DELFT UNIVERSITY OF TECHNOLOGY
REPORTS OF THE DEPARTMENT OF GEODESY
MATHEMATICAL AND PHYSICAL GEODESY

84.1

GENERALIZED INVERSES, ADJUSTMENT THE DATUM PROBLEM AND S-TRANSFORMATIONS

Preprint

P.J.G. Teunissen

Lecture Notes
International School of Geodesy
3rd Course: Optimization and Design of Geodetic Networks
Erice-Trapani-Sicily: 25 April - 10 May 1984

TABLE OF CONTENTS

PREFACE

Many problems in physical science involve the estimation of a number of unknown parameters which bear a linear (or linearized) relationship to a set of experimental data. The data may be contaminated by (systematic or random) errors, insufficient to determine the unknowns, redundant, or all of the above and consequently, questions as existence, uniqueness, stability, approximation and the physical description of the set of solutions are all of interest.
In econometrics, for instance, the problem of insufficient data is discussed under the heading of "multicollinearity" and the consequent lack of determinability of the parameters from the observations, is known there as the "identification problem". In geophysics, where the physical interpretation of an anomalous gravitational field involves deduction of the mass distribution which produces the anomalous field, there is a fundamental non uniqueness in potential field inversion, such that, for instance, even complete, perfect data on the earth's surface cannot distinguish between two buried spherical density anomalies having the same anomalous mass but different radii.
Also in geodesy one is confronted with similar problems. In physical geodesy, for instance, the fact that the data are generally measured only at discrete points, leaves one with the problem of determining a continuous unknown function from a finite set of data. And in geometric geodesy the non uniqueness in coordinate system definitions, plays a fundamental role when identifying, interpreting, qualifying and comparing results from geodetic network adjustments.

All the above mentioned problems are very similar and even formally equivalent, if they are described in terms of a linear model $E\{y\}_{mx1} = A_{mxn}\, x_{nx1}$, with rank $A < n$. And these problems of solving systems of linear equations with arbitrary size and degeneracy are readily handled via the concept of a generalized inverse.
In chapter one of these lecture notes we will therefore present the basics of the theory of generalized inverses. Contrary, however, to the algebraic approach taken in the many textbooks available on generalized inverses, we we will approach the problem of inverting matrices of arbitrary order and rank rather geometrically and show, amongst other things, how one can characterize an arbitrary generalized inverse uniquely.
Also the relation between generalized inverses and systems of

linear equations will become clear then (cf. e.g. sections I.2 and I.4).
In fact, our geometric approach enabled us to obtain some results concerning the theory of generalized inverses. To facilitate reference we have summarized the basic results in section six.
In chapter two the problem of free network adjustments, which essentially is a problem of inverse mapping, is treated. In this chapter we will discuss the datum problem and derive a general expression for S-transformations, which allows one to transform from one datum to another. Also the relation with the theory of generalized inverses is shown.

NOTATIONS AND PRELIMINARIES

A subset V of a vectorspace W, $V \subset W$, is a subspace of W, if V - with the same definition of vector addition and scalar multiplication in W - is a vectorspace. V is called a proper subspace if $V \subset W$ and $V \neq W$ and $V \neq \{0\}$.
Let W be a vectorspace and U and V subspaces of W. The intersection of U and V, denoted by $U \cap V$, is the set of vectors which are common to both U and V:

$$U \cap V = \{y \in W \mid y \in U \wedge y \in V\}$$

The intersection of two subspaces is again a subspace.
The sum of U and V, denoted by U + V, is the set of vectors which can be expressed as the sum of an element of U and an element of V:

$$U + V = \{y \in W \mid y = u + v,\ u \in V,\ v \in V\}$$

Also the sum of two subspaces is again a subspace.
If U and V are finite dimensional (throughout the sequel we will only be concerned with finite dimensional vectorspaces) we have the dimensional relation:

$$\dim U + \dim V = \dim (U \cap V) + \dim (U + V)$$

If U and V are subspaces of the vectorspace W such that $W = U + V$ and $U \cap V = \{0\}$, then W is called the direct sum of U and V, which we denote by $W = U \oplus V$. U and V are then called complementary subspaces of W.

For a matrix A of order mxn the linear space spanned by the columns of A is called the column space or range space of A and denoted by R(A). The rowspace of A, defined analogously, can therefore be denoted by $R(A^t)$.
R^m denotes the vectorspace of all m-tuples with real coordinates. Since

$R(A)$ consists precisely of those vectors in R^m which can be written as Ax for some x in R^n we have:

$$R(A) = \{y \mid y = Ax \text{ for some } x \text{ in } R^n\}$$

The nullspace of A, denoted by $N(A)$, is the set of all vectors in R^n that are mapped into the nullvector in R^m under A, i.e.

$$N(A) = \{x \mid Ax = 0\}$$

The dimension of the subspace spanned by the columns of matrix A is called the column rank r_c of A and the dimension of the subspace spanned by the rows of A the rowrank r_r. Thus $\dim R(A) = r_c$ and $\dim R(A^t) = r_r$. Since $\dim N(A) = n-r_r$, it follows from the dimensional relation

$$\dim R^n = n = \dim R(A) + \dim N(A)$$

that rowrank equals columnrank, i.e. $r_c = r_r = \text{rank } A$.

If we speak of orthogonality we mean, unless stated otherwise, orthogonality with respect to the canonical innerproduct. Thus two vectors y_1 and y_2 in R^m are orthogonal if $y_1^t y_2 = 0$.
The orthogonal complement of $R(A)$ is denoted by $R(A)^\perp$. Thus $R^m = R(A) \oplus R(A)^\perp$ and $y_1^t y_2 = 0 \; \forall \; y_1 \in R(A)$, $y_2 \in R(A)^\perp$. $A^\perp$ denotes a matrix such that $R(A^\perp) = R(A)^\perp$. Unless it is otherwise clear from the context, the columns of $A^\perp$ are assumed to be linearly independent. A matrix A is called orthogonal if $A^tA = I_n$ and $AA^t = I_m$.

In view of the close relationship between projectors and generalized inverses we give here the definition of a projector and some elementary results.
Let the two subspaces U and V of R^m be complementary, i.e. $R^m = U \oplus V$.
Consider an arbitrary vector $y \in R^m = U \oplus V$ and express $y = y_1 + y_2$ such that $y_1 \in U$ and $y_2 \in V$, where y_1 and y_2 are unique.
The mapping $P: y \to y_1$ is called the projector on U and along V. I-P, with I the identity matrix, is then the projector on V and along U. (see figure 1).
Thus the subspace U can be identified with the range of P, $R(P)$, and the subspace V with the nullspace of P, $N(P)$.

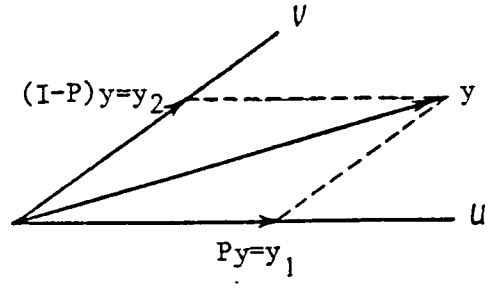


Figure 1

Now let the columns of the full rank matrix U span the subspace U. i.e. $R(U) = U$, and the columns of the full rank matrix V span the subspace V, i.e. $R(V) = V$. The projector which projects on U and along V is then given by

$$P_{U,V} = U\,[(V^{\perp})^{t}U]^{-1}(V^{\perp})^{t} = I - V\,[(U^{\perp})^{t}V]^{-1}(U^{\perp})^{t}$$

Note that the projector $P_{U,V}$ is independent of the matrix representations U and V for the subspaces U and V. The only conditions which need to be satisfied are $R(U) = U$ and $R(V) = V$.
A necessary and sufficient condition for matrix P to be a projector is that PP = P (idempotence) holds.

A vector y of order mx1 will usually denote in the sequel a data vector or vector of observables, and a vector x of order nx1 the parameter vector or vector of unknowns. The operator $E\{.\}$ denotes the mathematical expectation and the full rank matrices Q_y and $Q_{\hat{x}}$ the variance-covariance matrices of respectively the observables y and estimated unknowns $\hat{x}$.

I. GENERALIZED INVERSES, A GEOMETRIC APPROACH

I.1. Characterization of a set of linear equations

Many problems in physical science involve the estimation of a number of unknowns $\underset{nx1}{x}$, which bear a linear (or linearized) relationship to a set of experimental data $\underset{mx1}{y}$:

$$(1.1) \qquad \underset{mx1}{y} \doteq \underset{mxn}{A}\ \underset{nx1}{x}$$

The data may be contaminated by (random or systematic) errors, insufficient to determine the unknowns, redundant, or all of the above.
The first question that arises is whether a solution to (1.1) exists at all, i.e. whether the vector y can be written as a linear combination of the columns of matrix A. If this is the case we call the system *consistent*.
The system is certainly consistent if the rank of matrix A, rank A=r, equals the number of rows of A, i.e. r=m. In this case namely, the space spanned

by the columns of matrix A, R(A), equals R^m and therefore $y \in R^m = R(A)$. In all other cases, $r < m$, however, consistency is no longer guaranteed. To see this, observe that rank $A = r = \dim R(A)$. From $r < m$ then follows that $r = \dim R(A) < \dim R^m = m$, i.e. R(A) is a proper subspace of R^m, $R(A) \subset R^m$. It would thus be a mere coincidence if the vector $y \in R^m$ lies in the smaller dimensioned subspace $R(A) \subset R^m$. Consistency is thus guaranteed if $y \in R(A)$ or, equivalently, if y is orthogonal to the orthogonal complement of R(A), $R(A)^{\perp}$.
If we assume the subspace $R(A)^{\perp}$ to be spanned by the columns of a matrix of order mx(m-r), say U_1, consistency is guaranteed if

$$U_1^t y = 0, \quad \text{with} \quad R(U_1) = R(A)^{\perp}. \tag{1.2}$$

Assuming consistency, the next question one might ask is whether the solution to (1.1) is unique or not, i.e. whether the data gathered in the vector y are sufficient for determining the unknowns x. If not, the system is said to be *underdetermined.*
The solution is only unique if the rank of matrix A, rank A=r, equals the number of columns of A, i.e. r=n. To see this, assume x_1 and $x_2 \neq x_1$ to be two solutions to (1.1). Then $Ax_1 = Ax_2$ or $A(x_1 - x_2) = 0$ must hold. But this can only be the case if some of the columns of matrix A are linearly dependent, which contradicts our assumption of full column rank r=n. In all other cases, $r < n$, there will be more than one solution.

From the above considerations follows that it is the relation of r to m and n which decides the general character of a linear system. And a conceptually very nice theorem which makes this characterization precise, is the so-called Singular Value Decomposition theorem. It was first established by [Eckhart and Young, 1939 A Principal Axis Transformation for Non-Hermitian Matrices, Bull. Amer. Math. Soc., Vol. 45, pp. 118-121].

The Singular Value Decomposition theorem:

Let A be a real matrix of order mxn, with rank $A = r \leq \min(m,n)$. Then there exist orthogonal matrices $\underset{mxm}{U} = (\underset{mxr}{U_o} \vdots \underset{mx(m-r)}{U_1})$ and $\underset{nxn}{V} = (\underset{nxr}{V_o} \vdots \underset{nx(n-r)}{V_1})$ such that

$$A = (U_o \vdots U_1) \begin{pmatrix} \Lambda_r^{\frac{1}{2}} & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} V_o^t \\ V_1^t \end{pmatrix} = U_o \Lambda_r^{\frac{1}{2}} V_o^t, \tag{1.3}$$

with $\Lambda_r^{\frac{1}{2}} = \text{diag}\ (\sigma_1,\ldots,\sigma_r)\ ,\quad \sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_r > 0$

Proof:

Since A^tA is a real symmetric semi-positive definite matrix, it follows that its eigenvalues are non-negative. Denoting these eigenvalues by σ_i^2, i=1,...,n, we can arrange that $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_r > 0 = \sigma_{r+1} = \ldots = \sigma_n$. The corresponding orthogonal eigenvectors are denoted by $(v_1,\ldots,v_n)$ and we separate them into $V_o = (v_1,\ldots,v_r)$ and $V_1 = (v_{r+1},\ldots,v_n)$. With $\Lambda_r^{\frac{1}{2}} = \text{diag}\ (\sigma_1,..,\sigma_r)$ we thus have $A^tAV_o = V_o\Lambda_r$ or

$$\Lambda_r^{-\frac{1}{2}}V_o^tA^tAV_o\Lambda_r^{-\frac{1}{2}} = I_r\ , \tag{1.4}$$

and $A^tAV_1 = 0$ or,

$$AV_1 = 0 \tag{1.5}$$

The columns of the matrix

$$U_o = AV_o\Lambda_r^{-\frac{1}{2}} \tag{1.6}$$

are thus all of unit length and mutually orthogonal, i.e. $U_o^tU_o = I_r$. We can now choose a matrix U_1 such that the matrix $U = (U_o \vdots U_1)$ becomes orthogonal, i.e. $U^tU = I_m$ and $UU^t = I_m$. With (1.6) we therefore have that

$$U_o^tAV_o = \Lambda_r^{\frac{1}{2}} \quad \text{and} \quad U_1^tAV_o = 0 \tag{1.7}$$

The matrix

$$U^tAV = \begin{pmatrix} U_o^tAV_o & U_o^tAV_1 \\ U_1^tAV_o & U_1^tAV_1 \end{pmatrix},$$

therefore reduces with the aid of (1.5) and (1.7) to

$$U^tAV = \begin{pmatrix} \Lambda_r^{\frac{1}{2}} & 0 \\ 0 & 0 \end{pmatrix},$$

and premultiplication with U and postmultiplication with V^t finally gives the desired expression (1.3). Q.E.D.

