Estimation with singular inverses

By ARNE BJERHAMMAR, *Royal Institute of Technology, Stockholm, Sweden.*

(Manuscript received October 7, 1971)

ABSTRACT

The classical estimation procedure according to Gauss uses the inversion of non-singular matrices. The author has earlier presented a number of papers concerning the stochastical estimation with the use of singular inverses (Bjerhammar, 1948–1958). The present paper gives a review of the earlier contributions and an application to selected problems. For a geodetic network all points can be considered unknown and an estimate with a generalized inverse gives the minimum variance of the observations as well as the unknowns. This estimate is *invariant* with respect to rotations and translations, and therefore of special interest for the study of the *optimum network*. It is furthermore of interest because it eliminates the dependence on the coordinate system. A number of similar applications can be found in most sciences.

Gauss presented his famous theory of least squares without using any matrices, but several of his papers include operations which anticipate some type of matrix algebra. The classical matrix algebra was developed by Cayley who gave the definitions for addition, subtraction and multiplication of arbitrary matrices. For a non-singular matrix A (determinant not equal to zero $|A| \neq 0$) he used the following definition of an inverse

$$AA^{-1} = A^{-1}A = I \text{ (= unit matrix)}$$

This definition is not valid for singular matrices and therefore, it is of limited value for statistical applications.

Two independent generalizations of the matrix inverse have been published. Moore presented in 1935 the definition of a uniquely defined inverse for any matrix. The presentation was rather sophisticated and the paper received practically no attention for some time.

An independent study with a generalized approach was presented a few years later (Bjerhammar, 1948, 1949, 1955, 1957, 1958). In this new presentation any matrix has at least one inverse and all singular matrices have an infinite number of inverses. The inverse (A^{-1}) of the matrix A is defined by the relation

$$AA^{-1}A = A \tag{1}$$

These studies which include the first application of generalized inverses on statistical problems will be used for our following presentation. (For abstract rings see v. Neumann, 1936.)

Basic definitions

1. Unit matrix: A product of any matrix and its inverse is a unit matrix. Examples: $A^{-1}A$ and AA^{-1}
2. Regular unit matrix: Square matrix with all non-diagonal elements equal to zero and all diagonal elements equal to 1. (The principal diagonal from upper left corner to lower right corner.)

   Symbol I. $|I| = 1$ (Determinant equal to one.)

3. Singular unit matrix: Any unit matrix which is not equal to I. Symbol A^0. $|A^0| = 0$. (Any matrix with zero determinant is singular.)
4. Order of matrix: For a square matrix the number of elements in a row or a column defines the

order of the matrix. For rectangular matrices the minimum order is the smallest number of elements found in a row or column and the maximum order is the corresponding largest number.

5. Idempotence: Any matrix A satisfying the equation $AA = A$ is idempotent.
6. Transposed matrix: A matrix with the elements A_{ij} gives the transposed matrix A^* with the elements $(A^*)_{ij} = A_{ji}$.
7. Latent roots: The latent roots λ of a matrix A are obtained from the equation

$$|I\lambda - A| = 0.$$

8. Eigenvector: The eigenvector X for the latent root λ is defined by the equation $AX = \lambda X$.
9. Rank: The rank of a matrix A is equal to the order of the largest non-singular matrix that can be formed inside A. Rank of a matrix $A = r(A)$. Note:

$$r(A+B) \leqslant r(A) + r(B)$$

$$r(AB) \leqslant r(A) \quad \text{and} \quad r(AB) \leqslant r(B) \quad r(A) = r(A^{-1}A) = r(AA^{-1}) = r(A^*A) = r(AA^*)$$

Sylvester's law for the rank of square matrices

$$r(AB) \geqslant r(A) + r(B) - n \quad \text{when } n = \text{order of } A \text{ and } B$$

10. Trace: Sum of the elements A_{ii} of a matrix A. (Sum of the elements of the principal diagonal.)

Definition of a unique inverse for any non-zero matrix A. (From Bjerhammar, 1951, page 4)

The matrix A is bordered by independent orthogonal columns and rows until a square non-singular matrix U includes a unique inverse $\underline{A}^{-1}$ as a submatrix of U^{-1}. This submatrix is situated in "transposed position" relative the primary matrix. Cf. below.

We first use this bordering when the rank of the matrix $\underset{nm}{A}$ is equal to m and $n > m$. Then we have for $U = [AD]$ with $A^*D = 0$ (ortogonality) and $|D^*D| \neq 0$ (indipendence)

$$UU^{-1} = \begin{bmatrix} \underset{nm}{A} & \underset{nt}{D} \end{bmatrix} \begin{bmatrix} \underset{mn}{\underline{A}^{-1}} \\ \underset{tn}{\underline{D}^{-1}} \end{bmatrix} = \underset{nn}{I} \qquad \begin{aligned} r(A) &= m \neq 0 \\ r(D) &= n - m = t \\ r(U) &= n \end{aligned} \tag{2}$$

$$\begin{bmatrix} \underline{A}^{-1} \\ \underline{D}^{-1} \end{bmatrix} [A\,D] = \underset{nn}{I} \tag{3}$$

We rewrite these two equations and obtain

$$A\underline{A}^{-1} + D\underline{D}^{-1} = I \tag{2 a}$$

$$\left.\begin{aligned} \underset{mn}{\underline{A}^{-1}}\underset{nm}{A} &= \underset{mm}{I} \qquad & \underset{mn}{\underline{A}^{-1}}\underset{nt}{D} &= \underset{mt}{0} \\ \underset{tn}{\underline{D}^{-1}}\underset{nm}{A} &= \underset{tm}{0} \qquad & \underset{tn}{\underline{D}^{-1}}\underset{nt}{D} &= \underset{tt}{I} \end{aligned}\right\} \tag{3 a}$$

Any solution of $\underline{A}^{-1}$ (and $\underline{D}^{-1}$) that satisfies (3a) or (2a) must be unique because $U = [AD]$ is non-singular.

We introduce the solution

$$\underline{A}^{-1} = (A^*A)^{-1}A^* \quad |A^*A| \neq 0 \tag{4}$$

$$\underline{D}^{-1} = (D^*D)^{-1}D^* \qquad |D^*D| \neq 0 \tag{4a}$$

and find that it satisfies all the equations of (3a). Thus we have proved the uniqueness. We easily verify the following equations (cf. Penrose, 1955 and Bjerhammar, 1949)

$$\begin{matrix} A\underline{A}^{-1}A = A & \underline{A}^{-1}A\underline{A}^{-1} = \underline{A}^{-1} \\ \underline{A}^{-1}A = (\underline{A}^{-1}A)^* & (A\underline{A}^{-1}) = (A\underline{A}^{-1})^* \end{matrix} \qquad (\underline{A}^{-1} \text{ unique}) \tag{5}$$

The null matrix 0 has the "unique" inverse $\underline{0}^{-1} = 0$. (Definition)

Example:

$$A = \begin{bmatrix} 2 & 1 \\ 1 & 1 \\ 0 & 2 \end{bmatrix} \quad U = [AD] = \left[\begin{array}{cc|c} 2 & 1 & 2 \\ 1 & 1 & -4 \\ 0 & 2 & 1 \end{array}\right] \quad U^{-1} = \frac{1}{21}\left[\begin{array}{ccc} 9 & 3 & -6 \\ -1 & 2 & 10 \\ \hline 2 & -4 & 1 \end{array}\right] = \begin{bmatrix} \underline{A}^{-1} \\ \\ \underline{D}^{-1} \end{bmatrix}$$

or

$$\underline{A}^{-1} = \frac{1}{21}\begin{bmatrix} 9 & 3 & -6 \\ -1 & 2 & 10 \end{bmatrix} \quad A^*D = 0 \quad \underline{A}^{-1}A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad A\underline{A}^{-1} = \frac{1}{21}\begin{bmatrix} 17 & +8 & -2 \\ +8 & 5 & +4 \\ -2 & +4 & 20 \end{bmatrix}$$

In the general case when A^*A is singular we can use the corresponding orthogonal bordering of A until we have a square matrix which can be inverted

It is still simpler to use the following technique.

