AFCRL-TR-73-0432

Reports of the Department
of Geodetic Science, No. 20

STEPWISE AND SEQUENTIAL COLLOCATION

by

Helmut Moritz

The Ohio State University
Research Foundation
Columbus, Ohio 43212

Contract No. F19628-72-C-0120
Project No. 8607
Task No. 860701
Work Unit No. 86070101

Scientific Report No. 10

July, 1973

Contract Monitor: Bela Szabo
Terrestrial Sciences Laboratory

Prepared for

Air Force Cambridge Research Laboratories
Air Force Systems Command
United States Air Force
Bedford, Massachusetts 01730

ABSTRACT

In stepwise collocation, the estimation procedure is split up into two steps, corresponding to a partitioning of the covariance matrix to be inverted. Formulas for signal and parameter estimates and for their error covariances are derived and interpreted statistically. Stepwise adjustment by parameters and by conditions is exhibited as special cases of stepwise collocation. Sequential collocation is an extension of the procedure so as to comprise several steps. Finally, the relation between sequential collocation and Kalman filtering is considered in some detail.

FOREWORD

This report was prepared by Helmut Moritz, Professor, Technische Hochschule Graz, and Adjunct Professor, Department of Geodetic Science of The Ohio State University, under Air Force Contract No. F19628-72-C-0120, OSURF Project No. 3368A1, Project Supervisor, Urho A. Uotila, Professor, Department of Geodetic Science. The contract covering this research is administered by the Air Force Cambridge Research Laboratories, Air Force Systems Command, Laurence G. Hanscom Field, Bedford, Massachusetts, with Mr. Bela Szabo, Project Scientist.

CONTENTS

1. Stepwise Collocation

It is frequently convenient to split up the estimation by collocation into two steps, just as in ordinary least-squares adjustment. This may be done to reduce the size of matrices to be inverted; another application is the use of additional observations to improve the original estimates.

We start with the basic equations of collocation:

$$X = (A^T\bar{C}^{-1}A)^{-1}A^T\bar{C}^{-1}x \ , \tag{1-1}$$

$$s_P = C_P^T\bar{C}^{-1}(x - AX) \ ; \tag{1-2}$$

cf. (Moritz, 1972, eqs. (2-35) and (2-38)). In these expressions, x is the vector comprising all observations, s_P is a signal to be estimated, X is the vector of parameters to be estimated, $\bar{C}$ is the covariance matrix of the observation vector x, C_P is the covariance matrix between the vector x and the signal s_P, and T denotes the transpose; thus C_P is a column vector and C_P^T is the corresponding row vector. The matrix A is the "sensitivity matrix" characterizing the effect of the parameters X on the observed values x according to

$$x = AX + s + n \ . \tag{1-3}$$

(Moritz, 1972, eq. (1-2)).

Now we split up the observations x into two parts, the first part making up the vector x_1, and the second part forming the vector x_2. Thus the observation vector x is partitioned as follows.

$$x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \tag{1-4}$$

(note that x_1 and x_2 are themselves vectors!). The matrices $\bar{C}$ and C_P^T are partitioned accordingly:

$$\bar{C} = \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix} , \tag{1-5}$$

$$C_P^T = \begin{bmatrix} C_{P1} & C_{P2} \end{bmatrix} ; \tag{1-6}$$

e.g., C_{12} denotes the covariance matrix between the vector x_1 and the vector x_2 , and C_{P1} denotes the covariance matrix between s_P and x_1 . In the same way we partition the sensitivity matrix:

$$A = \begin{bmatrix} A_1 \\ A_2 \end{bmatrix} ; \tag{1-7}$$

thus A_1 characterizes the effect of X on x_1 , and similarly for A_2 .

Using this partitioning we wish to split up the estimation by (1-1) and (1-2) into two steps. Let us first consider the estimation of the parameters X by (1-1). For this purpose we need the partitioned inverse matrix $\bar{C}^{-1}$. Writing

$$\bar{C}^{-1} = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} \tag{1-8}$$

we have the well-known relations (cf. Faddeeva , 1959, § 14):

$$B_{22} = (C_{22} - C_{21}C_{11}^{-1}C_{12})^{-1},$$

$$B_{12} = -C_{11}^{-1}C_{12}B_{22} , \qquad B_{21} = -B_{22}C_{21}C_{11}^{-1} , \qquad (1\text{-}9)$$

$$B_{11} = C_{11}^{-1} - C_{11}^{-1}C_{12}B_{21} = C_{11}^{-1} + C_{11}^{-1}C_{12}B_{22}C_{21}C_{11}^{-1} .$$

Using these relations we find

$$A^T\bar{C}^{-1}A = \begin{bmatrix} A_1^T & A_2^T \end{bmatrix} \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} \begin{bmatrix} A_1 \\ A_2 \end{bmatrix} =$$

$$= A_1^TB_{11}A_1 + A_1^TB_{12}A_2 + A_2^TB_{21}A_1 + A_2^TB_{22}A_2$$

$$= A_1^TC_{11}^{-1}A_1 + (A_2^T-A_1^TC_{11}^{-1}C_{12})B_{22}(A_2-C_{21}C_{11}^{-1}A_1) .$$

With the abbreviations

$$\bar{A}_2 = A_2 - C_{21}C_{11}^{-1}A_1 , \qquad (1\text{-}10)$$

$$P_1 = A_1^TC_{11}^{-1}A_1 \qquad (1\text{-}11)$$

this becomes

$$A^T\bar{C}^{-1}A = P_1 + \bar{A}_2^TB_{22}\bar{A}_2 . \qquad (1\text{-}12)$$

The inverse matrix is then given by

$$(A^T\bar{C}^{-1}A)^{-1} = P_1^{-1} - P_1^{-1}\bar{A}_2^T(B_{22}^{-1}+\bar{A}_2P_1^{-1}\bar{A}_2^T)^{-1}\bar{A}_2P_1^{-1} . \qquad (1\text{-}13)$$

This is readily proved by multiplying the right-hand sides of (1-12) and (1-13): after some straightforward algebra the unit matrix results as it should be.

With the new abbreviation

$$\bar{C}_{22} = C_{22} - C_{21}C_{11}^{-1}C_{12} + \bar{A}_2 P_1^{-1}\bar{A}_2^T \tag{1-14}$$

(1-13) becomes

$$(A^T\bar{C}^{-1}A)^{-1} = P_1^{-1} - P_1^{-1}\bar{A}_2^T\bar{C}_{22}^{-1}\bar{A}_2 P_1^{-1} \ . \tag{1-15}$$

We further have

$$A^T\bar{C}^{-1}x = \begin{bmatrix} A_1^T & A_2^T \end{bmatrix} \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$= A_1^T B_{11} x_1 + A_1^T B_{12} x_2 + A_2^T B_{21} x_1 + A_2^T B_{22} x_2 \ ,$$

and substituting (1-9) and using (1-10) we obtain

$$A^T\bar{C}^{-1}x = A_1^T C_{11}^{-1} x_1 + \bar{A}_2^T B_{22}(x_2 - C_{21}C_{11}^{-1}x_1) \ . \tag{1-16}$$

The substitution of (1-15) and (1-16) into (1-1) gives

$$X = (A^T\bar{C}^{-1}A)^{-1}A^T\bar{C}^{-1}x$$

$$= P_1^{-1}A_1^T C_{11}^{-1}x_1 - P_1^{-1}\bar{A}_2^T\bar{C}_{22}^{-1}\bar{A}_2 P_1^{-1}A_1^T C_{11}^{-1}x_1 +$$

$$+ (P_1^{-1} - P_1^{-1}\bar{A}_2^T\bar{C}_{22}^{-1}\bar{A}_2 P_1^{-1})\bar{A}_2^T B_{22}(x_2 - C_{21}C_{11}^{-1}x_1) \ . \tag{1-17}$$

The first term on the right-hand side is

$$X_1 = P_1^{-1}A_1^TC_{11}^{-1}x_1 = (A_1^TC_{11}^{-1}A_1)^{-1}A_1^TC_{11}^{-1}x_1 \ , \tag{1-18}$$

which is nothing else but the least-squares collocation estimate of the vector X on the basis of the partial observation vector x_1 only, as the comparison with (1-1) shows.