The numbers σ_i, i=1, ..., n, are called the singular values of matrix A and they are the square roots of the non-negative eigenvalues of A^tA. The corresponding orthonormal eigenvectors v_i, i=1, ..., n, of A^tA are called the right singular or right eigenvectors of matrix A. The orthonormal eigenvectors of AA^t are given by the columns of U and they are called the left singular or left eigenvectors of matrix A.
From decomposition (1.3) follows that the columns of U_o form an orthonormal basis of the range space of matrix A and because of the orthogonality of U, the columns of U_1 constitute an orthonormal basis of the orthogonal complement of R(A). In the same way we see that the columns of V_o and V_1 form orthonormal bases of respectively $R(A^t)$ and $R(A^t)^\perp$. Summarizing, we have (see figure 2):

$$(1.8) \qquad \begin{cases} R(U_o) = R(A), \; R(U_1) = R(A)^\perp = N(A^t) \\ R(V_o) = R(A^t), \; R(V_1) = R(A^t)^\perp = N(A) \end{cases}$$

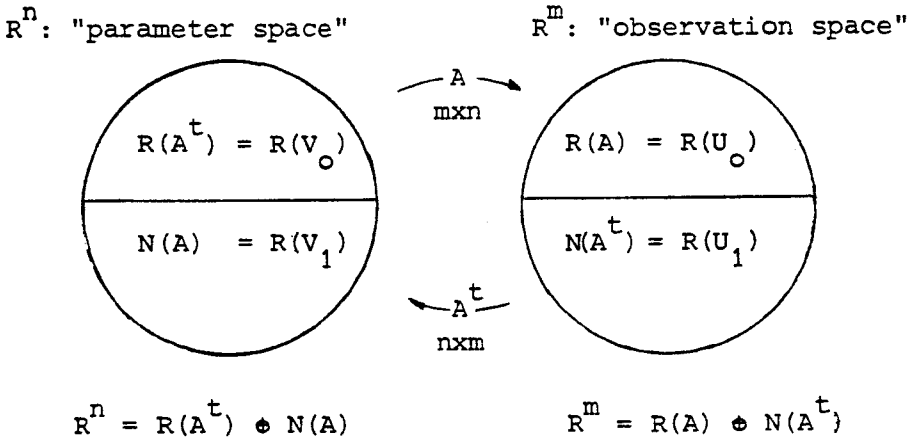


Figure 2

I.2. A unique characterization of an arbitrary generalized inverse

We know that for a square and regular matrix A a unique matrix inverse B exists, with the properties:

$$(2.1) \qquad AB = I \quad \text{and} \quad BA = I$$

For singular and rectangular matrices A of order mxn, however, in general no matrix B can be found for which (2.1) holds. For such matrices then a more relaxed inverse property is used, namely:

(2.2) $ABA = A$

This matrix equation follows from the idea that an inverse-like matrix B should solve the consistent set of equations $y = Ax$, with $y \in R(A)$. That is, matrix B should furnish a solution $x = By$ such that $y = ABy$ holds for all $y \in R(A)$, i.e. $ABA = A$.
Matrices B which satisfy (2.2) are called *generalized inverses of A*.
Expression (2.2) is, however, not a very illuminating one. In particular it does not tell us how generalized inverses look like or how they can be computed. (This is also a point of criticism we have against the many textbooks which deal with the theory of generalized inverses. In many of these textbooks, it seems that, algebraic manipulations and the stacking of theorems, lemma's and corollaries, and what have you, together with the sometimes superfluous naming of special types of generalized inverses, are preferred to a clear geometric exposition of what expression (2.2) actually tells us about the matrices B). In the following we will therefore rewrite expression (2.2) in such a form that it becomes relatively easy to understand the mapping characteristics of generalized inverses.
Fundamental is the following theorem:

Theorem 2.1

(2.3) $ABA = A \Leftrightarrow BAx=x, \quad \forall x \in R(S)$,
where the subspace $R(S) \subset R^n$ is arbitrary,
provided that $\dim R(S) = \dim R(A)$ and $R(S) \cap N(A) = \{0\}$,
i.e. $R^n = R(S) \oplus N(A)$.

(2.4) $ABA = A \Leftrightarrow ABy = y, \forall y \in R(A)$

Proof of (2.3)

($\Rightarrow$) From $ABA=A$ follows $BABA=BA$ meaning that BA is idempotent and thus a projector.
From $ABA=A$ also follows that $N(BA) = N(A)$.
To see this, consider $x \in N(BA)$. Then $BAx=0$ or $ABAx = Ax = 0$, which means that $x \in N(A)$. Thus $N(BA) \subset N(A)$. If on the other hand $x \in N(A)$, then $Ax = 0$ or $BAx = 0$, which means that $x \in N(BA)$. Thus $N(A) \subset N(BA)$.
Since BA is a projector with $N(BA) = N(A)$ we can write $BA = P_{R(BA),N(A)}$.
Thus $R(BA)$ is complementary to $N(A)$. By denoting the subspace

R(BA) by R(S) we can write BAx=x, $\forall x \in R(S)$. And the complementarity of R(S) and N(A) can be expressed as dim R(S) = dim R(A) and $R(S) \cap N(A) = \{0\}$.

($\Leftarrow$) $R^n = R(S) \oplus N(A)$ or dim R(S) = dim R(A) and $R(S) \cap N(A) = \{0\}$ means that R(S) is complementary to N(A). We can therefore construct the projector $P_{R(S),N(A)} = I - P_{N(A),R(S)}$. Thus we can now replace $BAx = x$, $\forall x \in R(S)$, by $BAP_{R(S),N(A)}\, x = P_{R(S),N(A)}\, x$, $\forall x \in R^n$. With $AP_{R(S),N(A)} = A(I-P_{N(A),R(S)}) = A$ we get $BAx = P_{R(S),N(A)}x$, $\forall x \in R^n$ or finally $ABAx = Ax$, $\forall x \in R^n$.

Proof of (2.4)

The proof is omitted since it is trivial.

So what does theorem 2.1 tell us about generalized inverses of A? First of all it tells us, and this is essential, that the only condition a generalized inverse B of A has to fulfil is that it maps the subspace R(A) onto a subspace R(S) complementary to N(A). Or stated differently: Every matrix B which satisfies

(2.5) $\quad BAS = S,$

with R(S) complementary to N(A), is a generalized inverse of A. Thus every generalized inverse of A determines a one-to-one relation between the subspace R(AS) = R(A) and R(S) (see figure 3).

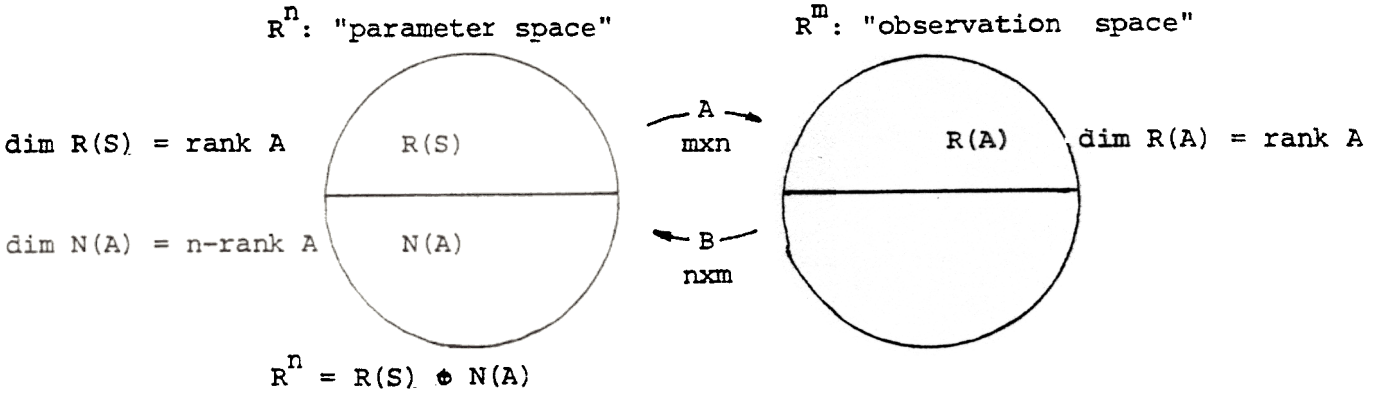


Figure 3

Secondly, theorem 2.1 gives us a neat way of uniquely characterizing any generalized inverse of A. To see this, consider expression (2.5). Expression (2.5) shows how matrix B maps a basis of the subspace R(A), namely AS, onto the subspace R(S). But since AS is only a basis of the subspace R(A) and not a basis of the total domain space R^m of B, expression (2.5) is not sufficient for determining B uniquely. Thus in order to compute a particular generalized inverse B of A we in addition to (2.5) have to say how B maps a basis of a subspace complementary to R(A). If we denote a subspace complementary to R(A) by $R(C^{\perp})$, we thus need to specify how B maps $R(C^{\perp})$; say:

$$BC^{\perp} = D \tag{2.6}$$

Since AS is a basis of R(A) and $C^{\perp}$ a basis of a subspace complementary to R(A) we have $R(AS:C^{\perp}) = R^m$. Thus if we take expressions (2.5) and (2.6) together,

$$B(AS:C^{\perp}) = (S:D), \tag{2.7}$$

we can uniquely determine B as

$$B = (S:D)(AS:C^{\perp})^{-1} \tag{2.8}$$

With $(AS:C^{\perp})^{-1} = \begin{pmatrix} (C^tAS)^{-1}C^t \\ [((AS)^{\perp})^tC^{\perp}]^{-1}((AS)^{\perp})^t \end{pmatrix}$, expression (2.8) then becomes

$$B = S(C^tAS)^{-1}C^t + D[((AS)^{\perp})^tC^{\perp}]^{-1}((AS)^{\perp})^t \tag{2.9}$$

So far we silently ignored our choice for D in (2.6). In principal, of course, the choice for matrix D is completely free. But, as we will see, one cam impose an extra condition, namely that $R(D) \subset R(V_1) = N(A)$, without affecting the generality of expression (2.9).
Assume that we have chosen matrices $\bar{D}$ and $\bar{C}^{\perp}$, with $R(\bar{C}^{\perp})$ complementary to R(A), such that

$$\bar{D} = B\bar{C}^{\perp} \tag{2.10}$$

With the projector identity $P_{R(S),N(A)} = I - P_{N(A),R(S)}$ we can then write (2.10) as

$$(P_{R(S),N(A)} + P_{N(A),R(S)})\,\bar{D} = B\bar{C}^{\perp}, \text{ or as}$$

$$(2.11) \qquad P_{N(A),R(S)}\bar{D} = B\bar{C}^{\perp} - P_{R(S),N(A)}\bar{D}$$

And with $BA = P_{R(S),N(A)}$, expression (2.11) becomes

$$(2.12) \qquad P_{N(A),R(S)}\bar{D} = B(\bar{C}^{\perp} - A\bar{D})$$

But this expression shows that if we put $D = P_{N(A),R(S)}\bar{D}$ and $C^{\perp} = \bar{C}^{\perp} - A\bar{D}$ we are back at (2.6), but now with the extra condition that $R(D) \subset N(A)$.

As a conclusion we thus have that any generalized inverse B of A is of the form

$$(2.13) \qquad \boxed{B = (S:D)(AS:C^{\perp})^{-1} = S(C^{t}AS)^{-1}C^{t} + D[((AS)^{\perp})^{t}C^{\perp}]^{-1}((AS)^{\perp})^{t}}$$

with R(S) complementary to N(A), $R(C^{\perp})$ complementary to R(A) and $R(D) \subset N(A)$. (see figure 4).

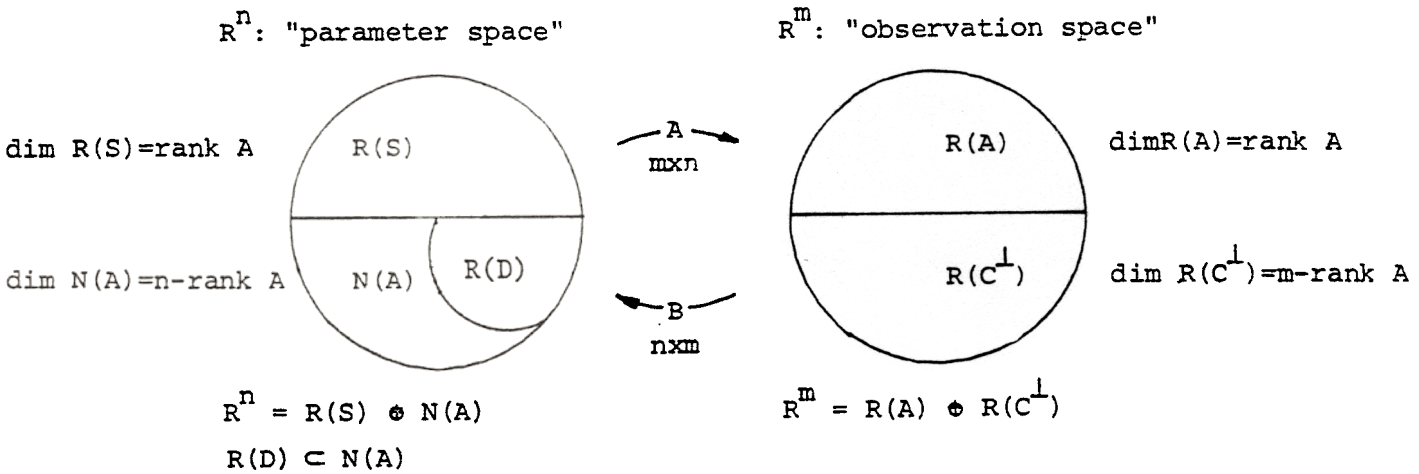


Figure 4

I.3. Right- and left inverses

From theorem 2.1 of the previous section we learned that, for any generalized inverse B of matrix A of order mxn, the matrices BA and AB behave like identity-matrices on respectively the subspaces R(S) and R(A). Thus in the special case that rank A = r = n the generalized inverses of A become left inverses, since then $BA = I_{n=r}$. Similarly, the generalized inverses of A become right inverses if rank A = r = m, because then $AB = I_{m=r}$ holds.
Let us now first, in order to give an interpretation of the subspace R(S), concentrate on the special case rank A = r = m.
If rank A = r = m then $R(A) = R^m$ and therefore the subspaces complementary to R(A) reduce to $R(C^{\perp}) = \{0\}$. With (2.6) we then also have $R(D) = \{0\}$. (see figure 5).

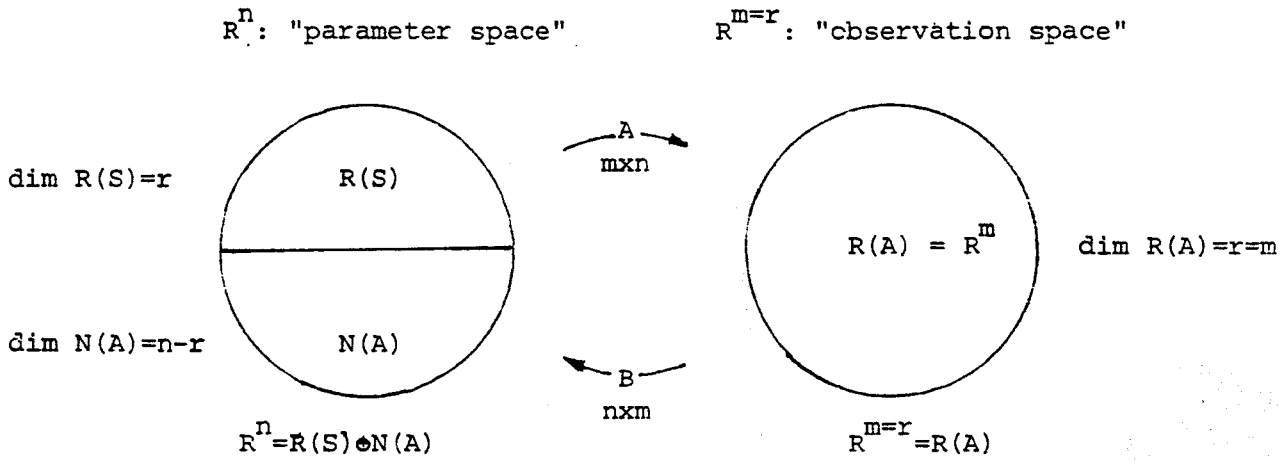


Figure 5

From (2.13) follows, with $R(C^{\perp}) = \{0\}$ or $R(C) = R^m$ and $R(D) = \{0\}$, that the general expression of right inverses is given by

(3.1) $B = S(AS)^{-1}$, with R(S) complementary to N(A).

