COVARIANCE MATRIX REPRESENTATION IN LINEAR FILTERING

Christopher C. Paige†

Abstract. The two usual mathematical approaches to estimating the state vector in dynamic linear models are Kalman filtering and information filtering. If there is no information on part of the initial state vector then the Kalman filtering model can not be used directly, while if part is known exactly then the information filtering model does not hold. If one part of the initial state vector is known accurately and there is no knowledge on some other part, then neither of these two approaches apply in theory. Altering the models to handle these cases can lead to numerically unreliable results. This is because poor information results in very large covariance matrices in Kalman filtering, while very good information leads to very large matrices in information filtering, and these can cause numerical difficulties. Thus even when these extreme cases do not hold, the two approaches while being correct mathematically may not be ideal numerically.

Here we suggest a general representation in which both the covariance and information matrices are implicitly defined. This gives a unified theoretical approach to the problem and fills in the gaps where Kalman and information filtering do not apply. It also allows the development of numerically reliable algorithms in all cases, and these may be desirable when accuracy is required in difficult cases. An example is given to show how to develop such algorithms.

1. Introduction. The filtering problem for discrete linear dynamic systems assumes we have a *time* equation relating the state x at time $k+1$ to the state at time k, and a *measurement* equation relating the output y to the state, both at time k. The time equation is

$$x_{k+1} = \underbrace{A_k}_{n\times n} x_k + \underbrace{B_k}_{n\times m} w_k, \qquad k = 1, 2, \ldots, \tag{1.1}$$

where the noise w_k satisfies

$$\mathcal{E}(w_k) = 0, \qquad \mathcal{E}(w_j w_k^T) = \delta_{jk} Q_k. \tag{1.2}$$

Here $\delta_{jk} = 1$ if $j = k$, but is zero otherwise. We will write (1.2) more briefly as

$$w_k \frown (0, Q_k), \tag{1.3}$$

where unless otherwise stated we will assume such sequences as $w_1, w_2, \ldots$, have uncorrelated vectors w_j.

The measurement equation is

$$y_k = \underbrace{C_k}_{p\times n} x_k + v_k, \qquad v_k \frown (0, R_k), \tag{1.4}$$

and unless otherwise stated we will assume such vectors as u_i and v_j are uncorrelated. The matrices A_k, B_k, C_k are assumed known, and the aim is to use the known y_j in order to estimate x_k. The approach

1980 Mathematics Subject Classification. 93E10, 65F20, 62J05.

† Supported by Natural Sciences and Engineering Research Council of Canada Grant A8652.

used is to find the best linear unbiased estimate where best means the minimum variance estimate, see [2]. Let $x_{k|j}$ be the minimum variance estimate of x_k given

$$y_1, \ldots, y_j. \tag{1.5}$$

We are also interested in the covariance of the error of the estimates, and write

$$\tilde{x}_{k|j} \equiv x_{k|j} - x_k, \tag{1.6}$$

$$P_{k|j} \equiv \mathcal{E}\{[\tilde{x}_{k|j} - \mathcal{E}(\tilde{x}_{k|j})][\tilde{x}_{k|j} - \mathcal{E}(\tilde{x}_{k|j})]^T\}. \tag{1.7}$$

Usually we are interested in unbiased estimators and so

$$\mathcal{E}(\tilde{x}_{k|j}) = 0. \tag{1.8}$$

There are two main approaches in the literature for obtaining the solution to this estimation problem, the key difference being the treatment of the covariance matrix. The first approach, called *covariance filtering*, effectively obtains the covariance matrix $P_{k|k}$ from $P_{k|k-1}$ by incorporating the measurement step (1.4), and then obtains $P_{k+1|k}$ from $P_{k|k}$ using the time step (1.1). The other approach obtains $P^{-1}_{k|k}$ from $P^{-1}_{k|k-1}$ using (1.4), and $P^{-1}_{k+1,k}$ from $P^{-1}_{k|k}$ using (1.1). This second approach is often called *information filtering* because of the relation of the covariance inverse to the *Fisher information matrix* ,see [12, pp.240-241]. This reference points out that *inverse covariance filtering* is a more precise terminology, and this will be used here.

These two approaches are fundamentally mathematically different in that there are problems which one can handle but the other cannot. For example if $P_{1|0} = O$ (we use O for zero matrices, 0 for zero vectors), corresponding to x_1 being known exactly, then the inverse of $P_{1|0}$ does not exist and inverse covariance filtering is not defined, but covariance matrix filtering is. See [2, p.140] for the reverse case.

For any given theoretical approach to obtaining estimates, there are many different ways of computing these estimates. In this area there are two broad classes of computational approaches. The first is to compute the actual covariance matrix, or its inverse, directly from the previous one. This is generally referred to as *filtering*. The second approach uses the fact that a covariance matrix, or its inverse, is symmetric nonnegative definite, and so can be factorized, as for example in

$$P_{k|k-1} = S_k S_k^T. \tag{1.9}$$

The so called *square root filtering* approach computes the factor such as S_k, or its inverse, directly from previous factor, or inverse. Note that the positive square root of a symmetric nonnegative definite matrix is also symmetric nonnegative definite, and as S_k in (1.9) is usually triangular, the terminology *factor filtering* is preferable, and will be used here. Factor filtering is preferable to updating full covariance matrices because it is numerically more accurate and the product $S_k S_k^T$ can never be indefinite or unsymmetric, while rounding errors could cause an updated P to be so, see for example [3].

The original filter in dynamic systems was introduced by Kalman [9] and is a covariance filter. This assumes x_1 comes from a distribution

$$x_1 \frown (x_{1|0}, P_{1|0}) \tag{1.10}$$

where this mean and covariance are known. The Kalman formulae are then as follows, see for example [2, p.39].

