(0) - \hat{\underline{x}}_0)(\underline{x}(0) - \hat{\underline{x}}_0)^T] = P_0$
Other Assumptions	$E[\underline{w}(t_1)\underline{v}^T(t_2)] = 0$ for all t_1, t_2; $R^{-1}(t)$ exists
Forward Filter	$\dot{\hat{\underline{x}}}(t) = F(t)\hat{\underline{x}}(t) + P(t)H^T(t)R^{-1}(t)[\underline{z}(t) - H(t)\hat{\underline{x}}(t)], \quad \hat{\underline{x}}(0) = \hat{\underline{x}}_0$
Error Covariance Propagation	$\dot{P}(t) = F(t)P(t) + P(t)F^T(t) + G(t)Q(t)G^T(t) - P(t)H^T(t)R^{-1}(t)H(t)P(t), \quad P(0) = P_0$
Backward Filter ($\tau = T - t$)	$\frac{d}{d\tau}\underline{s}(T-\tau) = [F^T(T-\tau) - P_b^{-1}(T-\tau)G(T-\tau)Q(T-\tau)G^T(T-\tau)]\underline{s}(T-\tau) + H^T(T-\tau)R^{-1}(T-\tau)\underline{z}(T-\tau), \quad \underline{s}(0) = \underline{0}$
Error Covariance Propagation ($\tau = T - t$)	$\frac{d}{d\tau}P_b^{-1}(T-\tau) = P_b^{-1}(T-\tau)F(T-\tau) + F^T(T-\tau)P_b^{-1}(T-\tau) - P_b^{-1}(T-\tau)G(T-\tau)Q(T-\tau)G^T(T-\tau)P_b^{-1}(T-\tau) + H^T(T-\tau)R^{-1}(T-\tau)H(T-\tau), \quad P_b^{-1}(0) = 0$
Optimal Smoother	$\hat{\underline{x}}(t\|T) = P(t\|T)[P^{-1}(t)\hat{\underline{x}}(t) + \underline{s}(t)] = [I + P(t)P_b^{-1}(t)]\hat{\underline{x}}(t) + P(t\|T)\underline{s}(t)$
Error Covariance Propagation	$P(t\|T) = [P^{-1}(t) + P_b^{-1}(t)]^{-1} = P(t) - P(t)P_b^{-1}(t)[I + P(t)P_b^{-1}(t)]^{-1}P(t)$

$$\dot{\hat{\underline{x}}}(t|T) = F\hat{\underline{x}}(t|T) + GQG^TP^{-1}(t)[\hat{\underline{x}}(t|T) - \hat{\underline{x}}(t)] \tag{5.2-14}$$

$$\dot{P}(t|T) = [F + GQG^TP^{-1}(t)]P(t|T) + P(t|T)[F + GQG^TP^{-1}(t)]^T - GQG^T \tag{5.2-15}$$

Equations (5.2-14) and (5.2-15) are integrated *backwards* from $t = T$ to $t = 0$, with starting conditions given by $\hat{\underline{x}}(T|T) = \hat{\underline{x}}(T)$ and $P(T|T) = P(T)$. Figure 5.2-1 is a block diagram of the optimal smoother. Note that the operation which produces the smoothed state estimate does *not* involve the processing of actual measurement data. It does utilize the *complete* filtering solution, however, so that problem must be solved first. Thus, fixed-interval smoothing cannot be done real-time, on-line. It must be done after all the measurement data are collected. Note also that $P(t|T)$ is a continuous time function even where $P(t)$ may be discontinuous, as can be seen from Eq. (5.2-15).

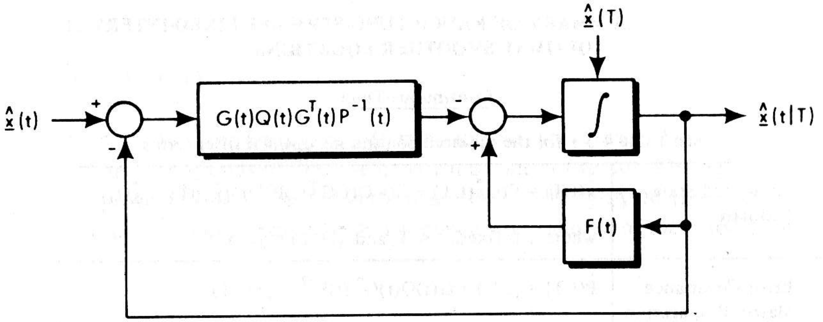

Figure 5.2-1 Diagram of Rauch-Tung-Striebel Fixed-Interval Continuous Optimal Smoother ($t \leq T$)

All smoothing algorithms depend, in some way, on the forward filtering solution. Therefore, accurate filtering is prerequisite to accurate smoothing. Since fixed-interval smoothing is done off-line, after the data record has been obtained, computation speed is not usually an important factor. However, since it is often necessary to process long data records involving many measurements, computation error due to computer roundoff is an important factor. Hence, it is desirable to have recursion formulas that are relatively insensitive to computer roundoff errors. These are discussed in Chapter 8 in connection with optimal filtering; the extension to optimal smoothing is straightforward (Ref. 1).

The continuous-time and corresponding discrete-time (Ref. 3) fixed-interval Rauch-Tung-Striebel optimal smoother equations are summarized in Table 5.2-2. In the discrete-time case the intermediate time variable is k, with final time denoted by N. $P_{k|N}$ corresponds to $P(t|T)$, $\hat{\underline{x}}_{k|N}$ corresponds to $\hat{\underline{x}}(t|T)$ and the single subscripted quantities $\hat{\underline{x}}_k$ and P_k refer to the discrete optimal filter solution. Another, equivalent fixed-interval smoother is given in Ref. 4 and the case of correlated measurement noise is treated in Ref. 5.

