JOURNAL OF GUIDANCE, CONTROL, AND DYNAMICS
Vol. 20, No. 1, January–February 1997

Covariance Projection Methods for Filters and Smoothers

Stephen R. McReynolds
Lockheed Martin Corporation, Valley Forge, Pennsylvania 19101

The covariance projection equations (error budgets), which express the error covariance as the sum of covariances due to different error sources, are derived from the sensitivity analysis equations of optimal estimators for batch processors, filters, and smoothers for discrete and continuous problems. The sensitivity analysis of optimal estimation covariances are shown to be the same as previously derived suboptimal estimation sensitivities, which characterize the loss of estimator performance due to the use of nonoptimal values for error covariances. Projection equations for steady-state filter and smoother values are obtained. Some steady-state filter examples are considered.

Nomenclature

E = expectation operator
F = state partials for continuous dynamics
H = measurement Jacobian for batch problem
H_k = measurement Jacobian of discrete measurement at t_k
K = Kalman gain
K_k = Kalman gain at t_k
P = covariance of errors in $\hat{x}$
$P^{f,ss}$ = steady-state filter covariance
$P_i^{f,ss}$ = steady-state filter covariance due to ith error source
$P^{(m)}$ = projection covariance of error in $\hat{x}$ due to errors in measurements
$P_i^{(m)}$ = projection covariance of error in $\hat{x}$ due to errors in measurements type i
P_{k+1}^{p} = covariance of in the errors in the predicted estimate
P^s = smoother covariance
$P^{s,ss}$ = steady-state smoother covariance
$P_i^{s,ss}$ = steady-state smoother covariance due to ith error source
P_0 = covariance of errors in $\hat{x}_0$
$P^{(0)}$ = projection covariance of error in $\hat{x}$ due to errors a priori in estimate
P_k^f = covariance of errors in the filtered estimated
Q_k = process noise covariance at time t_k
$Q_k(j)$ = process noise covariance of jth error source at time t_k
$R_i(t)$ = measurement noise covariance of ith error source at time t
R_k = measurement noise covariance at time t_k
$R_k(j)$ = measurement noise covariance of jth error source at time t_k
t_k = discrete time
u_k = process noise vector at time t_k
$u_k(j)$ = process noise vector of jth error source at time t_k
v = random measurement error vector
v_k = random measurement error vector at t_k
$v_k(j)$ = random measurement error vector of jth error source at t_k
x = state vector
$\hat{x}$ = optimal estimate of the state vector after the measurements have been added
$\tilde{x}$ = error in the optimal estimate, $\hat{x} - x$
$\tilde{x}(j)$ = error in the optimal estimate due to jth error source
x_k = system state at time t_k
$\hat{x}_k^f$ = filtered estimate at t_k
$\hat{x}_k^p$ = predicted estimate at t_k
$\hat{x}_k^s$ = smoothed estimate at t_k
$\hat{x}_0$ = a priori estimate
z = measurement vector
Γ_k = coefficient matrix of process noise for discrete system transition
λ_i = scale factor for the ith error source
Φ_k = transition matrix for discrete system
ψ = state vector

Introduction

FOR large-scale estimation problems, many modeling assumptions are required in constructing an estimation algorithm. Some of these assumptions may be more important than others. In developing an estimation algorithm, one would like to quantify the impacts of the individual error sources on the final solution to support either the design or application of estimation algorithms. Two methods have been developed for assessing impacts: error budgets and sensitivity analysis.

In error budget analysis, the error of the estimator (batch or sequential fitler/smoother) is expressed as the sum of individual, statistically independent error sources. This implies that the covariance is given by the sum of individual covariances (hence our use of the term projection). The error budget analysis for discrete sequential filters was proposed by Gelb (Ref. 1, p. 260). The equations for propagating the various components of projection decomposition paralleled those of the filter themselves. These equations grew out of suboptimal sensitivity analysis.

Suboptimal sensitivity analysis originated with least squares estimation. The consider option in Jet Propulsion Laboratory's (JPL's)

Stephen R. McReynolds, a Senior Systems Engineer with Lockheed Martin, received his Ph.D. in Applied Mathematics from Harvard University in 1966. He has worked at various aerospace companies, including the Jet Propulsion Laboratory. He specialized in control and estimation applications. He coauthored (with Peter Dyer) *The Computation and Theory of Optimal Control*, published by Academic Press in 1970. With A.E. Bryson, he originated the successive sweep method for nonlinear optimization. He also originated (with Peter Dyer) the square-root information filter. He is a Senior Member of AIAA.

Received April 25, 1996; presented as Paper 96-3739 at the AIAA Guidance, Navigation, and Control Conference, San Diego, CA, July 29–31, 1996; revision received Sept. 8, 1996; accepted for publication Sept. 30, 1996.

Single Precision Orbit Determination Program[2] helped to evaluate the impact of unsolved-for parameters on the solution covariance. Such a capability is also implemented within JPL's Double Precision Orbit Determination Program and The Aerospace Corporation's TRACE software. In these programs two covariances are calculated: a computed covariance reflecting the modeled error sources and an additional covariance reflecting the errors in unestimated parameters. The true covariance is equal to the sum of both covariances. By examining the contribution of the various error sources, using sensitivity and perturbation matrices,[3] the relative importance of different error sources was determined. Soong[4] developed suboptimal sensitivity equations for a batch processor to assess errors in the a priori information.

Within a decade of the formulation of optimal sequential filter smoothers,[5–7] suboptimal sensitivity methods were developed for estimators. Early work by Nishimura[8,9] and Heffes[10] derived equations for actual covariances using incorrect a priori statistics. Related to these papers is the work of Friedland,[11] who developed covariance calculations for an incorrect gain. Mehra[12] extended suboptimal analysis to smoothers. Along these lines are the papers by Griffis and Sage,[13–15] which treat filtering and smoothing and introduce the notation of sensitivity functions, which represent incremental losses of performance due to the suboptimal conditions. They also developed error analysis for errors in the system equation. Other methods were developed by Joseph,[16] Meditch,[17] Aoki,[18] and Aoki and Huddle,[19] and Huddle and Wismer[20] focused on developing techniques for developing reduced-state filters. Their work was influenced by Luenberger's[21] observer theory. Gelb (Ref. 1, p. 250) summarizes sensitivity equations for a suboptimal filter.

