[10] N. K. Sinha and W. Pille, "On-line parameter estimation using the matrix pseudo-inverse," *Proc. Inst. Elec. Eng.*, vol. 118, pp. 1041–1046, 1971.
[11] A. Dvoretzky, "On stochastic approximation," in *Proc. 3rd Berkeley Symp. on Mathematical Statistics and Probability*, vol. 1. Los Angeles, CA: Univ. California Press, 1956, pp. 39–56.
[12] R. N. Pandya, "A class of bootstrap estimators and their relationship to the generalized two stage least squares estimators," *IEEE Trans. Automat. Contr.*, vol. AC-19, pp. 831–835, Dec. 1974.

The Cramér–Rao Estimation Error Lower Bound Computation for Deterministic Nonlinear Systems

JAMES H. TAYLOR

***Abstract*—For continuous-time nonlinear deterministic system models with discrete nonlinear measurements in additive Gaussian white noise, the extended Kalman filter (EKF) convariance propagation equations *linearized about the true unknown trajectory* provide the Cramér–Rao lower bound to the estimation error covariance matrix. A useful application is establishing the optimum filter performance for a given nonlinear estimation problem by developing a simulation of the nonlinear system and an EKF linearized about the true trajectory.**

Introduction and Proposition

The problem of estimating unknown deterministic variables (in particular, the state variables generated by a nonlinear time-varying state vector differential equation driven by deterministic input variables) from discrete nonlinear observations in additive Gaussian white noise arises in many applications of modern estimation theory. Two typical examples are: tracking a deterministically maneuvering spacecraft from a platform that follows a known trajectory (e.g., estimating the target's position and velocity in a relative coordinate frame from noisy measurements of the platform/target line-of-sight angle); and estimating unknown nonrandom parameters in the nonlinear dynamic model of a system subject to known inputs, from nonlinear measurements of selected systems variables corrupted by additive Gaussian noise.

Since the practical estimation algorithms that are generally developed for such tasks are rarely optimal, a very real concern is determining "the best that can be done," and comparing a given algorithm's performance with that estimation error lower bound to see if the filter is adequate or to determine if seeking a more effective algorithm is worthwhile. Establishing the performance of a filter algorithm generally involves developing a computer simulation of the nonlinear system and observation models to provide realistic measurement sequences that can be processed by the algorithm which is also mechanized in the simulation. The system and observation models used in this simulation are generally called the "truth model"; it is often more detailed and realistic than the model used as a basis for deriving the filter algorithm (the so-called "filter model").

A powerful result in such a performance assessment is the Cramér–Rao inequality (cf. [1]). In essence, defining P to be the estimation error covariance matrix corresponding to any unbiased estimator[1] of the unknown deterministic variables, then the inequality can be stated as[2]

$$P \geqslant P^* \triangleq J^{-1} \tag{1}$$

where J is the Fisher information matrix. Thus, P^* defines the best that can be done, in the above sense.

In many instances, obtaining P^* is difficult (cf. [2]–[5]). For the case considered in this paper, however, the following result will be proven.

Proposition: The inverse Fisher information matrix (P^*) corresponding to a *dynamic system* modeled by a nonlinear time-varying state vector differential equation with deterministic inputs and *nonlinear time-varying observations* of the state variables corrupted by additive Gaussian white noise sequences *propagates according to the same equations as the filter covariance matrix for an extended Kalman filter (EKF) linearized about the true (unknown) trajectory*.

This exact result permits Cramér–Rao lower bound analyses to be performer very directly for this class of problems, with no additional analytic effort, and often with little programming effort in studies where the EKF has already been mechanized in a computer simulation as described above. In the latter case, it is simply necessary to change the EKF linearization from $\hat{x}$, the EKF's estimates of x (the normal "realizable" EKF [6]), to x (an "idealized" EKF) to obtain the Cramér–Rao lower bound.

Proof of the Proposition

Consider the following nonlinear continuous-discrete estimation problem. The continuous-time system dynamics obey

$$\dot{x} = f(x,u,t) \tag{2}$$

where x is the n-vector of state variables, u is a deterministic r-vector of inputs, and $f(\cdot,\cdot,t)$ denotes a general[3] nonlinear time-varying formulation of the system dynamics. The unknown deterministic trajectory to be estimated is then

$$x(t) = x(t; x_0, u(\tau): 0 \leqslant \tau \leqslant t)$$

where x_0 is also unknown. The value of x at each "sampling time" $t = kT_s$ is denoted

$$x_k \triangleq x(kT_s; x_0, u(\tau): 0 \leqslant \tau \leqslant kT_s), \qquad k = 1,2,\cdots$$

and discrete observations having the form

$$\begin{aligned} z_k &= h(x_k,k) + v_k \\ &\triangleq h_k + v_k, \qquad v_k \sim N(0,R_k) \end{aligned} \tag{3}$$