Measurement update:

$$P_{k|k} = P_{k|k-1} - P_{k|k-1}C_k^T[R_k + C_k P_{k|k-1} C_k^T]^{-1} C_k P_{k|k-1}. \tag{1.11}$$

Here

$$K_k \equiv P_{k|k-1}C_k^T[R_k + C_k P_{k|k-1}C_k^T]^{-1} \tag{1.12}$$

is also used in

$$x_{k|k} = x_{k|k-1} + K_k[y_k - C_k x_{k|k-1}]. \tag{1.13}$$

The matrix $A_k K_k$ is called the *gain matrix* in [2, p.140].

Time update:

$$x_{k+1|k} = A_k x_{k|k} \tag{1.14}$$

$$P_{k+1|k} = A_k P_{k|k} A_k^T + B_k Q_k B_k^T. \tag{1.15}$$

There are many realistic models where this approach can break down. One case is where nothing is known about x_1, or part of x_1, resulting in $P_{1|0}$ being nonexistent and the approach being inapplicable. Another occurs when the inverse in (1.11) does not exist, however see the comments on the use of the psuedo-inverse in [2, pp.25, 40 & 322]. Difficulty also occurs if Q_k in (1.2) or R_k in (1.4) does not exist. These difficulties can sometimes be circumvented by use of the inverse covariance filter of Fraser [7]. For the case where all inverses exist the formulae for covariance inverses can be obtained from the Kalman formulae (1.11) and (1.15) by use of the symmetric form of the *matrix inversion lemma* given by Duncan [5]. See [13, p.190] for some history on this well known result.

$$(H + GMG^T)^{-1} = H^{-1} - H^{-1}G(G^T H^{-1} G + M^{-1})^{-1} G^T H^{-1}, H = H^T, M = M^T. \tag{1.16}$$

It follows from this and (1.11) that

$$P_{k|k}^{-1} = P_{k|k-1}^{-1} + C_k^T R_k^{-1} C_k. \tag{1.17}$$

for the inverse covariance measurement update, while applying (1.16) to (1.15) gives

$$P_{k+1|k}^{-1} = M_k - M_k B_k (B_k^T M_k B_k + Q_k^{-1})^{-1} B_k^T M_k, \tag{1.18}$$

$$M_k \equiv A_k^{-T} P_{k|k}^{-1} A_k^{-1}, \tag{1.19}$$

for the inverse covariance time update. The corresponding expressions for the estimates are derived for example in [2, p. 140].

It is clear that the inverse covariance formulae cannot be used directly if $P_{1|0}$ or R_k or Q_k or A_k or the term $(\ldots)$ in (1.18) is singular. Even if none of these is singular, there can still be serious numerical problems if any are ill-conditioned with respect to solution of equations, see for example [8]. A similar comment holds for the original Kalman covariance filter. Such numerical difficulties can be allayed to some extent if the factored versions of these filters are used, see Dyer and McReynolds [6], Kaminski Bryson and Schmidt [10], Carlson [4], Bierman [3], and Maybeck [12]. Nevertheless not all possibilities are covered by these approaches, for example x_1 could have some elements unknown and others known exactly. Nor do any of these approaches necessarily lead to numerically stable algorithms for certain classes of problems.

Here we advocate a more liberal approach to the representation of covariance matrices, arguing that the use of covariance matrices only, or inverse covariance matrices only, is unnecessarily restrictive. In section 2 we show how an obvious filter which uses both covariance and inverse covariance matrices is both simple to derive and state, and is quite fast to compute. This suggests a more flexible approach to covariance matrix representation, so in Section 3 we suggest a more general representation of the covariance structure of a random vector, and indicate a very powerful overall approach to solving any such problem.

2. An Efficient Mixed Filter. A very simple set of filter equations arises if we allow both covariance and inverse covariance matrices in the one formulation. We can use the inverse covariance matrix measurement update,

$$P_{k|k}^{-1} = P_{k|k-1}^{-1} + C_k^T R_k^{-1} C_k, \tag{2.1}$$

which is (1.17), and the covariance matrix time update

$$P_{k+1|k} = A_k P_{k|k} A_k^T + B_k Q_k B_k^T, \tag{2.2}$$

which is (1.15). There is no problem in going from the triangular factor of a nonsingular matrix to the factor of its inverse, and since the factored versions are also numerically preferable we will derive these afresh, since this gives an introduction to ideas used later. We will use the notation

$$P_{k|k}^{-1} = U_k^T U_k, \quad P_{k|k-1} = S_k S_k^T, \quad R_k^{-1} = (R_k^{if})^T R_k^{if}, \quad Q_k = Q_k^f (Q_k^f)^T, \tag{2.3}$$

where *f* stands for *factor*, *if* for *inverse factor*, and here

$$U_k, \quad S_k, \quad R_k^{if}, \quad Q_k^f \quad \text{are upper triangular.} \tag{2.4}$$

Suppose we have $x_{1|0}$ and S_1 nonsingular, where it is known that x_1 comes from a distribution with

$$x_1 \frown (x_{1|0}, S_1 S_1^T).$$

We then have,see [1]

$$x_{1|0} = x_1 + S_1 u_1, \qquad u_1 \frown (0, I), \tag{2.5}$$

which with the first measurement equation (1.4) and

$$\tilde{v}_k = R_k^{if} v_k \frown (0, I) \tag{2.6}$$

gives

$$\begin{bmatrix} S_1^{-1} x_{1|0} \\ R_k^{if} y_1 \end{bmatrix} = \begin{bmatrix} S_1^{-1} \\ R_k^{if} C_1 \end{bmatrix} x_1 + \begin{bmatrix} u_1 \\ \tilde{v}_1 \end{bmatrix}, \qquad \begin{bmatrix} u_1 \\ \tilde{v}_1 \end{bmatrix} \frown (0, I). \tag{2.7}$$