The last term on the right-hand side may be transformed as follows:

$$(P_1^{-1}-P_1^{-1}\bar{A}_2^T\bar{C}_{22}^{-1}\bar{A}_2P_1^{-1})\bar{A}_2^TB_{22} =$$

$$= P_1^{-1}\bar{A}_2^T(I-\bar{C}_{22}^{-1}\bar{A}_2P_1^{-1}\bar{A}_2^T)B_{22} =$$

$$= P_1^{-1}\bar{A}_2^T\bar{C}_{22}^{-1}(\bar{C}_{22}-\bar{A}_2P_1^{-1}\bar{A}_2^T)B_{22} =$$

$$= P_1^{-1}\bar{A}_2^T\bar{C}_{22}^{-1}B_{22}^{-1}B_{22} =$$

$$= P_1^{-1}\bar{A}_2^T\bar{C}_{22}^{-1} \ ; \tag{1-19}$$

here we have used (1-9).

In view of (1-18) and (1-19), eq. (1-17) reduces to

$$X = X_1 + P_1^{-1}\bar{A}_2^T\bar{C}_{22}^{-1}(x_2-C_{21}C_{11}^{-1}x_1-\bar{A}_2X_1) \ . \tag{1-20}$$

This is the required equation for the parameter vector X. The first term on the right-hand side represents the estimate (1-18) on the basis of the first part, x_1, of the observations x; the second term expresses the improvement of the estimate by using, in addition, the second part, x_2, of the observations.

Now we shall effect a similar transformation for the signal estimate (1-2). Putting

$$z = x - AX \ , \qquad z = \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = \begin{bmatrix} x_1 - A_1X \\ x_2 - A_2X \end{bmatrix} \tag{1-21}$$

and using (1-5), (1-6), and (1-8) we find

$$s_P = \begin{bmatrix} C_{P1} & C_{P2} \end{bmatrix} \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix}$$

$$= (C_{P1}B_{11}+C_{P2}B_{21})z_1 + (C_{P1}B_{12}+C_{P2}B_{22})z_2 \ ,$$

whence, by (1-9),

$$s_P = C_{P1}C_{11}^{-1}z_1 + (C_{P2}-C_{P1}C_{11}^{-1}C_{12})B_{22}(z_2-C_{21}C_{11}^{-1}z_1). \tag{1-22}$$

By (1-21) we have

$$z_1 = x_1 - A_1X = x_1 - A_1X_1 - A_1(X-X_1) \ , \tag{1-23}$$

$$z_2 - C_{21}C_{11}^{-1}z_1 = x_2 - A_2X - C_{21}C_{11}^{-1}x_1 + C_{21}C_{11}^{-1}A_1X$$

$$= x_2 - C_{21}C_{11}^{-1}x_1 - \bar{A}_2X$$

$$= x_2 - C_{21}C_{11}^{-1}x_1 - \bar{A}_2X_1 - \bar{A}_2(X-X_1), \tag{1-24}$$

using (1-10).

We note that

$$s_{P,1} = C_{P1}C_{11}^{-1}(x_1 - A_1X_1) \tag{1-25}$$

is the estimate of the signal s_P on the basis of the partial observation vector x_1 only, as the comparison with (1-2) shows. Using (1-23), (1-24) and (1-25), eq. (1-22) then becomes

$$s_P = s_{P,1} - C_{P1}C_{11}^{-1}A_1(X-X_1) +$$

$$+ (C_{P2}-C_{P1}C_{11}^{-1}C_{12})B_{22}(x_2-C_{21}C_{11}^{-1}x_1-\bar{A}_2X_1)$$

$$- (C_{P2}-C_{P1}C_{11}^{-1}C_{12})B_{22}\bar{A}_2(X-X_1) .$$

The substitution of $X-X_1$ by (1-20) gives, I denoting the unit matrix,

$$s_P = s_{P,1} + \Big[(C_{P2}-C_{P1}C_{11}^{-1}C_{12})B_{22}(I-\bar{A}_2P_1^{-1}\bar{A}_2^T\bar{C}_{22}^{-1})$$

$$- C_{P1}C_{11}^{-1}A_1P_1^{-1}\bar{A}_2^T\bar{C}_{22}^{-1}\Big](x_2-C_{21}C_{11}^{-1}x_1-\bar{A}_2X_1)$$

$$= s_{P,1} + R\bar{C}_{22}^{-1}(x_2-C_{21}C_{11}^{-1}x_1-\bar{A}_2X_1) , \tag{1-26}$$

where we have put

$$R = (C_{P2}-C_{P1}C_{11}^{-1}C_{12})B_{22}(\bar{C}_{22}-\bar{A}_2P_1^{-1}\bar{A}_2^T) -$$

$$- C_{P1}C_{11}^{-1}A_1P_1^{-1}\bar{A}_2^T .$$

By (1-9) and (1-14) we find

$$B_{22}(\bar{C}_{22} - \bar{A}_2 P_1^{-1} \bar{A}_2^T) = I ,$$

so that

$$R = C_{P2} - C_{P1} C_{11}^{-1} C_{12} - C_{P1} C_{11}^{-1} A_1 P_1^{-1} \bar{A}_2^T .$$

Hence (1-26) becomes

$$s_P = s_{P,1} +$$

$$+ (C_{P2} - C_{P1} C_{11}^{-1} C_{12} - C_{P1} C_{11}^{-1} A_1 P_1^{-1} \bar{A}_2^T) \bar{C}_{22}^{-1} (x_2 - C_{21} C_{11}^{-1} x_1 - \bar{A}_2 X_1) . \quad (1\text{-}27)$$

This is the required equation for stepwise estimation of the signal, together with (1-25). It is analogous to the corresponding equation for X and has the same interpretation.

2. Covariances

We shall now transform the expressions for the covariances in a similar way, by partitioning the original expressions given in (Moritz, 1972), eqs. (3-33), (3-37), and (3-39). Using a slightly changed notation, these expressions read:

$$E_{XX} = (A^T \bar{C}^{-1} A)^{-1} = P^{-1} , \quad (2\text{-}1)$$

$$E_{XP} = -P^{-1} A^T \bar{C}^{-1} C_P , \quad (2\text{-}2)$$

$$\sigma_{PQ} = C_{PQ} - C_P^T \bar{C}^{-1} C_Q + C_P^T \bar{C}^{-1} A P^{-1} A^T \bar{C}^{-1} C_Q . \quad (2\text{-}3)$$

Here E_{XX} denotes the error covariance matrix of the parameter vector X, E_{XP} is the error covariance matrix between X and the signal s_P, and σ_{PQ} denotes the error covariance between the signals s_P and s_Q. The notations on the right-hand sides are the same as in eqs. (1-1) and (1-2).

The transformation of (2-1) by partitioning may be written down immediately using (1-15): we have

$$E_{XX} = P^{-1} = E_{XX,1} - P_1^{-1}\bar{A}_2^T\bar{C}_{22}^{-1}\bar{A}_2P_1^{-1} , \tag{2-4}$$

where

$$E_{XX,1} = P_1^{-1}. \tag{2-5}$$

The corresponding transformation of (2-2) is also easy if we note that, apart from the minus sign, (2-2) differs from (1-1) only by

$$x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \qquad \text{being replaced by} \qquad C_P = \begin{bmatrix} C_{1P} \\ C_{2P} \end{bmatrix} \quad ;$$

cf. (1-4) and (1-6), with $C_{1P} = C_{P1}^T$, $C_{2P} = C_{P2}^T$. Write (1-20) in the form

$$X = X_1 + P_1^{-1}\bar{A}_2^T\bar{C}_{22}^{-1}(x_2-C_{21}C_{11}^{-1}x_1-\bar{A}_2P_1^{-1}A_1^TC_{11}^{-1}x_1)$$

and perform the replacement indicated, also changing the sign. This gives at once

$$E_{XP} = E_{XP,1} - P_1^{-1}\bar{A}_2^T\bar{C}_{22}^{-1}(C_{2P}-C_{21}C_{11}^{-1}C_{1P}-\bar{A}_2P_1^{-1}A_1^TC_{11}^{-1}C_{1P}) , \tag{2-6}$$

where, by (2-2)

$$E_{XP,1} = -P_1^{-1}A_1^TC_{11}^{-1}C_{1P} . \qquad (2\text{-}7)$$

The partitioning of the remaining equation (2-3) is more laborious. Consider first the second term on the right-hand side of (2-3). Using (1-9) we find

$$-C_P^T\bar{C}^{-1}C_Q = -\begin{bmatrix} C_{P1} & C_{P2} \end{bmatrix} \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} \begin{bmatrix} C_{1Q} \\ C_{2Q} \end{bmatrix}$$

$$= -C_{P1}C_{11}^{-1}C_{1Q} - (C_{P2}-C_{P1}C_{11}^{-1}C_{12})B_{22}(C_{2Q}-C_{21}C_{11}^{-1}C_{1Q}) . \qquad (2\text{-}8)$$