All of the equations developed in the referenced suboptimal sensitivity literature are for suboptimal estimators. These equations examine the impacts on performance when actual statistics differ from those used by the estimator. Another concept of sensitivity would be that of optimal sensitivity, i.e., the change of performance of the optimal estimator due to a change in statistical assumptions. The author is not familiar with any references on optimal sensitivity. In this paper, it will be shown that suboptimal sensitivity and optimal sensitivity, although different concepts, result in the same equations. The advantage of using optimal sensitivity equations is that it is easier to develop these equations than suboptimal sensitivity equations, particularly for the smoother. Optimal sensitivity equations can be developed by directly differentiating the optimal estimator equations.

The concepts of sensitivity analysis and error budgets (covariance projections) are shown to be closely related. Covariance projections can be developed by way of developing optimal sensitivity equations for both batch and sequential filters/smoothers, continuous and discrete. It will be shown how the optimal covariance can be expressed as the sum of covariances, projective components, each reflecting a different modeled error source, with equations for these individual contributions. These equations for the projective factors will possess similar structure to that of the estimators. Some examples will be used to demonstrate how covariance projection methods can provide insight into the performance of optimal estimation algorithms.

One word of caution. Traditional applications of error budget concepts often involve global prediction, i.e., the total error can be predicted based on arbitrary changes of the error components. Whereas global prediction holds for suboptimal estimators, which do not adapt to the actual error covariances, global prediction does not apply to optimal filters, which do adapt to actual error covariances.

Covariance Projection Methods and Error Budget Allocations

The concept of error budget allocations, which has been applied to estimation systems for some time, follows naturally from the mathematics of linear systems and Gaussian errors. Let us assume that a single random error q is expressed as a sum of p zero-mean, independent Gaussian variables ϵ_i, such that

$$q = \sum_{i=1}^{p} \epsilon_i \tag{1}$$

Then q is a zero-mean Gaussian random variable with variance

$$\sigma_q^2 = \sum_{i=1}^{p} \sigma_i^2 \tag{2}$$

The set of numbers $[\sigma_i^2]$ represent the error budget allocation for the single random error q. Rather than use the term "error budget allocation," we shall use the term "covariance projection." Covariance projection is the representation of a total error in terms of the individual projection components. The term "projection" is used in Euclidean geometry to define components of a vector within given linear subspaces. If the linear subspaces are orthogonal, the projection components are orthogonal. If the linear spaces add up to the entire space, the sum of the projective components add up to the entire vector. In this case we are applying projection notions to covariance, expressing the covariance as a sum of terms, each representing the contribution for a specific statistically independent error source.

The projection representation enables one to perform a parameter sensitivity analysis. Suppose the variance of the ith error source is scaled by λ_i for all error sources, then the ith projection term would be given by $\lambda_i \sigma_i^2$. The total error would be given by

$$\sigma_q^2(\lambda_i) = \sum_{i=1}^{p} \lambda_i \sigma_i^2 \tag{3}$$

From the preceding equation, we see that

$$\frac{\partial \sigma_q^2}{\partial \lambda_i}(\lambda_i) = \sigma_i^2 \qquad \left.\frac{\partial \sigma_q}{\partial \sigma_i}\right|_{\lambda_i = 1} = \frac{\sigma_i}{\sigma_q} \tag{4}$$

Thus, the projective analysis enables us to calculate the covariance for an arbitrary change in component budget or the sensitivity of the standard deviation of the total error with respect to small changes in the standard deviations of the components.

Covariance Projection Methods for Least Squares Estimation Problems

In the preceding section, the mathematics was not directly applied to an estimation example. In this section it will be shown how error budgets can be applied to optimal least square estimators. Let us consider the simple case of projection analysis of a batch estimation with a priori information. The measurements shall be given by

$$z = Hx + v \tag{5}$$

where

$$E[v] = 0; \qquad E[vv^T] \doteq R \tag{6}$$

It will be assumed that v is uncorrelated with the errors in the a priori information. Borrowing from Kalman filter theory, the optimal estimate is given by

$$K \doteq P_0 H^T \left(H P_0 H^T + R\right)^{-1} \tag{7}$$

$$\hat{x} = \hat{x}_0 + K(z - H\hat{x}_0) \tag{8}$$

The optimal covariance is given by

$$P = (I - KH)P_0 = P_0(I - KH)^T \tag{9}$$

Substituting Eq. (5) into Eq. (8) yields

$$\hat{x} = \hat{x}_0 + K(Hx - H\hat{x}_0) + Kv \tag{10}$$

or

$$\hat{x} - x = (I - KH)(\hat{x}_0 - x) + Kv \tag{11}$$

Multiplying Eq. (11) by its transpose and taking expectations yields

$$P = (I - KH)P_0(I - KH)^T + KRK^T \tag{12}$$

In the preceding equation the first term on the right-hand side represents the contribution of the a priori errors to the error of the optimal estimate and, thus, is the projective component associated with the

a priori information. The last term is the projective component with respect to the measurement errors. Thus, the projective component for a priori errors is given by

$$P^{(0)} = (I - KH)P_0(I - KH)^T \tag{13}$$

and the projective component for measurement errors is given by

$$P^{(m)} = KRK^T \tag{14}$$

Using Eq. (9), Eq. (13) can be written as

$$P^{(0)} = P(I - KH)^T = PP_0^{-1}P_0(I - KH)^T = PP_0^{-1}P \tag{15}$$

The following relationship also holds:

$$K \doteq PH^TR^{-1} \tag{16}$$

Thus, substituting Eq. (16) into Eq. (14) gives

$$P^{(m)} = PH^TR^{-1}HP \tag{17}$$

With the correspondence $P_0^{-1} \leftrightarrow H^TR^{-1}H$, one can see that Eqs. (15) and (17) have the same form. Both quantities are the information matrices with their respective information.

