International Journal of Applied Mathematics

Volume 25 No. 1 2012, 33-39

SEQUENTIAL ASSIMILATION OF OBSERVATION DATA

Yuri N. Skiba

Centro de Ciencias de la Atmósfera
Universidad Nacional Autónnoma de México
Av. Universidad 3000, CU/UNAM
México, D.F., C.P. 04510, MÉXICO
e-mail: skiba@unam.mx

Abstract: An alternative recursive method of sequential assimilation of measurements in the Kalman filtering problem is proposed. It is proved, that if the dimensionality of the observation vector is equal to r, then through r steps, the method gives the same result as the standard algorithm of Kalman filtration, i.e. it gives the optimal estimate of the system state, see [1]-[3]. The merit of new algorithm is the direct method of realization, which does not require the application of approximate iterative procedures for the inversion of the Kalman filter matrix.

AMS Subject Classification: 93E11, 60G35
Key Words: Kalman filtering, sequential data assimilation

1. Kalman Filter

Consider a discrete linear mathematical model of n-dimensional observable stochastic dynamical system:

$$u(k+1) = A(k)u(k) + v(k) \ , \quad k \geq k_0, \tag{1}$$

$$z(k) = M(k)u(k) + w(k), \tag{2}$$

where $u(k) = \{u_1(k), u_2(k), ..., u_n(k)\}^T$ is the n-dimensional random column vector of the phase state of system at time t_k; $Eu(k_0) = 0$, E is the symbol of mathematical expectation; $z(k) = \{z_1(k), z_2(k), ..., z_r(k)\}^T$ is the r-dimensional random column vector of observations received at moment t_k; $A(k)$ is the known

Received: October 3, 2011

transition matrix of order n, which contains the system parameters and depends on the time t_k; $M(k)$ is the known transform matrix of the state vector in the measurement process, and $u^T(k)$ is transposed to $u(k)$.

The observed system is under the influence of n-dimensional vector of random perturbations $v(k) = \{v_1(k), v_2(k), ..., v_n(k)\}^T$ and r-dimensional vector of random errors of measurements $w(k) = \{w_1(k), w_2(k), ..., w_r(k)\}^T$ representing random Gaussian white-noise discrete processes with known symmetric nonnegative-definite covariance matrices $Q(k)$ and $W(k)$:

$$Ev(k) = 0 \ , \quad Ev(k)v^T(l) = Q(k)\delta_{kl}, \tag{3}$$

$$Ew(k) = 0 \ , \quad Ew(k)w^T(l) = W(k)\delta_{kl} \tag{4}$$

(δ_{kl} is the Kronecker symbol).

It is assumed that the initial state $u(k_0)$ is a random vector with zero mean value and covariance matrix $P(k_0)$, besides the initial state, the measurement errors and the perturbation vectors are mutually uncorrelated:

$$Eu(k_0)v^T(k) = 0 \ , \quad k \geq k_0, \tag{5}$$

$$Eu(k_0)w^T(k) = 0 \ , \quad k \geq k_0, \tag{6}$$

$$Ev(k)w^T(l) = 0. \tag{7}$$

It is known [1]-[3] that the linear unbiased estimate $u^*(k+1)$ with minimum mean square error is given by the Kalman filtering algorithm, consisting of the following steps:

$$\widehat{u}(k) = u^*(k) + H(k)[z(k) - M(k)u^*(k)] \ , \quad u^*(k_0) = 0, \tag{8}$$

$$u^*(k+1) = A(k)\widehat{u}(k), \tag{9}$$

besides the covariance matrix of estimation error

$$P^*(k+1) = E[u(k+1) - u^*(k+1)][u(k+1) - u^*(k+1)]^T$$

is calculated through the gain matrix $H(k)$ and covariance matrix of errors $\widehat{P}(k) = E[u(k) - \widehat{u}(k)][u(k) - \widehat{u}(k)]^T$:

$$H(k) = P^*(k)M^T(k)[M(k)P^*(k)M^T(k) + W(k)]^{-1}, \tag{10}$$

$$\widehat{P}(k) = P^*(k) - H(k)M(k)P^*(k) \ , \ P^*(k_0) = P(k_0), \tag{11}$$

$$P^*(k+1) = A(k)\widehat{P}(k)A^T(k) + Q(k). \tag{12}$$

It should be noted that for a linear model, the matrices of Kalman filtering (10)-(12) can be calculated irrespective of vectors (8) and (9).

