Likelihood Sensitivity and the Cramér–Rao Bound

WILLIAM A. GARDNER, MEMBER, IEEE

***Abstract*—The Cramér–Rao bound on the error variance of an unbiased estimator of an unknown parameter is given an intuitively pleasing interpretation in terms of the sensitivity of the likelihood function to changes in parameter value.**

Likelihood Sensitivity and the Cramér–Rao Bound

It has been said that the Cramér–Rao bound (1) on the error variance of an unbiased estimator $\hat{\theta}$ of an unknown parameter θ evades intuitive interpretation (e.g., [1, p. 232]). The purpose of this note is to bring to light an intuitively pleasing interpretation of the bound in terms of the sensitivity of the likelihood function of θ to changes in the value of θ.

Let $\hat{\theta}(X)$ be an unbiased estimator of an unknown parameter θ that is based on a random sample of data X, and let $p(\cdot|\theta)$ be the probability density function for X. The Cramér–Rao bound (normalized by θ) is defined to be the right member of the inequality

$$\frac{\operatorname{var}\{\hat{\theta}(X)-\theta\}}{\theta^2} \geqslant \left(\theta^2 E\left\{\left|\frac{d}{d\theta}\ln p(X|\theta)\right|^2\right\}\right)^{-1} \triangleq B. \quad (1)$$

This inequality was first stated by Fisher [2, sec. 7] and proved by Dugué [3], but is generally attributed to Cramér (cf. [4, sec. 32.3]) and Rao (cf. [5, p. 83]). The incremental sensitivity of the likelihood function of θ, $p(x|\cdot)$, to changes in the value of θ is defined by

$$\tilde{S}_\theta^p(x) \triangleq \left(\frac{\Delta p(x|\theta)}{p(x|\theta)}\right)\left(\frac{\Delta\theta}{\theta}\right)^{-1} \quad (2)$$

and is the ratio of the resultant percentage change in $p(x|\theta)$ to the causal percentage change in θ, evaluated at θ. The limiting value (as $\Delta\theta\to 0$) of (2) is given by

$$S_\theta^p(x) \triangleq \left(\frac{d}{d\theta}p(x|\theta)\right)\left(\frac{\theta}{p(x|\theta)}\right)$$
$$= \theta\frac{d}{d\theta}\ln p(x|\theta). \quad (3)$$

Therefore, the normalized bound (1) can be characterized by the sensitivity (3):

$$B = \left(E\left\{|S_\theta^p(X)|^2\right\}\right)^{-1}. \quad (4)$$

This yields the intuitively pleasing interpretation that the attainable mean square accuracy of an unbiased estimator is lower-bounded by the inverse mean square sensitivity of the likelihood function. If an unbiased estimator with error-variance as small as the Cramér–Rao bound exists, it is said to be efficient [2], and it is the maximum-likelihood estimator [2] defined by

$$p(x|\hat{\theta}) \geqslant p(x|\theta), \qquad \text{for all } \theta. \quad (5)$$

Thus, if the sensitivity of the likelihood function is high, the error variance of an efficient estimator is low and vice versa. This agrees with intuition.

A more concrete intuitive interpretation in terms of sensitivities can be given to the Cramér–Rao bound and its tightness for some specific models for $p(x|\theta)$, as exemplified in [6, p. 70] for the model where $X|\theta$ is a vector of independent identically distributed normal random variables with mean equal to a known function $f(\cdot)$ of an unknown parameter θ.

Manuscript received October 6, 1978.

The author is with the Department of Electrical Engineering, University of California, Davis, CA 95616.

References

[1] J. L. Melsa and D. L. Cohn, *Decision and Estimation Theory*. New York: McGraw-Hill, 1978.
[2] R. A. Fisher, "On the mathematical foundations of theoretical statistics," *Phil. Trans. Roy. Soc.*, London, vol. 222, pp. 309–368, 1922.
[3] D. Dugué, "Application des propriétés de la limite au sens du calcul des probabilités a l'étude de diverses questions d'estimation," *Jour. de l'Ec. Polytechn.*, pp. 305–372, 1937.
[4] H. Cramér, *Mathematical Methods of Statistics*. Princeton, NJ: Princeton Univ. Press, 1946.
[5] C. R. Rao, "Information and accuracy attainable in the estimation of statistical parameters," *Bull. Calcutta Math. Soc.*, vol. 37, pp. 81–91, 1945.
[6] H. L. Van Trees, *Detection, Estimation, and Modulation Theory, Part I*. New York: Wiley, 1968.

A Comparison of Some Kalman Estimators

DONALD M. LESKIW, MEMBER, IEEE, AND KENNETH S. MILLER

***Abstract*—Various estimators of a dynamical state vector are optimally combined to obtain new estimators. The performance of these new estimators is evaluated by comparing the traces of their covariance matrices. A nontrivial example is given to illustrate the techniques.**

I. Introduction

The Kalman theory develops a particular approach to the problem of estimating a dynamical state vector x. A system model is used to propagate an estimate of x throughout a time interval, say $[t_0, t)$, and a measurement model is used to update the estimate at time t. Suppose at a given point in time we have two unbiased estimates of x, say $\hat{x}$ and $\hat{y}$. For example, $\hat{x}$ could be an estimate obtained from a forward Kalman filter and $\hat{y}$ an estimate obtained from a backward Kalman filter. Then we may combine them in a linear fashion, say

$$\hat{\zeta} = \Gamma\hat{x} + \Delta\hat{y}$$

to obtain a new estimate $\hat{\zeta}$ of x. If the matrices Γ and Δ are so chosen that $\hat{\zeta}$ is an unbiased estimate of x, and also such that the trace of the covariance matrix of $\hat{\zeta}$ is minimized, then we call $\hat{\zeta}$ a smoothed estimate of x.

The specific problem we wish to address in this correspondence is the following. Suppose $\hat{x}(-)$ and $\hat{y}(-)$ are unbiased estimates of x at time t, and z is an independent observation at time t obtained from the measurement model. Using this information we wish to construct a better estimate of x at time t. We have available at least five options:

i) smooth $\hat{x}(-)$ and $\hat{y}(-)$, and then update with z;
ii) update $\hat{x}(-)$ with z to obtain $\hat{x}(+)$, and then smooth $\hat{x}(+)$ and $\hat{y}(-)$;
iii) update $\hat{y}(-)$ with z to obtain $\hat{y}(+)$, and then smooth $\hat{x}(-)$ and $\hat{y}(+)$;
iv) simultaneously process $\hat{x}(-)$, $\hat{y}(-)$, and z;

Manuscript received July 7, 1978, revised December 4, 1978. This work was supported in part by the Department of the Army under a subcontract of M.I.T. Lincoln Laboratory. Any views and conclusions contained in this document are those of the contributors and should not be interpreted as necessarily representing the official policies either expressed or implied of the United States government.

The authors are with the Riverside Research Institute, 80 West End Ave., New York, NY 10023.

0018-9448/79/0700-0491$00.75