Obviously, this result can be generalized to the cases where the measurements sets are partitioned. Different data sets may be appropriate because 1) they are different data types, 2) they belong to different sensors, and/or 3) they were collected at different times. Because it is assumed that the random measurement errors are independent, one can write

$$H^TR^{-1}H = \sum_{i=1}^{p} H^{(i)T}\left(R^{(i)}\right)^{-1}H^{(i)} \tag{18}$$

where the superscript (i) denotes the ith data set. Substituting Eq. (18) into Eq. (17) gives

$$P^{(m)} = \sum_{i=1}^{p} PH^{(i)T}R^{(i)-1}H^{(i)}P = \sum_{i=1}^{p} P_i^{(m)} \tag{19}$$

where

$$P_i^{(m)} = PH^{(i)T}R^{(i)-1}H^{(i)}P \tag{20}$$

Likewise, if the a priori information is uncorrelated, i.e., there is no correlation between different sets of parameters, then separate expressions can be developed for the different a priori components.

Sensitivity Methods for Least Squares Estimation Problems

In this section it will be shown that projection factors are related to suboptimal sensitivity analysis and that suboptimal and optimal sensitivity give the same values. This will be done for a least squares estimation problem. Since all estimation problems can be formulated as least squares problems, the general result can be applied to sequential filter/smoother algorithms.

Let us investigate the suboptimal sensitivity analysis. Assume that the true values of all error sources are scaled (the ith source is scaled by λ_i), but that the estimator uses the same error covariances as in the unscaled case, and the gain computed in the previous least squares example would not change. Then the ith projective factor of the true covariance is merely scaled by λ_i, giving the actual projective factor of the measurement errors as

$$P_a^{(m)}(i) = \lambda_i PH^{(i)T}R^{(i)-1}H^{(i)}P = \lambda_i P_i^{(m)} \tag{21}$$

The subscript a denotes actual, in contrast to the computed performance. With multiple data sets, the total actual covariance is given by the sum of all actual projective factors, i.e.,

$$P_a^{(m)} = \sum_{i=1}^{p} \lambda_i P_i^{(m)} \tag{22}$$

In the optimal sensitivity analysis, the actual error covariances used in the filter are changed to reflect the true values. Let us assume that the scale factor λ is applied to the measurement covariance in the previous batch problem. In this case the computed gain will change, as well as the final covariance. From Eq. (7), the new gain is given by

$$K(\lambda) \doteq P_0H^T\left(HP_0H^T + \lambda R\right)^{-1} \tag{23}$$

Likewise, the covariance will change to

$$P(\lambda) = (I - KH)P_0(I - KH)^T + \lambda KRK^T \tag{24}$$

Let us calculate the partial derivative $\partial P(\lambda)/\partial\lambda$. Taking the partial derivative of Eq. (24) yields

$$\left.\frac{\partial P(\lambda)}{\partial \lambda}\right|_{\lambda=1} = KRK^T|_{\lambda=1} + A + A^T$$

where, letting $K_\lambda = \partial K/\partial\lambda$,

$$A = \left[-(I - KH)P_0H^T + KR\right]K_\lambda^T$$
$$= \left[-P_0H^T + K\left(HP_0H^T + R\right)\right]K_\lambda^T$$

The term in the brackets in the last expression can be seen to be zero with the substitution Kalman filter expression $K = P_0H^T(HP_0H^T + R)^{-1}$. Thus, using Eq. (14), one has

$$\left.\frac{\partial P(\lambda)}{\partial \lambda}\right|_{\lambda=1} = KRK^T|_{\lambda=1} = P^{(m)} \tag{25}$$

Thus, the projective factor equals the sensitivity of the optimal covariance to a scale factor associated with the covariance of the corresponding error source!

Covariance Projection Methods for Sequential Estimation Problems for Stochastic Systems

A stochastic system has a time varying state, driven by process noise. Let us consider the discrete time case, and let the state at time t_k be denoted by $\boldsymbol{x}_k$. The basic dynamic modeling equation is given by

$$\boldsymbol{x}_{k+1} = \Phi_k\boldsymbol{x}_k + \Gamma_k\boldsymbol{u}_k \tag{26}$$

The sequence $\{\boldsymbol{u}_k\}$ is discrete white noise. Let us assume $\boldsymbol{u}_k$ is a zero-mean Gaussian random variable with covariance given by

$$E\left[\boldsymbol{u}_k\boldsymbol{u}_k^T\right] = Q_k \tag{27}$$

With the introduction of process noise, it makes sense to define projective factors with respect to process noise. For most applications one will want to calculate separate projective factors for process noise and measurement errors. However, to simplify the derivation, we shall assume each error source has some contribution to both the measurement error and process noise. To apply this equation to a particular type of error, i.e., process noise, where only one type of error is present, certain matrices used in the calculations will contain only zeros.

The equations for the projective components will be developed, paralleling those of the Kalman filter. Separate equations will be developed for the time prediction, followed by those for a measurement update.