Our new U-matrix is denoted

$$U = \begin{bmatrix} \underset{nm}{A} \\ \underset{tm}{D} \end{bmatrix} \qquad \begin{matrix} r(A) = r \quad r(D) = m - r = t \\ r(U) = m \end{matrix} \tag{6}$$

where

$$|A^*A + D^*D| \neq 0$$
$$DA^* = 0$$
$$|DD^*| \neq 0$$

Then we obtain the generalized inverse

$$\underset{mn}{\underline{A}^{-1}} = (\underset{mn}{A^*}\underset{nm}{A} + \underset{mt}{D^*}\underset{tm}{D})^{-1}\underset{mn}{A^*} \qquad |A^*A + D^*D| \neq 0 \tag{7}$$

This inverse of A is always unique for any matrix simply because $(A^*A + D^*D)$ is non-singular.

For $n = m$ we can use this formula without any difficulties, but for $m > n$ we might transpose the matrix before the computation and then transpose the final answer. We can also choose the following procedure. We denote this type of matrix B. Then we have

$$\underset{nk}{\underline{B}^{-1}} = \underset{nk}{B^*}(\underset{kn}{B}\underset{nk}{B^*} + \underset{kt}{D}\underset{tk}{D^*})^{-1} \qquad n > k \quad r(B) = r \quad r(D) = k - r = t \tag{8}$$

where

$$|BB^* + DD^*| \neq 0$$
$$D^*B = 0$$
$$|D^*D| \neq 0$$

Numerical example:

$$A = \begin{bmatrix} 2 & 2 \\ 1 & 1 \\ 2 & 2 \end{bmatrix}$$

Orthogonalization:
1st column; $2x + 1y + 2z = l_1 \quad x = 0.5\, l_1 - 0.5y - z$
2nd column: $2(0.5l_1 - 0.5y - z) + 1y + 2z = l_2$
Condition: $1l_1 - 1l_2 = 0$

$$D = [+1 \quad -1]$$

$$U = \begin{bmatrix} 2 & 2 \\ 1 & 1 \\ 2 & 2 \\ 1 & -1 \end{bmatrix} \quad A^{-1} = (A^*A + D^*D)^{-1}A^* = \frac{1}{18}\begin{bmatrix} 2 & 1 & 2 \\ 2 & 1 & 2 \end{bmatrix}$$

Greville (1960) defined the unique inverse of equation (7) in the following way

$$\underline{A}^{-1} = C^*(CC^*)^{-1}(B^*B)^{-1}B^* \tag{9}$$

where

$$\underset{mn}{A} = \underset{mr}{B}\,\underset{rn}{C} \qquad r(A) = r$$

Finally we use bordering with orthogonal rows and columns for the definition of an inverse

$$U = \begin{bmatrix} A & D_1 \\ D_2 & 0 \end{bmatrix} \qquad |U| \neq 0 \tag{10}$$

using the definition of the non-singular inverse we can write

$$UU^{-1}U = U \quad U^{-1}UU^{-1} = U^{-1}$$
$$UU^{-1} = (UU^{-1})^* \quad U^{-1}U = (U^{-1}U)^* \tag{11}$$

These equations give again our previous equation (5) for the determination of the inverse of A. Penrose (1955) used as a reference the publication (Bjerhammar, 1951) and gave this implicit definition of the inverse. It is obvious that the definition is identical with that of Bjerhammar when using bordering.

We give a few numerical examples of the computation of generalized inverses. First we use the primary matrix

$$A = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 3 & 6 & 3 \end{bmatrix}$$

This matrix has the rank 1 and we make an expansion with independent orthogonal columns and rows until we obtain a non-singular matrix. We choose the following combination

$$U = \left[\begin{array}{ccc|cc} 1 & 2 & 1 & -1 & -3 \\ 2 & 4 & 2 & 5 & -6 \\ 3 & 6 & 3 & -3 & 5 \\ \hline 5 & -2 & -1 & 0 & 0 \\ -1 & 1 & -1 & 0 & 0 \end{array}\right]$$

The primary matrix has only one independent row and one independent column. We add two orthogonal rows and two orthogonal columns.

This U matrix is non-singular and has the inverse

$$U^{-1}=\frac{1}{84}\left[\begin{array}{rrr|rr} 1 & 2 & 3 & 14 & 0 \\ 2 & 4 & 6 & 0 & 28 \\ 1 & 2 & 3 & -14 & -56 \\ \hline -24 & 12 & 0 & 0 & 0 \\ -18 & 0 & +6 & 0 & 0 \end{array}\right]$$

This means that we have found the generalized inverse

$$\underline{A}^{-1}=\frac{1}{84}\begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 1 & 2 & 3 \end{bmatrix}$$

satisfying

$$A\underline{A}^{-1}A=A \quad \underline{A}^{-1}A\underline{A}^{-1}=\underline{A}^{-1} \quad (A\underline{A}^{-1})=(A\underline{A}^{-1})^* \quad \underline{A}^{-1}A=(\underline{A}^{-1}A)^*$$

Discussion

We have here bordered the matrix A in a square matrix U with the use of orthogonal matrices. We are now going to prove that this orthogonalization gives a non-singular U-matrix. In the simple case when the first m-rows of the A-matrix give a non-singular matrix we split this matrix into two submatrices

$$A=\begin{bmatrix} \underset{mm}{E} \\ \underset{tm}{F} \end{bmatrix} \qquad m+t=n \quad r(A)=r(E)=r=m \quad \text{Degenerations: } n-r \tag{12}$$

Here the rank of A is equal to the rank of E. We use an orthogonalization in the following way.

$$U=[AD]=\left[\begin{array}{c|c} E & (E^{-1})^*F^* \\ \hline F & -I \end{array}\right] \quad \text{where } A^*D=0 \tag{13}$$

For each degeneration of A we have included an independent column orthogonal of A.
Now we have

$$r(U)=r(U^*U)=r\begin{bmatrix} E^*E+F^*F & 0 \\ 0 & I+FE^{-1}(E^{-1})^*F^* \end{bmatrix} \tag{14}$$

and

$$r(E^*E+F^*F)=m \quad r(I+FE^{-1}(E^{-1})^*F^*)=n-m \tag{15}$$

Then

$$r(U^*U)=r(U)=n \tag{16}$$

Thus we have proved that the orthogonal bordering gives a non-singular matrix U.