Consider now a system of linear equations

(3.2) $\underset{mx1}{y} = \underset{mxn}{A}\ \underset{nx1}{x}$, with rank A = r = m.

This system is clearly consistent for all $y \in R^m$. With a particular generalized inverse (right inverse), say B, of A, we can write the solution to (3.2) as

(3.3) $\{x\} = \{x \mid x = By + V_1\alpha \ , \ \forall\alpha \in R^{n-r}\}, \ \text{with } R(V_1) = N(A).$

And by choosing $\alpha=\alpha_1$ we get as a particular solution $x_1 \in \{x\}$:

(3.4) $x_1 = By + V_1\alpha_1 \ ,$

where α_1, so to say, contributes the extra information, which is lacking in y, to determine x_1.
Since R(B) = R(S) (see (3.1)) it follows from (3.4) that

(3.5) $\underset{(n-r)xn \ nx1}{(S^{\perp})^t x_1} = \underset{(n-r)x(n-r)}{[(S^{\perp})^t V_1]} \underset{(n-r)x1}{\alpha_1} \overset{\text{call}}{=} \underset{(n-r)x1}{c_1}$

But since α_1 or c_1 contributes the extra information, which is lacking in y, to determine x_1, the solution of the uniquely solvable system

(3.6) $\underset{(m+n-r)x1}{\begin{pmatrix} y \\ c_1 \end{pmatrix}} = \underset{(m+n-r)xn}{\begin{pmatrix} A \\ (S^{\perp})^t \end{pmatrix}} \underset{nx1}{x}, \ \text{with rank } A = r = m \ \text{ and } R^n = R(S) \oplus N(A),$

is precisely x_1:

(3.7) $x_1 = \begin{pmatrix} A \\ (S^{\perp})^t \end{pmatrix}^{-1} \begin{pmatrix} y \\ c_1 \end{pmatrix} = \left(S(AS)^{-1} \vdots V_1[(S^{\perp})^t V_1]^{-1} \right) \begin{pmatrix} y \\ c_1 \end{pmatrix} .$

Thus in order to find a particular solution to (3.2), say x_1, we merely need to extend the system of linear equations (3.2) to (3.6) by introducing the additional equations $c_1 = (S^{\perp})^t x$, so that the extended matrix

$$\begin{pmatrix} A \\ (S^{\perp})^t \end{pmatrix}$$

becomes square and regular. And this will be the case if R(S) is complementary to N(A); i.e. if $R^n = R(S) \oplus N(A)$. Furthermore all right inverses (3.1) of A are obtainable from

$$\begin{pmatrix} A \\ (S^{\perp})^t \end{pmatrix}^{-1} = \left(S(AS)^{-1} \vdots V_1[(S^{\perp})^t V_1]^{-1} \right), \ \text{with } R^n = R(S) \oplus N(A).$$

Let us now consider the case rank A=r=n. In this case all generalized inverses of A become left inverses. If matrix A is of full column rank the null space of A reduces to $N(A) = \{0\}$, meaning that $R(D) = \{0\}$ and $R(S) = R^n$. (see figure 6).

$R^{n=r}$: "parameter space" R^m: "observation space"

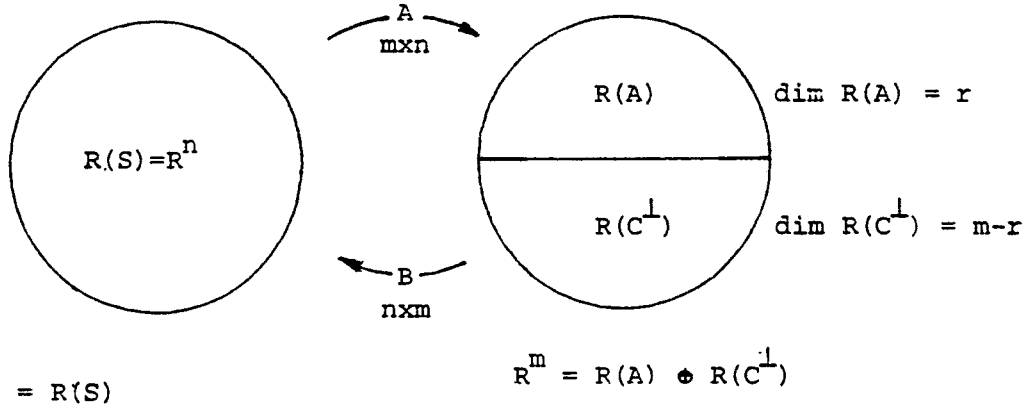


Figure 6

With $R(S) = R^n$ and $R(D) = \{0\}$ it follows from (2.13) that the general expression of left inverses is given by

(3.8) $B = (C^tA)^{-1}C^t$, with $R(C^\perp)$ complementary to $R(A)$.

The with the full column rank matrix A associated system of linear equations reads

(3.9) $y \doteq \underset{mx1}{A} \quad \underset{mxn}{x} \quad \underset{nx1}{}$, with rank A = r = n

And this system is inconsistent in general, i.e. $y \notin R(A) = R(U_o)$ or with $R_1^t y \neq 0$ with $R(U_1) = R(A)^\perp$. One way to remove the inconsistency is by subtracting from y its components which lie in the subspace $R(U_1)$: $y - U_1U_1^ty = U_oU_o^ty$.
(see figure 7)

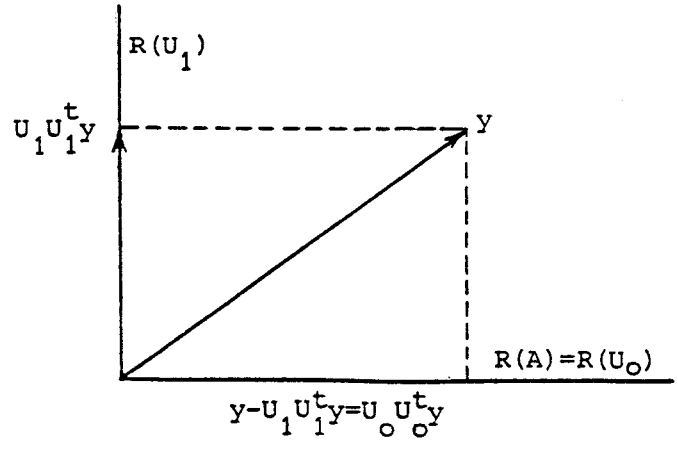


Figure 7

But one will admit that this is a rather arbitrary way of restoring consistency. In general one can therefore say that consistency is restored by subtracting from y its components which lie in a still freely chooseable subspace, $R(C^{\perp})$, complementary to R(A): $y - C^{\perp}(U_1^t C^{\perp})^{-1} U_1^t y = U_o (C^t U_o)^{-1} C^t y$ (see figure 8)

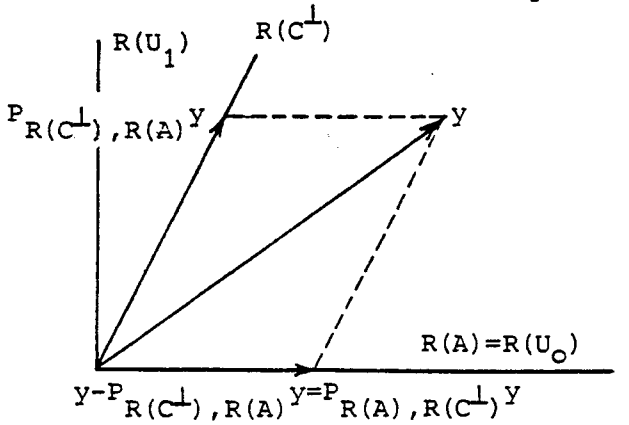


Figure 8

Thus the inconsistent system (3.9) is then replaced by the consistent system

$$(3.10) \qquad y - C^{\perp}(U_1^t C^{\perp})^{-1} U_1^t y = \underset{mxn}{A}\ \underset{nx1}{x}\ , \quad \text{with rank } A=r=n \text{ and } R^m = R(C^{\perp}) \oplus R(A).$$

But this system is identical to

$$(3.11) \qquad \underset{mx1}{y} = \underset{mxn}{(\ A} \vdots \underset{mx(m-r)}{C^{\perp}\ )} \quad \underset{(n+m-r)x1}{\begin{pmatrix} x \\ \lambda \end{pmatrix}}$$

To see this, pre-multiply (3.11) with U_1^t. We then get $U_1^t y = (U_1^t C^{\perp})\lambda$ or $\lambda = (U_1^t C^{\perp})^{-1} U_1^t y$. Thus the solution for x of (3.10) and (3.11) are identical.

Because of the full column rank of A in (3.11) and the complementarity of $R(C^{\perp})$ with R(A) it follows that the matrix $(A:C^{\perp})$ is square and regular. Thus the solution of (3.11) is therefore:

$$(3.12) \qquad \begin{pmatrix} x \\ \lambda \end{pmatrix} = (A:C^{\perp})^{-1} y = \begin{pmatrix} (C^t A)^{-1} C^t \\ (U_1^t C^{\perp})^{-1} U_1^t \end{pmatrix} y$$

Thus in order to make the in general inconsistent system (3.9) consistent, we merely need to extend the system of linear equations (3.9) to (3.11) by introducing additional unknowns, so that the extended matrix $(A:C^{\perp})$ becomes square and regular. And this will be the case if $R(C^{\perp})$ is complementary to R(A), i.e. if $R^m = R(A) \oplus R(C^{\perp})$. Furthermore all left inverses (3.8) of A are obtainable from

$$(A:C^{\perp})^{-1} = \begin{pmatrix} (C^tA)^{-1}C^t \\ (U_1^tC^{\perp})^{-1}U_1^t \end{pmatrix}$$

I.4. An arbitrary system of linear equations and arbitrary generalized inverses

In the previous section we saw how a particular solution of an underdetermined but otherwise consistent system of linear equations could be obtained by extending the matrix A rowwise. Especially the principal role played by the subspace R(S) complementary to N(A) in removing the underdeterminancy was demonstrated.
Similarly we saw how consistency of an inconsistent, but otherwise uniquely determined, system of linear equations was restored by extending the matrix A columnwise. And here the subspace $R(C^{\perp})$ complementary to R(A) played the decisive role.
It is therefore natural to try to apply a similar approach to an arbitrary system of linear equations which is possibly inconsistent and underdetermined at the same time. Let us assume such a system to be given by

$$(4.1) \qquad \underset{m\times 1}{y} \doteq \underset{m\times n}{A}\ \underset{n\times 1}{x}\ , \quad \text{rank } A = r \leq \min(m,n)$$

From the possible rank deficiency of matrix A in (4.1) follows that the unknowns x cannot be determined uniquely, even if $y \in R(A)$. Thus the information contained in y is not sufficient to determine x uniquely. One can overcome this problem by adding the minimum information needed to determine x uniquely. Thus by replacing (4.1) by

$$(4.2) \qquad \underset{(m+n-r)\times 1}{\begin{pmatrix} y \\ c \end{pmatrix}} \doteq \underset{(m+n-r)\times n}{\begin{pmatrix} A \\ (S^{\perp})^t \end{pmatrix}}\ \underset{n\times 1}{x}\ , \quad \text{with } R^n = R(S) \oplus N(A).$$

But although the extended matrix of (4.2) has full column rank, the system can still be inconsistent. Now we know from the previous section that inconsistency is removed by extending the matrix of (4.2) columnwise, so that the resulting matrix becomes square and regular. Since $R^m = R(A) \oplus R(C^{\perp})$, the range space of

$$\underset{(m+n-r)\times(m-r)}{\begin{pmatrix} C^{\perp} \\ 0 \end{pmatrix}}$$

is complementary to $R\begin{pmatrix} A \\ (S^{\perp})^t \end{pmatrix}$. But also $R\begin{pmatrix} C^{\perp} \\ X \end{pmatrix}$, with X an arbitrary matrix

of order $(n-r)x(m-r)$, is complementary to $R\left(\begin{pmatrix} A \\ (S^{\perp})^t \end{pmatrix}\right)$. Inconsistency is thus removed if we replace (6.2) by the uniquely solvable system:

$$(4.3)\qquad \underset{(m+n-r)x1}{\begin{pmatrix} y \\ c \end{pmatrix}} = \underset{(m+n-r)x(m+n-r)}{\begin{pmatrix} A & C^{\perp} \\ (S^{\perp})^t & X \end{pmatrix}} \underset{(m+n-r)x1}{\begin{pmatrix} x \\ \lambda \end{pmatrix}}, \quad \text{with } R^n = R(S) \oplus N(A) \text{ and } R^m = R(A) \oplus R(C^{\perp})$$

And the solution of (4.3) is given by

$$(4.4)\qquad \begin{pmatrix} x \\ \lambda \end{pmatrix} = \begin{pmatrix} A & C^{\perp} \\ (S^{\perp})^t & X \end{pmatrix}^{-1} \begin{pmatrix} y \\ c \end{pmatrix} =$$

$$= \left(\begin{array}{c|c} S(C^tAS)^{-1}C^t - V_1[(S^{\perp})^tV_1]^{-1}X[U_1^tC^{\perp}]^{-1}U_1^t & V_1[(S^{\perp})^tV_1]^{-1} \\ \hline [U_1^tC^{\perp}]^{-1}U_1^t & 0 \end{array}\right)\begin{pmatrix} y \\ c \end{pmatrix},$$

with $R(U_1) = R(A)^{\perp}$, $R(V_1) = N(A)$,
in which we recognize, if we put $-V_1[(S^{\perp})^tV_1]^{-1}X = D$ or $X = -(S^{\perp})^tD$, our general expression

$$(2.13)\qquad B = (S:D)(AS:C^{\perp})^{-1} = S(C^tAS)^{-1}C^t + D[U_1^tC^{\perp}]^{-1}U_1^t$$

of an arbitrary generalized inverse.