We now carry out the transformation

$$T_1 \left[\begin{array}{cc|c} S_1^{-1} & | & S_1^{-1} x_{1|0} \\ R_1^{if} C_1 & | & R_1^{if} y_1 \end{array}\right] = \left[\begin{array}{cc|c} U_1 & | & b_1 \\ O & | & r_1 \end{array}\right], \tag{2.8}$$

where T_1 is orthogonal and U_1 is upper triangular and necessarily nonsingular. This transforms (2.7) to

$$\begin{bmatrix} b_1 \\ r_1 \end{bmatrix} = \begin{bmatrix} U_1 \\ O \end{bmatrix} x_1 + \begin{bmatrix} \tilde{u}_1 \\ r_1 \end{bmatrix}, \qquad \text{where } \begin{bmatrix} \tilde{u}_1 \\ r_1 \end{bmatrix} = T_1 \begin{bmatrix} u_1 \\ \tilde{v}_1 \end{bmatrix} \frown (0, I). \tag{2.9}$$

The required estimate $x_{1|1}$ of x_1 is then the solution of

$$U_1 x_{1|1} = b_1, \qquad U_1 \quad \text{nonsingular,} \tag{2.10}$$

and then

$$U_1 x_{1|1} = b_1 = U_1 x_1 + \tilde{u}_1, \qquad \tilde{u}_1 \frown (0, I), \tag{2.11}$$

$$U_1 (x_{1|1} - x_1) = \tilde{u}_1, \qquad x_{1|1} = x_1 + U_1^{-1} \tilde{u}_1, \tag{2.12}$$

$$U_1 P_{1|1} U_1^T = I, \qquad P_{1|1} = U_1^{-1} U_1^{-T}, \qquad P_{1|1}^{-1} = U_1^T U_1. \tag{2.13}$$

Note that (2.12) has similar form to (2.6). When we combine (2.12) with the time update (1.1) and (1.2) and

$$w_k = -Q_k^f \tilde{w}_k, \qquad \tilde{w}_k \frown (0, I), \tag{2.14}$$

We obtain

$$x_2 = x_{2|1} - A_1 U_1^{-1} \tilde{u}_1 - B_1 Q_1^f \tilde{w}_1, \qquad x_{2|1} = A_1 x_{1|1}, \tag{2.15}$$

where $\tilde{u}_1$, $\tilde{w}_1$ have zero mean, unit covariance, and are uncorrelated with each other. It follows that

$$P_{2|1} = S_2 S_2^T = (A_1 U_1^{-1}, B_1 Q_1^f)(A_1 U_1^{-1}, B_1 Q_1^f)^T, \tag{2.16}$$

and so we compute S_2 from the orthogonal factorization

$$(A_1 U_1^{-1}, B_1 Q_1^f)\tilde{T}_1 = (O, S_2). \tag{2.17}$$

With

$$\tilde{T}_1^T \begin{bmatrix} \tilde{u}_1 \\ \tilde{w}_1 \end{bmatrix} = \begin{bmatrix} \tilde{r}_1 \\ u_2 \end{bmatrix} \frown (0, I),$$

we have from (2.15) and (2.17)

$$x_{2|1} = x_2 + S_2 u_2, \qquad u_2 \frown (0, I). \tag{2.18}$$

This has the identical form to (2.5), and so the computation can be continued for the next step. The computation is made up of the orthogonal transformations (2.8) and (2.17), with the state estimates given by (2.10) and (2.15). The equation (2.1) follows immediately from (2.8), while (2.2) follows from (2.17). Note that since we are dealing with the *factors* S_k and U_k there is no difficulty in switching from covariance matrices to inverse covariance matrices and back. The result is a filter that is easy to derive and understand, and is computationally efficient. However, like the filters in Section 1 it will not handle all cases, for example if $P_{1|0}$ is singular, and will not necessarily give good numerical results when some matrices are ill-conditioned for solution of equations.

3. General Covariance Structure. In section 2 we made use of different ways of representing the covariance structure of vectors. In (2.5) we used

$$x_1 = x_{1|0} - S_1 u_1, \qquad u_1 \frown (0, I), \tag{3.1}$$

corresponding to the traditional representation

$$x_1 \frown (x_{1|0}, S_1 S_1^T). \tag{3.2}$$

In (2.6) we used

$$R_k^{if} v_k = \tilde{v}_k \frown (0, I), \tag{3.3}$$

which corresponds to

$$v_k \frown (0, R_k), \qquad R_k^{-1} = (R_k^{if})^T R_k^{if} \tag{3.4}$$

when the inverse exists. In (2.11) we used

$$U_1 x_1 = b_1 - \tilde{u}_1, \qquad \tilde{u}_1 \frown (0, I), \qquad U_1 x_{1|1} = b_1, \tag{3.5}$$

which when U_1 has an inverse corresponds to the more usual

$$x_1 = x_{1|1} - u_1', \qquad \text{where } u_1' \equiv U_1^{-1} \tilde{u}_1 \frown (O, U_1^{-1} U_1^{-T}), \tag{3.6}$$

while in (2.14) we used

$$w_k = -Q_k^f \tilde{w}_k, \qquad \tilde{w}_k \frown (0, I) \tag{3.7}$$

instead of

$$w_k \frown (0, Q_k^f (Q_k^f)^T). \tag{3.8}$$

Every odd numbered equation above expresses a vector of interest in terms of a zero mean unit covariance vector, as this is the natural representation using covariance matrix factors, and leads to good computational algorithms. We see that (3.3) is essentially (3.5) with the vector of interest having zero mean, with a similar comment for (3.7) viz a viz (3.1).