For $r(A)=m$ but $r(E)<m$ we first consider a change of the order of the rows until we obtain an upper matrix E where $r(E)=r(A)$. After the computation of the inverse we restore the original

order. Thus our proof will also be useful here. For $r(A) < m$ we simply add the $m-r$ orthogonal rows and the new rectangular matrix is of rank m, and therefore our proof is still valid.

Existence. Any matrix A has at least one inverse. This is obvious from the non-singularity of the U-matrices.

The null matrix $\mathbf{0}$ has the inverse $\mathbf{0}$.

Orthogonalization. The orthogonal D-matrices can be obtained with the use of classical methods for orthogonalization. We have a system of linear equations. For example

$$A^*D = 0 \quad D = [I - (A^*)^{-1}A^*]M \qquad M = \text{arbitrary matrix of adequate order}$$

In most applications we have no **a priori** information concerning the order of the matrix D and therefore the following procedure is recommended for $AD^* = 0$.

$$A = \begin{bmatrix} a & a \\ ab & ab \\ ad & ad \end{bmatrix} \qquad r(A) = 1$$

The first column gives in a system of linear equations

$$ax + aby + adz = c_1$$

$$x = c_1/a - by - dz$$

The second column gives now

$$c_1 - aby - adz + aby + adz = c_2$$

Thus we obtain the condition

$$1\,c_1 - 1\,c_2 = 0 \quad \text{or} \quad D = [1 \; -1]$$

We easily verify that $AD^* = 0$. This technique can be extended to a matrix of any order.

We summarize:

One and only one unique inverse $\underline{A}^{-1}$ is obtained from any of the following transformations. (Cf. Bjerhammar, 1951.)

Non-singular representation of the singular pseudoinverse:

$$\underline{A}^{-1} = (A^*A + D^*D)^{-1}A^* \qquad AD^* = 0 \tag{17}$$

$$\underline{A}^{-1} = A^*(AA^* + DD^*)^{-1} \qquad A^*D = 0 \tag{18}$$

$$\underline{A}^{-1} = (A^*A + D^*D)^{-1}A^* \qquad A^*AD^* = 0 \tag{19}$$

$$\underline{A}^{-1} = A^*(AA^* + DD^*)^{-1} \qquad AA^*D = 0 \tag{20}$$

$$\underline{A}^{-1} = T^*(TT^*)^{-1}(TT^*)^{-1}TA^* \qquad r(T) = r(A) \qquad \underset{mr}{T^*}\,\underset{rm}{T} = \underset{mn}{A^*}\,\underset{nm}{A} \tag{21}$$[1]

$$\underline{A}^{-1} = A^*T^*(TT^*)^{-1}(TT^*)^{-1}T \qquad r(T) = r(A) \qquad \underset{kr}{T^*}\,\underset{rk}{T} = \underset{kn}{A}\,\underset{nk}{A^*} \tag{22}$$[1]

$$\begin{bmatrix} A & D_1 \\ D_2 & 0 \end{bmatrix}^{-1} = \begin{bmatrix} \underline{A}^{-1} & \underline{D}_2^{-1} \\ \underline{D}_1^{-1} & 0 \end{bmatrix} \quad A^*D_1 = 0 \qquad AD_2^* = 0 \tag{23}$$

---

[1] T is a submatrix of H, obtained after exclusion of all rows with $(H)_{ii} = 0$ when $H^*H = A^*A$ and $(H)_{ij} = 0$ for $i > j$.

All these seven formulas allow a straight forward computation of this inverse. (Cf. also Greville, 1960.) They are also useful for the implicit type of solution. For example

$$(A^*A + D^*D)X = A^*L \tag{24}$$

$$X = (A^*A + D^*D)^{-1}A^*L = A^{-1}L \tag{25}$$

Implicit singular definitions

$$\underline{A}^{-1} = (\underline{A^*A})^{-1}A^* \qquad (\underline{A^*A})^{-1} = \text{pseudoinverse of } (A^*A) \tag{26}$$

$$\underline{A}^{-1} = A^*(\underline{AA^*})^{-1} \qquad (\underline{AA^*})^{-1} = \text{pseudoinverse of } (AA^*) \tag{27}$$

$$\begin{cases} A\underline{A}^{-1}A = A \\ \underline{A}^{-1} = A^*(\underline{AA^*})^{-1}A(\underline{A^*A})^{-1}A^* \end{cases} \qquad \text{(Cf. Bjerhammar, 1958)} \tag{28}$$

$$A\underline{A}^{-1}A = A \quad \underline{A}^{-1}A\underline{A}^{-1} = \underline{A}^{-1} \quad A\underline{A}^{-1} = (A\underline{A}^{-1})^* \quad \underline{A}^{-1}A = (\underline{A}^{-1}A^*) \tag{29}$$

Penrose (1955) proved that $\underline{A}^{-1}$ here is unique.

A reciprocal inverse was defined in the following way. (Cf. Bjerhammar, 1957, 1958 and Oktaba, 1969.)

$$AA^{-1}A = A \quad A^{-1}AA^{-1} = A^{-1}$$

This inverse is not generally unique.

After replacing A^* by A^*P in the formula (17), (19), (21) and (26) we obtain one and only one reciprocal inverse for any $P = R^*R$ when $r(R) = n$. After replacing A^* by $P^{-1}A^*$ in the formulas (18), (20), (22) and (27) we obtain one and only one reciprocal inverse for any $P = R^*R$ when $r(R) = n$.

We now want to verify that our expression of the inverse (17) satisfies equation (1). A and D are orthogonal and thus we have

$$A^*AA^*A = (A^*A + D^*D)A^*A \quad AD^* = 0 \tag{30}$$

Then

$$A(A^*A + D^*D)^{-1}A^*AA^*A = AA^*A \tag{31}$$

$$A(A^*A + D^*D)^{-1}A^*A(A^*A + D^*D) = A(A^*A + D^*D) \tag{32}$$

The expressions inside the brackets are non-singular and cancel. Thus

$$A(A^*A + D^*D)^{-1}A^*A = A \tag{33}$$

$$AA^{-1}A = A \quad \text{Q.E.D.}$$

We can prove in a similar way that the inverse (17) satisfies

$$A^{-1}AA^{-1} = A^{-1}$$

However, it is simple to use the solution $U^{-1}UU^{-1} = U^{-1}$ where A and D are submatrices of U. In our next proof we start from the relation

$$(A^*A + D^*D)A^*A = A^*A(A^*A + D^*D) \tag{34}$$

Thus

$$A^*A(A^*A + D^*D)^{-1} = (A^*A + D^*D)^{-1}A^*A \tag{35}$$

Then we have proved

$$(A^{-1}A)^* = A^{-1}A \tag{36}$$

In a similar way $$(AA^{-1})^* = AA^{-1} \tag{36 a}$$

Thus we have proved the equivalence of the approach of the Bjerhammar (1951) and Penrose (1955).