SMOOTHABILITY

A state is said to be *smoothable* if an optimal smoother provides a state estimate superior to that obtained when the final optimal filter estimate is extrapolated backwards in time. In Ref. 2, it is shown that *only those states which are controllable by the noise driving the system state vector are smoothable.* Thus, constant states are not smoothable, whereas randomly time-varying states are smoothable. This smoothability condition is explored below.

Consider the case where there are no system disturbances. From Eq. (5.2-14), we find ($Q = 0$)

$$\dot{\hat{\underline{x}}}(t|T) = F\hat{\underline{x}}(t|T) \tag{5.2-16}$$

TABLE 5.2-2 SUMMARY OF RAUCH-TUNG-STRIEBEL FIXED-INTERVAL OPTIMAL SMOOTHER EQUATIONS

Continuous-Time

(See Table 4.3-1 for the required continuous optimal filter terms)

Smoothed State Estimate	$\dot{\hat{\underline{x}}}(t\|T) = F(t)\hat{\underline{x}}(t\|T) + G(t)Q(t)G^T(t)P^{-1}(t)[\hat{\underline{x}}(t\|T) - \hat{\underline{x}}(t)]$ where T is fixed, $t \leq T$, and $\hat{\underline{x}}(T\|T) = \hat{\underline{x}}(T)$.
Error Covariance Matrix Propagation	$\dot{P}(t\|T) = [F(t) + G(t)Q(t)G^T(t)P^{-1}(t)]P(t\|T) + P(t\|T)[F(t) + G(t)Q(t)G^T(t)P^{-1}(t)]^T - G(t)Q(t)G^T(t)$ where $P(T\|T) = P(T)$

Discrete-Time

(See Table 4.2-1 for the required discrete optimal filter terms)

Smoothed State Estimate	$\hat{\underline{x}}_{k\|N} = \hat{\underline{x}}_k(+) + A_k[\hat{\underline{x}}_{k+1\|N} - \hat{\underline{x}}_{k+1}(-)]$ where $A_k = P_k(+)\Phi_k^T P_{k+1}^{-1}(-)$, $\hat{\underline{x}}_{N\|N} = \hat{\underline{x}}_N(+)$ for $k = N-1$.
Error Covariance Matrix Propagation	$P_{k\|N} = P_k(+) + A_k[P_{k+1\|N} - P_{k+1}(-)]A_k^T$ where $P_{N\|N} = P_N(+)$ for $k = N-1$

The solution is ($\hat{\underline{x}}(T|T) = \hat{\underline{x}}(T)$)

$$\hat{\underline{x}}(t|T) = \Phi(t,T)\hat{\underline{x}}(T) \tag{5.2-17}$$

Thus, the optimal fixed-interval smoother estimate, when $Q = 0$, is the final optimal filter estimate extrapolated backwards in time. The corresponding smoothed state error covariance matrix behavior is governed by

$$\dot{P}(t|T) = FP(t|T) + P(t|T)F^T \tag{5.2-18}$$

for which the solution is [$P(T|T) = P(T)$]

$$P(t|T) = \Phi(t,T)\,P(T)\,\Phi^T(t,T) \tag{5.2-19}$$

If, in addition, $F = 0$, it follows that $\Phi(t,T) = I$ and hence, that

$$\hat{\underline{x}}(t|T) = \hat{\underline{x}}(T) \tag{5.2-20}$$

and

$$P(t|T) = P(T) \tag{5.2-21}$$

for all $t \leq T$. That is, *smoothing offers no improvement over filtering when $F = Q = 0$.* This corresponds to the case in which a constant vector is being estimated with no process noise present. Identical results clearly apply to the m constant states of an n^{th} order system ($n > m$); hence, the validity of the smoothability condition.

Example 5.2-1

This spacecraft tracking problem was treated before in Example 4.3-2. The underlying equations are:

$$\dot{x} = w, \qquad w \sim N(0,q)$$

$$z = x + v, \qquad v \sim N(0,r)$$

and the steady-state optimal filter solution was shown to be $p(t) = \alpha$, where $\alpha = \sqrt{rq}$. Examine the steady-state, fixed-interval optimal smoother both in terms of (1) forward-backward optimal filters, and (2) Rauch-Tung-Striebel form.

Part 1 – The forward filter Riccati equation is (f=0, g=h=1)

$$\dot{p} = q - p^2/r$$

which, in the steady state ($\dot{p}$=0), yields $p = \sqrt{rq} = \alpha$. The backward filter Riccati equation is from Eq. (5.2-7)

$$\frac{d}{d\tau}p_b = q - p_b^2/r$$

which has the steady state $p_b = \sqrt{rq} = \alpha$. Thus, we find, for the smoothed covariance

$$p(t|T) = \left(p^{-1}(t) + p_b^{-1}(t)\right)^{-1}$$

$$= \frac{\alpha}{2}$$

which is half the optimal filter covariance. Consequently,

$$\hat{x}(t|T) = p(t|T)\,[\hat{x}(t)/p(t) + \hat{x}_b(t)/p_b(t)]$$

$$= \frac{1}{2}\,[\hat{x}(t) + \hat{x}_b(t)] \tag{5.2-22}$$

The smoothed estimate of x is the *average* of forward plus backward estimates, in steady state.

Part 2 – In Rauch-Tung-Striebel form, the steady-state smoothed covariance matrix differential equation (Table 5.2-2, T fixed, $t \leq T$) is