Thus we have used two distinctly different representations of the covariance structure of a random vector x, that is

$$x = \bar{x} + Su, \qquad u \frown (0, I), \tag{3.9}$$

as in (3.1), and

$$Ux = b + u, \qquad u \frown (0, I), \tag{3.10}$$

as in (3.5). When $S = U^{-1}$ these are equivalent, but (3.9) allows linear combinations of the elements of x to be constant (known *a priori*) by having an S which does not have full row rank, while (3.10) allows the possibility when U has less than full column rank of having no information at all on the part of x in $\mathcal{N}(U)$, the null space of U.

The two representations are mathematically quite distinct except when S and U are square and non-singular, but even then they will probably lead to numerically different results. For example when U has norm 1 but is ill-conditioned with respect to solution of equations, then $S = U^{-1}$ will have a very large norm. The obvious generalization is to replace both (3.9) and (3.10) by

$$Ux = b + Su, \qquad u \frown (0, I), \tag{3.11}$$

where ideally the norms of S and U are of reasonable size. This has the form of (3.9) if $U = I$, and the form of (3.10) if $S = I$, but it also covers cases which neither (3.9) nor (3.10) can handle. For example taking

$$U = \begin{bmatrix} 1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \qquad S = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

in (3.11) we can have the sum of the first two elements of x being random, their difference constant, with no information on the third element.

It is important to have such freedom in representing all available knowledge of the initial vector x_1 in filtering, and thus to have one expression

$$\hat{b}_1 = \hat{U}_1 x_1 + \hat{S}_1 \hat{u}_1, \qquad \hat{u}_1 \frown (0, I), \tag{3.12}$$

which can be used to include the initial conditions in covariance filtering, inverse covariance filtering, and any feasible combination of both. However the noise vectors w_k and v_k in (1.1) and (1.4) also have different representations in covariance filtering, which uses the representation in (3.7) or (3.8), and inverse covariance filtering, which uses (3.3). To provide a general representation covering any feasible combination of these cases, and the possibility of correlation between v_k and w_k we could write

$$\begin{bmatrix} R_k^l & \\ N_k^l & Q_k^l \end{bmatrix} \begin{bmatrix} v_k \\ w_k \end{bmatrix} = - \begin{bmatrix} R_k^r & \\ N_k^r & Q_k^r \end{bmatrix} \begin{bmatrix} \tilde{v}_k \\ \tilde{w}_k \end{bmatrix}, \qquad \begin{bmatrix} \tilde{v}_k \\ \tilde{w}_k \end{bmatrix} \frown (0, I) \tag{3.13}$$

instead of (1.3) and (1.4), where l and r stand for left and right. Of course this would introduce far more flexibility than is usually needed, but it is included here for generality. Also different representations of (3.13) will lead to different algorithms with different efficiencies.

This rather general discrete linear dynamic model could then be expressed in one large structured model as

$$\begin{bmatrix} \hat{b}_1 \\ y_1 \\ 0 \\ 0 \\ d_1 \\ y_2 \\ 0 \\ \cdot \\ \cdot \end{bmatrix} = \begin{bmatrix} \hat{U}_1 & & & & & & \\ C_1 & I & & & & & \\ & R_1^l & & & & & \\ & N_1^l & Q_1^l & & & & \\ A_1 & & B_1 & E_1 & & & \\ & & & C_2 & I & & \\ & & & & R_2^l & & \\ & & & & \cdot & \cdot & \\ & & & \cdot & & \cdot & \cdot \end{bmatrix} \begin{bmatrix} x_1 \\ v_1 \\ w_1 \\ x_2 \\ v_2 \\ \cdot \\ \cdot \end{bmatrix} + \begin{bmatrix} \hat{S}_1 & & & & & \\ O & & & & & \\ & R_1^r & & & & \\ & N_1^r & Q_1^r & & & \\ & & O & & & \\ & & O & & & \\ & & & R_2^r & & \\ & & & \cdot & \cdot & \\ & & & & \cdot & \end{bmatrix} \begin{bmatrix} \hat{u}_1 \\ \tilde{v}_1 \\ \tilde{w}_1 \\ \tilde{v}_2 \\ \cdot \end{bmatrix} \tag{3.14}$$

with all the $\hat{u}_1, \tilde{v}_j, \tilde{w}_j$ having zero means, unit covariances, and being uncorrelated. With obvious notation this can be written

$$y = Fx + Lu, \qquad u \frown (0, I). \tag{3.15}$$

Note that (3.14) includes known inputs d_j to the system equations, which can be in descriptor form and have $E_j \neq I$, where these E_j could be singular. Such a model can simply be altered to include any other correlations that are present.

The important point of (3.15) is not that F and L need ever be stored as full matrices, but that (3.14) shows the full structure of the problem. This facilitates the derivation of efficient algorithms for solving the problem. In fact since no products of matrices appear in (3.14) it has as its subblocks the original data, and a numerically stable algorithm for finding the optimal estimate for x in (3.15) will be numerically stable for the original problem. But numerically stable methods for computing optimal estimates for x in (3.15) have already been given in [11], [14] and [15], and these approaches can be used here, only now care must be taken to design algorithms which take advantage of the structure to ensure efficiency. This approach was used in [18] for the special case of inverse covariance filtering where $B_k = I$ in (1.1). Smoothed estimates were then seen to be available with minimal extra computation, and the covariance factors of these smoothed estimates could be obtained quite easily.