In equations (21) and (26) we use $(\underline{A^*A})^{-1}$. Here we put

$$Q = A^*A = T^*T \quad (\underline{A^*A})^{-1} = Q^{-1} = T^*(TT^*)^{-1}(TT^*)^{-1}T \tag{37}$$

Then we have

$$QQ^{-1}Q = Q \quad Q^{-1}QQ^{-1} = Q^{-1} \quad Q^{-1}Q = (Q^{-1}Q)^* \quad QQ^{-1} = (QQ^{-1})^*$$

Thus we have proved that Q^{-1} is a pseudoinverse.

In order to verify complete expressions like (26) we write

$$\underline{A}^{-1} = (A^*AA^*A + D^*D)^{-1}A^*AA^* \quad A^*AD^* = 0 \tag{38}$$

Then the previous proof for equation (17) can be used once more.

We now introduce two generalized inverses using the auxilliary matrices R_1 and R_2 of adequate order

$$(A^{-1})_1 = (R_1A)^{-1}R_1 \quad r(R_1A) = r(A) = r(R_1) \tag{39}$$

$$(A^{-1})_2 = R_2(AR_2)^{-1} \quad r(AR_2) = r(A) = r(R_2) \tag{40}$$

If the rank is lower than the minimum order then these two expressions are non-unique inverses of equation (1).

It has been proved that for any fixed matrices R_1 and R_2 we obtain one and only one unique inverse of any (non-zero) matrix A (Bjerhammar, 1957, 1958)

$$A^{-1} = (A^{-1})_2A(A^{-1})_1 \tag{41}$$

This means that we here have a general definition giving one and only one inverse for any selected R_1 and R_2. As special cases we select a few examples of such unique inverses with a statistical application.

For $R_1 = R_2 = A^*$ we obtain

$$\underline{A^{-1}} = A^*(AA^*)^{-1}A(A^*A)^{-1}A^* = (\underline{A^*A})^{-1}A^* = (A^*A + D^*D)^{-1}A^* \tag{42}$$

For $R_1 = A^*P \quad R_2 = A^*$ we find

$$\underline{A^{-1}} = A^*(AA^*)^{-1}A(A^*PA)^{-1}A^*P = (\underline{A^*PA})^{-1}A^*P = (A^*PA + D^*D)^{-1}A^*P \quad AD^* = 0 \tag{43}$$

$R_2 = QB^*$ gives finally for a matrix B

$$\underline{B^{-1}} = QB^*(BQB^*)^{-1}(B^*B)^{-1}B^* = QB^*(\underline{BQB^*})^{-1} = QB^*(BQB^* + DD^*)^{-1} \quad B^*D = 0 \tag{44}$$

Notations: When not otherwise stated then $\underline{A}^{-1}$ is defined according to (42).

For generalized inverses we have:

1. All singular unit matrices are idempotent.

   Examples: $A^0A^0 = A^0 \quad (I - A^0)\ (I - A^0) = I - A^0$

   The proof is obvious from the definition of A^{-1} and A^0.
2. If the matrix A^0 has the rank r and the order n then $(I - A^0)$ has the rank $n - r$. Proof: Sylvesters law gives when considering $r(A + B) \leqslant r(A) + r(rB)$

$$r(AB) \geqslant r(A) + r(B) - n \quad \text{For } A^0 = A \text{ and } I - A^0 = B \quad \text{we obtain}$$
$$r\{(I - A^0)A^0\} \geqslant r(I - A^0) + r(A^0) - n \quad \Leftrightarrow (n - r) + (r) - (n) = 0$$

3. The latent roots of A^0 or $I - A^0$ are all equal to 1 or 0 and the sum of the latent roots is equal to the rank.

   (Note $\lambda X = A^0X$ and $\lambda X = A^0A^0X$. Thus $\lambda^2X = \lambda X$ and $\lambda = 1$ or $\lambda = 0$.)

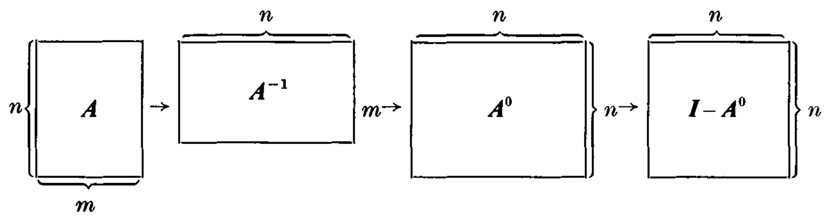
4. Let ϵ represent a vector of normally distributed and independent stochastic variables with zero means $E\{\epsilon\} = 0$ and unit variance $N(0, 1)$, then $\epsilon^* A^0 \epsilon$ is equivalent to the sum of the squares of r independent stochastic variables.

Proof. All non-diagonal elements vanish after transcribing with the latent roots:

$\epsilon^* A^0 \epsilon = \eta_1^2 \lambda_1 + \eta_2^2 \lambda_2 + \dots + \eta_r^2 \lambda_r$ where $\eta_1, \dots, \eta_r$ are independent stochastic variables and $\lambda_1 = \lambda_2 \dots =$

$\lambda_r = 1$. Thus $\epsilon^* A^0 \epsilon$ will have a χ^2-distribution with r degrees of freedom. In the same way $\epsilon^*(I - A^0)\epsilon$ can be transcribed as a sum of $n - r$ independent stochastic variables.

Flow diagram



Flow diagram for a computation of a generalized inverse when using triangular decomposition

$$\underset{A}{\begin{bmatrix} 1 & 0 & 3 \\ 2 & 0 & 6 \\ 3 & 0 & 9 \\ 2 & 1 & 2 \end{bmatrix}} \rightarrow \underset{A^*A}{\begin{bmatrix} 18 & 2 & 46 \\ 2 & 1 & 2 \\ 46 & 2 & 130 \end{bmatrix}} \rightarrow \underset{T}{\begin{bmatrix} \frac{18}{\sqrt{18}} & \frac{2}{\sqrt{18}} & \frac{46}{\sqrt{18}} \\ 0 & \frac{7}{3\sqrt{7}} & -\frac{28}{3\sqrt{7}} \end{bmatrix}}$$

$$\underset{TT^*}{\begin{bmatrix} \frac{1222}{9} & -\frac{1274}{3\sqrt{126}} \\ -\frac{1274}{3\sqrt{126}} & \frac{833}{63} \end{bmatrix}} \rightarrow \underset{(TT^*)^{-1}}{\frac{1}{364}\begin{bmatrix} \frac{833}{63} & \frac{1274}{3\sqrt{126}} \\ \frac{1274}{3\sqrt{126}} & \frac{1222}{9} \end{bmatrix}}$$

$$\frac{1}{26}\begin{bmatrix} 17 & 12 & 3 \\ 12 & 10 & -4 \\ 3 & -4 & 25 \end{bmatrix} = A^{-1}A$$

$$\frac{1}{364}\begin{bmatrix} -7 & -14 & -21 & 168 \\ -8 & -16 & -24 & 140 \\ 11 & 22 & 33 & -56 \end{bmatrix}$$

$$A^{-1} = T^*(TT^*)^{-1}(TT^*)^{-1}TA^*$$

$$\frac{1}{14}\begin{bmatrix} 1 & 2 & 3 & 0 \\ 2 & 4 & 6 & 0 \\ 3 & 6 & 9 & 0 \\ 0 & 0 & 0 & 14 \end{bmatrix} = AA^{-1}$$

2. Complete set of solutions

Definition. For an arbitrary matrix A we define the inverse A^{-1}.

$$AA^{-1}A = A \tag{1}$$

It can be proved that a matrix has at least one inverse.