I.5. Transformation properties and some special types of generalized inverses

With the aid of the singular value decomposition $A = U_o\Lambda_r^{\frac{1}{2}}U_o^t$ (see (1.3)) we can write for our general expression (2.13) of generalized inverses

$$(5.1)\qquad B = (S:D)(AS:C^{\perp})^{-1} = S(C^tU_o\Lambda_r^{\frac{1}{2}}V_o^tS)^{-1}C^t + D[U_1^tC^{\perp}]^{-1}U_1^t \ , \quad \text{or}$$

$$(5.2)\qquad B = [S(V_o^tS)^{-1}V_o^t][V_o\Lambda_r^{-\frac{1}{2}}U_o^t][U_o(C^tU_o)^{-1}C^t] + D[U_1^tC^{\perp}]^{-1}U_1^t \ ,$$

In this last expression we recognize the projectors $P_{R(S),N(A)} = S(V_o^tS)^{-1}V_o^t$, projecting onto $R(S)$ and along $N(A)$, and $P_{R(A),R(C^{\perp})} = U_o(C^tU_o)^{-1}C^t$, projecting on $R(A)$ and along $R(C^{\perp})$. We can therefore obtain *any* particular generalized inverse $\bar{B}$, uniquely characterized by, say $R(\bar{S})$, $R(\bar{C}^{\perp})$ and $\bar{D}$, from arbitrary generalized inverses B, by applying the transformation rule:

$$(5.3)\qquad \bar{B} = [\bar{S}(V_o^t\bar{S})^{-1}V_o^t][B][U_o^t(\bar{C}^tU_o)^{-1}\bar{C}^t] + \bar{D}[U_1^t\bar{C}^{\perp}]^{-1}U_1^t$$

Let uw now consider some special types of generalized inverses and see what role is played by the subspaces $R(S)$, $R(C^{\perp})$ and $R(D)$.

- *Least squares generalized inverses* -

Assume given the inconsistent system of linear equations

$$(5.4) \qquad \underset{mx1}{y} \doteq \underset{mxn}{A}\ \underset{nx1}{x} \qquad \text{rank } A = r < m$$

From the least squares criterium $\min_x [y-Ax]^t Q_y^{-1} [y-Ax]$, with Q_y^{-1} the weight matrix of the datavector y, follows

$$(5.5) \qquad A^t Q_y^{-1} (y-Ax) = 0$$

and with (4.3) this gives $A^t Q_y^{-1} C^{\perp} \lambda = 0 \ \forall \ \lambda \in R^{m-r}$ or $R(C^{\perp}) = R(Q_y U_1)$.
The corresponding least squares generalized inverses are therefore obtained by choosing $R(C^{\perp}) = R(Q_y U_1)$, while $R(S)$ and $R(D)$ may still be chosen arbitrarily.

- *Minimum norm generalized inverses* -

Consider the consistent system of linear equations

$$(5.6) \qquad \underset{mx1}{y} = \underset{mxn}{A}\ \underset{nx1}{x} \qquad \text{rank } A = r < n$$

The set of solutions to (5.6) is given by

$$(5.7) \qquad \{x\} = \{x \mid x = By + V_1\alpha, \ \forall \alpha \in R^{n-r}\} \ , \ R(V_1) = N(A) \ .$$

with B an arbitrary generalized inverse of A.

We now want to find a solution such that

$$(5.8) \qquad \min_{\alpha} [By+V_1\alpha]^t Q_x^{-1} [B_y+V_1\alpha]$$

From this condition follows $\alpha = - [V_1^t Q_x V_1]^{-1} V_1^t Q_x By$ and the unique minimum norm solution is thus given by

(5.9) $\quad x = (I-V_1[V_1^t Q_x V_1]^{-1} V_1^t Q_x)\, By = P_{R(Q_x^{-1}V_o),N(A)}\, By$

Since $y \in R(A)$, it follows from an expression like (5.2) that By is independent of the choices made for $R(C^{\perp})$ and R(D). It therefore follows from (5.9) that the minimum norm generalized inverses are obtained by choosing $R(S) = R(Q_x^{-1}V_o)$, while leaving $R(C^{\perp})$ and R(D) open for arbitrary choices.

- *Maximum rank generalized inverses* -

From expression (5.1), $B = (S:D)(AS:C^{\perp})^{-1}$, follows that R(B) = R(S:D). Since dim R(S) = dim R(A) = rank A, we see that dim R(B) ≥ dim R(A) or rankB ≥ rank A. Thus the rank of any generalized inverse of A is greater than or equal to the rank of A. Furthermore R(B) = R(S:D) shows that the rank of B is completely determined by the choice for D and B will have maximum rank, rank B = min(m,n), if one can choose D such that dim R(D) = min(m,n)-r.

- *Reflexive generalized inverses (minimum rank inverses)* -

Generalized inverses B of A which have minimum rank, i.e. rank B = rank A, are called reflexive generalized inverses. And from our interpretation of R(D) we know that the minimum rank property can only be obtained by choosing R(D) = {0}. Taking R(D) = {0} in (2.13) and (5.3), the general expression of reflexive generalized inverses becomes

(5.10) $\quad B = S(C^tAS)^{-1}C^t = [S(V_o^tS)^{-1}V_o^t]\,\bar{B}\,U_o[(C^tU_o^{-1})C^t]$,

with $R^n = R(S) \oplus N(A)$, $R^m = R(A) \oplus R(C^{\perp})$ and where $\bar{B}$ can be any arbitrary generalized inverse of A.
The reflexive or reciprocal character of these generalized inverses follows from the fact that the matrices A and B of (5.10) are generalized inverses of each other, i.e. ABA = A and BAB = B holds.

- *The minimum norm least squares generalized inverse (Pseudo-inverce)* -

From the least squares criterium follows that we must choose $R(C^{\perp}) = R(Q_y U_1)$. And the minimum norm condition gives us an equation like (5.9):

(5.11) $\quad x = P_{R(Q_x^{-1}V_o),N(A)} \cdot By$, with B being a least squares inverse of A.

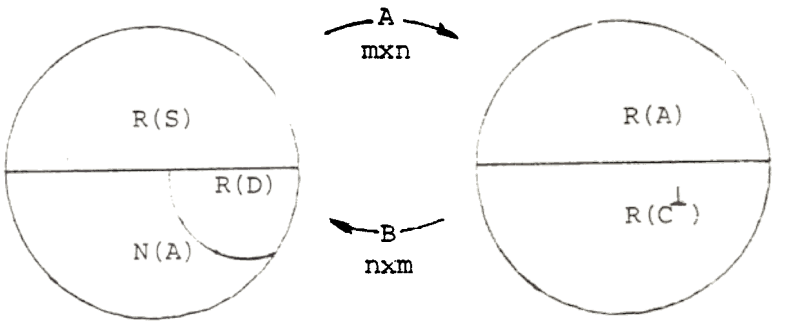
Contrary to (5.9), however, we now may have a data vector y for which $y \notin R(A)$. It therefore follows from (5.11) that the unique minimum norm least squares

generalized inverse is found by choosing $R(C^{\perp}) = R(Q_y U_1)$, $R(S) = R(Q_x^{-1} V_o)$ and $R(D) = \{0\}$. In the special case that $Q_y = I_m$ and $Q_x = I_n$ this inverse is known as the Moore-Penrose or Pseudo-inverse.

I.6. Summary

(6.1) $$\boxed{ABA = A \;\Leftrightarrow\; B = S(C^t AS)^{-1} C^t + D\left[((AS)^{\perp})^t C^{\perp}\right]^{-1} ((AS)^{\perp})^t}$$

R^n: "parameter space" R^m: "observation space"



$R^n = R(S) \oplus N(A)$ $\qquad$ $R^m = R(A) \oplus R(C^{\perp})$

$R(D) \subset N(A)$

A $m \times n$ rank $A = r$	INVERSES	INVERSELIKE PROPERTIES	SOLUTIONS OF LIN. SYST. OF EQUATIONS	
$m = n = r$	Caley inverse: $B = A^{-1}$	$AB = I, \quad BA = I$	$y = Ax$	$x = A^{-1}y = By$
$m = r$	Right inverse: $B = S(AS)^{-1}$	$AB = I_m, \quad BA = P_{R(S),N(A)}$	$\begin{pmatrix} y \\ c \end{pmatrix} = \begin{pmatrix} A \\ (S^{\perp})^t \end{pmatrix} x$	$x = \begin{pmatrix} A \\ (S^{\perp})^t \end{pmatrix}^{-1} \begin{pmatrix} y \\ c \end{pmatrix} = \left(B \vdots V_1 [(S^{\perp})^t V_1]^{-1} \right) \begin{pmatrix} y \\ c \end{pmatrix}$
$n = r$	Left inverse: $B = (C^t A)^{-1} C^t$	$AB = P_{R(A),R(C^{\perp})}, \quad BA = I_n$	$y = (A \vdots C^{\perp}) \begin{pmatrix} x \\ \lambda \end{pmatrix}$	$\begin{pmatrix} x \\ \lambda \end{pmatrix} = (A \vdots C^{\perp})^{-1} y = \begin{pmatrix} B \\ [U_1^t C^{\perp}]^{-1} U_1^t \end{pmatrix} y$
$r \leq \min(m,n)$	Generalized inverse: $B = S(C^t AS)^{-1} C^t + D[((AS)^{\perp})^t C]^{-1} ((AS)^{\perp})^t$	$AB = P_{R(A),R(C^{\perp})}, \quad BA = P_{R(S),N(A)}$	$\begin{pmatrix} y \\ c \end{pmatrix} = \left(\begin{array}{c:c} A & C^{\perp} \\ \hdashline (S^{\perp})^t & -(S^{\perp})^t D \end{array}\right) \begin{pmatrix} x \\ \lambda \end{pmatrix}$	$\begin{pmatrix} x \\ \lambda \end{pmatrix} = \left(\begin{array}{c:c} A & C^{\perp} \\ \hdashline (S^{\perp})^t & -(S^{\perp})^t D \end{array}\right)^{-1} \begin{pmatrix} y \\ c \end{pmatrix} = \left(\begin{array}{c:c} B & V_1[(S^{\perp})^t V_1]^{-1} \\ \hdashline [U_1^t C^{\perp}]^{-1} U_1^t & 0 \end{array}\right) \begin{pmatrix} y \\ c \end{pmatrix}$

Least squares inverses: $R(C^{\perp}) = R(Q_y A^{\perp})$

Minimum norm inverses: $R(S) = R(Q_x^{-1} A^t)$

Reflexive inverses: $R(D) = \{0\}$

Minimum norm least squares inverse: $R(C^{\perp}) = R(Q_y A^{\perp})$; $R(S) = R(Q_x^{-1} A^t)$; $R(D) = \{0\}$

II. ON S-TRANSFORMATIONS

"Where he looked for a flower, he discovers a whole garden. But only one flower in the garden gives him immortality, and nobody knows which it is. So what to do? There is only one thing - go back and describe the garden"

[P.C. Sabatier, 1979]*

II.1. Introduction

In chapter one we have seen how to characterize an arbitrary generalized inverse of matrix A_{mxn} uniquely. In particular the principal role played by the subspaces $R(S)$, complementary to $N(A)$, and $R(C^{\perp})$, complementary to $R(A)$, was demonstrated. The choice for $R(C^{\perp})$ determines the way in which the inconsistent system of linear equations $y \doteq Ax$ is made consistent and by choosing $R(S)$, with $x \in R(S)$, one gets around the difficulty of underdeterminancy. And, as we know, consistency is guaranteed if $y \in R(A)$ or, equivalently, if y is orthogonal to the orthogonal complement of $R(A)$, $R(A)^{\perp}$. Or in terms of linear equations: if $y = Ax$ for some x, or if $U_1^t y = 0$, with $R(U_1) = R(A)^{\perp}$.

In the frame-work of adjustment theory these two types of linear equations correspond to the so-called second standard problem - or Gauss-Markov model formulation, $E\{\underline{y}\} = Ax$, and the first standard problem - or adjustment-by-conditions formulation, $U_1^t E\{\underline{y}\} = 0$ (The underscore indicates the stochastic character of the observables). And since consistency is merely related to the choice for the subspace $R(C^{\perp})$ of the observation space R^m, and thus has nothing to do with the parameter space R^n and the subspace $R(S)$, the least squares estimators of $E\{\underline{y}\}$ will be identical for both model formulations $E\{\underline{y}\} = Ax$ and $U_1^t E\{\underline{y}\} = 0$. With a Gauss-Markov formulation, however, one aims at more than only adjustment. One then also aims at finding an estimate of x which corresponds to the least squares estimate of $E\{\underline{y}\}$. And here is where the problem of underdeterminancy, with the many possible choices for $R(S)$, might manifest itself, i.e. where the flower might turn out to be a garden.

In the theory of geodetic networks the problem of underdeterminancy is encountered as a consequence of the fundamental non-uniqueness in the relation

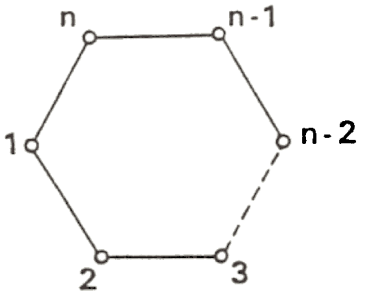
* P.C. Sabatier; Geophys. J.R. astr. Soc. (1979) 58, 523-524.

between geodetic observables, like height differences, angles, distances etc., and coordinates. In a levelling network, for instance, absolute heights cannot be computed if only height differences are measured. That is, for computing heights, one additional needs information on the absolute vertical datum as a *conditio sine qua non*. Similarly, one cannot obtain position, orientation and scale of a triangulation network if only angles are measured. Such networks for which the observational data are insufficient to determine either the (horizontal and/or vertical) position, orientation or scale, are called *free networks*.*

As a consequence of the underdeterminancy, the design matrix A of the linear(ized) Gauss-Markov model $E\{\underline{y}\} = Ax$ will have a rank deficiency. Therefore no unbiased linear estimator $\underline{\hat{x}} = B\underline{y}$ of x exists, since this would require $E\{\underline{\hat{x}}\} = BAx = x$ for all x, or $BA = I$, which is impossible since the rank of a product of two matrices cannot exceed the rank of either factor. But although x is not unbiased estimable, there exist linear functions of x which are. And an essential part of this second chapter focusses on the question tractable choices of such linear functions exist and how to interpret them.

Coordinates and datum definitions

Let us commence, in order to fix our minds, with the simple example of a levelling network (see figure 1).



1

After adjusting the network we obtain a consistent set of adjusted height differences satisfying the condition: $h_{12} + h_{23} + \ldots h_{(n-2)(n-1)} + h_{(n-1),n} + h_{n,1} = 0$. And it is clear that we cannot compute the heights of the net-

* Here and in the remaining part of this chapter we disregard configuration defects, since in our opinion they are merely the result of poor surveying

work points from these adjusted height differences. So how are we then to present our results of adjustment? One way is of course, to just list all the adjusted height differences with their variances and covariances and leave it at that. But, we only need a set of height differences which are mutually independent. For instance, the result of adjustment is completely described by listing the adjusted height differences $h_{12}, h_{23}, \ldots, h_{(n-2)(n-1)}, h_{(n-1)n}$ and their variances and covariances. But then again also the set $h_{23}, \ldots, h_{(n-2)(n-1)}, h_{(n-1)n}, h_{n,1}$ suffices and in fact many more choices are possible. So which to choose?