In general the formulation (3.14) and (3.15) gives great ease in understanding the problem and in formulating effective and numerically stable algorithms, as well as encompassing the covariance matrix filtering problem and the inverse covariance matrix filtering problem, and being more general than both.

Note that (3.15) has exactly the same form as (3.12), which is a natural and desirable consistency. It is shown in [11] that the minimum variance estimate for x in (3.15) is the solution of the generalized linear least squares problem

$$\underset{x,\,u}{\text{minimize}} \quad u^T u \qquad \text{subject to} \qquad y = Fx + Lu, \tag{3.16}$$

and it is shown how to solve this in a numerically stable way. Note since (3.12) has the form of (3.15), the initial estimate $x_{1|0}$ for x_1 is the solution of

$$\underset{x_1,\,\hat{u}_1}{\text{minimize}} \quad \hat{u}_1^T \hat{u}_1 \qquad \text{subject to} \qquad \hat{b}_1 = \hat{U}_1 x_1 + \hat{S}_1 \hat{u}_1, \tag{3.17}$$

but if $\hat{U}_1$ has full row rank this satisfies

$$\hat{U}_1 x_{1|0} = \hat{b}_1. \tag{3.18}$$

It turns out that numerically stable solutions to the filtering problem for (3.14) and (3.15) lead to expressions of the form (3.12) for each x_k, and so the above comments apply.

It is also shown in [17] how the geometric structure of such problems as (3.15) and (3.16) is revealed by the generalized singular value decomposition (GSVD) of F and L, see [19]. Although GSVD algorithms would rarely be used on large F and L in (3.15), they may well give useful information on the individual state vector in for example (3.12).

4. A General Covariance Matrix Factor Filter. The model (3.15) with structure (3.14) is far more general than would usually be needed, and a numerically stable and efficient algorithm for this would be inefficient for most practical uses. As a result we will only give algorithms for some of the more likely specializations of (3.14). By following these ideas practitioners will be able to design stable and efficient algorithms for their own specialization of (3.14).

Here we give an algorithm for the model with the fully general initial condition (3.12), but with only covariance matrix factors of the noises v_k and w_k supplied. This will correspond to $R_k^l = I$, $N_k^l = O$, $Q_k^l = I$ in (3.14), and these unit matrices could then be eliminated. It is more straightforward to set up the model without them initially, but working from (3.14) illustrates an important aspect of the numerical stability of the approach used.

To eliminate the third (block) row of (3.14) when $R_1^l = I$ we transform the nonzero matrices in the second and third rows as follows

$$\begin{bmatrix} O & I \\ I & -I \end{bmatrix} \begin{bmatrix} y_1 & | & C_1 & I & | & O \\ 0 & | & O & I & | & R_1^r \end{bmatrix} = \begin{bmatrix} 0 & | & O & I & | & R_1^r \\ y_1 & | & C_1 & O & | & -R_1^r \end{bmatrix}. \tag{4.1}$$

Since $N_1^l = O$ the unit matrix in the right hand side of (4.1) has only zero matrices above and below it in the transformed (3.14), and the new second row in (3.14) will not affect the solution of (3.16) and can be dropped, along with the column of F in which this I lies.

To eliminate the original fourth row of (3.14) when $N_1^l = O$, $\quad Q_1^l = I$ we transform the nonzero matrices in the original fourth and fifth rows as follows

$$\begin{bmatrix} -B_1 & I \\ I & O \end{bmatrix} \begin{bmatrix} 0 & | & O & I & O & | & N_1^r & Q_1^r \\ d_1 & | & A_1 & B_1 & E_1 & | & O & O \end{bmatrix} = \begin{bmatrix} d_1 & | & A_1 & O & E_1 & | & -B_1 N_1^r & -B_1 Q_1^r \\ 0 & | & O & I & O & | & N_1^r & Q_1^r \end{bmatrix}. \tag{4.2}$$

With an identical argument to that for (4.1), the column and row containing the unit matrix on the right of (4.2) can be dropped.

The same method can be used to eliminate all the other $R_j^l = I, N_j^l = O, Q_j^l = I$, resulting in the transformed version of (3.14)

$$\begin{bmatrix} \hat{b}_1 \\ y_1 \\ d_1 \\ y_2 \\ d_2 \\ \cdot \\ \cdot \\ y_k \\ d_k \end{bmatrix} = \begin{bmatrix} \hat{U}_1 & & & & \\ C_1 & & & & \\ A_1 & E_1 & & & \\ & C_2 & & & \\ & A_2 & E_2 & & \\ & & \cdot & & \\ & & \cdot & \cdot & \\ & & & C_k & \\ & & & A_k & E_k \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \cdot \\ x_k \\ x_{k+1} \end{bmatrix} + \begin{bmatrix} \hat{S}_1 & & & & & & & & \\ & R_1^f & & & & & & & \\ & N_1 & Q_1^f & & & & & & \\ & & & R_2^f & & & & & \\ & & & N_2 & Q_2^f & & & & \\ & & & & & \cdot & & & \\ & & & & & \cdot & \cdot & & \\ & & & & & & & R_k^f & \\ & & & & & & & N_k & Q_k^f \end{bmatrix} \begin{bmatrix} \hat{u}_1 \\ v_1 \\ w_1 \\ v_2 \\ w_2 \\ \cdot \\ \cdot \\ v_k \\ w_k \end{bmatrix}$$

or

$$y = Fx + Lu, \qquad u \frown (0, I) \tag{4.3}$$

where for simplicity we have made the notational changes

$$\tilde{v}_i \to v_i, \quad \tilde{w}_i \to w_i, \quad -R_i^r \to R_i^f, \quad -B_i N_i^r \to N_i, \quad -B_i Q_i^r \to Q_i^f \quad . \tag{4.4}$$

Initially it appears that forming and using $B_i Q_i^r$ and $B_i N_i^r$ could contribute to numerical instability in the algorithm, but in the derivation here this is seen to be part of the transformation (4.2). Now

$$\begin{bmatrix} -B & I_n \\ I_m & O \end{bmatrix} \begin{bmatrix} O & I_m \\ I_n & B \end{bmatrix} = \begin{bmatrix} I_n & O \\ O & I_m \end{bmatrix}, \tag{4.5}$$

and so the *transformation matrix* is well conditioned if $\|B\|$ is not too large, no matter what the conditioning of B. But it was shown in [16] that numerically stable solution of (3.16) only requires of such left transformations that they be well conditioned, and so $\|B_i\| \sim 1$ is sufficient to ensure (4.2) does not contribute numerical instability.