Theorem: An arbitrary linear matrix equation

$$AX = L \tag{2}$$

with at least one solution X_0, has the complete set of solutions $\underset{\sim}{X}$. (From Bjerhammar, 1951 p. 9.)

$$\boxed{\underset{\sim}{X} = A^{-1}L + (I - A^{-1}A)M} \tag{3}$$

where

I = unit matrix

A^{-1} = any particular inverse that satisfies (1)

M = the total set of matrices that can be added to X

Proof. Equation (2) is satisfied by

$$X = X_0 + (I - A^{-1}A)M \tag{4}$$

We put $M = A^{-1}AA^{-1}L - X_0$ and obtain

$$A^{-1}L \in \underset{\sim}{X} \tag{5}$$

This means that $A^{-1}L$ is a particular solution of (2). Thus we can write

$$A^{-1}L + (I - A^{-1}A)M \subset \underset{\sim}{X} \tag{6}$$

Here we replace M by $\underset{\sim}{X}$ and obtain

$$\underset{\sim}{X} = A^{-1}L + (I - A^{-1}A)M \tag{7}$$

Thus we have

$$\underset{\sim}{X} \subset A^{-1}L + (I - A^{-1}A)M \subset \underset{\sim}{X} \quad \text{Q.E.D} \tag{8}$$

A consistency condition is obtained from (2) and (3)

$$\boxed{AA^{-1}L = L} \tag{9}$$

According to (3) we have the solution for the *homogeneous* equations

$$\underset{\sim}{X} = (I - A^{-1}A)M \tag{10}$$

Equations (9) and (10) give

$$(A - AA^{-1}A)M = 0 \tag{11}$$

Thus we find that the second part of the right member of equation (3) gives the solution of the homogeneous equation. Equations (3) and (1) give the complete set of inverses

$$\boxed{\underset{\sim}{A}^{-1} = A^{-1}AA^{-1} + N(I - AA^{-1}) + (I - A^{-1}A)M} \tag{12}$$

where

M, N = the total sets of matrices that can be added to A^{-1}.

From (2) and (9) we obtain for *inhomogeneous* equations

$$\boxed{\underset{\sim}{X} = \underset{\sim}{A}^{-1} L} \quad (L \neq 0) \tag{13}$$

3. The complete set of linear unbiased estimates

Theorem: The theoretical means ξ and λ satisfy the inhomogeneous equation $(\lambda \neq 0)$

$$\underset{nm}{A}\ \underset{m1}{\xi} = \underset{n1}{\lambda} \tag{1}$$

with the corresponding observation equations for the outcomes

$$\underset{nm}{A}\ \underset{m1}{X} \simeq \underset{n1}{L} \qquad \textit{with } E\{L\} = \lambda \tag{2}$$

where $\{l_1, l_2, l_3 \dots l_n\}^* = L^*$ *is the set of observations (outcomes).*
Then the complete set of unbiased (linear) estimates is given by

$$\boxed{\underset{m1}{\underset{\sim}{X}} = \underset{mn}{\underset{\sim}{A}^{-1}}\ \underset{n1}{L}} \tag{3}$$

We here note that we are using an arbitrary matrix A which means that we also accept a matrix with a lower rank than the smallest number of elements in the vertical or horizontal direction. (Non-stochastic errors are excluded.)

Proof. Equation (1) defines our theoretical mean ξ in the following way

$$\xi = \underset{\sim}{A}^{-1}\lambda \quad (\text{unique for } A^{-1}A = I) \tag{4}$$

In our estimate we replace λ by L and obtain

$$\underset{\sim}{X} = \underset{\sim}{A}^{-1}L \tag{5}$$

Thus we find the expectation for any inverse

$$E\{\underset{\sim}{X}\} = E\{\underset{\sim}{A}^{-1}L\} = \underset{\sim}{A}^{-1}E\{L\} = \underset{\sim}{A}^{-1}\lambda = \xi \quad \text{Q.E.D.} \tag{6}$$

The completeness of our solution is obvious from equation (3) in the previous chapter. For

$$E\{L\} \neq \lambda \quad \text{is } E\{X\} \neq \xi \tag{7}$$

Theorem: If the theoretical means ξ and λ satisfy the homogeneous equation $(\lambda = 0)$

$$\underset{nm}{A}\ \underset{m1}{\xi} = \underset{n1}{\lambda} = \underset{n1}{0} \quad (\xi = (I - A^{-1}A)M) \tag{1 a}$$

and have the observation equations

$$AX = L \quad (\textit{consistent or inconsistent}) \tag{2 a}$$

Then the following estimates are unbiased

$$\boxed{\underset{\sim}{X} = A^{-1}L + (I - A^{-1}A)M} \tag{3 a}$$

The proof is obvious from the preceeding study.

In the most general case the rank of A will be lower than the minimum dimension of the matrix Then the complete set of unbiased estimates is obtained from (3) and (2:12)

$$\boxed{\underset{\sim}{X} = A^{-1}AA^{-1}L + N(I - AA^{-1})L + (I - A^{-1}A)ML} \tag{9}$$

Because of the inconsistency of our observation equations we are not entitled to make use of the relations

$$A^{-1}AA^{-1}L = A^{-1}L$$

and

$$N(I - AA^{-1})L = N(I - AA^{-1})AX = 0$$

which are only valid for the corresponding consistent equation (2:13).

It is of interest that for each N and M of equation (9) we have an infinite set of unbiased solutions which are all valid for equation (2).

We also note the equivalent expression for the complete set of linear unbiased estimates.

$$X = (RA)^{-1}RL \tag{10}$$

where R includes the set of matrices that satisfy rank RA = rank A.

Proof.

$$\underset{\sim}{A}^{-1}L \subset (\underset{\sim}{IA})^{-1}IL \subset (\underset{\sim}{RA})^{-1}RL \subset \underset{\sim}{A}^{-1}L$$

From our complete set of solutions (9) we select a subset which is valid when

$$A^{-1}A = I \tag{11}$$

Then we have the complete set of linear unbiased estimates

$$\boxed{\underset{\sim}{X} = A^{-1}L + N(I - AA^{-1})L} \tag{12}$$

or when using equation (10) for $(RA)^{-1}RA = I$

$$\boxed{\underset{\sim}{X} = (RA)^{-1}RL = (A^*R^*RA)^{-1}A^*R^*RL = (A^*PA)^{-1}A^*PL} \quad (P = R^*R) \tag{12 a}$$

Proof.

$$\underset{\sim}{A}^{-1}L \subset II\underset{\sim}{A}^{-1}L \subset [A^*(\underset{\sim}{A}^{-1})^*\underset{\sim}{A}^{-1}A]^{-1}A^*(\underset{\sim}{A}^{-1})^*\underset{\sim}{A}^{-1}L \subset (A^*R^*RA)^{-1}A^*R^*RL \subset \underset{\sim}{A}^{-1}L$$

Equation (12) can be written

$$\underset{\sim}{X} = (A^*A)^{-1}A^*L + N[I - A(A^*A)^{-1}A^*]L \tag{12b}$$

In this way we present the *complete set of linear unbiased estimates* which are all closely linked with the generalized inverse of (1:1). This means that if we repeat the measurements the final estimate will approach the theoretical mean.