In the case of large values of r, the largest computational load in the Kalman filtering (8)-(12) accounted for the calculation of the gain matrix $H(k)$ (see formula (10)). Note that both matrices H^T and MP^* have r rows and n columns. At each time step, the calculation of H requires to solve the matrix Wiener-Hopf equation [4]:

$$BH^T = MP^* \ , \quad B = MP^*M^T + W \tag{13}$$

that is, to find the inverse matrix B^{-1}. Consequently, with increasing the number r, the computer time required for calculating the matrix B^{-1} grows approximately as r^3.

2. Algorithm of Sequential Assimilation

For convenience, let us rewrite again the algorithm of recursive estimation with a minimum mean square error (8), (10) and (11), separately from the extrapolation formulas (9) and (12), omitting for simplicity the time index k and the symbol *:

$$\widehat{u} = u + H[z - Mu], \tag{14}$$

$$H = PM^T[MPM^T + W]^{-1}, \tag{15}$$

$$\widehat{P} = P - HMP. \tag{16}$$

We now give a recursive algorithm of sequential estimation (one after another) of r measurements z_i $(i = 1, ..., r)$ that solves the problem of optimal assimilation of observations (14)-(16) exactly through r steps. The merit of the new algorithm is a direct method of implementation, which does not require the use of approximate iterative procedure for the inversion of matrix B in the Wiener-Hopf equation (13).

The method is economical and convenient in practice. If the dimension of observation vector (the order of matrix B) varies in time, the application of the recursion method makes it possible to avoid the use of dynamic arrays in computer programs and thus to prevent undesirable memory fragmentation and excessive use of computer resources. The proposed algorithm can also be applied in problems of optimal interpolation.

In what follows, $C_{i\cdot}$ and $C_{\cdot i}$ denote the row vector and column vector formed by the ith row and ith column of matrix C, respectively. We now prove the following assertion.

Theorem. *Let $u^{(0)} = u$ and $P^{(0)} = P$ where u is the state vector, P is the covariance matrix from (14)-(16), and let r be the dimension of observation vector z of system (1),(2). Then, after r applications of the recurrence formulas*

$$h^{(i)} = P^{(i-1)}(M_{i\cdot})^T \tag{17}$$

$$g^{(i)} = h^{(i)}/[M_{i\cdot}h^{(i)} + W_{ii}] \tag{18}$$

$$u^{(i)} = u^{(i-1)} + g^{(i)}[z_i - M_{i\cdot}u^{(i-1)}] \tag{19}$$

$$P^{(i)} = P^{(i-1)} - g^{(i)}[h^{(i)}]^T \quad , \quad (i = 1, ..., r) \tag{20}$$

the vector $u^{(r)}$ and symmetric matrix $P^{(r)}$ will identically coincide with the state vector$\widehat{u}$ and covariance matrix $\widehat{P}$ calculated with formulas (14)-(16).

In (17)-(20), for each i , $h^{(i)}$, $g^{(i)}$ and $u^{(i)}$ denote the column vectors after ith iteration, z_i is the ith component of observation vector z, and W_{ii} is the ith diagonal element of matrix W $(i = 1, ..., r)$. Note that the order of square matrix $P^{(r)}$ is n.

Proof. The proof is by using the method of mathematical induction.

1) Let $r = 1$. Then obviously, $M_{1\cdot} \equiv M$, $h^{(1)} = PM^T$, $g^{(1)} = H$, and the assertion is true.