Time Prediction

The standard Kalman prediction formulas, relating the filtered estimate $\hat{\boldsymbol{x}}_k^f$ and covariance P_k^f at t_k to the one step predicted estimate $\hat{\boldsymbol{x}}_{k+1}^p$ and its covariance P_{k+1}^p at t_{k+1}, are given by

$$\hat{\boldsymbol{x}}_{k+1}^p = \Phi_k\hat{\boldsymbol{x}}_k^f \tag{28}$$

$$P_{k+1}^p = \Phi_kP_k^f\Phi_k^T + \Gamma_kQ_k\Gamma_k^T \tag{29}$$

Now let the process noise vector be expressed as the sum of independent projective components,

$$\boldsymbol{u}_k = \sum_{j=1}^{p} \boldsymbol{u}_k(j) \tag{30}$$

where

$$Q_k(j) = E\left[\boldsymbol{u}_k(j)\boldsymbol{u}_k(j)^T\right] \tag{31}$$

$$Q_k = \sum_{j=1}^{p} Q_k(j) \tag{32}$$

The expression of the total error in terms of its projective components is given by

$$\tilde{\boldsymbol{x}}_k^f = \sum_{j=1}^{n} \tilde{\boldsymbol{x}}_k^f(j) \tag{33}$$

By assumption, all these errors are independent. Let us define

$$P_k^f(j) = E\left\{\tilde{\boldsymbol{x}}_k^f(j)\left[\tilde{\boldsymbol{x}}_k^f(j)\right]^T\right\} \tag{34}$$

Also, by definition

$$\tilde{\boldsymbol{x}}_{k+1}^p(j) = \Phi_k\tilde{\boldsymbol{x}}_k^f(j) - \Gamma_k\boldsymbol{u}_k(j) \tag{35}$$

Squaring and taking expectations yields

$$P_{k+1}^p(j) = \Phi_k P_k^f(j)\Phi_k^T + \Gamma_k Q_k(j)\Gamma_k^T \tag{36}$$

The preceding derivation followed the methods used in the suboptimal estimation sensitivity analysis. Note that Eq. (36) can be derived directly from Eq. (29) by a perturbation approach.

It is easy to show that Eq. (36) implies Eq. (29). The assumption of statistically independent errors implies the projective decompositions by P_k^f and P_{k+1}^p are given by

$$P_k^f = \sum_{j=1}^{p} P_k^f(j) \tag{37}$$

$$P_{k+1}^p = \sum_{j=1}^{p} P_{k+1}^p(j) \tag{38}$$

Now summing Eq. (36) over j and using Eqs. (37) and (38) yields Eq. (29).

Measurement Update

This development extends the example in the least squares example to the case where the a priori state estimate also depends on measurement errors. In the discrete case, the measurement update equation is given by

$$\boldsymbol{z}_k = H_k\boldsymbol{x}_k + \boldsymbol{v}_k \tag{39}$$

where $\{\boldsymbol{v}_k\}$ is a zero-mean white noise sequence with statistical properties

$$E[\boldsymbol{v}_k] = 0; \qquad E\left[\boldsymbol{v}_k\boldsymbol{v}_k^T\right] \doteq R_k \tag{40}$$

As with the process noise, we shall assume that the measurement noise contains all components of the error sources. Let

$$\boldsymbol{v}_k = \sum_{j=1}^{p} \boldsymbol{v}_k(j) \tag{41}$$

and

$$E\left[\boldsymbol{v}_k(j)\boldsymbol{v}_k(j)^T\right] = R_k(j) \tag{42}$$

This implies projective decomposition of the measurement covariance is given by

$$R_k = \sum_{j=1}^{p} R_k(j) \tag{43}$$

The key equation to derive the time update is the discrete equivalent of Eq. (11),

$$\hat{\boldsymbol{x}}_k^f - \boldsymbol{x}_k = (I - K_kH_k)\left(\hat{\boldsymbol{x}}_k^p - \boldsymbol{x}_k\right) + K_k\boldsymbol{v}_k \tag{44}$$

where

$$K_k \doteq P_k^p H_k^T\left(H_k P_k^p H_k^T + R_k\right)^{-1} \tag{45}$$

Now let us assume the predicted error can be expressed as a sum of statistically independent error sources that correspond to the measurement errors,

$$\tilde{\boldsymbol{x}}_k^p = \sum_{j=1}^{p} \tilde{\boldsymbol{x}}_k^p(j) \tag{46}$$

Now Eq. (42) and the definition of projective components imply

$$\tilde{\boldsymbol{x}}_k^f(j) = (I - K_kH_k)\tilde{\boldsymbol{x}}_k^p(j) + K_k\boldsymbol{v}_k(j) \tag{47}$$

Multiplying Eq. (47) on the right by the transpose of itself and taking expectations gives

$$P_k^f(j) = (I - K_kH_k)P_k^p(j)(I - K_kH_k)^T + K_kR_k(j)K_k^T \tag{48}$$

If the jth error source does not have a measurement noise contribution, the last term on the right side is zero. Summing Eq. (48) over j yields an equation characterizing a form of the measurement update equivalent to Eq. (12), namely,

$$P_k^f = (I - K_kH_k)P_k^p(I - K_kH_k)^T + K_kR_kK_k^T \tag{49}$$

If measurements of different types have statistically independent errors, simultaneous measurements may be processed sequentially. If one wishes to obtain a projective decomposition showing different data types for sequential measurement updates, Eq. (48) will be repeated twice for each component. On the first step, the last term will be zero for the second data type; on the second step, the last term will be zero for the first data type.

The preceding proof of Eq. (48) still followed the methods used in suboptimal sensitivity.

Summary of Projective Equation for Sequential Filters

In this summary the preceding update steps are summarized, indicating special cases for measurement errors, process noise, and a priori parameters.

Time Update

Process noise parameters

$$P_{k+1}^p(j) = \Phi_k P_k^f(j)\Phi_k^T + \Gamma_k Q_k(j)\Gamma_k^T \tag{50}$$

All others (measurement errors and a priori)

$$P_{k+1}^p(j) = \Phi_k P_k^f(j)\Phi_k^T \tag{51}$$

Measurement Update

Measurement error parameters, for jth error source,

$$P_k^f(j) = (I - K_kH_k)P_k^p(j)(I - K_kH_k)^T + K_kR_k(j)K_k^T \tag{52}$$

All others (process noise, a priori, other types of measurement errors)

$$P_k^f(j) = (I - K_kH_k)P_k^p(j)(I - K_kH_k)^T \tag{53}$$

Smoothing

For sequential estimation problems, covariance projection methods can also be applied to smoothing. Let the smoother state and covariance at t_k be designated by $\hat{\boldsymbol{x}}_k^s$ and P_k^s, respectively. The equations for the Rauch–Tung–Striebel (R–T–S) smoother are given by