Having made the transformation (4.2) we then apparently drop the row containing the numerically uncorrupted information on N_1^r and Q_1^r, and this may raise some doubts. However we have not really dropped it, it is just not needed in solving (3.16), but when this has been solved this row is now available for computing the original w_1 in (3.14), if it is needed. So it appears that when the norms of the B_i are reasonable, this elimination step is numerically quite safe.

To find best estimates for the state in (4.3) we solve (3.16) in a sequential manner. The transformation in the first measurement step has the form

$$\begin{bmatrix} \hat{T}_1 & O \\ O & I \end{bmatrix} \begin{bmatrix} \hat{U}_1 & \hat{S}_1 & O \\ C_1 & O & R_1^f \\ A_1 & O & N_1 \end{bmatrix} \begin{bmatrix} I & O \\ O & \hat{P}_1 \end{bmatrix} = \begin{bmatrix} O & * & * \\ \bar{U}_1 & * & * \\ A_1 & O & N_1 \end{bmatrix} \begin{bmatrix} I & O \\ O & \hat{P}_1 \end{bmatrix} = \begin{bmatrix} O & \tilde{S}_1 & O \\ \bar{U}_1 & \tilde{R}_1 & \bar{R}_1 \\ A_1 & \tilde{N}_1 & \bar{N}_1 \end{bmatrix} \tag{4.6}$$

where $\bar{U}_1$ has full row rank and $\tilde{S}_1$ has full column rank. $\hat{T}_1$ can be an orthogonal matrix, or the product of stabilized elimination matrices for greater efficiency. $\hat{P}_1$ must be orthogonal, and can be applied along with $\hat{T}_1$ to maintain triangular form throughout. To illustrate this suppose $\hat{U}_1$ and $\hat{S}_1$ are 3×3 lower triangular, C_1 is 2×3, and R_1^f is 2×2 lower triangular. We illustrate the first elimination (1) in $\hat{T}_1$, the first rotation (2) in $\hat{P}_1$, and the final form of the matrices.

$$\begin{bmatrix} \hat{U}_1 & \hat{S}_1 & O \\ C_1 & O & R_1^f \\ A_1 & O & N_1 \end{bmatrix} = \begin{bmatrix} \times & & & \times & & & & \\ \times & \times & & \times & \times & & & \\ \times & \times & \otimes & \times & \times & \times & 1 & \\ \times & \times & \times & 1 & 1 & 1 & \times & \\ \times & \times & \times & & & 2 & \times & \times \\ \times & \times & \times & & & 2 & \times & \times \\ \times & \times & \times & & & 2 & \times & \times \\ \times & \times & \times & & & 2 & \times & \times \end{bmatrix} \to \begin{bmatrix} & & & \times & & & & \\ & & & \times & \times & & & \\ \times & & & \times & \times & \times & & \\ \times & \times & & \times & \times & \times & \times & \\ \times & \times & \times & \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times & \times & \times & \times \end{bmatrix} \tag{4.7}$$

The elements 1 are introduced by the first elimination, and the elements 2 by the first rotation. Note how the $\bar{N}_1$ block has a different dimension from the N_1 block. For $n \times n$ $A_1, \hat{U}_1$ and $\hat{S}_1$, $p \times p$ R_1^f, and $p \times n$ C_1 and N_1^T the total cost of this measurement step using stabilized eliminations from the left and 4 multiplication rotations from the right is about

$$7n^2p + 5np^2/2 \qquad \text{multiplications,} \tag{4.8}$$

or if $N_1 = O$, about

$$3n^2p + 5np^2/2 \qquad \text{multiplications.} \tag{4.9}$$

We also apply $\hat{T}_1$ to y, and $\hat{P}_1^T$ in theory to the noise to give

$$\hat{T}_1 \begin{bmatrix} \hat{b}_1 \\ y_1 \end{bmatrix} = \begin{bmatrix} r_1 \\ \tilde{b}_1 \end{bmatrix}, \qquad \hat{P}_1^T \begin{bmatrix} \hat{u}_1 \\ v_1 \end{bmatrix} = \begin{bmatrix} \tilde{v}_1 \\ \bar{u}_1 \end{bmatrix} \frown (0, I), \tag{4.10}$$

so the transformed model is

$$\begin{bmatrix} r_1 \\ \tilde{b}_1 \\ d_1 \\ \cdot \end{bmatrix} = \begin{bmatrix} O & & \\ \bar{U}_1 & & \\ A_1 & E_1 & \\ & \cdot & \cdot \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \cdot \end{bmatrix} + \begin{bmatrix} \tilde{S}_1 & & & \\ \tilde{R}_1 & \bar{R}_1 & & \\ \tilde{N}_1 & \bar{N}_1 & Q_1^f & \\ & & & \cdot \end{bmatrix} \begin{bmatrix} \tilde{v}_1 \\ \bar{u}_1 \\ w_1 \\ \cdot \end{bmatrix} \quad . \tag{4.11}$$