2) Assume that the theorem is true for $r = m$, i.e., $u^{(m)} = \widehat{u}$ and $P^{(m)} = \widehat{P}$. Let us further denote the matrix $P^{(m)}$, M, W and H obtained for $r = m$ by the symbols $\mathbf{P}$, $\mathbf{M}$, $\mathbf{W}$ and $\mathbf{H}$, respectively, and vector z - by the symbol $\mathbf{z}$. Thus, in this case we have

$$u^{(m)} \equiv \widehat{u} = u + \mathbf{H}(\mathbf{z} - \mathbf{M}u), \tag{21}$$

$$\mathbf{P} \equiv \widehat{P} = P - \mathbf{HM}P, \tag{22}$$

$$\mathbf{H} = P\mathbf{M}^T[\mathbf{M}P\mathbf{M}^T + \mathbf{W}]^{-1}. \tag{23}$$

3) We now prove the theorem for $r = m + 1$, i.e. we show that

$$u^{(r)} \equiv \widehat{u} = u + H[z - Mu], \tag{24}$$

$$P^{(r)} \equiv \widehat{P} = P - HMP. \tag{25}$$

First prove (24). By virtue of (19) and (21) we get

$$u^{(r)} = u^{(m)} + g^{(r)}[z_r - M_{r\cdot}u^{(m)}]$$

$$= u + \mathbf{H}(\mathbf{z} - \mathbf{M}u) + g^{(r)}\{z_r - M_{r\cdot}[u + \mathbf{H}(\mathbf{z} - \mathbf{M}u)]\}$$
$$= u + [\mathbf{H} - g^{(r)} M_{r\cdot}\mathbf{H}](\mathbf{z} - \mathbf{M}u) + g^{(r)}(z_r - M_{r\cdot}u).$$

Thus,

$$u^{(r)} = u + G(z - Mu) \tag{26}$$

where

$$z = \begin{pmatrix} \mathbf{z} \\ z_r \end{pmatrix}, \qquad M = \begin{pmatrix} \mathbf{M} \\ M_{r\cdot} \end{pmatrix}, \qquad G = \left(\mathbf{H} - g^{(r)} M_{r\cdot}\mathbf{H} \;\middle|\; g^{(r)}\right) \tag{27}$$

are the r-dimensional column vector, $(r \times n)$-matrix, and $(n \times r)$-matrix, respectively. Thus, formulas (24) and (26) are identical if and only if

$$G = H. \tag{28}$$

We now show the validity of (28). By virtue of (15),

$$H = PM^T B^{-1}, \quad B = MPM^T + W. \tag{29}$$

It is easy to see that the symmetric matrix B can be represented as

$$B = \begin{pmatrix} \mathbf{B} & s \\ s^T & b_{rr} \end{pmatrix},$$

where s is the m-dimensional column vector, $\mathbf{B}$ is the $(m \times m)$-matrix, and b_{rr} is a number, besides,

$$s = \mathbf{M}P(M_{r\cdot})^T, \tag{30}$$

$$\mathbf{B} = \mathbf{MPM}^T + \mathbf{W}, \tag{31}$$

$$b_{rr} = M_{r\cdot}P(M_{r\cdot})^T + W_{rr}. \tag{32}$$

Applying the bordering method [5] for calculating the symmetric inverse matrix B, we obtain

$$B^{-1} = \begin{pmatrix} \mathbf{B}^{-1} + \alpha^{-1}\mathbf{B}^{-1}ss^T\mathbf{B}^{-1} & -\alpha^{-1}\mathbf{B}^{-1}s \\ \alpha^{-1}s^T\mathbf{B}^{-1} & \alpha^{-1} \end{pmatrix}, \tag{33}$$

where

$$\alpha = b_{rr} - s^T\mathbf{B}^{-1}s. \tag{34}$$

It follows from formulas (29), (33) that

$$H = \left(P\mathbf{M}^T \;\middle|\; P(M_{r\cdot})^T\right) B^{-1} = \left(\mathbf{H} - \varphi s^T\mathbf{B}^{-1} \;\middle|\; \varphi\right), \tag{35}$$

where

$$\varphi = \alpha^{-1}[P(M_{r\cdot})^T - \mathbf{H}s] \tag{36}$$

is the n-dimensional column vector. Comparing (27) and (35), we see that (28) is satisfied if and only if

$$M_{r\cdot}\ \mathbf{H} = s^T\mathbf{B}^{-1} \quad \text{and} \quad g^{(r)} = \varphi. \tag{37}$$

But due to (30) and (23),

$$s^T\mathbf{B}^{-1} = M_{r\cdot}P\mathbf{M}^T\mathbf{B}^{-1} = M_{r\cdot}\mathbf{H}.$$