It seems reasonable, although we know we cannot really compute absolute heights, to look for a set of which the height differences at least resemble some of the properties of height coordinates. The advantage of working with coordinates in general is namely, that they all have *one and the same* reference in common. With coordinates the relative position of any two points in a network is easily obtained without that one needs to bother about the way in which these two network points are connected by the measured elements. Coordinates are also very tractable for drawing maps or making profiles of the whole or parts of the network.

Thus we like to have a set of height differences from which one can easily obtain the relative vertical position of any two points of the network. A set like $h_{12}, h_{23}, \ldots, h_{(n-2)(n-1)}, h_{(n-1)n}$ does not really suffice, because in order to get a picture of the height difference between, say the points P_2 and P_{n-2}, we need to draw a profile of almost the whole network. An appropriate set is, however, $h_{12}, h_{13}, \ldots, h_{1,(n-2)}, h_{1,(n-1)}, h_{1,n}$. In this case we just can take point P_1 as reference and mark out the height differences h_{12} and $h_{1,(n-2)}$. Also note that these height differences are very close to being height coordinates, since we merely have to adopt the reference point P_1 as origin or give it an arbitrary height, say h_1. By adopting the arbitrary height h_1 for reference point P_1, we can write

$$h_i = h_{1i} + h_1 = h_i^{(1)} + h_1 \qquad (2.1$$

And this expression shows that the set of height differences $h_{1i} = h_i^{(1)}$ can indeed be considered to be a set of height coordinates. But instead of taking point P_1 as reference we could also have taken point P_5. We then get

$$h_i = h_{5i} + h_5 = h_i^{(5)} + h_5 \qquad (2.2)$$

In fact we can take any point of the network as point of reference. We can even take the centre of "gravity" or the main-point P_M, with the adopted height $h_M = \frac{1}{n}(h_1 + h_2 + \ldots + h_n)$, as point of reference:

$$(2.3) \qquad h_i = \frac{1}{n}(h_{1i} + h_{2i} + \ldots + h_{ni}) + h_M = h_i^{(M)} + h_M .$$

Thus all sets of height differences like $h_i^{(1)}$, $h_i^{(5)}$ or $h_i^{(M)}$ can be seen as sets of height coordinates. And it is now not more a question of which set to choose, because any set will do.

It is important, however, to observe that the statistical properties, the first and second moments, of the heights $h_i^{(1)}$, $h_i^{(5)}$ or $h_i^{(M)}$, very much depend on the choice of reference point. From $E\{h_3^{(1)}\} = E\{h_{12} + h_{23}\}$ and $E\{h_3^{(5)}\} = E\{-h_{34} - h_{45}\}$ follows for instance that $E\{h_3^{(1)}\} \neq E\{h_3^{(5)}\}$. Also their variances and covariances differ. Thus if one wants to compare two sets of heights, where the two sets are computed from two different and independent observational campaigns - for instance for the purpose of a deformation analysis - it is essential that these heights are defined with respect to the same reference.

Now in order to get all heights in the same reference system one needs to be able to transform from one system to another. From writing (2.1) as

$$(2.4) \qquad \begin{pmatrix} h_1 \\ h_2 \\ \vdots \\ h_n \end{pmatrix} = \begin{pmatrix} \vdots \\ h_i^{(1)} \\ \vdots \end{pmatrix} + \left(\begin{array}{c|ccc} 1 & 0 & \ldots & 0 \\ \hline 1 & & & \\ \vdots & & 0 & \\ 1 & & & \end{array}\right) \begin{pmatrix} h_1 \\ h_2 \\ \vdots \\ h_n \end{pmatrix} ,$$

follows that

$$(2.5') \qquad \begin{pmatrix} \vdots \\ h_i^{(1)} \\ \vdots \end{pmatrix} = \left(\begin{array}{c|ccc} 0 & 0 & \ldots & 0 \\ \hline -1 & 1 & & \\ \vdots & & \ddots & 0 \\ -1 & & 0 & 1 \end{array}\right) \begin{pmatrix} h_1 \\ h_2 \\ \vdots \\ h_n \end{pmatrix} ,$$

and substitution of (2.2) then gives

$$(2.5'') \qquad \begin{pmatrix} \vdots \\ h_i^{(1)} \\ \vdots \end{pmatrix} \left(\begin{array}{c|ccc} 0 & 0 & \ldots & 0 \\ \hline -1 & 1 & & \\ \vdots & & \ddots & 0 \\ -1 & & 0 & 1 \end{array}\right) \begin{pmatrix} \vdots \\ h_i^{(5)} \\ \vdots \end{pmatrix}$$

Expressions (2.5) thus show how one can transform from any height system, e.g. $h_i^{(5)}$, to the height system defined by taking point P_1 as reference. In a similar way we obtain from (2.3) and (2.1) the transformation

$$(2.6) \qquad \begin{pmatrix} \vdots \\ h_i^{(M)} \\ \vdots \end{pmatrix} = \frac{1}{n} \begin{pmatrix} n-1 & -1 & \cdots & -1 \\ -1 & n-1 & -1 & \cdots \\ \vdots & \ddots & \ddots & \ddots \\ -1 & \cdots & -1 & n-1 \end{pmatrix} \begin{pmatrix} \vdots \\ h_i^{(1)} \\ \vdots \end{pmatrix}$$

Transformations like (2.5) and (2.6) which transform one reference system into another are called *S-transformations*.

Let us now consider a two dimensional planar triangulation network in which only angles are measured (see figure 2).

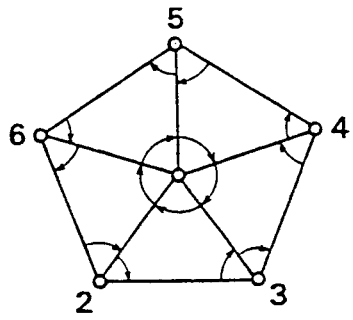


Fig. 2

After adjusting the network we obtain a consistent set of adjusted angles, which determines the shape of the network. It determines the shape and that is all it determines. The position, orientation and scale of the network are still unknown or still free to be chosen arbitrarily.
Now in order to describe the shape of the network we have, just like in case of the levelling network, many possibilities at hand. Each set of mutually independent adjusted angles, for instance, will do. But on the grounds of the earlier mentioned arguments we are in favour of coordinates. However, in order to compute coordinates we first need to fix some reference, i.e. we need to fix the position, orientation and scale of the network. One way to accomplish this is by fixing two points of the network, i.e. by assigning arbitrary and non-stochastic coordinates to two points of the network. For instance, we can start by fixing the points P_1 and P_2 and then compute, with the aid of the adjusted angles, the coordinates of the points P_1, P_4, P_5 and P_6 (see figure 3a). Or we can fix the points P_3 and P_1 and then compute the

points P_4, P_5, P_6 and P_2 (see figure 3b).

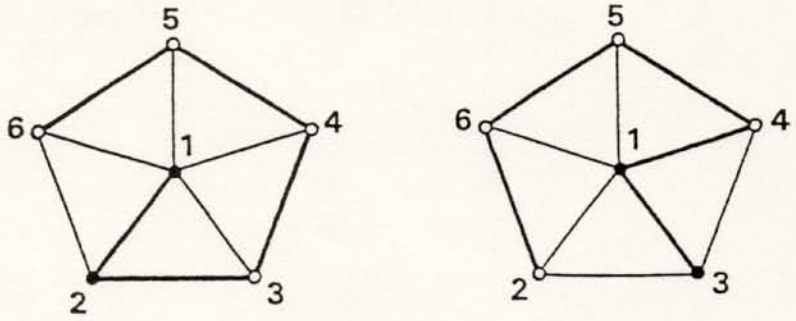


Fig. 3 (a) (b)

Let us for the moment, however, leave in the middle which two points we fix. Let's just call them P_r and P_s. We then can write (see figure 4).

$$(2.7)\qquad \begin{cases} x_j = x_r + l_{rs} \sin A_{rs} + l_{sj} \sin (A_{rs} + \pi + \alpha_{rsj}) \\ y_j = y_r + l_{rs} \cos A_{rs} + l_{sj} \cos (A_{rs} + \pi + \alpha_{rsj}) \end{cases}$$

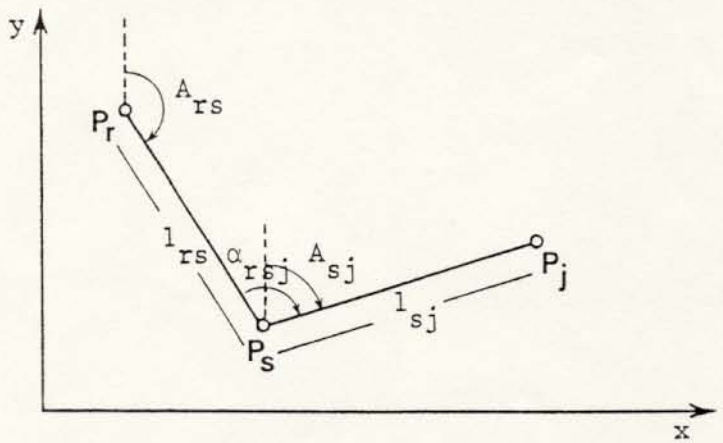


figure 4

Linearization of (2.7) gives*

$$(2.8)\qquad \begin{cases} \Delta x_j = \Delta x_r + x^o_{rs}\Delta \ln l_{rs} + y^o_{rs}\Delta A_{rs} + x^o_{sj}\,\Delta\ln l_{sj} + y^o_{sj}\,\Delta A_{rs} + y^o_{sj}\,\Delta\alpha_{rsj} \\ \Delta y_j = \Delta y_r + y^o_{rs}\Delta \ln l_{rs} - x^o_{rs}\Delta A_{rs} + y^o_{sj}\,\Delta\ln l_{sj} - x^o_{sj}\,\Delta A_{rs} - x^o_{sj}\,\Delta\alpha_{rsj}, \end{cases}$$

which we can write as

$$\begin{pmatrix} \Delta x_j \\ \Delta y_j \end{pmatrix} = \begin{pmatrix} y^o_{sj} & x^o_{sj} \\ -x^o_{sj} & y^o_{sj} \end{pmatrix} \begin{pmatrix} \Delta\alpha_{rsj} \\ \Delta\ln \frac{l_{sj}}{l_{sr}} \end{pmatrix} + \begin{pmatrix} 1 & 0 & y^o_{rj} & x^o_{rj} \\ 0 & 1 & -x^o_{rj} & y^o_{rj} \end{pmatrix} \begin{pmatrix} \Delta x_r \\ \Delta y_r \\ \Delta A_{rs} \\ \Delta \ln l_{rs} \end{pmatrix}$$

(2.9)

---

* The upperindex "o" indicates the approximate values.

Note that the first term on the right hand side of this expression only contains observed quantities, since $\Delta \ln \frac{l_{sj}}{l_{sr}}$ can be computed from the angles in the triangle $P_r P_s P_j$ by means of the sine-rule.
From inverting the relation

$$\begin{pmatrix} \Delta x_{rs} \\ \Delta y_{rs} \end{pmatrix} = \begin{pmatrix} y^o_{rs} & x^o_{rs} \\ -x^o_{rs} & y^o_{rs} \end{pmatrix} \begin{pmatrix} \Delta A_{rs} \\ \Delta \ln l_{rs} \end{pmatrix}$$

follows

$$\begin{pmatrix} \Delta A_{rs} \\ \Delta \ln l_{rs} \end{pmatrix} = \frac{1}{(l^o_{rs})^2} \begin{pmatrix} y^o_{rs} & -x^o_{rs} \\ x^o_{rs} & y^o_{rs} \end{pmatrix} \begin{pmatrix} \Delta x_{rs} \\ \Delta y_{rs} \end{pmatrix}$$

And substitution of this expression into (2.9) gives

$$(2.10) \quad \begin{pmatrix} \Delta x_j \\ \Delta y_j \end{pmatrix} = \begin{pmatrix} y^o_{sj} & x^o_{sj} \\ -x^o_{sj} & y^o_{sj} \end{pmatrix} \begin{pmatrix} \Delta \alpha_{rsj} \\ \Delta \ln \frac{l_{sj}}{l_{sr}} \end{pmatrix} +$$

$$+ \underbrace{\begin{pmatrix} 1 - \frac{x^o_{rj}x^o_{rs} + y^o_{rj}y^o_{rs}}{(l^o_{rs})^2} & \frac{-x^o_{rj}y^o_{rs} + x^o_{rs}y^o_{rj}}{(l^o_{rs})^2} & \frac{x^o_{rj}x^o_{rs} + y^o_{rj}y^o_{rs}}{(l^o_{rs})^2} & \frac{x^o_{rj}y^o_{rs} - y^o_{rj}x^o_{rs}}{(l^o_{rs})^2} \\ \frac{x^o_{rj}y^o_{rs} - x^o_{rs}y^o_{rj}}{(l^o_{rs})^2} & 1 - \frac{x^o_{rj}x^o_{rs} + y^o_{rj}y^o_{rs}}{(l^o_{rs})^2} & \frac{-x^o_{rj}y^o_{rs} + y^o_{rj}x^o_{rs}}{(l^o_{rs})^2} & \frac{x^o_{rj}x^o_{rs} + y^o_{rj}y^o_{rs}}{(l^o_{rs})^2} \end{pmatrix}}_{\text{call } M^{r,s}_j} \begin{pmatrix} \Delta x_r \\ \Delta y_r \\ \Delta x_s \\ \Delta y_s \end{pmatrix}$$

If we now take points P_r and P_s as reference- or base points by assigning to them the non-stochastic approximate coordinates x^o_r, y^o_r and x^o_s, y^o_s ($\Delta x_r = \Delta y_r = \Delta x_s = \Delta y_s = 0$), the coordinates of any other point P_j of the network are computed as

$$(2.11) \quad \begin{pmatrix} \Delta x_j \\ \Delta y_j \end{pmatrix}^{(r,s)} = \begin{pmatrix} y^o_{sj} & x^o_{sj} \\ -x^o_{sj} & y^o_{sj} \end{pmatrix} \begin{pmatrix} \Delta \alpha_{rsj} \\ \Delta \ln \frac{l_{sj}}{l_{rs}} \end{pmatrix},$$

where the upperindices (r,s) indicate that these coordinates are computed with respect to the base points P_r and P_s. These coordinates are thus just like the height coordinates $h_i^{(1)}$, $h_i^{(5)}$ and $h_i^{(M)}$, all linear(ized) functions of the observables.

