AIAA 80-1787R

Marginal Optimization of Observation Schedules

Angus Andrews*
Rockwell International Science Center, Thousand Oaks, Calif.

Introduction

THIS Note presents a method for optimizing the selection of observations to be used in estimating the state of a linear dynamic system. Problems of this type arise in numerous applications, including space navigation[1,2] and calibration[3,4] and alignment[5] of inertial navigation systems. In these problems, the associated loss function, which is to be minimized in selecting the observations, depends upon the propagation in time of the estimation errors. For example, in calibration and alignment of inertial navigation systems, the objective is to optimize the subsequent performance of the navigation system. Battin[1] and Denham and Speyer[2] derived differential improvement methods for the observation schedule, for the case that the quadratic loss function depends upon the estimation errors at a fixed time, and the objective is to define the optimum values of certain parameters, with respect to which the measurement sensitivity matrices are differentiable.

A concise representation of generalized time-dependent quadratic loss functions was presented in a previous Note.[6] In this Note, the equivalent representation is developed for a discrete-time system. This results in a recursion formula for defining the loss functions, and a computationally efficient algorithm for selecting observations in the order of maximum marginal improvement of the loss function.

An Ill-Posed Optimization Problem

Let x be the state of a linear multistage process

$$x_k = \Phi_k x_{k-1} + G_k u_k \qquad (k=1,2,3,\ldots,N) \tag{1}$$

with process noise u_k of known covariance Q. Suppose that, for each k, for some index set J_k, for each $j \in J_k$ there is a possible observation

$$y(j) = H(j) x_k + v(j) \tag{2}$$

with noise $v(j)$ of known covariance $R(j)$. In different applications, J_k may be a finite set or a compact connected set.

One would like to optimize the choices of $j \in J_k$ in this sequence of observations, with respect to a suitable loss function. The loss functions that appear most natural to the problem depend upon the covariances P_k of estimation uncertainty. One can readily show that the resulting optimization problem is "ill-posed," in the sense that the solution is not necessarily unique.

Consider the case that $\Phi \equiv 0$, $G \equiv 0$ (constant states) and the loss function depends only upon the final covariance P_N of estimation uncertainty. But

$$P_N = \left[P_0^{-1} + \sum_{k=1}^{N} H_k^T R_k^{-1} H_k \right]^{-1} \tag{3}$$

where P_0 is the a priori covariance of estimation uncertainty. Note that P_N, and therefore the loss function, will be invariant under all permutations of the observation sequence. It is not difficult to construct examples for which an optimal observation sequence is not invariant under permutations of order. In these cases, then, the optimal solution is not unique.

Although this optimization problem has been posed for some time,[1,2] and there have been allusions to the existence of solutions,[5] there are no general solution methods known to the author. The following is a general solution method for optimization in a restricted sense.

Marginal Optimization

A sequence of observations $\{(H_k, R_k),\ k=1,2,3,\ldots\}$ will be called "marginally optimal" with respect to a loss function $\mathcal{L}$ if, for each $i>0$, given the previous observations, the ith observation is chosen to minimize $\mathcal{L}$, assuming no subsequent observations. Such sequences are not necessarily optimal in the broader sense (except for $N=1$), and it is not difficult to construct examples for which they are strictly suboptimal. However, one can define a general method for marginal optimization with respect to a relatively broad class of quadratic loss functions of estimation errors.

Marginal Quadratic Loss Functions

We need a quadratic loss function for the case that no observations be made after the ith stage. These loss functions are of the form

$$\mathcal{L}_i = \sum_{k=i}^{N} w_k E\langle |M_k x_k|^2 \rangle \qquad (i=0,1,2,\ldots,N) \tag{4}$$

where x_k is the n-vector of estimation uncertainty at the kth stage and $E\langle\ \rangle$ denotes the expected value. The scalars w_k and matrices M_k can be tailored to the problem at hand. It was shown in a previous Note[6] how these can be defined for most measures of performance of inertial navigation systems. One of the principal results of this Note is that these loss functions allow a "trace factorization" into statistical and dynamical factors:

$$\mathcal{L}_i = \text{trace}[P_i W_i] + \text{trace}[Q V_i] \tag{5}$$

where the trace of a square matrix is the sum of its diagonal elements and

$$W_i = B_i^{-T} \left(\sum_{k=i}^{N} w_k B_k^T M_k^T M_k B_k \right) B_i^{-1} \tag{6}$$

$$V_i = \sum_{k=i+1}^{N} G_k^T W_k G_k \tag{7}$$

$$B_k = \Phi_k B_{k-1} \qquad (k>i) \tag{8}$$

The proof of Eqs. (5-8) depends upon the following well-known formula for the dynamical update of the covariance of uncertainty:

$$P_k = \Phi_k P_{k-1} \Phi_k^T + G_k Q G_k^T \tag{9}$$

Using this, one can readily prove, by recursion on $k \geq i$, that

$$P_k = B_k \left[B_i^{-1} P_i B_i^{-T} + \sum_{j=i+1}^{k} B_j^{-1} G_j Q G_j^T B_j^{-T} \right] B_k^T \tag{10}$$

Presented as Paper 80-1787 at the AIAA Guidance and Control Conference, Danvers, Mass., Aug. 11-13, 1980; submitted Oct. 24, 1980; revision received Aug. 6, 1981.

*Technical Staff. Member AIAA.

for B_k defined by Eq. (8). Equation (10) will be used in Eq. (14) below. In the following sequence of equations, we will use the facts that $|z|^2 = \text{trace}[zz^T]$ for any vector z, and the trace of a matrix product is invariant under cyclic permutations of the factors:

$$\mathcal{L}_i = \sum_{k=i}^{N} w_k E\langle \text{trace}[(M_k x_k)(M_k x_k)^T]\rangle \tag{11}$$

$$= \sum_{k=i}^{N} w_k E\langle \text{trace}[(x_k x_k^T)(M_k^T M_k)]\rangle \tag{12}$$

$$= \text{trace} \sum_{k=i}^{N} w_k P_k M_k^T M_k \tag{13}$$

$$= \text{trace}\left[P_i B_i^{-T}\left(\sum_{k=i}^{N} w_k B_k^T M_k^T M_k B_k\right) B_i^{-1}\right]$$

$$+ \text{trace}\left[Q \sum_{k=i+1}^{N} w_k \sum_{j=i+1}^{k} G_j^T B_j^{-T} B_k^T M_k^T M_k B_k B_j^{-1} G_j\right] \tag{14}$$

$$= \text{trace}[P_i W_i] + \text{trace}\left[Q \sum_{j=i+1}^{N} G_j^{-T} B_j^{-T}\right.$$

$$\left.\times\left(\sum_{k=j}^{N} w_k B_k^T M_k^T M_k B_k\right) B_j^{-1} G_j\right] \tag{15}$$

$$= \text{trace}[P_i W_i] + \text{trace}[Q V_i] \tag{16}$$

which completes the proof.

Benefit Function for Kalman Filtering

If we assume that the Kalman filter will be used for estimating the system state, then the change in the covariance of estimation uncertainty that would result from an observation will be given by the following well-known replacement formula:

$$P_i := P_i - P_i H_i^T (H_i P_i H_i^T + R_i)^{-1} H_i P_i \tag{17}$$

The corresponding change in the loss functions will be the following:

$$\mathcal{L}_i := \mathcal{L}_i - \text{trace}[P_i H_i^T (H_i P_i H_i^T + R_i)^{-1} H_i P_i W_i] \tag{18}$$

The last term will be called the "benefit function." It represents the decrease in the loss function resulting from the choice of the observation represented by (H_i, R_i). Note that it does not depend upon V_i. It can also be expressed in the following form:

$$\mathcal{B}_i(H,R) = \text{trace}[(HP_iH^T + R)^{-1}(HP_i W_i P_i H^T)] \tag{19}$$

This represents the benefit as a differentiable function of H and R. The problem of maximizing the benefit by choosing H and R is then a relatively straightforward optimization problem. In the case that a finite number of choices are available, it can be solved by enumeration. The trace factorization is particularly efficient in this case, because it requires only n^2 multiplies to compute the trace of the product of two $n \times n$ matrices. In the case that H and R are differentiable functions of real parameters $j_1, j_2, \ldots, j_m$, it can be solved by differential (e.g., Newton's) methods and evaluations at boundaries of the parameter domain J_k.