We now make use of a selected unbiased estimate from (12b)

$$\hat{X} = A^{-1}L \tag{13}$$

Then we obtain the variance

$$s^2 = \frac{(A\hat{X} - L)^*(A\hat{X} - L)}{n} \quad \text{(biased estimate of } \sigma^2\text{)} \tag{14}$$

The present approach will be further explained with a numerical model. Observation equations

$$\begin{aligned} 2x_1 + 1x_2 &= 6 \\ 1x_1 + 1x_2 &= 4 \\ 0x_1 + 2x_2 &= 5 \end{aligned} \qquad A = \begin{bmatrix} 2 & 1 \\ 1 & 1 \\ 0 & 2 \end{bmatrix} \qquad L = \begin{bmatrix} 6 \\ 4 \\ 5 \end{bmatrix}$$

We can use any generalized inverse of A for the computation of an unbiased estimate.

We start with the following inverse

$$A^{-1} = \begin{bmatrix} 1 & -1 & 0 \\ -1 & 2 & 0 \end{bmatrix}$$

which satisfies the relation $AA^{-1}A = A$ (and $A^{-1}A = I$) and is obtained simply by computing the Cayley inverse of the upper four elements and adding zeros.

Then we obtain the unbiased estimate of x_1 and x_2

$$x_1 = 2 \quad x_2 = 2 \quad \text{with } s^2 = \frac{0^2 + 0^2 + 1^2}{3} = 0.333$$

As an alternative approach we use the inverse

$$A^{-1} = \begin{bmatrix} 0 & 1 & -0.5 \\ 0 & 0 & 0.5 \end{bmatrix}$$

and we obtain the unbiased estimate of x_1 and x_2

$$x_1 = 1.5 \quad x_2 = 2.5 \quad \text{with } s^2 = \frac{0.5^2 + 0^2 + 0^2}{3} = 0.083$$

Finally we make use of the inverse

$$A^{-1} = (A^*A)^{-1}A^* = \frac{1}{21}\begin{bmatrix} 9 & 3 & -6 \\ -1 & 2 & 10 \end{bmatrix}$$

which gives the unbiased estimate of x_1 and x_2

$$x_1 = 1\frac{15}{21} \quad x_2 = 2\frac{10}{21} \qquad s^2 = \frac{(-2)^2 + 4^2 + (-1)^2}{1323} = 0.016$$

If we compare these three unbiased solutions we have at present no means that make it possible to decide which method will be closest to the true answer. We can only state that the first solution resulted in the largest variance and the last one gave the smallest variance. We obtain the complete set of unbiased estimates.

$$X = \frac{1}{21}\left[\begin{bmatrix} 9 & 3 & -6 \\ -1 & 2 & 10 \end{bmatrix} + \begin{bmatrix} n_{11} & n_{12} & n_{13} \\ n_{21} & n_{22} & n_{23} \end{bmatrix}\begin{bmatrix} 4 & -8 & 2 \\ -8 & 16 & -4 \\ 2 & -4 & 1 \end{bmatrix}\right]\begin{bmatrix} l_1 \\ l_2 \\ l_3 \end{bmatrix}$$

We can select any finite values of $n_{11} \ldots n_{23}$ and the limiting value of X will be identical for any set of unbiased observations. In an example with 100 sets of observations we obtained the following results:

Selected n-values						*Arithmetic mean of 100 solutions*	
$n_{11}=0$	$n_{12}=0$	$n_{13}=0$	$n_{21}=0$	$n_{22}=0$	$n_{23}=0$	$x_1=1.11$	$x_2=2.52$
$n_{11}=1$	$n_{12}=1$	$n_{13}=1$	$n_{21}=1$	$n_{22}=1$	$n_{23}=1$	$x_1=1.10$	$x_2=2.53$
$n_{11}=1$	$n_{12}=0$	$n_{13}=1$	$n_{21}=0$	$n_{22}=1$	$n_{23}=1$	$x_1=1.11$	$x_2=2.52$

Apparently the three sets of solutions converge to the same results. All three solutions are unbiased. When we use a very large number of observations almost identical results can be expected. For smaller sets of observations it is very important to find the "best" estimate.

4. Generalized method of least squares

In the following investigation we will make use of a linear stochastic model where the observations define the unknown parameters indirectly by a system of linear equations. The stochastic model is denoted in the following way

$$\begin{aligned} &a_{11}\xi_1 + a_{12}\xi_2 + a_{13}\xi_3 + \dots + a_{1m}\xi_m = \lambda_1 \\ &a_{21}\xi_1 + a_{22}\xi_2 + a_{23}\xi_3 + \dots + a_{2m}\xi_m = \lambda_2 \quad m<n \\ &\text{— — — — — — — — — —} \\ &a_{n1}\xi_1 + a_{n2}\xi_2 + a_{n3}\xi_3 + \dots + a_{nm}\xi_m = \lambda_n \end{aligned} \tag{1}$$

where

a_{ij} = non-stochastic parameters
λ_i = theoretical mean of an observed stochastic variable
ξ_j = theoretical mean of an unknown stochastic variable

This stochastic model is transcribed using matrix symbols.

$$\underset{nm}{A}\ \underset{m1}{\boldsymbol{\xi}} = \underset{n1}{\boldsymbol{\lambda}} \quad r(A)>0 \tag{1 a}$$

We assume that the stochastic model gives an adequate presentation of our set of observations.

The unknown theoretical means λ_i and ξ_j will be estimated by the observations. Then we obtain the following observation equations.

$$\begin{aligned} &a_{11}\xi_1 + a_{12}\xi_2 + a_{13}\xi_3 + \dots + a_{1m}\xi_m = l_1 - \varepsilon_1 \\ &a_{21}\xi_1 + a_{22}\xi_2 + a_{23}\xi_3 + \dots + a_{2m}\xi_m = l_2 - \varepsilon_2 \quad m<n \\ &\text{— — — — — — — — — —} \\ &a_{n1}\xi_1 + a_{n2}\xi_2 + a_{n3}\xi_3 + \dots + a_{nm}\xi_m = l_n - \varepsilon_n \end{aligned} \tag{1b}$$

where

l_i = direct observation of a stochastic variable
ε_i = observation error (accidental)

The expectation of an observation is denoted in the following way

$$E\{l_i\} = \lambda_i \quad (i = 1, 2 \dots n) \tag{1c}$$

The theoretical means will remain unknown and we have to replace them by suitable estimates. In the simplest case we will use a simple observation as our sample mean. In more advanced application we prefer to use the arithmetic mean of a set of observations. Then it will be possible to form an estimate of the variance of a single observation. The present study is restricted to a case where all observations are considered to have the same a priori variance. This means that all observations are equivalent or have equal weight. Our observation equations are now transcribed using matrix symbols.

$$\underset{nm}{A}\ \underset{ml}{\xi} = \underset{nl}{L} - \underset{nl}{\epsilon} \quad \text{with} \quad E\{\epsilon\epsilon^*\} = \sigma^2 I \tag{2}$$

After premultiplication by any selected inverse we obtain the general expression of ξ (cf. 1:3)

$$\underset{ml}{\xi} = \underset{mn}{A^{-1}}\underset{nl}{L} - \underset{mn}{A^{-1}}\underset{nl}{\epsilon} + (\underset{mm}{I} - \underset{mn}{A^{-1}}\underset{nm}{A})\underset{ml}{M} \tag{3}$$

This expression is only unique for $A^{-1}A = I$ and ϵ known.