$$\hat{\boldsymbol{x}}_k^s = \hat{\boldsymbol{x}}_k^f + K_k^s\left[\hat{\boldsymbol{x}}_{k+1}^s - \hat{\boldsymbol{x}}_{k+1}^p\right] \tag{54}$$

$$P_k^s = P_k^f + K_k^s\left[P_{k+1}^s - P_{k+1}^p\right]K_k^{s^T} \tag{55}$$

$$K_k^s \doteq P_k^f\Phi_k^T\left(P_{k+1}^p\right)^{-1} \tag{56}$$

Equation (54) implies

$$\tilde{x}_k^s(j) = \tilde{x}_k^f(j) + K_k^s\left[\tilde{x}_{k+1}^s(j) - \tilde{x}_{k+1}^p(j)\right] \tag{57}$$

Following the suboptimal sensitivity approach, we could take the expression in Eq. (57) and derive an equation for the jth projective factor of the smoothed covariance. However, all of the error terms on the right-hand side of Eq. (57) are correlated, necessitating the derivation of the cross covariances of the various terms. Rather than do this, an easier approach is provided by the earlier observation: the projective factor represents the sensitivity of the optimal covariance to a scale factor associated with the covariance of the corresponding error source. One could easily verify that this approach can be used to derive the equation for the preceding filter steps. (For the measurement update, refer to the section on batch processing). Applying the chain rule in taking the derivative of Eq. (55) gives

$$\begin{aligned}\frac{\partial P_k^s}{\partial \lambda_j} &= \frac{\partial P_k^f}{\partial \lambda_j} + K_k^s\left[\frac{\partial P_{k+1}^s}{\partial \lambda_j} - \frac{\partial P_{k+1}^p}{\partial \lambda_j}\right]K_k^{s^T}\\ &+ \frac{\partial K_k^s}{\partial \lambda_j}\left[P_{k+1}^s - P_{k+1}^p\right] + K_k^s\left[P_{k+1}^s - P_{k+1}^p\right]\frac{\partial K_k^s}{\partial \lambda_j}\end{aligned}$$

Now introducing the notation $P_k^s(j) = \partial P_k^s/\partial\lambda_j$, etc., this equation can be written as

$$P_k^s(j) = P_k^f(j) + K_k^s\left[P_{k+1}^s(j) - P_{k+1}^p(j)\right]K_k^{s^T} + A_k(j) + A_k^T(j) \tag{58}$$

where

$$A_k(j) \doteq K_k^s(j)\left[P_{k+1}^s - P_{k+1}^p\right]K_k^{s^T} \tag{59}$$

and $K_k^s(j)$ is used to denote the partial derivative of the smoother gain with respect to the jth parameter. To derive an expression for $K_k^s(j)$, we shall use the following relation for partial derivatives of a symmetric matrix:

$$\frac{\partial A^{-1}}{\partial x} = -A^{-1}\frac{\partial A}{\partial x}A^{-1} \tag{60}$$

Taking the partial derivative of Eq. (56) gives

$$\begin{aligned}K_k^s(j) &\doteq \frac{\partial K_k^s}{\partial \lambda_j} = P_k^f(j)\Phi_k^T\left(P_{k+1}^p\right)^{-1} - K_k^s P_{k+1}^p(j)\left(P_{k+1}^p\right)^{-1}\\ &= \left[P_k^f(j)\Phi_k^T - K_k^s P_{k+1}^p(j)\right]\left(P_{k+1}^p\right)^{-1}\end{aligned} \tag{61}$$

To show that the equations for the total error of the smoother can be easily recovered from the equations for the projective components, Eq. (58) is summed over j, giving

$$P_k^s = P_k^f + K_k^s\left[P_{k+1}^s - P_{k+1}^p\right]K_k^{s^T} + A_k + A_k^T \tag{62}$$

where

$$A_k = \sum_{i=1}^{p} A_k(i) = \left[\sum_{i=1}^{p} K_k^s(i)\right]\left[P_{k+1}^s - P_{k+1}^p\right]K_k^s \tag{63}$$

Now from Eq. (61),

$$\begin{aligned}\sum_{i=1}^{p} K_k^s(i) &= \sum_{i=1}^{p}\left[P_k^f(i)\Phi_k^T - K_k^s P_{k+1}^p(i)\right]\left(P_{k+1}^p\right)^{-1}\\ &= \left[P_k^f\Phi_k^T - K_k^s P_{k+1}^p\right]\left(P_{k+1}^p\right)^{-1} = 0\end{aligned} \tag{64}$$

Thus, the last two expressions on the right-hand side of Eq. (62) are zero, yielding Eq. (55). The equations needed to compute the projective components are given by Eqs. (58), (59), and (61). Note that these require quantities used in the standard smoother, the smoother covariance, and the projective component covariances from the filter. Each projective component of the smoother depends on the projective components of the filter of the same factor, and thus projective components can be calculated separately in a parallel implementation.