With (2.11) we can write (2.10) as

$$(2.12)\qquad \begin{pmatrix} \Delta x_j \\ \Delta y_j \end{pmatrix} = \begin{pmatrix} \Delta x_j \\ \Delta y_j \end{pmatrix}^{(r,s)} + \begin{pmatrix} M_j^{r,s} \end{pmatrix} \begin{pmatrix} \Delta x_r \\ \Delta y_r \\ \Delta x_s \\ \Delta y_s \end{pmatrix}$$

And this expression is the complete analogon of the expressions (2.1), (2.2) and (2.3) we found for the levelling network.
Also here we can transform from one system to another. The S-transformation that transforms any reference system into the (r,s)-system, for instance, follows from (2.12) as:

$$(2.13)\qquad \begin{pmatrix} \Delta x_r \\ \Delta y_r \\ \Delta x_s \\ \Delta y_s \\ \cdots \\ \Delta x_j \\ \Delta y_j \\ \cdots \\ \cdot \\ \cdot \\ \cdot \\ \cdot \end{pmatrix}^{(r,s)} = \begin{pmatrix} & \cdot & & \cdot & \\ 0 & \cdot & 0 & \cdot & \cdots \\ \cdots & \cdot & \cdots & \cdot & \cdots \\ & \cdot & & \cdot & \\ -M_j^{r,s} & \cdot & I_2 & \cdot & 0 \\ \cdots & \cdot & \cdots & \cdot & \cdots \\ \vdots & \cdot & 0 & \cdot & I \\ & \cdot & & \cdot & \end{pmatrix} \begin{pmatrix} \Delta x_r \\ \Delta y_r \\ \Delta x_s \\ \Delta y_s \\ \cdots \\ \Delta x_j \\ \Delta y_j \\ \cdots \\ \cdot \\ \cdot \\ \cdot \end{pmatrix}$$

In the above given derivation we implicitly made use of the model of similarity transformations. To see this, assume the origin of the reference system pertaining to the set of approximate coordinates to be positioned in point P_r. Then $x_r^o = y_r^o = 0$. The difference quantities Δx_r, Δy_r, ΔA_{rs} and $\Delta \ln l_{rs}$ in (2.9) can then be interpreted as translation (Δt_x, Δt_y), rotation ($\Delta\phi$), and scale ($\Delta\lambda$) parameters. And with (2.11), relation (2.9) then reduces to

$$(2.14)\qquad \begin{pmatrix} \Delta x_j \\ \Delta y_j \end{pmatrix} = \begin{pmatrix} \Delta x_j \\ \Delta y_j \end{pmatrix}^{(r,s)} + \begin{pmatrix} 1 & 0 & y_j^o & x_j^o \\ 0 & 1 & -x_j^o & y_j^o \end{pmatrix} \begin{pmatrix} \Delta t_x \\ \Delta t_y \\ \Delta\phi \\ \Delta\lambda \end{pmatrix},$$

which is exactly the result one would obtain when linearizing the two dimensional similarity transformation :

$$(2.15)\qquad \begin{pmatrix} x_j \\ y_j \end{pmatrix} = \lambda \begin{pmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{pmatrix} \begin{pmatrix} x_j \\ y_j \end{pmatrix}^{(r,s)} + \begin{pmatrix} t_x \\ t_y \end{pmatrix},$$

under the assumption $\lambda^o = 1$, $\phi^o = 0$, $t_x^o = t_y^o = 0$.
Thus we can derive transformations like (2.13) from the (differential) similarity transformation (2.14) by fixing two points P_r and P_s, i.e. by setting $\Delta x_r^{(r,s)} = \Delta y_r^{(r,s)} = \Delta x_s^{(r,s)} = \Delta y_s^{(r,s)} = 0$. The coordinates $\begin{pmatrix} \Delta x_j \\ \Delta y_j \end{pmatrix}^{(r,s)}$ may then be thought of as having resulted from a (differential) similarity transformation to the approximate coordinates x_r^o, y_r^o, x_s^o, y_s^o of the base points P_r and P_s.

Summarizing, we can say that despite the fact that many ways exist for presenting the results of a free network adjustment, only special types of presentations deserve to be named coordinates. And since these coordinates depend on the reference system chosen, one should exercise great care in future manipulations. In particular one should be aware of the fact that their statistical properties depend on the chosen reference system.
In the following sections we will generalize the above given results and show how one can characterize the group of S-transformations. Also the relation with the theory of generalized inverses will become clear then.

II.3. S-transformations

Consider the linear(ized) Gauss-Markov model

$$(3.1)\qquad \underset{mx1}{E\{\underline{y}\}} = \underset{mxn}{A}\ \underset{nx1}{x}\ .$$

Definition: A linear function $c^t x$ is said to be unbiased estimable under the linear model (3.1), if there exists a linear function $b^t \underline{y}$ such that $E\{b^t \underline{y}\} = c^t x$.

Since $E\{b^t \underline{y}\} = b^t A x = c^t x$ should hold for all $x \in R^n$, it follows that a necessary and sufficient condition for $c^t x$ to be unbiased estimable is $c \in R(A^t)$. Thus every linear function of x, including x, is unbiased estimable if rank $A = r = n$, i.e. if $R(A^t) = R^n$. Not so if rank $A = r < n$. Then x is certainly not unbiased estimable.

In case of an angular network the linearized observation equations

$$E\{\Delta\underline{\alpha}_{kij}\} =$$

$$= \begin{pmatrix} -\frac{y^o_{ik}}{(l^o_{ik})^2} & \vdots & \frac{x^o_{ij}}{(l^o_{ik})^2} & \vdots & -\frac{y^o_{ij}}{(l^o_{ij})^2} + \frac{y^o_{ik}}{(l^o_{ik})^2} & \vdots & \frac{x^o_{ij}}{(l^o_{ik})^2} - \frac{y^o_{ij}}{(l^o_{ij})^2} & \vdots & \frac{y^o_{ij}}{(l^o_{ij})^2} & \vdots & -\frac{x^o_{ij}}{(l^o_{ij})^2} \end{pmatrix} \begin{pmatrix} \Delta x_k \\ \Delta y_k \\ \Delta x_i \\ \Delta y_i \\ \Delta x_j \\ \Delta y_j \end{pmatrix}$$

(3.2)

constitute the linear model (3.1). And it is clear that the observed angles alone do not suffice to determine the unknown coordinates. The corresponding design matrix A will therefore have a rank defect, i.e. rank A = r < n. This situation is now typical for all free network adjustments, where the observables may consist of angles, distances, distance ratios etc. In all these cases the observations alone do not suffice to estimate the unknown coordinate increments unbiasedly.

From the condition $c \in R(A^t)$ follows, not surprisingly, that the linear functions Ax are unbiased estimable. But what we like to find are unbiased estimable linear functions of x, which still can be interpreted as coordinates. Let us denote these unbiased estimable coordinates by $\underset{n\times 1}{x^{(s)}}$. Unbiased estimability then implies that $x^{(s)}$ is unbiased estimable by linear functions of $\underline{y}$, say $\underset{n\times m}{B}\,\underset{m\times 1}{\underline{y}}$. We can then write

$$(3.3) \qquad x^{(s)} = E\{B\underline{y}\} \text{ , or}$$

$$(3.4) \qquad x^{(s)} = BAx.$$

In section two we already met some examples of expression (3.3), namely the height coordinates $h_i^{(1)}$, $h_i^{(5)}$ and $h_i^{(M)}$ which are linear functions of the observed height differences and the coordinates $(\Delta x_j^{(r,s)}, \Delta y_j^{(r,s)})$ which follow from expression (2.11).

Now if we stick to the example of a two dimensional angular network, the $x^{(s)}$ should be transformable to any other coordinate system, say x, by means of the (differential) similarity transformations (2.14):

$$(3.5) \qquad x = x^{(s)} + Mp \text{ ,}$$

$$\text{with} \qquad R(M) = R\left(\begin{pmatrix} \vdots & \vdots & \vdots & \vdots \\ 1 & 0 & y^o_j & x^o_j \\ & & & \\ 0 & 1 & -x^o_j & y^o_j \\ \vdots & \vdots & \vdots & \vdots \end{pmatrix}\right) \text{ , and } p = \begin{pmatrix} \Delta t_x \\ \Delta t_y \\ \Delta\phi \\ \Delta\lambda \end{pmatrix}$$

And since the angular observables are invariant for any networkshape preserving coordinate transformation, i.e. for any similarity transformation of the type (3.5), we have

$$(3.6) \qquad Ax = Ax^{(s)}$$

or with (3.4)

$$Ax = ABAx.$$

Since this expression should hold for all coordinate system definitions possible, i.e. $\forall x \in R^n$, it follows from theorem 2.1 of chapter one that BA is a projector, projecting onto a subspace R(S) complementary to N(A) and along the nullspace $N(A) = R(M) = R(V_1)$. We therefore have

$$(3.8) \qquad x^{(s)} = BAx = S(V_o^t S)^{-1} V_o^t x, \text{ with } R^n = R(S) \oplus N(A) \text{ and } N(A) = R\begin{pmatrix} \vdots & \vdots & \vdots & \vdots \\ 1 & 0 & y_j & x_j \\ 0 & 1 & -x_j & y_j \\ \vdots & \vdots & \vdots & \vdots \end{pmatrix}$$

And with the projector identity $S(V_o^t S)^{-1} V_o^t = I - V_1 [(S^\perp)^t V_1]^{-1} (S^\perp)^t$, this expression can be written as

$$(3.9) \qquad x = x^{(s)} + V_1 [(S^\perp)^t V_1]^{-1} (S^\perp)^t x$$

which is the generalization of the expressions (2.1), (2.2), (2.3) and (2.12) we found in section two. The (S)-system is thus defined by the restrictions

$$(3.10) \qquad (S^\perp)^t x = 0, \text{ with } R^n = R(S) \oplus N(A)$$

Since the coordinate system defining subspace R(S) only needs to fulfil the complementarity condition $R^n = R(S) \oplus N(A)$, it follows that there are many more (S)-systems possible than the ones considered in section two. Some more examples are given at the end of this section.
The general expression for an arbitrary S-transformation now readily follows from (3.8) as

$$= P_{R(S_i),N(A)} = S_i (V_o^t S_i)^{-1} V_o^t = I_n - V_1 [(S_i^\perp)^t V_1]^{-1} (S_i^\perp)^t \; ; \; R^n = R(S_i) \oplus N(A)$$

(3.11)

The last expression in (3.11) is for practical purposes the most manageable, since $R(V_1) = N(A)$ is given by the linearized similarity tranformation and $S_i^{\perp}$ is chosen in order to define the reference system. Note by the way, that the S_i-matrix only depends on the subspaces $R(S_i)$ and $N(A)$ and not on the matrix representations taken for these subspaces.
Because of the projector property of the S_i-matrix we have

(3.12) $\quad S_i S_i = S_i$ (idempotence)

And since all the S_i project along the same subspace N(A), we also have (see figure 5)

(3.13) $\quad S_i S_j = S_i$

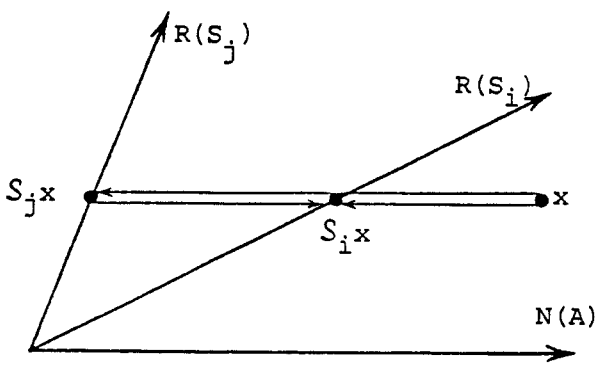


Figure 5

So far we only considered two dimensional networks of the angular type, but the situation for other type of free networks is very similar. The only difference lies in the nullspace N(A). For a two dimensional trilateration network, for instance, the nullspace of the design matrix reduces to

$$(3.14) \qquad N(A) = R\left(\begin{bmatrix} \vdots & \vdots & \vdots \\ 1 & 0 & y_j^o \\ 0 & 1 & -x_j^o \\ \vdots & \vdots & \vdots \end{bmatrix}\right)$$

i.e. the scale parameter is excluded from the similarity transformation. And if azimuths and distances are measured the nullspace becomes

$$(3.15) \qquad N(A) = R\left(\begin{bmatrix} \vdots & \vdots \\ 1 & 0 \\ 0 & 1 \\ \vdots & \vdots \end{bmatrix}\right).$$

For three dimensional free networks, we will need the model of the three dimensional similarity transformation:

$$(3.16)\qquad \begin{pmatrix} x_i \\ y_i \\ z_i \end{pmatrix} = \lambda R_x R_y R_z \begin{pmatrix} x_i \\ y_i \\ z_i \end{pmatrix} + \begin{pmatrix} t_x \\ t_y \\ t_z \end{pmatrix}$$

with: λ scale factor,

$$R_x = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\phi_x & \sin\phi_x \\ 0 & -\sin\phi_x & \cos\phi_x \end{pmatrix};\quad R_y = \begin{pmatrix} \cos\phi_y & 0 & -\sin\phi_y \\ 0 & 1 & 0 \\ \sin\phi_y & 0 & \cos\phi_y \end{pmatrix};\quad R_z = \begin{pmatrix} \cos\phi_z & \sin\phi_z & 0 \\ -\sin\phi_z & \cos\phi_z & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

and $\begin{pmatrix} t_x \\ t_y \\ t_z \end{pmatrix}$ the translation vector.

Linearization, under the assumption $\lambda^o = 1$, $\phi_x^o = \phi_y^o = \phi_z^o = 0$, and $t_x^o = t_y^o = t_z^o = 0$, then gives

$$\begin{pmatrix} \Delta x_i \\ \Delta y_i \\ \Delta z_i \end{pmatrix} = \begin{pmatrix} \Delta x_i \\ \Delta y_i \\ \Delta z_i \end{pmatrix}^{(s)} + \begin{pmatrix} 1 & 0 & 0 & 0 & -z_i^o & y_i^o & x_i^o \\ 0 & 1 & 0 & z_i^o & 0 & -x_i^o & y_i^o \\ 0 & 0 & 1 & -y_i^o & x_i^o & 0 & z_i^o \end{pmatrix} \begin{pmatrix} \Delta t_x \\ \Delta t_y \\ \Delta t_z \\ \Delta\phi_x \\ \Delta\phi_y \\ \Delta\phi_z \\ \Delta\lambda \end{pmatrix},$$

(3.17)

from which the nullspace of the design matrix of a free three dimensional network can be extracted.
In table 1 we have given the various matrices which characterize the nullspace of the design matrix for free networks.

Let us now consider some examples.