(where z_k is an m-vector) are available at each sampling time. Finally, it is assumed that some unbiased *a priori* statistical information about x_0 is available, of the form

$$\hat{x}_0 \sim N(x_0, S_0). \tag{4}$$

This information may be viewed as a "random filter initialization," i.e., $\hat{x}(0)$ may be provided to the estimator by other means (e.g., a boost-phase tracking system in the spacecraft tracking problem outlined above), or it may be considered to be an "additional fictitious measurement" at $t=0$,

$$z_0 = x_0 + v_0, \qquad v_0 \sim N(0, S_0)$$

which is of different dimensionality from $z_1, z_2, \cdots$. In either case, it is desired to determine how well x_K can be estimated given the data

$$\begin{aligned} Z_K &= \{\hat{x}_0, z_1, z_2, \cdots, z_K\} \\ &= \{z_0, z_1, \cdots, z_K\}. \end{aligned}$$

Manuscript received August 25, 1978. This work was supported by the Army Ballistic Missile Defense Advanced Technology Center, Huntsville, AL, under Contract DASG60-75-C-0053.

The author was with The Analytic Sciences Corporation, Reading, MA 01867. He is now with the School of Mechanical and Aerospace Engineering, Oklahoma State University, Stillwater, OK 74074.

[1] The most general result known to the author [8] when the estimate $\hat{x}$ exhibits a bias $b(x)$ is $P \geqslant P^* \triangleq (I+(\partial b/\partial x))J^{-1}(I+(\partial b/\partial x))^T$; see also [1]. However, the term $b(x)$ is generally not available in analytic form, so this result is of questionable value here.

[2] The matrix inequality $P \geqslant P^*$ is equivalent to stating that $(P-P^*)$ is positive semi-definite. Ensuring that J^{-1} exists is straightforward, and not considered in this development.

[3] The only restriction is that the associated Jacobian matrix $F \triangleq \partial f/\partial x$ must be continuous for all x and t in the domain of interest.

0018-9286/79/0400-0343$00.75 ©1979 IEEE

Following Van Trees [1], the conditional information matrix[4] for the unknown deterministic state vector x at $t=KT_s$ is given by

$$J_K=-E\left[\frac{\partial^2 \ln p_K}{\partial x_K^2}\,\middle|\, X_K\right] \tag{5}$$

where p_K is the conditional density of Z_K,

$$p_K=p_{Z_K|X_K}$$

and X_K is the set

$$X_K=\{x_0,x_1,\cdots,x_K\}.$$

Given J_K, the Cramér–Rao lower bound for an unbiased estimator is given by (1); in particular, for each element of x_K,

$$\sigma_{iK}^2 \triangleq E\left[(x_{Ki}-\hat{x}_{Ki}(Z_K))^2\right] \geq (J_K^{-1})_{ii} \triangleq P^*_{Kii}$$

where P^*_{Kii} denotes the (i,i) element of P^*_K in (1).

By assumption, (3) and (4) yield

$$p_{Z_K|X_K}=\frac{1}{(2\pi)^{n/2}|S_0|^{1/2}}e^{-1/2(x_0-\hat{x}_0)^T S_0^{-1}(x_0-\hat{x}_0)} \cdot \prod_{k=1}^{K}\frac{1}{(2\pi)^{m/2}|R_k|^{1/2}}e^{-1/2(z_k-h(x_k))^T R_k^{-1}(z_k-h(x_k))} \tag{6}$$

so

$$-\ln p_K=\text{constant}+\tfrac{1}{2}(x_0-\hat{x}_0)^T S_0^{-1}(x_0-\hat{x}_0) +\tfrac{1}{2}\sum_{k=1}^{K}(z_k-h(x_k))^T R_k^{-1}(z_k-h(x_k)).$$

Taking the expectation of the second partials of $-\ln p_K$ in accordance with (5) yields

$$J_K=\left(\frac{\partial x_0}{\partial x_K}\right)^T S_0^{-1}\left(\frac{\partial x_0}{\partial x_K}\right)+\sum_{k=1}^{K}\left(\frac{\partial h_k}{\partial x_K}\right)^T R_k^{-1}\left(\frac{\partial h_k}{\partial x_K}\right) \tag{7}$$

where several unstated terms have been eliminated by noting that $\partial x_0/\partial x_K$ and $\partial h(x_k)/\partial x_K$ are deterministic, so taking the expected values of terms of the form indicated below yields no contribution to (7):

$$E[(\text{deterministic factor})*(x_0-\hat{x}_0)]=0$$
$$E[(\text{deterministic factor})*(z_k-h(x_k))]=0.$$

Finally, (7) is reformulated by defining

$$H_k \triangleq \frac{\partial h(x_k,k)}{\partial x_k} \tag{8}$$

so, by the chain rule of partial differentiation,

$$J_K=\left(\frac{\partial x_0}{\partial x_K}\right)^T S_0^{-1}\left(\frac{\partial x_0}{\partial x_K}\right)+\sum_{k=1}^{K}\left(\frac{\partial x_k}{\partial x_K}\right)^T H_k^T R_k^{-1} H_k\left(\frac{\partial x_k}{\partial x_K}\right). \tag{9}$$