Continuous Estimation

The equations for continuous filtering and smoothing can be derived directly from the perturbations of the covariance equations. The Kalman–Bucy filter covariance propagation is given by

$$\begin{aligned}\dot{P}^f &= F(t)P^f + P^f F(t)^T + \Gamma(t)Q(t)\Gamma(t)^T\\ &- P^f(t)H(t)^T R^{-1}H(t)P^f(t)\end{aligned} \tag{65}$$

Now assume

$$P^f(t,\lambda) \doteq \sum_{i=1}^{p}\lambda_i P_i^f(t) \tag{66}$$

$$Q(t,\lambda) \doteq \sum_{i=1}^{p}\lambda_i Q_i(t) \tag{67}$$

$$R(t,\lambda) \doteq \sum_{i=1}^{p}\lambda_i R_i(t) \tag{68}$$

Substituting Eqs. (66–68) into Eq. (65) and taking the partial derivative with respect to λ_i yields the desired equations. To simplify the equations let us assume that every error source appears in the measurements, but that different error sources appear in different measurements. Let the measurements for the ith error source be denoted by

$$z^{(i)} = H^{(i)}x + v^{(i)} \tag{69}$$

Then

$$H^T R(t,\lambda)^{-1}H = \sum_{i=1}^{p} H^{(i)T}\frac{R_i^{-1}(t)}{\lambda_i}H^{(i)}(t) \tag{70}$$

The partial derivative of this expression with respect to λ_i is

$$\left.\frac{\partial H^T R(t,\lambda)^{-1}H}{\partial \lambda_i}\right|_{\lambda_j=1} = -H^{(i)T}R_i^{-1}(t)H^{(i)}(t) \tag{71}$$

Thus, the partial derivative of Eq. (65) with respect to λ_i (with the indicated substitutions) becomes

$$\begin{aligned}\dot{P}_i^f &= \bar{F}(t)P_i^f + P_i^f\bar{F}(t)^T + \Gamma(t)Q_i(t)\Gamma(t)^T\\ &+ P^f(t)H_i(t)^T R_i^{-1}H_i(t)P^f(t)\end{aligned} \tag{72}$$

where

$$\bar{F} = F - P(t)H^T(t)R^{-1}H(t) \tag{73}$$

Equation (72) will have three forms depending on whether the noise parameter is a process noise, measurement noise, or a priori. For process noise parameters, the last expression is zero. For measurement noise the second to the last term is zero. For other parameters (a priori), the last two terms will be zero.

The propagation of the continuous R–T–S smoother covariance is given by

$$\dot{P}^s(t) = F^s(t)P^s(t) + P^s(t)[F^s(t)]^T - \Gamma(t)Q(t)\Gamma(t)^T \tag{74}$$

where

$$F^s(t) \doteq F(t) - \Gamma(t)Q(t)\Gamma(t)^T[P^f(t)]^{-1} \tag{75}$$

Taking the partial derivative of Eq. (74) with respect to λ_i yields

$$\begin{aligned}\dot{P}_i^s(t) &= F^s(t)P_i^s(t) + P_i^s(t)[F^s(t)]^T\\ &- \Gamma(t)Q_i(t)\Gamma(t)^T + B(t) + B(t)^T\end{aligned} \tag{76}$$

where

$$B(t) = \left[\Gamma Q_i(t)\Gamma^t + \Gamma Q(t)\Gamma^T P^f(t)^{-1}P_i^f(t)\right][P^f(t)]^{-1}P^s(t) \tag{77}$$

Steady-State Conditions

For a large class of sequential estimation problems with constant coefficients, filter covariances will approach a steady-state condition in which the covariances will not change in time. Equations for steady-state projection components will be established in this section for continuous estimation problems. Let us define the steady-state filter covariances for the continuous problem by the following equation:

$$P^{f,\mathrm{ss}} = \sum_{i=1}^{p} P_i^{f,\mathrm{ss}} \tag{78}$$

The steady-state filter covariance must satisfy

$$0 = FP^{f,\mathrm{ss}} + P^{f,\mathrm{ss}}F^T + \Gamma Q \Gamma^T - P^{f,\mathrm{ss}} H^T R^{-1} H P^{f,\mathrm{ss}} \tag{79}$$

The steady-state projection components must satisfy

$$0 = \bar{F}^{\mathrm{ss}} P_i^{f,\mathrm{ss}} + P_i^{f,\mathrm{ss}} \bar{F}^{\mathrm{ss}^T} + \Gamma Q_i \Gamma^T + P^{f,\mathrm{ss}} H_i^T R_i^{-1} H_i P^{f,\mathrm{ss}} \tag{80}$$

where

$$\bar{F}^{\mathrm{ss}} = F - P^{f,\mathrm{ss}} H^T R^{-1} H \tag{81}$$

Note that although Eq. (80) is nonlinear in $P^{f,\mathrm{ss}}$ it is linear in $P_i^{f,\mathrm{ss}}$.

Let the smoother covariance be expressed by

$$P^{s,\mathrm{ss}} = \sum_{i=1}^{p} P_i^{s,\mathrm{ss}} \tag{82}$$

The steady-state smoother covariance must satisfy

$$0 = F^s P^{s,\mathrm{ss}} + P^{s,\mathrm{ss}}[F^s]^T - \Gamma Q \Gamma^T \tag{83}$$

where

$$F^s \doteq F - \Gamma Q \Gamma^T (P^{f,\mathrm{ss}})^{-1} \tag{84}$$

From Eq. (76) the steady-state solution for $P_i^{s,\mathrm{ss}}$ becomes

$$0 = F^s P_i^{s,\mathrm{ss}} + P_i^{s,\mathrm{ss}}[F^s]^T - \Gamma Q_i(t) \Gamma^T + B + B^T \tag{85}$$

where

$$B = \left[\Gamma Q_i \Gamma^t + \Gamma Q \Gamma^T P^{f,\mathrm{ss}-1} P_i^{f,\mathrm{ss}}\right][P^{f,\mathrm{ss}}]^{-1} P^{s,\mathrm{ss}} \tag{86}$$