Example 1: Levelling network

For a free levelling network we just have one translational degree of freedom, i.e. dim N(A) = 1. Thus in order to define an S-system we need to choose a

	decisive observational type	dim N(A)	N(A)			
			Translation	rotation		scale
levelling/ rel. gravity	height differences/ gravity differences	dim N(A)=1	1			
2 dim. planar network	distances and azimuths	dim N(A)=2	$\begin{matrix}1 & 0\\ 0 & 1\end{matrix}$			
	distances	dim N(A)=3	$\begin{matrix}1 & 0\\ 0 & 1\end{matrix}$	$\begin{matrix}y_j^o\\ -x_j^o\end{matrix}$		
	angles and/or distance ratios	dim N(A)=4	$\begin{matrix}1 & 0\\ 0 & 1\end{matrix}$	$\begin{matrix}y_j^o\\ -x_j^o\end{matrix}$		$\begin{matrix}x_j^o\\ y_j^o\end{matrix}$
3 dim. network	distances, azimuths astron. latitude and longitude	dim N(A)=3	$\begin{matrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{matrix}$			
	distances	dim N(A)=6	$\begin{matrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{matrix}$	$\begin{matrix}0 & -z_j^o & y_j^o\\ z_j^o & 0 & -x_j^o\\ -y_j^o & x_j^o & 0\end{matrix}$		
	angles and/or distance ratios	dim N(A)=7	$\begin{matrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{matrix}$	$\begin{matrix}0 & -z_j^o & y_j^o\\ z_j^o & 0 & -x_j^o\\ -y_j^o & x_j^o & 0\end{matrix}$		$\begin{matrix}x_j^o\\ y_j^o\\ z_j^o\end{matrix}$

Table 1

vector $S^{\perp}_{nx1}$, such that R(S) is complementary to $N(A) = R\begin{pmatrix}1\\ \vdots\\ 1\end{pmatrix}$. Let us take $R(S^{\perp}) = N(A)$. Using expression (3.11), the corresponding S-transformation matrix then becomes

$$(3.18)\qquad S_M = I_n - \begin{pmatrix}1\\ \vdots\\ 1\end{pmatrix} \cdot \frac{1}{n} \cdot [1 \ldots 1] = \frac{1}{n}\begin{pmatrix} n-1 & -1 & \ldots & -1\\ -1 & n-1 & & \\ \vdots & & \ddots & \\ -1 & & & n-1\end{pmatrix}$$

Example 2: two dimensional network of the angular type

In this case we have four degrees of freedom and the nullspace of the design matrix is given by

$$N(A) = R\left(\begin{bmatrix} \vdots & \vdots & \vdots & \vdots \\ 1 & 0 & y_j^o & x_j^o \\ 0 & 1 & -x_j^o & y_j^o \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix}\right) .$$

Let us now again choose matrix $S^{\perp}_{n\times 4}$, such that $R(S^{\perp}) = N(A)$. The four restrictions $(S^{\perp})^t_{4\times n}\, x_{n\times 1} = 0$ then become

$$(3.19)\quad \sum_{i=1}^{1/2n} \Delta x_i = 0,\ \sum_{i=1}^{1/2n} \Delta y_i = 0,\ \sum_{i=1}^{1/2n} (y_i^o \Delta x_i - x_i^o \Delta y_i) = 0 \ \text{ and } \sum_{i=1}^{1/2n} (x_i^o \Delta x_i + y_i^o \Delta y_i) = 0$$

Giving the corresponding S-system the upper index (M), we get from applying (3.11) the following transformation

$$(3.20)\quad \begin{bmatrix} \Delta x_j \\ \Delta y_j \end{bmatrix}^{(M)} = \begin{bmatrix} \Delta x_j \\ \Delta y_j \end{bmatrix} - \frac{2}{n(r-\bar{x}^2-\bar{y}^2)} \begin{bmatrix} (r-y_j^o\bar{y}-x_j^o\bar{x}) & (y_j^o\bar{x}+x_j^o\bar{y}) & (y_j^o-\bar{y}) & (x_j^o-\bar{x}) \\ (x_j^o\bar{y}-y_j^o\bar{x}) & (r-x_j^o\bar{x}+y_j^o\bar{y}) & (x_j^o-\bar{x}) & (\bar{y}+y_j^o) \end{bmatrix} \begin{bmatrix} \sum_{i=1}^{n/2} \Delta x_i \\ \sum_{i=1}^{n/2} \Delta y_i \\ \sum_{i=1}^{n/2} (y_i^o \Delta x_i - x_i^o \Delta y_i) \\ \sum_{i=1}^{n/2} (x_i^o \Delta x_i + y_i^o \Delta y_i) \end{bmatrix}$$

with: $r = \frac{2}{n} \sum_{i=1}^{n/2} ((x_i^o)^2 + (y_i^o)^2)$, $\bar{x} = \frac{2}{n} \sum_{i=1}^{n/2} x_i^o$, $\bar{y} = \frac{2}{n} \sum_{i=1}^{n/2} y_i^o$ and n/2 being the number of network points.

Example 3: two dimensional network with distance and azimuth observations.

In this case we have two translational degrees of freedom left. And the null-space of the design matrix is given by

$$N(A) = R\left(\begin{bmatrix} \vdots & \vdots \\ 1 & 0 \\ 0 & 1 \\ \vdots & \vdots \end{bmatrix}\right).$$

By fixing just one network point, say P_r, the corresponding S-transformation simply is

$$(3.21)\quad \begin{bmatrix} \Delta x_j \\ \Delta y_j \end{bmatrix}^{(r)} = \begin{bmatrix} \Delta x_j \\ \Delta y_j \end{bmatrix} - \begin{bmatrix} \Delta x_r \\ \Delta y_r \end{bmatrix} .$$

Example 4: other then cartesian coordinates

So far we assumed the unknowns in the observation equations to be cartesian coordinates. But the theory is of course also valid for other type of coordinates. Instead of taking cartesian coordinates one can for instance take polar-, spherical- or geodetic coordinates. Let us assume that $\underset{n\times 1}{x}$ denote cartesian coordinates and $\underset{n\times 1}{\bar{x}}$ an other coordinate type. With the linear(ized) transformation

$$\bar{x} = Tx \ ,$$

the original Gauss-Markov model

$$E\{\underline{y}\} = Ax, \quad \text{with } N(A) = R(M) = R\left(\begin{bmatrix} \vdots & \vdots & \vdots & \vdots \\ 1 & 0 & y_j^o & x_j^o \\ & & & \\ 0 & 1 & -x_j^o & y_j^o \\ \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix}\right)$$

then transforms to

$$E\{\underline{y}\} = AT^{-1}Tx = AT^{-1}\bar{x} = \bar{A}\bar{x}, \text{ with } N(\bar{A}) = R(TM)$$

Thus the columns of the matrix TM span the nullspace $N(\bar{A})$.

Let us take as an example spherical coordinates. We then have

$$(3.22) \qquad \begin{pmatrix} x_i \\ y_i \\ z_i \end{pmatrix} = \begin{pmatrix} r_i \cos\phi_i \cos\lambda_i \\ r_i \cos\phi_i \sin\lambda_i \\ r_i \sin\phi_i \end{pmatrix}$$

And linearization gives

$$\begin{pmatrix} \Delta x_i \\ \Delta y_i \\ \Delta z_i \end{pmatrix} = \begin{pmatrix} -r_i^o \sin\phi_i^o \cos\lambda_i^o & -r_i^o \cos\phi_i^o \sin\lambda_i^o & r_i^o \cos\phi_i^o \cos\lambda_i^o \\ -r_i^o \sin\phi_i^o \sin\lambda_i^o & r_i^o \cos\phi_i^o \cos\lambda_i^o & r_i^o \cos\phi_i^o \sin\lambda_i^o \\ r_i^o \cos\phi_i^o & 0 & -r_i^o \sin\phi_i^o \end{pmatrix} \begin{pmatrix} \Delta\phi_i \\ \Delta\lambda_i \\ \Delta \ln r_i \end{pmatrix}$$

or

$$(3.23)\quad \begin{pmatrix} \Delta\phi_i \\ \Delta\lambda_i \\ \Delta \ln r_i \end{pmatrix} = \frac{1}{r_i^o} \begin{pmatrix} -\sin\phi_i^o \cos\lambda_i^o & -\sin\phi_i^o \sin\lambda_i^o & \cos\phi_i^o \\ -\cos^{-1}\phi_i^o \sin\lambda_i^o & \cos^{-1}\phi_i^o \cos\lambda_i^o & 0 \\ \cos\phi_i^o \cos\lambda_i^o & \cos\phi_i^o \sin\lambda_i^o & \sin\phi_i^o \end{pmatrix} \begin{pmatrix} \Delta x_i \\ \Delta y_i \\ \Delta z_i \end{pmatrix}$$

Substitution of (3.17) in the above expression then finally gives

$$\begin{pmatrix} \vdots \\ \Delta\phi_i \\ \Delta\lambda_i \\ \Delta \ln r_i \\ \vdots \end{pmatrix} = \begin{pmatrix} \vdots \\ \Delta\phi_i \\ \Delta\lambda_i \\ \Delta \ln r_i \\ \vdots \end{pmatrix}^{(s)} + \begin{pmatrix} \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ -r_i^{o\,-1}\sin\phi_i^o\cos\lambda_i^o & -r_i^{o\,-1}\sin\phi_i^o\sin\lambda_i^o & r_i^{o\,-1}\cos\phi_i^o & -\sin\lambda_i^o & \cos\lambda_i^o & 0\ 0 \\ -r_i^{o\,-1}\cos^{-1}\phi_i^o\sin\lambda_i^o & r_i^{o\,-1}\cos^{-1}\phi_i^o\cos\lambda_i^o & 0 & \tan\phi_i^o\cos\lambda_i^o & \tan\phi_i^o\sin\lambda_i^o & -1\ 0 \\ r_i^{o\,-1}\cos\phi_i^o\cos\lambda_i^o & r_i^{o\,-1}\cos\phi_i^o\sin\lambda_i^o & r_i^{o\,-1}\sin\phi_i^o & 0 & 0 & 0\ 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots\ \vdots \end{pmatrix} \begin{pmatrix} \Delta t_x \\ \Delta t_y \\ \Delta t_z \\ \Delta\phi_x \\ \Delta\phi_y \\ \Delta\phi_z \\ \Delta\lambda \end{pmatrix}$$

(3.24)

And this expression enables us to derive any S-transformation in terms of spherical coordinates.

Example 5: orientation and scale unknowns included in the observation equations

In deriving S-transformations we so far assumed the unknowns x in the Gauss-Markov model $E\{y\} = Ax$ to consist solely of coordinates. This assumption is valid if the observables are functions of coordinates only, which is the case with angles, distance ratios, distances etc. In practice, however, one will often write down the observation equations in terms of directions instead of angles. Beside coordinates, one will then also have orientation unknowns. Similarly, one will additionally have scale unknowns if the observation equations are expressed in terms of pseudo-distances instead of distance ratios. Therefore, in practice, the Gauss-Markov model will be of the form

$$(3.25)\qquad E\{y\} = (A_1 A_2) \begin{pmatrix} x_1 \\ x_2 \end{pmatrix},$$

with x_1: orientation- and/or scale unknowns; x_2: coordinate unknowns. And for the corresponding normal equations we have

$$(3.26)\qquad \begin{pmatrix} A_1^t Q_y^{-1} A_1 & A_1^t Q_y^{-1} A_2 \\ A_2^t Q_y^{-1} A_1 & A_2^t Q_y^{-1} A_2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} A_1^t Q_y^{-1} y \\ A_2^t Q_y^{-1} y \end{pmatrix}$$

For the derivation of S-transformations it is now of interest whether one considers the unknowns x_1 to be merely nuisance parameters which are to be reduced from the normalequations or whether one intends to involve these unknowns in the many S-system definitions possible.
If one opts for the first approach, the reduced normalequations follow from premultiplying (3.26) with the matrix

$$\begin{pmatrix} I & 0 \\ -A_2^t Q_y^{-1} A_1 (A_1^t Q_y^{-1} A_1)^{-1} & I \end{pmatrix}$$

This gives

$$(3.27)\qquad \begin{pmatrix} A_1^t Q^{-1} A_1 & A_1^t Q_y^{-1} A_2 \\ 0 & A_2^t Q_y^{-1} (I - A_1 (A_1^t Q_y^{-1} A_1)^{-1} A_1^t Q_y^{-1}) A_2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} A_1^t Q_y^{-1} y \\ A_2^t Q_y^{-1} (I - A_1 (A_1^t Q_y^{-1} A_1)^{-1} A_1^t Q_y^{-1}) y \end{pmatrix}$$

And recognizing the matrix $I - A_1 (A_1^t Q_y^{-1} A_1)^{-1} A_1^t Q_y^{-1}$ as the projector $P_{R(Q_y A_1^{\perp}),R(A_1)} \overset{\text{call}}{=} P$, which projects onto the subspace $R(Q_y A_1^{\perp})$ and along the subspace $R(A_1)$, we can write the reduced normalequations as

$$(3.28)\qquad (PA_2)^t Q_y^{-1} (PA_2) x_2 = (PA_2)^t Q_y^{-1} y$$

Since x_2 now only contains coordinates we are back at our familiar situation. For instance, for a two-dimensional planar network of the angular type we have

$$(3.29)\qquad N((PA_2)^t Q_y^{-1} (PA_2)) = N(PA_2) = R\begin{pmatrix} \vdots & \vdots & \vdots & \vdots \\ 1 & 0 & y_j^o & x_j^o \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 1 & -x_j^o & y_j^o \\ \vdots & \vdots & \vdots & \vdots \end{pmatrix}$$

The approach of reducing for the unknowns x_1 is the one which is usually followed for two dimensional planar networks, since the orientation- and scale unknowns in this case do not have a special significance of their own. However, it is good to point out that in principal these unknowns can also be used in the definition of an S-system. In case of a two dimensional planar

network with direction- and pseudo-distance measurements for instance, one can scale and orientate the network by fixing one scale unknown and one orientation unknown. The translational degree of freedom is then taken care of by fixing e.g. one network point.

This brings us to the second approach where the unknowns x_1 are intended to be involved in the S-system definition. Since the null-space $N(A_1 : A_2)$ of the Gauss-Markov model (3.25) differs from (3.29) we see that some modification is needed, in order to find transformations like (3.17).

Let us first assume we have a two dimensional planar network with only direction measurements r_{ij}. In figure 6 a part of such a network is drawn. Also the theodolite frame in point P_s is shown by dashed lines; the direction P_sP_o being the direction of zero reading.