Equations (2) and (3) give for any inverse

$$(\underset{nn}{I} - \underset{nn}{A^0})\underset{nl}{\epsilon} = \underset{nl}{L} - \underset{nn}{A^0}\underset{nl}{L} \quad (\underset{nn}{A^0} = \underset{nm}{A}\ \underset{mn}{A^{-1}}) \tag{4}$$

When using $(I - A^0)^{-1} = I$ we obtain according to (2:3)

$$\underset{\sim}{\epsilon} = L - A^0L + A^0M \tag{5}$$

which is the complete set of error estimates that satisfies our observation equation (2). Any error estimate that satisfies the observation equation is here called a "consistent error estimate". When not otherwise stated we use the particular inverse

$$A^{-1} = (A^*A)^{-1}A^* \quad \text{with } (A^0)^* = A^0 \tag{6}$$

We have

$$\underset{\sim}{\epsilon^*\epsilon} = L^*L - L^*A^0L + M^*A^0A^0M \tag{7}$$

which has a unique minimum for $A^0M = 0$ giving the error estimate $\hat{\epsilon} = L - A^0L$.

Now we have found our estimate ($\hat{X}$) according to the method of least squares which gives the smallest variance. "Generalized least squares estimate".

$$s^2 = \hat{\epsilon}^*\hat{\epsilon}/Sp(I + A^0) \qquad \boxed{\underset{\sim}{\hat{X}} = (\underset{\sim}{A^*A})^{-1}A^*L} \qquad \text{Note: } E\{\hat{X}\} = \xi \qquad E(s^2) = \sigma^2 \tag{8}$$

For $A^{-1}A \neq I$ the Gaussian approach gives no solution. Our definitions give infinite solutions of X with a unique minimum of $\epsilon^*\epsilon$ for this case. We have already proved that the solution is unbiased. We easily find the alternative estimates

$$X = (\underset{\sim}{A^*PA})^{-1}A^*PL \quad \hat{\epsilon}^*P\hat{\epsilon} = \min \tag{8b}$$

Furthermore we have the *unique* estimates

$$\hat{X} = (\underline{A^*PA})^{-1}A^*PL \quad \hat{\epsilon}^*P\,\hat{\epsilon} = \min \quad \sum s_{x_i}^2 = \min \tag{8 c}$$

$$\hat{X} = (A^*PA + D^*D)^{-1}A^*PL \quad \hat{\epsilon}^*P\,\hat{\epsilon} = \min \quad \sum s_{x_i}^2 = \min \quad AD^* = 0 \tag{8 d}$$

$$\hat{X} = A^*(AA^*)^{-1}(A^*PA)^{-1}A^*PL \quad \hat{\epsilon}^*P\,\hat{\epsilon} = \min \quad \sum s_{x_i}^2 = \min \tag{8 e}$$

The proof of (8b) is obvious from the preceeding presentation. The three equations (8c), (8d) and (8e) are all equivalent and we choose to prove the validity of the equations in the following way.

We have the estimate

$$\hat{X} = \underline{A}^{-1}L \tag{9}$$

The error of ϵ is then

$$\hat{\epsilon} = (I - A\underline{A}^{-1})L \tag{10}$$

Furthermore we obtain the error of $\hat{X}$

$$\epsilon_{\hat{x}} = \underline{A}^{-1}\epsilon \tag{11}$$

With an arbitrary inverse we obtain the error estimate of $\hat{X}$

$$\epsilon_{\hat{x}} = \underline{A}^{-1}\epsilon + (I - \underline{A}^{-1}A)M \tag{12}$$

with

$$\epsilon_{\hat{x}}^* \epsilon_{\hat{x}} = \epsilon^*(\underline{A}^{-1})^*(\underline{A}^{-1})\epsilon + M^*(I-\underline{A}^{-1}A)^*(I-\underline{A}^{-1}A)M + 2\epsilon^*(\underline{A}^{-1})^*(I-\underline{A}^{-1}A)M \tag{13}$$

The last product vanishes and we obtain the minimum for $M = 0$.

Thus our estimate gives the minimum of the sum of variances from all estimated x-values because

$$E\{\epsilon_{\hat{x}}^* \epsilon_{\hat{x}}\} = \sigma^2 Sp\{(\underline{A}^{-1})^* \underline{A}^{-1}\} = Sp\sigma^2(\underline{A^*PA})^{-1} = \sum \sigma_{x_i}^2 = \min \tag{14}$$

For A^*PA non-singular we recognize the classical least squares solution of Gauss.

Equation (5) defines the complete set of errors which are consistent with the observation equations. This set of errors also includes biased estimates of the unknown parameters.

Equation (3:9) defines the complete set of unbiased estimates and we compute the corresponding set of consistent error estimates (consistent because they satisfy the equation (2:2)).

$$\epsilon = L - AX = L - AA^{-1}L = L - A^0L - AN(I - A^0)L \tag{15}$$

Finally the least squares solution makes use of the following estimates of the errors

$$\hat{\epsilon} = L - A(\utilde{A^*PA})^{-1}A^*PL = L - {}_{A}{}^{0}L \quad \hat{\epsilon}^*P\hat{\epsilon} = \min \tag{16}$$

This generalized least squares solution is rather useful for a number of applications. A unique solution is obtained for

$$\hat{\epsilon} = L - A(\underline{A^*PA})^{-1}A^*PL \quad \hat{\epsilon}^*P\hat{\epsilon} = \min \quad \sum s_{x_i}^2 = \min \tag{17}$$

5. Invariant standard deviation of coordinates

The error ellipses are often used when it is important to give a more detailed description of the quality of the coordinates in a network. Then the standard deviations of the coordinates are measured in the direction of the corresponding maxima and minima. This means that the solution is invariant with respect to rotations. These estimates are still dependent of the origin of the network which means that for a large system we get very different results when we put the two given points in different sections of the network. In our invariant approach we consider all coordinates unknown and then we obtain a singular system. For our solution we use the $\underline{A}^{-1}$ inverse and then we have an invariant expression for the sum of the variances.

$$\frac{1}{k}\sum_{i=1}^{k} s_{x_i}^2 + s_{y_i}^2 = \min \tag{1}$$

where

s_{x_i} = standard deviation of the ith x-coordinate
s_{y_i} = standard deviation of the ith y-coordinate
k = number of points