Examples

Discrete Example

Let us consider a simple discrete example of a one-dimensional random walk. The dynamics is given by

$$x_{k+1} = x_k + u_k \tag{87}$$

where $\{u_k\}$ is a white noise sequence with

$$E\left[u_k^2\right] = Q \tag{88}$$

The measurements are given by

$$z_k = x_k + v_k \tag{89}$$

where $\{v_k\}$ is a white noise sequence with

$$E\left[v_k^2\right] = R \tag{90}$$

The a priori information is given by

$$\hat{x}_0 \quad \text{with} \quad E\left[\tilde{x}_0^2\right] = P_0 \tag{91}$$

Equations for projective components of process noise, measurement noise, and a priori will be derived with the notation

$$P_k = P_k^Q + P_k^R + P_k^A \tag{92}$$

For initialization

$$P_0^{Q,p} = 0; \qquad P_0^{R,p} = 0; \qquad P_0^{A,p} = P_0 \tag{93}$$

For time update

$$P_{k+1}^{Q,p} = P_k^{Q,f} + Q \tag{94}$$

$$P_{k+1}^{R,p} = P_k^{R,f} \tag{95}$$

$$P_{k+1}^{A,p} = P_k^{A,f} \tag{96}$$

For measurement update

$$P_{k+1}^{Q,f} = \left[1 - P_{k+1}^p \big/ \left(R + P_{k+1}^p\right)\right]^2 P_{k+1}^{Q,f} \tag{97}$$

$$P_{k+1}^{A,f} = \left[1 - P_{k+1}^p \big/ \left(R + P_{k+1}^p\right)\right]^2 P_{k+1}^{A,p} \tag{98}$$

$$P_{k+1}^{R,f} = \left[1 - P_{k+1}^p \big/ \left(R + P_{k+1}^p\right)\right]^2 P_{k+1}^{R,p} + (P_{k+1}^p)^2 R \tag{99}$$

Continuous Example

The preceding example can be considered a discrete version of the following continuous example. Assume x is Wiener process (Brownian motion),

$$\dot{x} = u \tag{100}$$

where u is zero-mean white noise

$$E[u(t)u(\tau)] = Q\delta(t - \tau) \tag{101}$$

Measurements are given by

$$z = x + v \tag{102}$$

where v is zero-mean white noise

$$E[v(t)v(\tau)] = R\delta(t - \tau) \tag{103}$$

The equation for the optimal covariance is given by

$$\dot{P} = Q - (P^2/R) \tag{104}$$

The equations for the various projective components of the filter covariance are given by

$$\dot{P}_A = -(2PP_A/R) \tag{105}$$

$$\dot{P}_Q = Q - (2PP_Q/R) \tag{106}$$

$$\dot{P}_R = (P^2/R) - (2PP_R/R) \tag{107}$$

The preceding equations are nonlinear in P but are linear in the other components and thus can be easily solved if the initial covariance is started at its steady-state value

$$Q - [(P^{t,\mathrm{ss}})^2/R] = 0 \tag{108}$$

This implies

$$P^{f,\mathrm{ss}} = \sqrt{RQ} \tag{109}$$

Substituting for P into the preceding differential equations yields

$$\dot{P}_A = -2\sqrt{(Q/R)}P_A \tag{110}$$

$$\dot{P}_Q = Q - 2\sqrt{(Q/R)}P_Q \tag{111}$$

$$\dot{P}_R = Q - 2\sqrt{(Q/R)}P_R \tag{112}$$

The asymptotic solutions of the preceding equations are given by

$$P_Q = P_R = \tfrac{1}{2}\sqrt{RQ}; \qquad P_A = 0 \tag{113}$$

Note that one can compute the projection components by taking partial derivatives

$$P_Q = P_R = \frac{\partial P}{\partial \lambda} = \left.\frac{\partial \sqrt{\lambda RQ}}{\partial \lambda}\right|_{\lambda=1} = \frac{1}{2}\sqrt{RQ} \tag{114}$$

$$P_Q + P_R = \sqrt{RQ} = P \tag{115}$$

The specific equations for the projection components are given by

$$P_A(t) = P^{f,\mathrm{ss}}\exp\left(-\frac{2t}{\sqrt{R/Q}}\right)$$
$$P_Q(t) = P_R(t) = \frac{P^{f,\mathrm{ss}}}{2}\left[1 - \exp\left(-\frac{2t}{\sqrt{R/Q}}\right)\right] \tag{116}$$

As can be seen from these equations, the projective factors for process noise and measurement noise are equal and behave as an exponentially correlated process with correlation time $\sqrt{(R/Q)}$ and steady-state covariance $\sqrt{(RQ)}$.

Two-Dimensional Continuous Example

Define the two-dimensional matrices

$$F = \begin{bmatrix} 0 & 1 \\ -0.5 & 0 \end{bmatrix} \qquad H = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$
$$R = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \qquad Q = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \tag{117}$$

The equation for the steady-state filter covariance [Eq. (79)] is given by

$$0 = \begin{bmatrix} 0 & 1 \\ -0.5 & 0 \end{bmatrix} P^{f,\mathrm{ss}} + P^{f,\mathrm{ss}} \begin{bmatrix} 0 & -0.5 \\ 1 & 0 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - P^{f,\mathrm{ss}} P^{f,\mathrm{ss}} \tag{118}$$

These equations were solved using Newton's method to yield

$$P^{f\mathrm{ss}} = \begin{bmatrix} 1.1394 & 0.1623 \\ 0.1623 & 0.9008 \end{bmatrix} \tag{119}$$

The steady-state matrix $\bar{F}^{\mathrm{ss}} \doteq F - P^{f,\mathrm{ss}} H^T R^{-1} H$ [Eq. (81)] is

$$\bar{F}^{\mathrm{ss}} = \begin{bmatrix} -1.1394 & 0.8377 \\ -0.6623 & -0.9089 \end{bmatrix} \tag{120}$$

To obtain the projective components of the steady-state covariance, various versions of Eq. (80) must be solved. To solve for the components of the process noise, the following sets of equations must be solved:

$$0 = \bar{F}^{\mathrm{ss}} P_i^{f,\mathrm{ss}} + P_i^{f,\mathrm{ss}} \bar{F}^{\mathrm{ss}^T} + \Gamma Q_i \Gamma^T \tag{121}$$

For example, the projective component for the first component of the process noise is obtained from the solution of the equation

$$0 = \bar{F}^{\mathrm{ss}} P_i^{f,\mathrm{ss}} + P_i^{f,\mathrm{ss}} \bar{F}^{\mathrm{ss}^T} + \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \tag{122}$$