K auxiliary matrices[5] are needed to evaluate J_K according to (9):

$$M_k \triangleq \frac{\partial x_k}{\partial x_K}=\left(\frac{\partial x_K}{\partial x_k}\right)^{-1} \triangleq \Phi_{K,k}^{-1}, \qquad k=0,1,2,3,\cdots,K-1. \tag{10}$$

Under the general condition that the Jacobian matrix

$$F(t) \triangleq \frac{\partial f(x,u,t)}{\partial x} \tag{11}$$

is continuous in x and t, it can be shown [7] that $\Phi_{K,k}$ can be determined exactly by integrating

$$\dot{\Phi}=F(t)\Phi \tag{12}$$

from $t=kT_s$ to KT_s, subject to the initial condition

$$\Phi(kT_s,kT_s)=I.$$

These relations are identical to the transition matrix relations that form the basis for the discrete EKF [6], except that F is evaluated along the true state trajectory x rather than along $\hat{x}$ as in the usual EKF realization. In addition, the continuity property

$$\Phi_{K,k}=\Phi_{K,K-1}\Phi_{K-1,k}$$

permits (9) to be written in a recursive form; by inspection,

$$J_K=(\Phi_{K-1}^{-1})^T J_{K-1}\Phi_{K-1}^{-1}+H_K^T R_K^{-1} H_K \tag{13}$$

where the shorter standard notation $\Phi_{K-1} \triangleq \Phi_{K,K-1}$ is used for simplicity. In terms of P^*_K,

$$(P^*_K)^{-1}=(\Phi_{K-1}P^*_{K-1}\Phi_{K-1}^T)^{-1}+H_K^T R_K^{-1} H_K \tag{14}$$

which is identical in form to the EKF inverse covariance matrix propagation equation in the absence of process noise [6]. This observation completes the proof of the proposition.

Conclusion

The generality and simplicity of the proposition presented in this paper should prove to be of considerable value in assessing the performance of unbiased estimators for nonlinear systems in which the state variables can be assumed to be deterministic, i.e., random initial conditions and process noise can be neglected. The inclusion of estimator bias, if known, can be taken into account as well.[6] The result will be particularly useful in studies where the EKF is a candidate estimator, and a simulation with the EKF embedded in a "truth" or "real world" model that provides measurement data to the EKF has been developed to study its capabilities; the accompanying truth model provides the required linearization trajectory x. In such instances, the changes needed to linearize about the true state vector instead of $\hat{x}$ are generally simple, yielding P^* directly.

Acknowledgment

The author gratefully acknowledges the helpful discussions and comments of C. F. Price, T. H. Kerr, and J. I. Galdos, all colleagues at The Analytic Sciences Corporation.

References

[1] H. L. Van Trees, *Detection, Estimation, and Modulation Theory*, Part I. New York: Wiley, 1968.
[2] D. L. Snyder and I. B. Rhodes, "Filtering and control performance bounds with implications on asymptotic separation," *Automatica*, vol. 8, pp. 747–753, Nov. 1972.
[3] M. Zakai and J. Ziv, "Lower and upper bounds in the optimal filtering error of certain diffusion processes," *IEEE Trans. Inform. Theory*, vol. IT-18, pp. 325–331, May 1972.
[4] B. Z. Bobrovski and M. Zakai, "A lower bound in the estimation error for Markov processes," *IEEE Trans. Automat. Contr.*, vol. AC-20, pp. 785–788, Dec. 1975.
[5] I. B. Rhodes and A. S. Gilman, "Cone-bounded nonlinearities and mean square bounds—Estimation lower bound," *IEEE Trans. Automat. Contr.*, vol. AC-20, pp. 632–642, Oct. 1975.
[6] A. Gelb, Ed., *Applied Optimal Estimation*. Cambridge, MA: M.I.T. Press, 1974.
[7] G. Sansone and R. Conti, *Non-Linear Differential Equations*. Oxford, England: Pergamon, 1964.
[8] D. C. Rife, M. Goldstein, and R. R. Boorstyn, "A unification of Cramér–Rao type bounds," *IEEE Trans. Inform. Theory*, vol. IT-21, pp. 330–332, May 1975.

[4]Equation (5) is valid if the indicated partial derivatives exist and are absolutely integrable; these conditions are guaranteed by the problem statement; cf. p in (6).

[5]M_k is nonsingular if $x_K^k \triangleq x(KT_s; x_k, u(\tau)$: $kT_s \leq \tau \leq KT_s)$ is continuous with respect to x; this condition devolves from uniqueness, which is guaranteed if (2) is globally Lipschitzian [7]. Continuity of the Jacobian matrix is a stronger constraint.

[6]See footnote 1.