We can now compute the unique $\tilde{v}_1$ in

$$\tilde{S}_1 \tilde{v}_1 = r_1, \qquad \tilde{S}_1 \quad \text{full column rank}, \tag{4.12}$$

and eliminate it. If there is no solution then the model is inconsistent, so we will assume we have a correct model. Then we define

$$\bar{b}_1 = \tilde{b}_1 - \tilde{R}_1 \tilde{v}_1, \qquad \bar{d}_1 = d_1 - \tilde{N}_1 \tilde{v}_1, \tag{4.13}$$

to give for our model

$$\begin{bmatrix} \bar{b}_1 \\ \bar{d}_1 \\ \cdot \end{bmatrix} = \begin{bmatrix} \bar{U}_1 & & \\ A_1 & E_1 & \\ & \cdot & \cdot \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \cdot \end{bmatrix} + \begin{bmatrix} \bar{R}_1 & & \\ \bar{N}_1 & Q_1^f & \\ & & \cdot \end{bmatrix} \begin{bmatrix} \bar{u}_1 \\ w_1 \\ \cdot \end{bmatrix}. \tag{4.14}$$

The estimates for the first step follow from the first two rows of this, see the comments on (3.17) and (3.18). Since $\bar{U}_1$ has full row rank, $x_{1|1}$ must satisfy

$$\bar{b}_1 = \bar{U}_1 x_{1|1} \tag{4.15}$$

while if E_1 has full row rank $x_{2|1}$ satisfies

$$\bar{d}_1 = A_1 x_{1|1} + E_1 x_{2|1} \qquad . \tag{4.16}$$

Note that we now have the updated version of (3.12)

$$\bar{b}_1 = \bar{U}_1 x_1 + \bar{R}_1 \bar{u}_1, \qquad \bar{u}_1 \frown (0, I), \tag{4.17}$$

which includes the first measurement. With (4.15) the covariance matrix representation for the error is then

$$\bar{U}_1 (x_{1|1} - x_1) = \bar{R}_1 \bar{u}_1, \qquad \bar{u}_1 \frown (0, I). \tag{4.18}$$

In the time step we eliminate A_1 and $\bar{N}_1$ in (4.14)

$$\bar{T}_1 \left[\begin{array}{cc|cc} \bar{U}_1 & O & \bar{R}_1 & O \\ A_1 & E_1 & \bar{N}_1 & Q_1^f \end{array}\right] \begin{bmatrix} I & O \\ O & \bar{P}_1 \end{bmatrix} = \left[\begin{array}{cc|cc} U_1 & U_{12} & * & * \\ & \hat{U}_2 & * & * \end{array}\right] \begin{bmatrix} I & O \\ O & \bar{P}_1 \end{bmatrix} = \left[\begin{array}{cc|cc} U_1 & U_{12} & S_1 & \hat{S}_{12} \\ & \hat{U}_2 & & \hat{S}_2 \end{array}\right]$$

$$\bar{T}_1 \begin{bmatrix} \bar{b}_1 \\ \bar{d}_1 \end{bmatrix} = \begin{bmatrix} b_{1|1} \\ \hat{b}_2 \end{bmatrix}, \qquad \bar{P}_1^T \begin{bmatrix} \bar{u}_1 \\ w_1 \end{bmatrix} = \begin{bmatrix} u_1 \\ \hat{u}_2 \end{bmatrix}, \qquad U_1 \quad \text{full row rank}. \tag{4.19}$$

Continuing our earlier illustration we can take

$$U_1, \quad \hat{U}_2, \quad S_1 \quad \text{and} \quad \hat{S}_2 \quad \text{lower triangular} \tag{4.20}$$

by again using an elimination from the left followed by a rotation from the right, and so on. The first two such pairs are illustrated below for 3×3 $\bar{U}_1, A_1, E_1, \bar{R}_1, \bar{N}_1$, and 3×2 Q_1^f

$$\begin{array}{cccccc} \times & & & & & \\ \times & \times & & & & \\ \times & \times & \times & 1 & 3 & \\ \times & \times & \otimes & \times & & \\ \times & \times & \otimes & \times & \times & \\ \times & \times & \times & \times & \times & \times \end{array} \qquad \begin{array}{ccccc} \times & & & & \\ \times & \times & & & \\ \times & \times & \times & 1 & 3 \\ \times & \times & \otimes & \times & \\ \times & \times & \otimes & \times & \times \\ \times & \times & \times & \times & \times \end{array} \tag{4.21}$$

Note since $m = 2 < n = 3$ the next "rotation" from the right must move the last and second last columns forward by one, placing the third last at the end. As a result S_1 will be 3×2. We see we eliminate A_1 and $\bar{N}_1$ a column at a time.

We have illustrated the situation for lower triangular E_1. Usually $E_1 = I$, but if this is not so an initial transformation could be used to ensure E_1 is lower triangular, and this would certainly be worthwhile in the constant coefficient case.

For $n \times n$ matrices $\bar{U}_1, A_1, E_1, \bar{R}_1, \bar{N}_1$, and an $n \times m$ matrix Q_1^f, with E_1 lower triangular, this time step takes about

$$\frac{9}{2}(n^3 + n^2 m - n m^2) + \frac{3}{2}m^3 \qquad \text{multiplications} \tag{4.22}$$

using stable eliminations from the left and 4 multiplication rotations from the right. When $n = m$ this is $6n^3$ multiplications. The $9n^3/2$ term in (4.22) makes it very expensive, but the model (4.3) does include very general initial information and correlation between v_k and w_k, as well as handling descriptor systems.