Loss Function for *UD* Filtering

The marginal optimization procedure can also be posed in terms of the "UD" filter mechanization of Bierman.[7] In this case, the covariance of estimation uncertainty is represented by a unit upper triangular matrix U and a diagonal matrix D such that $P = UDU^T$. The update equations for U and D analogous to Eqs. (17) and (19) are given by Bierman. The analogous equation for the loss function is the following:

$$\text{trace}(PW) = \sum_{i=1}^{n} D_{ii} \sum_{j=1}^{i}\left[\sum_{k=j}^{i} U_{kj} L_{ki}\right]^2 \tag{20}$$

where U and D are evaluated after the observation, and where L is the lower triangular matrix such that $W = LL^T$. In this case, the dependence of U and D (and $\mathcal{L}$) upon H and R is given by an algorithm, and is not expressed as a differentiable function.

Recursive Generation of W_i and V_i

Note that all proofs depending upon Eq. (8) do not require any particular initial or final value of the B-matrices. If we further define B_N as the identity matrix I, then

$$W_N = w_N M_N^T M_N \tag{21}$$

and the W_i can be generated by backward recursion, using the following relations:

$$B_i = \Phi_{i+1}^{-1} B_{i+1} \qquad W_i = W_{i+1} + w_i B_i^T M_i^T M_i B_i \tag{22}$$

The V-matrices are not necessary for the optimization procedure. However, they can be generated by the following equations:

$$V_N = 0 \qquad V_{i-1} = V_i + G_i^T B_i^{-T} W_i B_i^{-1} G_i \tag{23}$$

The Marginal Optimization Procedure

Given the initial covariance of state uncertainty P_0, the matrices Φ_k, G_k, Q defining the state dynamics, and the scalars w_k and matrices M_k defining the loss function, the following procedure defines the marginally optimal observation sequence:

1) For $i = N, N-1, \ldots, 1$ generate the matrices W_i by using Eqs. (21-22).
2) For $i = 1, 2, \ldots, N$:
   2a) Update P_i using Eq. (9).
   2b) Select (H,R) to maximize $\mathcal{B}_i$, using Eq. (19), or to minimize $\mathcal{L}_i$, using Eq. (20).
   2c) Update P_i using Eq. (17).

References

[1] Battin, R.H., *Astronautical Guidance*, McGraw-Hill, New York, 1964, pp. 325-331.

[2] Denham, W.F. and Speyer, J.L., "Optimal Measurement and Velocity Correction Programs for Midcourse Guidance," *AIAA Journal*, Vol. 2, May 1964, pp. 896-907.

[3] Friedland, B., "On the Calibration Problem," *IEEE Transactions on Automatic Control*, Vol. AC-22, Dec. 1977, pp. 899-905.

[4] Andrews, A., "Optimum Calibration of Inertial Navigation Systems," *Optimization Days*, McGill University, Toronto, May 1979, pp. 25-26.

[5] Baziw, J. and Leondes, C.T., "In-flight Alignment and Calibration of Inertial Measurement Units—Part II: Experimental Results," *IEEE Transactions on Aerospace and Electronic Systems*, Vol. AES-8, July 1972, pp. 450-465.

[6] Andrews, A., "Method for Error Budgeting of Inertial Navigation Systems," *Journal of Guidance and Control*, Vol. 2, Sept. 1979, pp. 438-440.

[7] Bierman, G.J., *Factorization Methods for Discrete Sequential Estimation*, Academic Press, New York, 1977, pp. 76-90.