By (1-16), with C_P^T replacing A^T and A replacing x, we have

$$C_P^T\bar{C}^{-1}A = C_{P1}C_{11}^{-1}A_1 + (C_{P2}-C_{P1}C_{11}^{-1}C_{12})B_{22}\bar{A}_2 , \qquad (2\text{-}9)$$

and from (2-2), (2-6), and (2-7) we get

$$P^{-1}A^T\bar{C}^{-1}C_Q = P_1^{-1}A_1^TC_{11}^{-1}C_{1Q} + P_1^{-1}\bar{A}_2^T\bar{C}_{22}^{-1}\bar{C}_{2Q} \qquad (2\text{-}10)$$

with

$$\bar{C}_{2Q} = C_{2Q} - C_{21}C_{11}^{-1}C_{1Q} - \bar{A}_2P_1^{-1}A_1^TC_{11}^{-1}C_{1Q} . \qquad (2\text{-}11)$$

On multiplying (2-9) and (2-10) and taking (2-8) into account, eq. (2-3) becomes

$$\sigma_{PQ} = \sigma_{PQ,1} - (C_{P2}-C_{P1}C_{11}^{-1}C_{12})B_{22}(C_{2Q}-C_{21}C_{11}^{-1}C_{1Q}) +$$

$$+ (C_{P2}-C_{P1}C_{11}^{-1}C_{12})B_{22}(\bar{A}_2P_1^{-1}A_1^TC_{11}^{-1}C_{1Q}+\bar{A}_2P_1^{-1}\bar{A}_2^T\bar{C}_{22}^{-1}\bar{C}_{2Q}) +$$

$$+ C_{P1}C_{11}^{-1}A_1P_1^{-1}\bar{A}_2^T\bar{C}_{22}^{-1}\bar{C}_{2Q} \qquad (2\text{-}12)$$

where

$$\sigma_{PQ,1} = C_{PQ} - C_{P1}C_{11}^{-1}C_{1Q} + C_{P1}C_{11}^{-1}A_1P_1^{-1}A_1^T C_{11}^{-1}C_{1Q} \ . \qquad (2\text{-}13)$$

Using (2-11) we may transform (2-12) as follows:

$$\sigma_{PQ} = \sigma_{PQ,1} - (C_{P2} - C_{P1}C_{11}^{-1}C_{12})B_{22}(\bar{C}_{2Q} - \bar{A}_2P_1^{-1}\bar{A}_2^T\bar{C}_{22}^{-1}\bar{C}_{2Q}) +$$

$$+ C_{P1}C_{11}^{-1}A_1P_1^{-1}\bar{A}_2^T\bar{C}_{22}^{-1}\bar{C}_{2Q}$$

$$= \sigma_{PQ,1} - (C_{P2} - C_{P1}C_{11}^{-1}C_{12})B_{22}(\bar{C}_{22} - \bar{A}_2P_1^{-1}\bar{A}_2^T)\bar{C}_{22}^{-1}\bar{C}_{2Q} +$$

$$+ C_{P1}C_{11}^{-1}A_1P_1^{-1}\bar{A}_2^T\bar{C}_{22}^{-1}\bar{C}_{2Q} \ .$$

By (1-9) and (1-14),

$$B_{22}(\bar{C}_{22} - \bar{A}_2P_1^{-1}\bar{A}_2^T) = B_{22}B_{22}^{-1} = I \ ,$$

so that we finally get the simple result.

$$\sigma_{PQ} = \sigma_{PQ,1} - \bar{C}_{P2}\bar{C}_{22}^{-1}\bar{C}_{2Q} \ . \qquad (2\text{-}14)$$

Equations (2-4), (2-6), and (2-14) are the desired expressions for the error covariances if the collocation is performed in two steps.

3. Results and Statistical Interpretation

Let us summarize the results so far obtained. The following notations are introduced:

$$\bar{A}_2 = A_2 - C_{21}C_{11}^{-1}A_1 , \tag{3-1}$$

$$P_1 = A_1^T C_{11}^{-1} A_1 , \tag{3-2}$$

$$\bar{x}_2 = x_2 - C_{21}C_{11}^{-1}x_1 - \bar{A}_2 X_1 , \tag{3-3}$$

$$\bar{C}_{22} = C_{22} - C_{21}C_{11}^{-1}C_{12} + \bar{A}_2 P_1^{-1} \bar{A}_2^T , \tag{3-4}$$

$$\bar{C}_{P2} = C_{P2} - C_{P1}C_{11}^{-1}C_{12} - C_{P1}C_{11}^{-1}A_1 P_1^{-1} \bar{A}_2^T ; \quad \bar{C}_{2P} = \bar{C}_{P2}^T . \tag{3-5}$$

Then the estimates after the first step are:

$$X_1 = P_1^{-1} A_1^T C_{11}^{-1} x_1 , \tag{3-6}$$

$$s_{P,1} = C_{P1}C_{11}^{-1}(x_1 - A_1 X_1) , \tag{3-7}$$

and the corresponding error covariances

$$E_{XX,1} = P_1^{-1} , \tag{3-8}$$

$$E_{XP,1} = -P_1^{-1} A_1^T C_{11}^{-1} C_{1P} , \tag{3-9}$$

$$\sigma_{PQ,1} = C_{PQ} - C_{P1}C_{11}^{-1}C_{1Q} + C_{P1}C_{11}^{-1}A_1 P_1^{-1} A_1^T C_{11}^{-1} C_{1Q} . \tag{3-10}$$

The final estimates (1-20) and (1-27) may be written:

$$X = X_1 + P_1^{-1}\bar{A}_2^T\bar{C}_{22}^{-1}\bar{x}_2 \ , \qquad (3\text{-}11)$$

$$s_P = s_{P,1} + \bar{C}_{P2}\bar{C}_{22}^{-1}\bar{x}_2 \ , \qquad (3\text{-}12)$$

and their error covariances are

$$E_{XX} = E_{XX,1} - P_1^{-1}\bar{A}_2^T\bar{C}_{22}^{-1}\bar{A}_2 P_1^{-1} \ , \qquad (3\text{-}13)$$

$$E_{XP} = E_{XP,1} - P_1^{-1}\bar{A}_2^T\bar{C}_{22}^{-1}\bar{C}_{2P} \ , \qquad (3\text{-}14)$$

$$\sigma_{PQ} = \sigma_{PQ,1} - \bar{C}_{P2}\bar{C}_{22}^{-1}\bar{C}_{2Q} \ . \qquad (3\text{-}15)$$

These equations permit a very simple statistical interpretation. Consider least-squares prediction without systematic parameters. Denote the observation vector by y and the predicted signal vector by u. By (1-2) we have in the new notation, with $A = 0$,

$$u = C_{uy} \, C_{yy}^{-1} \, y \ . \qquad (3\text{-}16)$$

We assert that both (3-11) and (3-12) are of this form: for (3-11) we have

$$u = X - X_1 \ , \qquad (3\text{-}17)$$

$$C_{uy} = P_1^{-1}\bar{A}_2^T \ ; \qquad (3\text{-}18)$$

for (3-12),

$$u = s_P - s_{P,1} \, , \tag{3-19}$$

$$C_{uy} = \bar{C}_{P2} \, ; \tag{3-20}$$

and in both cases,

$$y = \bar{x}_2 \, , \tag{3-21}$$

$$C_{yy} = \bar{C}_{22} \, . \tag{3-22}$$

In order to prove this, we must first verify that

$$M\{\bar{x}_2\} = 0 \tag{3-23}$$

since (3-16) presupposes $M\{y\} = 0$. On substituting (3-1) and (3-6), eq. (3-3) becomes

$$\bar{x}_2 = x_2 - C_{21}C_{11}^{-1}x_1 - (A_2 - C_{21}C_{11}^{-1}A_1)P_1^{-1}A_1^T C_{11}^{-1}x_1 \, . \tag{3-24}$$

By eq. (3-5) of (Moritz, 1972) we have

$$M\{x_1\} = A_1\bar{X} \, , \qquad M\{x_2\} = A_2\bar{X} \, , \tag{3-25}$$

$\bar{X}$ denoting the true value of the parameter vector. Thus

$$\begin{aligned} M\{\bar{x}_2\} &= \left[A_2 - C_{21}C_{11}^{-1}A_1 - (A_2 - C_{21}C_{11}^{-1}A_1)P_1^{-1}A_1^T C_{11}^{-1}A_1 \right] \bar{X} \\ &= \left[A_2 - C_{21}C_{11}^{-1}A_1 - (A_2 - C_{21}C_{11}^{-1}A_1) \right] \bar{X} = 0 \, , \end{aligned}$$

which proves (3-23).