We have a singular system of observation equations

$$\underset{nm}{A}\ \underset{m1}{X} = \underset{n1}{L} - \underset{n1}{\epsilon} \quad r(A) < m \tag{2}$$

This system has a unique solution when using the $\underline{A}^{-1}$-inverse

$$\hat{X} = \underline{A}^{-1} L \tag{3}$$

This solution gives the error estimate

$$\hat{\epsilon}_{\hat{x}} = \underline{A}^{-1}\hat{\epsilon} \quad \text{with } \hat{\epsilon}^*\hat{\epsilon} = \min \tag{4}$$

When using an arbitrary inverse we obtain the new error estimate

$$\hat{\epsilon}_{\hat{x}} = \underline{A}^{-1}\hat{\epsilon} + (I - \underline{A}^{-1}A)M \tag{5}$$

with

$$\hat{\epsilon}^*_{\hat{x}}\hat{\epsilon}_{\hat{x}} = \hat{\epsilon}^*(\underline{A}^{-1})^*(\underline{A}^{-1})\hat{\epsilon} + M^*I - \underline{A}^{-1}A)^*(I - \underline{A}^{-1}A)M + 2\hat{\epsilon}^*(\underline{A}^{-1})^*(I - \underline{A}^{-1}A)M \tag{6}$$

where

$$\hat{\epsilon}^*(\underline{A}^{-1})^*(I - \underline{A}^{-1}A)M = 0 \tag{7}$$

Our sum of squares has a minimum for $M = 0$.

Thus we have proved

$$\hat{\epsilon}^*_{\hat{x}}\hat{\epsilon}_{\hat{x}} = \min \tag{8}$$

Furthermore we have for the $\underline{A}^{-1}$-inverse solution

$$E\{\epsilon^*_{\hat{x}}\epsilon_{\hat{x}}\} = Sp\sigma^2(\underline{A^*A})^{-1} = \sum_{i=1}^{m} \sigma^2_{x_i} + \sigma^2_{y_i} = \min \tag{9}$$

This means that the $\underline{A}^{-1}$-inverse solutions give minimum variance of the observations as well as minimum of the sum of the variances for the estimated unknowns. We already know that the sum of the variances of x and y is an invariant with respect to rotations. Our minimum of the sum of the coordinate variances is thus unique and invariant with respect to coordinate translations and rotations.

In an alternative approach we transcribe the $\underline{A}^{-1}$-inverse

$$\underline{A}^{-1} = (A^*A + D^*D)^{-1}A^* \tag{10}$$

where $A^*D = 0$ and $r(A^*, D^*) = m$.

When the covariance matrix of the observation is P^{-1} we use the inverse

$$\underline{A}^{-1} = (A^*PA + D^*D)^{-1}A^*P \tag{11}$$

And we obtain the general solution with

$$\hat{\epsilon}^*P\hat{\epsilon} = \min \tag{12}$$

and the estimated variance of the unknown

$$s^2Sp(\underline{A^*PA})^{-1} = \sum_{i=1}^{k} s^2_{x_i} + s^2_{y_i} = \min \tag{13}$$

It is of interest to note that in the classical presentation of the standard deviation of the coordinates rather confusing results are sometimes presented. Especially continental networks suffer considerably from a presentation where the whole network is linked to a couple of fixed points. With increasing distance from this origin we can expect large standard deviations for coordinates of points situated far away from the origin.

For the user of the coordinates these standard deviations can be very misleading. Using the present singular approach we get a standard deviation which is much more representative with respect to the surrounding points.

This invariant standard deviation of coordinates gives a very useful tool when comparing different networks. We can select some alternative definitions:

Optimum networks:

1. Network with

$$\frac{Sp\,\sigma^2(\underline{A^*PA})^{-1}}{k} = \min$$

2. Network with the smallest value of

$$(\sigma_{x_i}^2 + \sigma_{y_i}^2)_{\max}$$

REFERENCES

Baarda, W. 1949. Some remarks on the computation and adjustment of large systems of geodetic triangulation. Report to the I.U.G.G. at Rome, Netherlands Geodetic Commission.

Bjerhammar, A. 1948. Några synpunkter på matriskalkylens användning inom utjämningsräkningen. [Some considerations regarding the application of the calculus of matrices to the method of least squares.] *Svensk Lantmäteritidskrift.*

Bjerhammar, A. 1949. Metoder för förenkling av utjämningsräkningen enligt minsta kvadratprincipen. [Methods for simplification of adjustment by means of the method of least squares.] *Svensk Lantmäteritidskrift.*

Bjerhammar, A. 1951. Application of calculus of matrices to method of least squares with special references to geodetic calculations. *Trans. of the Royal Institute of Technology*, Stockholm, no. 49.

Bjerhammar, A. 1955. En ny matrisalgebra. *Svensk Lantmäteritidskrift, 5–6.*

Bjerhammar, A. 1958. A generalized matrix algebra. Stockholm.

Cramer, H. 1945. *Mathematical methods of statistics.* Uppsala.

Doolittle, M. H. 1878. On least squares solutions. U.S.G.G.S., Annual Report, App. 8, Paper 3, 115–20.

Fox. L., Huskey, H. D. & Wilkinson, J. H. 1948. Notes on the solution of algebraic linear simultaneous equations. *Mech. App. Maths. 1,* 149–73.

Greville, T. N. E. 1960. Some applications of the pseudoinverse of a matrix. *SIAM Review 2,* No. 1.

Grossmann, W. 1968. Grundzüge der Ausgleichungsrechnung, nach der Methode der kleinsten Quadrate nebst Anwendungen in der Geodäsie. Springer-Verlag, Berlin.

Helmert, F. R. 1872. Die Ausgleichungsrechnung nach der Methode der kleisten Quadrate. Leipzig.

Moore, E. H. 1935. General analysis. *Mem. Phil. Soc. 9.*

Moritz, H. 1967. Statistical Analysis of Gravity Anomalies. *Intern. Dict. of Geophys.*, p. 1436 ff., Oxford.

Oktaba, W. 1969. Generalized inverses. *Biom. Zeitschr.*, 11.

Penrose, R. A. 1955. A generalized inverse for matrices. *Proc. Camb. phil. Soc. 51.* 406–13.

Tienstra, J. M. 1947. An extension of the technique of the methods of least squares to correlated observations. *Bull. Geod.*, 301–5.

Wolf, H. 1961/68. Ausgleichung nach der Methode der kleinsten Quadrate. Hamburg–Bonn.

ОЦЕНКА С СИНГУЛЯРНЫМИ ОБРАЩЕНИЯМИ

В классической процедуре оценок по Гауссу используется обращение неособенных матриц. Ранее автор представил ряд статей о стохастических оценках с использованием сингулярных обращений (Бьерхаммер 1948–1958). В данной работе содержится обзор предыдущих исследований и дается их применение к избранным задачам. Для геодезической сети все точки могут рассматриваться неизвестными и оценка с обобщенным обращением дает минимум изменчивости наблюдаемых величин и неизвестных. Эта оценка инвариантна по отношению к вращениям и переносам и поэтому представляет специальный интерес для изучения оптимальной сети. Далее, она интересна, поскольку исключает зависимость от системы координат. Ряд подобных приложений может быть найден в большинстве разделов науки.