On the other hand, using (17), (18), (22), (30), (32) and (34), we have

$$g^{(r)} = \mathbf{P}(M_{r\cdot})^T[M_{r\cdot}\mathbf{P}(M_{r\cdot})^T + W_{rr}]^{-1}$$

$$= (P - \mathbf{HM}P)(M_{r\cdot})^T[M_{r\cdot}(P - \mathbf{HM}P)(M_{r\cdot})^T + W_{rr}]^{-1}$$

$$= [P(M_{r\cdot})^T - \mathbf{H}s][b_{rr} - s^T\mathbf{B}^{-1}s]^{-1}$$

$$= \alpha^{-1}[P(M_{r\cdot})^T - \mathbf{H}s] = \varphi.$$

Thus, the validity of (37), and hence, (28) is proved. In other words, (24) is fulfilled for $r = m + 1$.

We now prove (25) for $r = m + 1$. Indeed, by virtue of (20), (37), (23) and assumption (22) we obtain

$$P^{(r)} = \mathbf{P} - g^{(r)}[h^{(r)}]^T = P - \mathbf{HM}P - \varphi M_{r\cdot}(P - P\mathbf{M}^T\mathbf{H}^T)$$

$$= P - \mathbf{HM}P - \varphi M_{r\cdot}P + \varphi s^T\mathbf{B}^{-1}\mathbf{M}P = P - (\mathbf{H} - \varphi s^T\mathbf{B}^{-1})\mathbf{M}P - \varphi M_{r\cdot}P$$

$$= P - (\mathbf{H} - \varphi s^T\mathbf{B}^{-1} \mid \varphi)\begin{pmatrix} \mathbf{M}P \\ M_{r\cdot}P \end{pmatrix} = P - HMP = \widehat{P}.$$

Hence, the algorithm (17)-(20) is valid for any number of measurements r. The theorem is proved. □

Remark 1. The algorithm (17)-(20) can also be applied to the optimal interpolation procedure [6], one of the most commonly used methods of adaptation of observational and forecasting data [7]. Within the scope of designations used here, the algorithm of optimal interpolation can be written as follows:

$$\widehat{u} = Hz \tag{38}$$

$$H = PM^T[MPM^T + W]^{-1}, \tag{39}$$

$$\widehat{P} = P - HMP. \tag{40}$$

The theorem proved above remains valid in this case with the apparent change $u^{(0)} = 0$ (compare the formulas (38)-(40) and (14)-(16)).

Remark 2. In the particular case when M is the unit matrix, the algorithm (17)-(20) is simplified, reducing to

$$h^{(i)} = P_{\cdot i}^{(i-1)}, \tag{41}$$

$$g^{(i)} = h^{(i)}/[P_{ii}^{(i-1)} + W_{ii}], \tag{42}$$

$$u^{(i)} = u^{(i-1)} + g^{(i)}[z_i - u_i^{(i-1)}], \tag{43}$$

$$P^{(i)} = P^{(i-1)} - g^{(i)}[h^{(i)}]^T \quad (i = 1, ..., r). \tag{44}$$

References

[1] S.G. Mohinder, P.A. Angus, *Kalman Filtering: Theory and Practice Using MATLAB*, Wiley-IEEE Press, Hoboken, New Jersey (2008).

[2] P. Zarchan, H. Musoff, *Fundamentals of Kalman Filtering: A Practical Approach*, American Institute of Aeronautics and Astronautics, Inc., Reston, VA, USA (2005).

[3] K. Brammer, G. Siffling, *Kalman-Bucy-Filter, Deterministische Beobachtung und Stochastische Filterung*, R. Oldenbourg Verlag, München (1975).

[4] K. Leondes, Ed., *Filtering and Stochastic Control in Dynamic Systems*, McGraw-Hill, New York (1978).

[5] D.K. Faddeev, V.N. Faddeeva, *Computational Methods of Linear Algebra*, H.W. Freeman, San Francisco (1963).

[6] A.N. Kolmogorov, Interpolation and extrapolation of random stationary sequences, *Izv. Akad. Nauk SSSR, Ser. Mat.*, **5** (1941), 3-13.

[7] L.S. Gandin, *Objective Analysis of Meteorological Fields*, Gidrometeoizdat, Leningrad (1963), In Russian; English translation by: Israel Program for Scientific Translations, Jerusalem (1965).