Introduce now the "centered observations" z_1 and z_2 by

$$z_1 = x_1 - A_1\bar{X} ,$$
$$z_2 = x_2 - A_2\bar{X} , \qquad (3\text{-}26)$$

so that

$$M\{z_1\} = 0 , \qquad M\{z_2\} = 0 \qquad (3\text{-}27)$$

by (3-25). Then (3-24) becomes by means of (3-1),

$$\bar{x}_2 = z_2 + A_2\bar{X} - C_{21}C_{11}^{-1}(z_1+A_1\bar{X}) -$$
$$- (A_2-C_{21}C_{11}^{-1}A_1)P_1^{-1}A_1^TC_{11}^{-1}(z_1+A_1\bar{X}) =$$
$$= z_2 - C_{21}C_{11}^{-1}z_1 - \bar{A}_2P_1^{-1}A_1^TC_{11}^{-1}z_1 +$$
$$+ \bar{A}_2\bar{X} - \bar{A}_2P_1^{-1}A_1^TC_{11}^{-1}A_1\bar{X} .$$

The last two terms cancel in view of (3-2), and there remains

$$\bar{x}_2 = z_2 - (C_{21} + \bar{A}_2P_1^{-1}A_1^T)C_{11}^{-1}z_1 . \qquad (3\text{-}28)$$

Thus,

$$\bar{x}_2\bar{x}_2^T = \left[z_2 - (C_{21}+\bar{A}_2P_1^{-1}A_1^T)C_{11}^{-1}z_1\right]\left[z_2^T - z_1^TC_{11}^{-1}(C_{12}+A_1P_1^{-1}\bar{A}_2^T)\right]$$

$$= z_2 z_2^T - (C_{21} + \bar{A}_2 P_1^{-1} A_1^T) C_{11}^{-1} z_1 z_2^T -$$

$$- z_2 z_1^T C_{11}^{-1} (C_{12} + A_1 P_1^{-1} \bar{A}_2^T) +$$

$$+ (C_{21} + \bar{A}_2 P_1^{-1} A_1^T) C_{11}^{-1} z_1 z_1^T C_{11}^{-1} (C_{12} + A_1 P_1^{-1} \bar{A}_2^T) .$$

Forming the average M and taking into account that

$$M\{z_1 z_1^T\} = C_{11} , \quad \text{etc.} \tag{3-29}$$

we find

$$C_{yy} = M\{\bar{x}_2 \bar{x}_2^T\}$$

$$= C_{22} - (C_{21} + \bar{A}_2 P_1^{-1} A_1^T) C_{11}^{-1} C_{12} -$$

$$- C_{21} C_{11}^{-1} (C_{12} + A_1 P_1^{-1} \bar{A}_2^T) +$$

$$+ (C_{21} + \bar{A}_2 P_1^{-1} A_1^T) C_{11}^{-1} (C_{12} + A_1 P_1^{-1} \bar{A}_2^T) ,$$

which, after obvious cancellations, becomes

$$= C_{22} - C_{21} C_{11}^{-1} C_{12} + \bar{A}_2 P_1^{-1} A_1^T C_{11}^{-1} A_1 P_1^{-1} \bar{A}_2^T$$

$$= C_{22} - C_{21} C_{11}^{-1} C_{12} + \bar{A}_2 P_1^{-1} \bar{A}_2^T = \bar{C}_{22} ,$$

by (3-4). This proves (3-22).

We further have

$$(s_P - s_{P,1})\bar{x}_2^T = s_P\bar{x}_2^T - C_{P1}C_{11}^{-1}(x_1 - A_1X_1)\bar{x}_2^T .$$

Now,

$$x_1 - A_1X_1 = (I - A_1P_1^{-1}A_1^TC_{11}^{-1})x_1$$

$$= (I - A_1P_1^{-1}A_1^TC_{11}^{-1})z_1 +$$

$$+ (A_1 - A_1P_1^{-1}A_1^TC_{11}^{-1}A_1)\bar{X}$$

$$= (I - A_1P_1^{-1}A_1^TC_{11}^{-1})z_1 , \qquad (3\text{-}30)$$

by (3-26) and (3-2). Thus, by (3-28),

$$(s_P - s_{P,1})\bar{x}_2^T = s_Pz_2^T - s_Pz_1^TC_{11}^{-1}(C_{12} + A_1P_1^{-1}\bar{A}_2^T) -$$

$$- C_{P1}C_{11}^{-1}(I - A_1P_1^{-1}A_1^TC_{11}^{-1})z_1z_2^T +$$

$$+ C_{P1}C_{11}^{-1}(I - A_1P_1^{-1}A_1^TC_{11}^{-1})z_1z_1^TC_{11}^{-1}(C_{12} + A_1P_1^{-1}\bar{A}_2^T) .$$

Hence (3-20) becomes

$$C_{uy} = M\{(s_P - s_{P,1})\bar{x}_2^T\}$$

$$= C_{P2} - C_{P1}C_{11}^{-1}(C_{12} + A_1P_1^{-1}\bar{A}_2^T) -$$

$$- C_{P1}C_{11}^{-1}(I - A_1P_1^{-1}A_1^TC_{11}^{-1})C_{12} +$$

$$+ C_{P1}C_{11}^{-1}(I - A_1P_1^{-1}A_1^TC_{11}^{-1})(C_{12} + A_1P_1^{-1}\bar{A}_2^T) ,$$

which is readily seen to reduce to (3-5). This proves (3-20).

Finally we have to consider (3-17) and (3-18). The signal u is now $X-X_1$ or $\bar{X}-X_1$ if the true value $\bar{X}$ is to be emphasized. We have

$$\bar{X} - X_1 = \bar{X} - P_1^{-1}A_1^T C_{11}^{-1}(z_1 + A_1\bar{X})$$

$$= -P_1^{-1}A_1^T C_{11}^{-1} z_1 \ , \qquad (3\text{-}31)$$

as the other two terms cancel. Thus, by (3-28),

$$(\bar{X}-X_1)\bar{x}_2^T = -P_1^{-1}A_1^T C_{11}^{-1} z_1 z_2 + $$

$$+ P_1^{-1}A_1^T C_{11}^{-1} z_1 z_1^T C_{11}^{-1}(C_{12} + A_1 P_1^{-1}\bar{A}_2^T) \ .$$

Then (3-18) becomes

$$C_{uy} = M\{(\bar{X}-X_1)\bar{x}_2^T\}$$

$$= -P_1^{-1}A_1^T C_{11}^{-1} C_{12} + P_1^{-1}A_1^T C_{11}^{-1}(C_{12} + A_1 P_1^{-1}\bar{A}_2^T)$$

$$= P_1^{-1}A_1^T C_{11}^{-1} A_1 P_1^{-1}\bar{A}_2^T = P_1^{-1}\bar{A}_2^T \ ,$$

which was to be shown.

In this way we have seen that the estimation formulas for the second step, (3-11) and (3-12), can, in fact, be considered as special cases of the simple least-squares prediction formula (3-16). This is particularly striking for $X-X_1$, eq. (3-11), since X itself is certainly not estimated by such a formula (see, however, Appendix A).

The error covariance formula corresponding to (3-16) is given by

$$E_{uv} = C_{uv} - C_{uy}C_{yy}^{-1}C_{yv} \; ; \qquad (3\text{-}32)$$

cf. (2-3) with $A = 0$. Here E_{uv} represents the error covariance matrix for two signal vectors u and v that are both estimated from the same data y by (3-16), and C_{uv} the corresponding signal covariance matrix. Thus, E_{uv} is the "a posteriori" covariance matrix and C_{uv} is the "a priori" covariance matrix.