The equations for the projective components due to measurement errors must solve the equations

$$0 = \bar{F}^{\mathrm{ss}} P_i^{f,\mathrm{ss}} + P_i^{f,\mathrm{ss}} \bar{F}^{\mathrm{ss}^T} + P^{f\mathrm{ss}} H_i^T R_i^{-1} H_i P^{f,\mathrm{ss}} \tag{123}$$

The projective component due to the first component of process noise is obtained by solving the equation

$$0 = \bar{F}^{\mathrm{ss}} P_i^{f,\mathrm{ss}} + P_i^{f,\mathrm{ss}} \bar{F}^{\mathrm{ss}^T} + P^{f,\mathrm{ss}} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} P^{f,\mathrm{ss}} \tag{124}$$

In this example, the steady-state covariance and its steady-state projective components are given in Table 1. As can be seen in the example, the coupled dynamics impacts the steady-state solution nonsymmetrically. The impact of y on x increases the role of process noise, whereas the impact of x on y decreases the role of process noise relative to the case where the dynamics is not coupled, as in the first example.

Table 1 Budget allocation of steady-state covariance for 2 × 2 example

Variable	P(1, 1)	P(1, 2)	P(2, 2)
v_1	0.3708	−0.0925	0.0680
v_2	0.1088	0.1479	0.4463
		Measurement total	
	0.4796	0.0555	0.5143
u_1	0.5276	−0.0573	0.0568
u_2	0.1322	0.1641	0.3297
		Process noise total	
	0.6598	0.1068	0.3865
Total	1.1394	0.1623	0.9008

Conclusions

The equations for computing the error budgets/sensitivities for covariances of least squares estimators and sequential filters and smoothers relative to input error sources, including measurement noise, process noise, and a priori data were derived. It was shown that the optimal and suboptimal sensitivities are the same. Examples show how the covariance projection analysis gives insight into the factors that drive performance.

References

[1] Gelb, A., *Applied Optimal Estimation*, 11th printing, MIT Press, Cambridge, MA, 1989.

[2] Warner, M. R., Nead, M. W., and Hudson, R. H., "The Orbit Determination Program of the Jet Propulsion Laboratory," TM 33-168, California Inst. of Technology, Pasadena, CA, March 1964.

[3] Dyer, P., and McReynolds, S. R., "The Sensitivity Matrix Method for Orbit Determination with Application to a Mars Orbiter," *JPL Space Program Summary*, Vol. 2, Nos. 37–56, 1969.

[4] Soong, T. S., "On A Priori Statistics in Minimum Variance Estimation Problems," *Journal of Basic Engineering*, Vol. 87D, March 1965, pp. 109–204.

[5] Kalman, R. E., and Bucy, R. S., "New Results in Linear Filtering and Prediction Theory," *Journal of Basic Engineering*, Vol. 83D, March 1961, pp. 95–108.

[6] Kalman, R. E., "A New Approach to Linear Filtering and Prediction Theory," *Journal of Basic Engineering*, Vol. 82D, March 1960, pp. 35–45.

[7] Rauch, H. E., Tung, F., and Striebel, C., "Maximum Likelihood Estimates of Linear Dynamic Systems," *AIAA Journal*, Vol. 3, No. 8, 1965, pp. 1445–1450.

[8] Nishimura, T., "On the A Priori Information in Sequential Estimation Problems," *IEEE Transactions on Automatic Control*, Vol. AC-11, April 1966, p. 197.

[9] Nishimura, T., "Correction to and Extension of 'On the A Priori Information in Sequential Estimation Problems,'" *IEEE Transactions on Automatic Control*, Vol. AC-12, Feb. 1967, p. 123 (Correspondence).

[10] Heffes, J., "The Effect of Erroneous Models on the Kalman Filter Response," *IEEE Transactions on Automatic Control*, Vol. AC-11, July 1966, pp. 541–543.

[11] Friedland, B., "The Effect of an Incorrect Gain in the Kalman Filter," *IEEE Transactions on Automatic Control*, Vol. 12, 1966, p. 610.

[12] Mehra, R., "On Optimal and Suboptimal Linear Smoothing," *Proceedings of the National Electronics Conference*, Vol. 24, 1968.

[13] Griffis, R. E., and Sage, A. P., "Sensitivity Analysis of Discrete Filtering and Smoothing Algorithms," *AIAA Journal*, Vol. 7, No. 10, 1969, pp. 1890–1897.

[14] Griffis, R. E., and Sage, A. P., "Large and Small Scale Sensitivity Analysis of Optimum Estimation Algorithms," *IEEE Transactions on Automatic Control*, Vol. AC-13, Aug. 1968, pp. 320–392.

[15] Griffis, R. E., and Sage, A. P., "Sensitivity Analysis of Fixed Point Linear Smoothing Algorithms," *International Journal on Control*, Vol. 8, No. 4, 1968, pp. 321–337.

[16] Joseph, P. D., "Suboptimal Linear Filtering," Space Technology Lab. Inc., Tech. Rept., El Segundo, CA, Dec. 1963.

[17] Meditch, J. S., "Suboptimal Linear Filtering for Continuous Dynamic Process," Aerospace Corp., Tech. Rept., El Segundo, CA, July 1964.

[18] Aoki, M., "Control of Large Scale Dynamic Systems by Aggregation," *Joint Automatic Control Conference* (Philadelphia, PA), 1967, pp. 178–186.

[19] Aoki, M., and Huddle, D. A., "Estimation of the State Vector of a Linear Stochastic System with a Constrained Estimator," *IEEE Transactions on Automatic Control* (short papers), Vol. AC-12, Aug. 1967, pp. 432, 433.

[20] Huddle, J. R., and Wismer, D. A., "Degradation of Linear Filter Performance Due to Modeling Error," *IEEE Transactions on Automatic Control*, Vol. AC-13, No. 4, 1978, pp. 421–424.

[21] Luenberger, D. G., "Observing the State of a Linear System," *IEEE Transactions on Military Electronics*, Vol. MIL-8, April 1964, pp. 74–80.