After this time step the transformed model is

$$\begin{bmatrix} b_{1|1} \\ \hat{b}_2 \\ y_2 \\ d_2 \\ \cdot \end{bmatrix} = \begin{bmatrix} U_1 & U_{12} & & \\ & \hat{U}_2 & & \\ & C_2 & & \\ & A_2 & E_2 & \\ \cdot & & & \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \cdot \end{bmatrix} + \begin{bmatrix} S_1 & \hat{S}_{12} & & & \\ & \hat{S}_2 & & & \\ & & R_2^f & & \\ & & N_2 & Q_2^f & \\ & & & & \cdot \end{bmatrix} \begin{bmatrix} u_1 \\ \hat{u}_2 \\ v_2 \\ w_2 \\ \cdot \end{bmatrix}. \tag{4.23}$$

Since U_1 has full row rank we can delete the first row and column and solve the remaining problem, which we see has exactly the same form as the original problem (4.3), and so we can continue the process. The "initial condition" for x_2 is now

$$\hat{b}_2 = \hat{U}_2 x_2 + \hat{S}_2 \hat{u}_2, \qquad \hat{u}_2 \frown (0, I), \tag{4.24}$$

which is an alternate representation to that in (4.14). In the usual case E_1 in (4.19) has full row rank, and we saw $\bar{U}_1$ has full row rank, therefore $\hat{U}_2$ has full row rank and then $x_{2|1}$ satisfies

$$\hat{U}_2 x_{2|1} = \hat{b}_2. \tag{4.25}$$

Since the present general formulation allows general square E_1 the resulting $\hat{U}_2$ may not have full row rank, in which case the generalized linear least squares problem, see (3.17), for $x_{2|1}$ in (4.24) could be solved, or the computation could be continued without computing $x_{2|1}$, as it is not actually needed in later steps.

It is possible in such a general formulation as this that the estimates are not defined uniquely in for example (4.15) and (4.25). In such cases particular solutions can be chosen, for example minimum 2-norm solutions. Alternatively these solutions are not needed to produce later estimates, and need not be computed if that is acceptable in the physical problem. The point is that if the system and measurement equations are supplying sufficient information then after a certain number of steps the estimates will be uniquely defined, and can then be computed.

By considering these transformations applied to the complete model (4.3) it is also possible to derive numerically reliable algorithms for computing smoothed estimates $x_{1|j}, \ldots, x_{j-1|j}$, but this paper is already too long to deal with this.

Acknowledgement: Paul Van Dooren was very helpful in reading and improving the original manuscript.

Bibliography

[1] T.W. Anderson, *An Introduction to Multivariate Statistical Analysis*, John Wiley, New York, 1958, pp. 25-26.

[2] B.D.O. Anderson and J.B. Moore, *Optimal Filtering*, Prentice-Hall, Englewood Cliffs, New Jersey, 1979.

[3] G.J. Bierman, *Factorization Methods for Discrete Sequential Estimation*, Academic Press, New York, 1977.

[4] N.A. Carlson, "Fast triangular formulation of the square root filter", *AIAA J.*, 11 (1973), 1259-1265.

[5] W.J. Duncan, "Some devices for the solution of large sets of simultaneous linear equations", *The Lond., Ed., and Dublin Philos. Mag. and J. of Sci., Seventh Ser.*, 35 (1944), 660-670.

[6] P. Dyer and S. McReynolds, "Extension of square root filtering to include process noise", *J. Optimization Theory Appl.*, 3 (1969), 444-459.

[7] D.C. Fraser, "A new technique for the optimal smoothing of data", M.I.T. Instrumentation Lab., Report T-474, 1967.

[8] G.H. Golub and C.F. Van Loan, *Matrix Computations*, The Johns Hopkins University Press, Baltimore, Maryland, 1983.

[9] R.E. Kalman, "A new approach to linear filtering and prediction problems", *J. Basic Eng., Trans. ASME*, Series D, 82 (1960), 35-45.

[10] P.G. Kaminski, A.E. Bryson and S.F. Schmidt, "Discrete square root filtering: a survey of current techniques", *IEEE Trans. Automatic Control*, 16 (1971), 727-735.

[11] S. Kourouklis and C.C. Paige, "A constrained least squares approach to the general Gauss-Markov linear model", *J. Am. Statist. Assoc.*, 76 (1981), 620-625.

[12] P.S. Maybeck, *Stochastic Models, Estimation, and Control, Vol. 1*, Academic Press, New York, 1979.

[13] D.V. Ouellette, "Schur complements and statistics", *Linear Algebra and Its Applications*, 36 (1981), 187-295.

[14] C.C. Paige, "Numerically Stable Computations for General Univariate Linear Models", *Commun. Statist.*, B7 (1978), 437-453.

[15] C.C. Paige, "Computer solution and perturbation analysis of generalized linear least squares problems", *Maths. Comput.*, 33 (1979), 171-183.

[16] C.C. Paige, "Fast numerically stable computations for generalized linear least squares problems", *SIAM J. Numer. Anal.*, 16 (1979), 165-171.

[17] C.C. Paige, "The general linear model and the generalized singular value decomposition", submitted to *Linear Algebra and Its Applications*, Special Issue on Linear Algebra in Statistics.

[18] C.C. Paige and M.A. Saunders, "Least squares estimation of discrete linear dynamic systems using orthogonal transformations", *SIAM J. Numer. Anal.*, 14 (1977), 180-193.

[19] C.C. Paige and M.A. Saunders, "Towards a generalized singular value decomposition", *SIAM J. Numer. Anal.*, 18 (1981), 398-405.

SCHOOL OF COMPUTER SCIENCE
McGILL UNIVERSITY
805 SHERBROOKE ST WEST
MONTREAL, QUEBEC H3A 2K6
CANADA