One sees immediately that all the formulas (3-13), (3-14) and (3-15) are of the form (3-32).
For (3-13) we have

$$u = v = X - X_1 \; , \qquad C_{uy} = P^{-1}\bar{A}_2^T = C_{yv}^T \; ;$$

for (3-14),

$$u = X - X_1 \; , \qquad v = s_P - s_{P,1} \; , \qquad C_{yv} = \bar{C}_{2P} \; ;$$

and for (3-15),

$$u = s_P - s_{P,1} \; , \; v = s_Q - s_{Q,1} \; , \qquad C_{uy} = \bar{C}_{P2} \; .$$

The "a priori" covariances C_{uv} are the error covariances after the first step: $E_{XX,1}$, $E_{XP,1}$, and $\sigma_{PQ,1}$, respectively.

One might wonder why in expressions such as (3-12), the vector $\bar{x}_2$ plays the role of observation and not, e.g., x_2 itself. The reason is as follows. Eq. (3-7) depends on

$$\bar{x}_1 = x_1 - A_1 X_1$$

and the vector consisting of $\bar{x}_1$ together with $\bar{x}_2$ is equivalent to the original observation vector x_1 together with x_2 . However, whereas x_1 and x_2 will, in general, be correlated, <u>the new vectors $\bar{x}_1$ and $\bar{x}_2$ will be uncorrelated</u>. In fact,

$$\bar{x}_1\bar{x}_2^T = (x_1 - A_1 X_1)\bar{x}_2^T$$

$$= (I - A_1 P_1^{-1} A_1^T C_{11}^{-1}) z_1 z_2^T -$$

$$- (I - A_1 P_1^{-1} A_1^T C_{11}^{-1}) z_1 z_1^T C_{11}^{-1} (C_{12} + A_1 P_1^{-1} \bar{A}_2^T) ,$$

by (3-28) and (3-30), so that

$$M\{\bar{x}_1\bar{x}_2^T\} = (I - A_1 P_1^{-1} A_1^T C_{11}^{-1}) C_{12} -$$

$$- (I - A_1 P_1^{-1} A_1^T C_{11}^{-1})(C_{12} + A_1 P_1^{-1} \bar{A}_2^T)$$

$$= -(I - A_1 P_1^{-1} A_1^T C_{11}^{-1}) A_1 P_1^{-1} \bar{A}_2^T$$

$$= -A_1 P_1^{-1} \bar{A}_2^T + A_1 P_1^{-1} A_1^T C_{11}^{-1} A_1 P_1^{-1} \bar{A}_2^T$$

$$= 0 \qquad (3\text{-}33)$$

by (3-2).

Now, if $\bar{x}_1$ and $\bar{x}_2$ are uncorrelated, then their respective contributions to the estimate for s_P will simply add; to see this, put $C_{12} = 0$ in formulas such as (1-22).

Therefore, in (3-12), the contribution of $\bar{x}_2$,

$$\bar{C}_{P2}\bar{C}_{22}^{-1}\bar{x}_2 ,$$

is simply added to $s_{P,1}$, which represents the contribution of $\bar{x}_1$.

The dependence of (3-11) on $\bar{x}_2$ may be explained along similar lines, although the details are somewhat more involved.

In this way we can understand the structure of equations (3-11) through (3-15), which is conceptually very clear: formally they are completely covered already by the elementary theory of least-squares prediction as given, e.g., in (Heiskanen and Moritz, 1967, pp. 268-270).

<u>Applications of Stepwise Collocation</u>. - Two applications immediately present themselves:

1. If the size of the matrix $\bar{C}$ to be inverted is too large, then a stepwise procedure might be applicable since C_{11} and $\bar{C}_{22}$ are smaller matrices than $\bar{C}$.

2. Let signals and parameters have been estimated using observations x_1 . If new observations x_2 are available, then the original estimates can be improved by stepwise collocation.

Both procedures, breaking down the problem into smaller steps and adding a new group of observations, may be iterated. This may be called sequential collocation and will be described in sec. 5.

4. Stepwise Adjustment as a Special Case

Adjustment by Parameters. - The mathematical model for least-squares collocation is

$$x = AX + s + n ,$$

and for least-squares adjustment by parameters,

$$x = AX + n . \tag{4-1}$$

Thus, adjustment by parameters may be viewed as a special case of collocation if both the signal and all signal covariances are zero, so that now

$$\bar{C} = D , \tag{4-2}$$

where D is the error covariance matrix. Cf. (Moritz, 1972, secs. 1 and 2).

In this case we have the partitioning

$$D = \begin{bmatrix} D_{11} & 0 \\ 0 & D_{22} \end{bmatrix} , \tag{4-3}$$

if we further assume that the observations x_1 and x_2 are uncorrelated. Then

$$C_{11} = D_{11} , \quad C_{22} = D_{22} , \quad C_{12} = 0 , \quad C_{21} = 0 ,$$

so that

$$\bar{A}_2 = A_2 , \qquad P_1 = A_1^T D_{11}^{-1} A_1 , \tag{4-4}$$

$$\bar{x}_2 = x_2 - A_2 X_1 \ , \tag{4-5}$$

$$\bar{C}_{22} = D_{22} + A_2 P_1^{-1} A_2^T \ . \tag{4-6}$$

Thus (3-6) and (3-11) reduce to

$$X_1 = P_1^{-1} A_1^T D_{11}^{-1} x_1 \ , \tag{4-7}$$

$$X = X_1 + P_1^{-1} A_2^T (D_{22} + A_2 P_1^{-1} A_2^T)^{-1} (x_2 - A_2 X_1) \ , \tag{4-8}$$

and the error covariance matrix of X is given by

$$E_{xx,1} = P_1^{-1} \ , \tag{4-9}$$

$$E_{xx} = E_{xx,1} - P_1^{-1} A_2^T (D_{22} + A_2 P_1^{-1} A_2^T)^{-1} A_2 P_1^{-1} \ . \tag{4-10}$$

It will be recalled that X_1 is the estimate for the parameter vector on the basis of only the first part of the observations, forming the vector x_1, and that X is the estimate on the basis of the full vector x; thus the second term on the right-hand side of (4-8) represents the improvement due to the use of the second part of the observations, x_2. Similarly, $E_{xx,1}$ is the error covariance matrix of the estimate X_1, and E_{xx} is the error covariance of the estimate X, so that the second term in (4-10) represents the gain in accuracy due to the use of x_2 in addition to x_1.

Adjustment by Conditions. - The condition equations may be written in the matrix form (Hirvonen, 1971, p. 154)

$$B^T v + w = 0 \ , \qquad (4\text{-}11)$$

where v is the vector of residuals and w is the vector of misclosures. By equations (13.29), (13.30) and (13.32), loc. cit., the solution is given by

$$v = -P^{-1}B(B^T P^{-1} B)^{-1} w \ . \qquad (4\text{-}12)$$

Since both v and w are random variables with zero expectation, let us try to apply the simple formula for least-squares prediction (that is, collocation without systematic parameters):

$$v = C_{vw} C_{ww}^{-1} w \ . \qquad (4\text{-}13)$$

In the adjustment problem, the covariance matrix of v is the given error covariance matrix D :

$$M\{vv^T\} = D = P^{-1} \ . \qquad (4\text{-}14)$$

The covariance matrices C_{vw} and C_{ww} are to be found by error propagation: we have

$$vw^T = -vv^T B \ ,$$

$$ww^T = B^T vv^T B \ ,$$

whence,on forming the average M ,

$$C_{vw} = -DB \ , \qquad (4\text{-}15)$$

$$C_{ww} = B^T D B \ . \qquad (4\text{-}16)$$

On substituting the last two equations into (4-13) we obtain (4-12), which shows that adjustment by conditions may also be regarded as a special case of collocation.

More precisely, (4-13) is recognized as a special case of (1-2) on performing the following identifications:

$$\begin{aligned} s_P &= v , \qquad x = w , \\ C_P^T &= C_{vw} = -DB , \\ \bar{C} &= C_{ww} = B^T DB , \\ A &= 0 \end{aligned} \tag{4-17}$$

(the fact that s_P in (1-2) is a single quantity only, whereas v is a vector, is irrelevant in this context: this equation would be formally the same if s_P were a vector).

Let us now split up (4-11) into two systems of condition equations:

$$\begin{aligned} B_1^T v + w_1 &= 0 , \\ B_2^T v + w_2 &= 0 , \end{aligned} \tag{4-18}$$

so that

$$B = \begin{bmatrix} B_1 & B_2 \end{bmatrix} , \qquad w = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} . \tag{4-19}$$

Then, (3-7) and (3-12) together with (4-17) give:

$$v_1 = C_{P1}C_{11}^{-1}w_1 , \tag{4-20}$$

$$v = v_1 + \bar{C}_{P2}\bar{C}_{22}^{-1}(w_2 - C_{21}C_{11}^{-1}w_1) , \tag{4-21}$$

where

$$\begin{aligned} C_{P1} &= -DB_1 , & C_{P2} &= -DB_2 , \\ C_{11} &= B_1^T DB_1 , & C_{12} &= B_1^T DB_2 , \\ C_{21} &= B_2^T DB_1 , & C_{22} &= B_2^T DB_2 \end{aligned} \tag{4-22}$$

and

$$\begin{aligned} \bar{C}_{P2} &= C_{P2} - C_{P1}C_{11}^{-1}C_{12} , \\ \bar{C}_{22} &= C_{22} - C_{21}C_{11}^{-1}C_{12} . \end{aligned} \tag{4-23}$$

Equations (4-20) and (4-21), with (4-22) and (4-23), are easily recognized as the basic formulas for the Boltz partitioning procedure ("Boltzsches Entwicklungsverfahren") well-known from triangulation adjustment. In fact, apart from notation,(4-21) is identical to eq. (3442,7) of (Wolf, 1968, p. 363); the matrix

$$Z = C_{11}^{-1}C_{12} \tag{4-24}$$

constitutes Boltz' "Zwischenkorrelaten".

The error covariance matrix after adjustment is, of course, given by specializing (3-10) and (3-15) to the present case.

5. Sequential Collocation; Relation to Kalman Filtering

Two-step collocation may be iterated; this will be called sequential collocation.

In stepwise collocation (see, particularly, sec. 3) we have investigated how the estimation of parameters X and signals s_P may be done in two steps, the first step using a part x_1 of the observation vector x only and the second taking into account the remaining observations, forming the vector x_2. In other words, we have seen how the estimates X and s_P are improved if, in addition to a first group of observations, x_1, a second group, x_2, is used.

In sequential collocation, we may add a third group of observations, a fourth group, and so forth. At each step, the formulas of sec. 3 are applied.

This may be formulated mathematically as follows.

Let the vector x_k comprise all observations up to and including the k-th group, and let y_k comprise the observations of the k-th group itself. Thus,

$$x_k = \begin{bmatrix} y_1 \\ y_2 \\ \cdot \\ \cdot \\ \cdot \\ y_k \end{bmatrix} . \tag{5-1}$$

In the same way we partition the sensitivity matrix A_k, which characterizes the effect of the parameters X on x_k :

$$A_k = \begin{bmatrix} B_1 \\ B_2 \\ \cdot \\ \cdot \\ B_k \end{bmatrix} ;$$

thus, B_k represents the effect of X on y_k , so that

$$B_k X + s_k + n_k = y_k \; . \tag{5-2}$$

If we consider an additional group of observations, y_{k+1} , we obviously have

$$x_{k+1} = \begin{bmatrix} x_k \\ y_{k+1} \end{bmatrix} , \qquad A_{k+1} = \begin{bmatrix} A_k \\ B_{k+1} \end{bmatrix} , \tag{5-3}$$

and this partitioning corresponds to (1-4) and (1-7).

Corresponding to (1-5) and (1-6) we have

$$C_{k+1} = \begin{bmatrix} C_k & F_{k+1} \\ F^T_{k+1} & G_{k+1} \end{bmatrix} , \qquad C_{P,k+1} = \begin{bmatrix} C_{P,k} & G_{P,k+1} \end{bmatrix} . \tag{5-4}$$

Thus C_k and G_{k+1} represent the covariance matrices of x_k and y_{k+1} , respectively; and F_{k+1} represents the cross-covariance matrix between x_k and y_{k+1} . Similarly, $C_{P,k}$ denotes the covariance between s_P and x_k , and $G_{P,k+1}$ denotes the covariance between s_P and y_{k+1} .

Transcribing eqs. (3-1) through (3-5) to the present case and using the abbreviation

$$Z_{k+1} = F^T_{k+1} C_k^{-1} , \tag{5-5}$$

we have

$$\bar{B}_{k+1} = B_{k+1} - Z_{k+1}A_k \ , \tag{5-6}$$

$$P_k = A_k^T C_k^{-1} A_k \ , \qquad Q_k = P_k^{-1} \ , \tag{5-7}$$

$$\bar{y}_{k+1} = y_{k+1} - Z_{k+1}x_k - \bar{B}_{k+1}X_k \ , \tag{5-8}$$

$$\bar{G}_{k+1} = G_{k+1} - Z_{k+1}F_{k+1} + \bar{B}_{k+1}Q_k\bar{B}_{k+1} \ , \tag{5-9}$$

$$\bar{G}_{P,k+1} = G_{P,k+1} - C_{P,k}Z_{k+1}^T - C_{Pk}C_k^{-1}A_kQ_k\bar{B}_{k+1} \ . \tag{5-10}$$

Eqs. (3-11) and (3-12) become

$$X_{k+1} = X_k + K_{k+1}\bar{y}_{k+1} \ , \tag{5-11}$$

$$s_{P,k+1} = s_{P,k} + \bar{G}_{P,k+1}\bar{G}_{k+1}^{-1}\bar{y}_{k+1} \ , \tag{5-12}$$

where we have put

$$K_{k+1} = Q_k\bar{B}_{k+1}^T\bar{G}_{k+1}^{-1} \ . \tag{5-13}$$

X_k and X_{k+1} denote the estimates for X on the basis of the observations x_k and x_{k+1}, respectively; similarly for $s_{P,k+1}$ and $s_{P,k}$.

The error covariances after collocation are, according to (3-13), (3-14), and (3-15), given by

$$E_{XX,k+1} = Q_{k+1} = Q_k - K_{k+1}\bar{B}_{k+1}Q_k \ , \tag{5-14}$$

$$E_{XP,k+1} = E_{XP,k} - K_{k+1}\bar{G}^T_{P,k+1} , \tag{5-15}$$

$$\sigma_{PQ,k+1} = \sigma_{PQ,k} - \bar{G}_{P,k+1}\bar{G}^{-1}_{k+1}\bar{G}^T_{Q,k+1} . \tag{5-16}$$

Using these formulas we may, starting with $k = 1$, successively compute the estimates X_k and $s_{P,k}$ and their error covariances for $k = 2, 3, \ldots$ The initial estimates for $k = 1$ are, of course, to be obtained from (3-6) and (3-7) and their accuracy, from (3-8), (3-9), and (3-10).

In particular, after an initial estimate of X and s_P from a sufficient number of observations, each consecutive group y_k may consist of one observation only. In this method it is remarkable that, after computing the initial estimate, no matrix inversion is needed any more. In fact, $\bar{G}_k$ now reduces to a single element, so that its inverse is simply the reciprocal of this element, and Q_k is computed successively from (5-14) without having to use the direct expression (5-7).

This advantage is, however, compensated by a great increase of matrix multiplications, so that the total computational effort is, in general,not reduced. In fact, stepwise collocation may be compared to matrix inversion by partitioning, and sequential collocation, using groups of one observation each, then corresponds to matrix inversion by bordering (Faddeeva, 1959, § 14 and 15).

Discrete Kalman Filtering and Prediction. - Let a dynamical system be described by a parameter vector $X(t)$ which is a function of the time t. Consider equally spaced discrete instants $t_1, t_2, t_3, \ldots$, and let

$$X^k = X(t_k) \tag{5-17}$$

be the parameter vector at time t_k. It is assumed that the parameter vectors at consecutive instants are linearly related:

$$X^{k+1} = \Phi_{k+1} X^k + u_{k+1} , \tag{5-18}$$

where Φ_{k+1} is a regular square matrix and u_k is a random vector of zero mean, representing "internal noise" of the dynamical system; u_k and u_ℓ are uncorrelated if $t_k \neq t_\ell$.

The observations y_k at time t_k are assumed to be linear functions of X^k:

$$y_k = B_k X^k + n_k , \tag{5-19}$$

where n_k is a random vector of zero mean representing the measuring error or "external noise". Obviously, (5-19) is identical to the observation equations in adjustment by parameters. The vector y_k is assumed to be of the same kind (in particular, to have the same dimension) for each instant t_k. The noise vectors n_k of the individual groups are assumed to be uncorrelated.

Denote by X^k_ℓ the estimate of X^k (the parameter vector at time t_k) on the basis of all observations $y_1, y_2, \ldots$ up to and including y_ℓ. Thus, after determining X^k on the basis of the observations up to (and including) t_k, that is, X^k_k, we may predict X^{k+1} at time t_k by

$$X^{k+1}_k = \Phi_{k+1} X^k_k . \tag{5-20}$$

This follows from (5-18) since u_{k+1} is uncorrelated to all quantities referring to the preceding instants t_1, t_2, . . ., t_k and unknown, so that its estimate will be zero. The notation X_k^{k+1} signifies that X^{k+1} is estimated on the basis of the observation up to t_k .

Denoting by Q_k^{k+1} and Q_k^k the error covariance matrices of X_k^{k+1} and X_k^k , respectively, we get from (5-18) by error propagation

$$Q_k^{k+1} = \Phi_{k+1} Q_k^k \Phi_{k+1}^T + R_{k+1} , \tag{5-21}$$

where

$$R_{k+1} = M\{u_{k+1} u_{k+1}^T\} \tag{5-22}$$

is the covariance matrix of internal noise.

The problem now is to find X_{k+1}^{k+1} , that is to improve the estimate (5-20) of X^{k+1} by including the observations of the (k+1)th group:

$$y_{k+1} = B_{k+1} X^{k+1} + n_{k+1} . \tag{5-23}$$

We may use X^{k+1} as parameter vector also for the preceding groups of number 1, 2, . . ., k as we can successively express X^k, X^{k-1}, . . . by means of X^{k+1} using

$$X^k = \Phi_{k+1}^{-1} X^{k+1} - \Phi_{k+1}^{-1} u_{k+1} \tag{5-24}$$

(by (5-18)) and the analogous relations for X^{k-1}, . . .

In this way we find a relation of the form

$$A_k X^{k+1} + z_k = x_k \ , \tag{5-25}$$

where x_k comprises $y_1, y_2, \ldots, y_k$ in agreement with (5-1) and the random vector z_k is a linear combination of observational noise $n_1, n_2, \ldots, n_k$ and internal noise $u_1, u_2, \ldots, u_{k+1}$.

This relation is to be considered together with (5-23), so that (5-3) holds. In (5-4), G_{k+1} represents the covariance matrix of n_{k+1} :

$$G_{k+1} = M\{n_{k+1} n_{k+1}^T\} \ ;$$

and

$$F_{k+1} = 0 \tag{5-26}$$

because n_{k+1} is not correlated to z_k. Thus also

$$Z_{k+1} = 0 \ ,$$

and (5-6), (5-8), and (5-9) become

$$\bar{B}_{k+1} = B_{k+1} \ , \tag{5-27}$$

$$\bar{y}_{k+1} = y_{k+1} - B_{k+1} X_k^{k+1} , \tag{5-28}$$

$$\bar{G}_{k+1} = G_{k+1} + B_{k+1} Q_k^{k+1} B_{k+1} \ , \tag{5-29}$$

where Q^{k+1} takes the place of Q since the parameter vector is now

$$X = X^{k+1} . \tag{5-30}$$

Eq. (5-11) gives

$$X_{k+1}^{k+1} = X_k^{k+1} + K_{k+1}(y_{k+1} - B_{k+1}X_k^{k+1}) \tag{5-31}$$

with

$$K_{k+1} = Q_k^{k+1}B_{k+1}^T\bar{G}_{k+1}^{-1} , \tag{5-32}$$

and (5-14) becomes

$$Q_{k+1}^{k+1} = Q_k^{k+1} - K_{k+1}B_{k+1}Q_k^{k+1} . \tag{5-33}$$

Let us compile the principal formulas. If the estimates X_k^k and Q_k^k are known, then first estimates for the quantities X^{k+1} and Q^{k+1} for the next stage are

$$\begin{aligned} X_k^{k+1} &= \Phi_{k+1}X_k^k , \\ Q_k^{k+1} &= \Phi_{k+1}Q_k^k\Phi_{k+1}^T + R_{k+1} ; \end{aligned} \tag{5-34}$$

these estimates are updated by

$$\begin{aligned} X_{k+1}^{k+1} &= X_k^{k+1} + K_{k+1}(y_{k+1} - B_{k+1}X_k^{k+1}) , \\ Q_{k+1}^{k+1} &= Q_k^{k+1} - K_{k+1}B_{k+1}Q_k^{k+1} . \end{aligned} \tag{5-35}$$

These relations are evaluated recursively. Let an initial estimate X_o^o be given, as well as its accuracy estimate Q_o^o. By (5-34) we can compute X_o^1 and Q_o^1:

$$X_o^1 = \Phi_1 X_o^o ,$$
$$Q_o^1 = \Phi_1 Q_o^o \Phi_1^T + R_1 , \qquad (5\text{-}36)$$

and then we get X_1^1 and Q_1^1 by (5-35). Then (5-34) gives X_1^2 and Q_1^2, and (5-35) provides X_2^2 and Q_2^2. In this way we can successively compute all estimates for X^k and Q^k.

In deriving Kalman filtering from stepwise collocation we have followed a similar approach as Deutsch (1965) and Koch and Lauer (1971), who have derived it from stepwise adjustment. The usual treatment is in terms of Markov processes (Bucy and Joseph, 1968; Bjerhammar, 1971). In fact, the relation (see (5-18))

$$X^{k+1} - \Phi_{k+1} X^k = u_{k+1} , \qquad (5\text{-}37)$$

with uncorrelated random variables u_{k+1}, is characteristic for wide-sense Markov processes; cf. (Papoulis, 1965, sec. 11-6).

<u>Relation to Collocation and Least-Squares Prediction</u>. - Let us express all observations y_k in terms of one and the same parameter vector X, for which we may take the vector X^o. Successive application of (5-18) for $k = 0, 1, 2, \ldots$ gives

$$X^1 = \Phi_1 X + u_1 \ ,$$
$$X^2 = \Phi_2 X_1 + u_2 = \Phi_2 \Phi_1 X + \Phi_2 u_1 + u_2 \ ,$$
$$X^3 = \Phi_3 X_2 + u_3 = \Phi_3 \Phi_2 \Phi_1 X + \Phi_3 \Phi_2 u_1 + \Phi_3 u_2 + u_3 \ ,$$
$$\vdots$$
$$X^k = M_k X + \sum_{\ell=1}^{k} N_{k\ell} u_\ell \ , \tag{5-38}$$

where

$$M_k = \Phi_k \Phi_{k-1} \cdot \cdot \cdot \Phi_1 \ ,$$
$$N_{k\ell} = \Phi_k \Phi_{k-1} \cdot \cdot \cdot \Phi_{\ell+1} \ . \tag{5-39}$$

This is inserted into (5-19) with the result

$$B_k M_k X + B_k \sum_{\ell=1}^{k} N_{k\ell} u_\ell + n_k = y_k$$

or

$$B_k' X + s_k + n_k = y_k \ , \tag{5-40}$$

where

$$B_k' = B_k M_k \tag{5-41}$$

and

$$s_k = B_k \sum_{\ell=1}^{k} N_{k\ell} u_\ell \ . \tag{5-42}$$

This is precisely of the form (5-2) where, by (5-42), the place of the signal s_k is taken by a linear combination of internal noise u_ℓ .

Thus, Kalman filtering is equivalent to least-squares collocation if the signal s_k is composed of uncorrelated random vectors u_ℓ in the form (5-42). This imposes an essential restriction on the s_k and on their covariances. For instance, it is well known that in the stationary case and for one-dimensional u_ℓ , the covariance of s_k must be of the form

$$\operatorname{cov}(s_k, s_{k\pm r}) = C_1 e^{-C_2 r} , \tag{5-43}$$

with suitable constants C_1 and C_2 (stationary Markov covariance).

However, if the signal s_k may be decomposed into the form (5-42), then the Kalman approach, while leading to the same result as collocation, presents considerable advantages from a computational point of view, which may be outlined as follows.

The approach through sequential collocation would successively estimate X and s_k (at all computation points), first from y_1 , then from y_1 and y_2 , then from y_1, y_2, y_3 , etc., the results being improved at every step. Then the "total signal" (observation y_k with noise n_k removed) would be estimated by

$$t_k = B_k' X + s_k , \tag{5-44}$$

in view of (5-40).

In the Kalman approach, we estimate X^k instead, using (5-34) and (5-35). Then we simply have

$$t_k = B_k X_k^k \, . \tag{5-45}$$

Since now u_k takes the place of s_k, we have $F_{k+1} = 0$ because u_k is not correlated to the preceding u_ℓ, whereas the use of s_k implies a nonzero F_{k+1}.

Thus, for the special case under consideration, the Kalman method is computationally very appropriate. However, its scope is much more restricted than the scope of collocation. Consider, for instance, the problem of one-dimensional filtering and prediction (Fig. 1). Provided the signal s_k

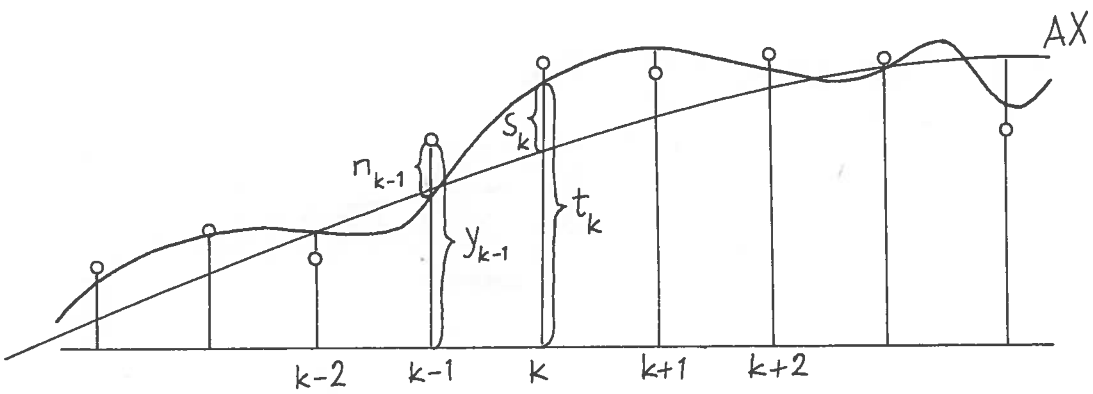

Figure 1

is of Markov form, then the Kalman method gives the total signal t_k (5-45) at the point k, using observations up to point k only (the observations to the right of point k are disregarded); furthermore these observations up to k may be used to predict t_{k+1}, and also t_{k+2}, t_{k+3}, . . . since

$$X_k^{k+r} = \Phi_{k+r}\Phi_{k+r-1} \cdot \cdot \cdot \Phi_{k+1} X_k^k \tag{5-46}$$

by an extension of the argument leading to (5-34). We do not, however, obtain estimates for t_k using observations beyond k, and we do not obtain the total signal t at points between the measuring points, say, between k and $k+1$. Thus we can, to a certain extent, filter and extrapolate by means of the Kalman approach, but we cannot interpolate.

The underlying Markov model also implies considerable restrictions. As an example, in the one-dimensional continuous case, (5-43) is equivalent to a covariance function

$$C(x) = C_1 e^{-C_2|x|} , \tag{5-47}$$

which is not differentiable at $x = 0$ and, therefore, implies sample functions that are, in general, not differentiable. Thus, such sample functions are not sufficiently smooth to adequately represent, for instance, gravity anomalies.

Thus, the Kalman approach is eminently appropriate to certain applications "in real time", e.g., to space navigation, but its scope is restricted to such applications. On the other hand, least-squares filtering and prediction and their extension, least-squares collocation, are considerably more general.

APPENDIX

Prediction and Adjustment by Parameters

In sections 3 and 4 we have seen that almost all estimation problems can be reduced to an application of the simple least-squares prediction formula (3-16). This holds for adjustment by conditions--cf. (4-13)--as well as for stepwise collocation: not only as regards the signals, s_{P1} and $s_P - s_{P1}$, but even as regards the improvements in the parameter values, $X - X_1$, as we have seen in sec. 3. Only the parameter values themselves, X and X_1, were not of the form (3-16).

However, even the parameter estimation may be reduced to the prediction formula if a suitable limiting procedure is applied, the idea being to consider the parameters as random variables of very small a-priori weight; this idea was introduced by Schmid (1965).

The observation equations for adjustment by parameters are

$$AX + n = x \; ; \qquad (A-1)$$

cf. (4-1). If X is regarded as a random vector of zero mean, then it may be estimated by (3-16):

$$X = C_{Xx} C_{xx}^{-1} x \; . \qquad (A-2)$$

The occurring covariance matrices may be determined as follows. We have

$$Xx^T = XX^TA^T + Xn^T ,$$

$$xx^T = AXX^TA^T + AXn^T + nX^TA^T + nn^T .$$

Denoting the autocovariance matrices of X and n by C and D, respectively, and assuming X to be uncorrelated with s, we find by forming the average M as usual:

$$C_{Xx} = CA^T , \tag{A-3}$$

$$C_{xx} = ACA^T + D . \tag{A-4}$$

Thus (A-2) becomes

$$X = CA^T(ACA^T+D)^{-1}x . \tag{A-5}$$

We now take the matrix C to be a diagonal matrix the elements of which are equal and very large (corresponding to very small weights):

$$C = \frac{1}{\varepsilon} I \tag{A-6}$$

where ε is very small and I is the unit matrix. Then (A-5) becomes

$$X = A^T(AA^T+\varepsilon D)^{-1}x . \tag{A-7}$$

This may be transformed as follows:

$$X = \left[I+\varepsilon(A^TD^{-1}A)^{-1}\right]^{-1}\left[I+\varepsilon(A^TD^{-1}A)^{-1}\right]A^T(AA^T+\varepsilon D)^{-1}x$$

$$= \left[I+\varepsilon(A^TD^{-1}A)^{-1}\right]^{-1}\left[A^T+\varepsilon(A^TD^{-1}A)^{-1}A^T\right](AA^T+\varepsilon D)^{-1}x$$

$$= \left[I+\varepsilon(A^TD^{-1}A)^{-1}\right]^{-1}(A^TD^{-1}A)^{-1}A^TD^{-1}(AA^T+\varepsilon D)(AA^T+\varepsilon D)^{-1}x$$

$$= \left[I+\varepsilon(A^TD^{-1}A)^{-1}\right]^{-1}(A^TD^{-1}A)^{-1}A^TD^{-1}x \quad . \tag{A-8}$$

For $\varepsilon \to 0$ this reduces to

$$X_o = (A^TD^{-1}A)^{-1}A^TD^{-1}x \ , \tag{A-9}$$

which is the usual estimate obtained by least-squares adjustment by parameters.

This shows that the basic prediction formula (A-2) covers even this case if a suitable passage to the limit is used.

As a byproduct we find the relation between the collocation solution (A-8), involving a finite covariance matrix C of form (A-6), and usual adjustment by parameters (A-9):

$$X = (I + \varepsilon P^{-1})^{-1}X_o \ , \tag{A-10}$$

where

$$P = A^TD^{-1}A \tag{A-11}$$

is the weight matrix of the parameters estimated by (A-9).

This is rather analogous to the case considered in (Moritz, 1970); the relation (6-6), loc. cit., is the analogue of our present expression (A-10).

REFERENCES

Bjerhammar, A. (1971) Prediction and filtering of non-stationary stochastic processes in geodesy. Presented at XV General Assembly, IUGG, Moscow.

Bucy, R.S., and P.D. Joseph (1968) Filtering for Stochastic Processes with Applications to Guidance. Interscience Publ., New York.

Deutsch, R. (1965) Estimation Theory. Prentice-Hall, Englewood Cliffs, N.J.

Faddeeva, V.N. (1959) Computational Methods of Linear Algebra (transl. from Russ.) Dover Publ., New York.

Heiskanen, W.A., and H. Moritz (1967) Physical Geodesy. Freeman, San Francisco.

Hirvonen, R.A. (1971) Adjustment by Least-Squares in Geodesy and Photogrammetry. Ungar, New York.

Koch, K.R., and S. Lauer (1971) Automation der Isoliniendarstellung mit Hilfe des Wiener- und des Kalman-Filters. Publ. No. 2, Institute of Theoretical Geodesy, University of Bonn.

Moritz, H. (1970) Combination of satellite harmonics and gravimetry. Report No. 146, Dept. of Geodetic Science, Ohio State University.

Moritz, H. (1972) Advanced least-squares methods. Ibid., Report No. 175.

Papoulis, A. (1965) Probability, Random Variables, and Stochastic Processes. McGraw-Hill, New York.

Schmid, H.H. (1965) Ein allgemeiner Ausgleichungs-Algorithmus zur Auswertung von hybriden Meßanordnungen. Bildmessung und Luftbildwesen, vol. 33, p. 93.

Wolf, H. (1968) Ausgleichungsrechnung nach der Methode der kleinsten Quadrate. Dümmler, Bonn.