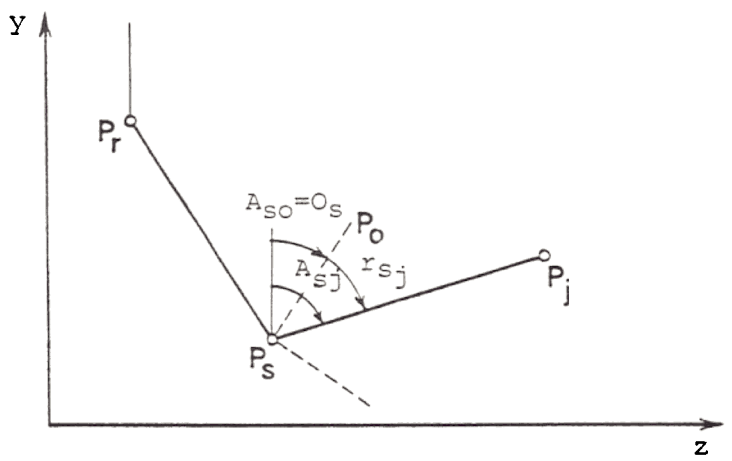


Figure 6

The idea now is to temporarily assume point P_o to be another ordinary network point. Its coordinates are then given by

$$\begin{aligned} x_o &= x_s + l_{so} \sin O_s \\ y_o &= y_s + l_{so} \cos O_s \end{aligned} \tag{3.30}$$

With the orientation unknown O_s being the azimuth of line P_sP_o. By interpreting P_o as an ordinary network point, we see that the direction observable r_{sj} can be interpreted as the angle observable α_{osj}. But this means that transformation (2.14) applies:

$$(3.31)\qquad \begin{pmatrix} \vdots \\ \Delta x_s \\ \Delta y_s \\ \Delta x_o \\ \Delta y_o \\ \vdots \end{pmatrix} = \begin{pmatrix} \vdots \\ \Delta x_s \\ \Delta y_s \\ \Delta x_o \\ \Delta y_o \\ \vdots \end{pmatrix}^{(s)} + \begin{pmatrix} \vdots & \vdots & \vdots & \vdots \\ 1 & 0 & y_s^o & x_s^o \\ 0 & 1 & -x_s^o & y_s^o \\ 1 & 0 & y_o^o & x_o^o \\ 0 & 1 & -x_o^o & y_o^o \\ \vdots & \vdots & \vdots & \vdots \end{pmatrix} \begin{pmatrix} \Delta t_x \\ \Delta t_y \\ \Delta\phi \\ \Delta\lambda \end{pmatrix} .$$

We can now use the linearized version of (3.30) in order to express the transformation (3.31) in terms of the coordinates $(\ldots \Delta x_s\ \Delta y_s\ \Delta o_s\ \Delta \ln l_{so} \ldots)^t$. This is similar to what we have done in example 4.
From (3.30) we obtain

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ 0 & -1 & 0 & 1 \end{pmatrix} \begin{pmatrix} \Delta x_s \\ \Delta y_s \\ \Delta x_o \\ \Delta y_o \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & l_{so}^o \cos o_s^o & l_{so}^o \sin o_s^o \\ 0 & 0 & -l_{so}^o \sin o_s^o & l_{so}^o \cos o_s^o \end{pmatrix} \begin{pmatrix} \Delta x_s \\ \Delta y_s \\ \Delta o_s \\ \Delta \ln l_{so} \end{pmatrix} ,$$

or

$$(3.32)\qquad \begin{pmatrix} \Delta x_s \\ \Delta y_s \\ \Delta o_s \\ \Delta \ln l_{so} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \dfrac{-\cos o_s^o}{l_{so}^o} & \dfrac{\sin o_s^o}{l_{so}^o} & \dfrac{\cos o_s^o}{l_{so}^o} & \dfrac{-\sin o_s^o}{l_{so}^o} \\ \dfrac{-\sin o_s^o}{l_{so}^o} & \dfrac{-\cos o_s^o}{l_{so}^o} & \dfrac{\sin o_s^o}{l_{so}^o} & \dfrac{\cos o_s^o}{l_{so}^o} \end{pmatrix} \begin{pmatrix} \Delta x_s \\ \Delta y_s \\ \Delta x_o \\ \Delta y_o \end{pmatrix}$$

Substitution of (3.31) into the above expression then gives

$$(3.33)\qquad \begin{pmatrix} \vdots \\ \Delta x_s \\ \Delta y_s \\ \Delta o_s \\ \Delta \ln l_{so} \\ \vdots \end{pmatrix} = \begin{pmatrix} \vdots \\ \Delta x_s \\ \Delta y_s \\ \Delta o_s \\ \Delta \ln l_{so} \\ \vdots \end{pmatrix}^{(s)} + \begin{pmatrix} \vdots & \vdots & \vdots & \vdots \\ 1 & 0 & y_s^o & x_s^o \\ 0 & 1 & -x_s^o & y_s^o \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ \vdots & \vdots & \vdots & \vdots \end{pmatrix} \begin{pmatrix} \Delta t_x \\ \Delta t_y \\ \Delta\phi \\ \Delta\lambda \end{pmatrix}$$

However, since the direction measurements only determine the direction $P_s P_o$ we can delete the Δlnl_{so}-row from (3.33). Thus the null-space of the Gauss-Markov model (3.25) becomes, if only directions are measured:

$$(3.34) \qquad N(A_1 \vdots A_2) = R \begin{pmatrix} \vdots & \vdots & \vdots & \vdots \\ 1 & 0 & y_j^o & x_j^o \\ 0 & 1 & -x_j^o & y_j^o \\ 0 & 0 & 1 & 0 \\ \vdots & \vdots & \vdots & \vdots \end{pmatrix}$$

If only pseudo-distances are measured, we obtain in a similar way

$$(3.35) \qquad N(A_1 \vdots A_2) = R \begin{pmatrix} \vdots & \vdots & \vdots & \vdots \\ 1 & 0 & y_j^o & x_j^o \\ 0 & 1 & -x_j^o & y_j^o \\ 0 & 0 & 0 & 1 \\ \vdots & \vdots & \vdots & \vdots \end{pmatrix}$$

Although the results (3.34) and (3.35) could have easily been predicted, we have given a so detailed derivation because the same reasoning applies to the more complicated situation which arises in case of three dimensional networks.

In figure 7 we have generalized the situation of figure 6 to three dimensions.

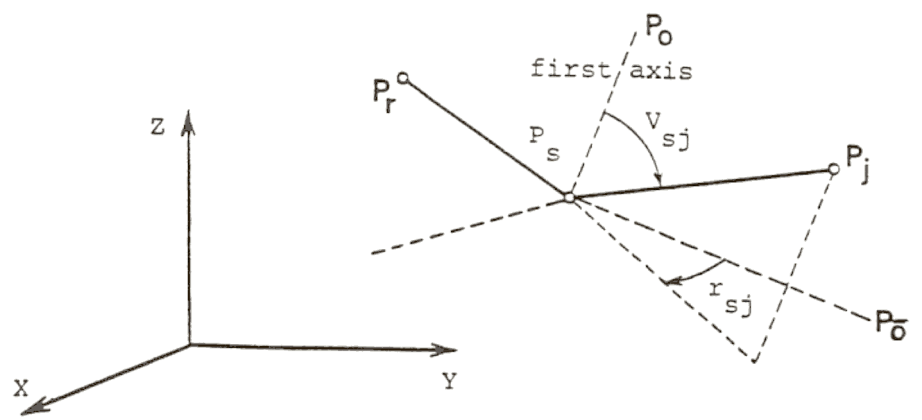


Figure 7

Since the local coordinate differences

$$\begin{pmatrix} x_{sj} \\ y_{sj} \\ z_{sj} \end{pmatrix} = \begin{pmatrix} l_{sj} \sin v_{sj} \sin r_{sj} \\ l_{sj} \sin v_{sj} \cos r_{sj} \\ l_{sj} \cos v_{sj} \end{pmatrix}$$

in the theodolite frame are related to the coordinate differences $(X_{sj}\ Y_{sj}\ Z_{sj})^t$ by the transformation

$$\begin{bmatrix} X_{sj} \\ Y_{sj} \\ Z_{sj} \end{bmatrix} = \begin{bmatrix} -\sin\Theta_{2,s} & -\sin\Theta_{1,s}\cos\Theta_{2,s} & \cos\Theta_{1,s}\cos\Theta_{2,s} \\ \cos\Theta_{2,s} & -\sin\Theta_{1,s}\sin\Theta_{2,s} & \cos\Theta_{1,s}\sin\Theta_{2,s} \\ 0 & \cos\Theta_{1,s} & \sin\Theta_{1,s} \end{bmatrix} \begin{bmatrix} \cos\Theta_{3,s} & \sin\Theta_{3,s} & 0 \\ -\sin\Theta_{3,s} & \cos\Theta_{3,s} & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_{sj} \\ y_{sj} \\ z_{sj} \end{bmatrix},$$

we get for the points P_o and P_{-o}:

$$\begin{bmatrix} X_o \\ Y_o \\ Z_o \end{bmatrix} = \begin{bmatrix} X_s \\ Y_s \\ Z_s \end{bmatrix} + l_{so} \begin{bmatrix} \cos\Theta_{1,s} & \cos\Theta_{2,s} \\ \cos\Theta_{1,s} & \sin\Theta_{2,s} \\ \sin\Theta_{1,s} & \end{bmatrix}; \begin{bmatrix} X_{-o} \\ Y_{-o} \\ Z_{-o} \end{bmatrix} = \begin{bmatrix} X_s \\ Y_s \\ Z_s \end{bmatrix} + l_{so} \begin{bmatrix} -\sin\Theta_{2,s}\sin\Theta_{3,s} & -\sin\Theta_{1,s}\cos\Theta_{2,s}\cos\Theta_{3,s} \\ \cos\Theta_{2,s}\sin\Theta_{3,s} & -\sin\Theta_{1,s}\sin\Theta_{2,s}\cos\Theta_{3,s} \\ \cos\Theta_{1,s}\cos\Theta_{3,s} & \end{bmatrix}$$

(3.36)

Following the same reasoning as before and assuming that pseudo-distances and directions, horizontal as well as vertical, are measured we then finally get the transformation:

$$\begin{bmatrix} \vdots \\ \Delta X_s \\ \Delta Y_s \\ \Delta Z_s \\ \Delta\Theta_{1,s} \\ \Delta\Theta_{2,s} \\ \Delta\Theta_{3,s} \\ \Delta\lambda_s \\ \vdots \end{bmatrix} = \begin{bmatrix} \vdots \\ \Delta X_s \\ \Delta Y_s \\ \Delta Z_s \\ \Delta\Theta_{1,s} \\ \Delta\Theta_{2,s} \\ \Delta\Theta_{3,s} \\ \Delta\lambda_s \\ \vdots \end{bmatrix}^{(s)} + \begin{bmatrix} \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & 0 & 0 & 0 & -Z_s^o & Y_s^o & X_s^o \\ 0 & 1 & 0 & Z_s^o & 0 & -X_s^o & Y_s^o \\ 0 & 0 & 1 & -Y_s^o & X_s^o & 0 & Z_s^o \\ 0 & 0 & 0 & -\sin\Theta_{2,s} & \cos\Theta_{2,s} & 0 & 0 \\ 0 & 0 & 0 & \tan\Theta_{1,s}^o\cos\Theta_{2,s}^o & \tan\Theta_{1,s}^o\sin\Theta_{2,s}^o & -1 & 0 \\ 0 & 0 & 0 & \cos^{-1}\Theta_{1,s}^o\cos\Theta_{2,s}^o & \cos^{-1}\Theta_{1,s}^o\sin\Theta_{2,s}^o & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix} \begin{bmatrix} \\ \Delta t_x \\ \Delta t_y \\ \Delta t_z \\ \Delta\phi_x \\ \Delta\phi_y \\ \Delta\phi_z \\ \Delta\lambda \\ \end{bmatrix},$$

(3.37)

$\Theta_{1,s}$ and $\Theta_{2,s}$ are the orientation unknowns which determine the direction of the first theodolite axis with respect to the XYZ-frame and $\Theta_{3,s}$ is the horizontal orientation unknown. Thus if the theodolite has been levelled, the angles $\Theta_{1,s}$ and $\Theta_{2,s}$ determine the direction of the gravity vector in point P_s with respect to the XYZ-frame.

II.4. The relation with generalized inverses

Consider again the Gauss-Markov model

$$(4.1) \qquad E\{\underset{mx1}{y}\} = \underset{mxn}{A}\ \underset{nx1}{x}\ , \quad \text{rank } A = r < n, \quad m \geq r\ , \quad \underset{mxm}{Q_y}$$

From chapter one we know that the set of weighted least squares estimates $\hat{x}$ of x is given by

$$\{\hat{x}\} = \{\hat{x} \mid \hat{x} = B_o y + V_1 \alpha,\ \forall \alpha \in R^{n-r},\ R(V_1) = N(A)\}$$

with B_o being a particular weighted least squares generalized inverse of A, uniquely characterized by $R(C^{\perp}) = R(Q_y A^{\perp})$ and arbitrary choices for R(S) and R(D). We now also know that in order to obtain an unbiased least squares estimate of the coordinates $x^{(s)}$, we need to transform the in general biased estimate $\hat{x}$ by means of an appropriate S-transformation:

$$\hat{x}^{(s)} = [S(V_o^t S)^{-1} V_o^t]\hat{x}\ , \quad Q_{\hat{x}(s)} = [S(V_o^t S)^{-1} V_o^t]\, Q_{\hat{x}}\, [S(V_o^t S)^{-1} V_o^t]^t$$

Once again we see here that the first- and second moments, $E\{\underset{\sim}{\hat{x}}^{(s)}\}$ and $Q_{\hat{x}(s)}$, very much depend on the chosen S-system.
Substitution of one of the estimates out of the set (4.2) into (4.3) gives

$$(4.4) \qquad \hat{x}^{(s)} = [S(V_o^t S)^{-1} V_o^t] B_o y$$

And it is readily seen from our general expression (2.13) or (5.3) in chapter one of generalized inverses of A that (4.4) reduces to

$$(4.5) \qquad \hat{x}^{(s)} = S(C^t A S)^{-1} C^t y$$

with $\quad R(C^{\perp}) = R(Q_y A^{\perp})$.

Thus *any* reflexive weighted least squares generalized inverse of A will *always* give us unbiased estimates of the type $\hat{x}^{(s)}$:

$$(4.6) \qquad \hat{x}^{(s)} = S(S^t A^t Q_y^{-1} A S)^{-1} S^t A^t Q_y^{-1} y.$$

Expressions (4.5) and (4.6) indicate that we obtain the estimate $\hat{x}^{(s)}$ by simply solving for the extended Gauss-Markov model:

(4.7) $E\{\underset{mx1}{\underline{y}}\} = \underset{mxn}{A}\ \underset{nx1}{x}$, rank $A = r < n$, $m \geq r$, $\underset{mxm}{Q_y}$; under the restrictions

$(S^{\perp})^t x = 0$, with $R^n = R(S) \oplus N(A)$

In practice this is probably also the easiest way of computing the estimate $\hat{x}^{(s)}$.

REFERENCES

Books in the mathematical literature which deal with the theory of generalized inverses are:

Ben-Israel, A. and T.N.E. Greville (1974). *Generalized Inverses: Theory and Applications*. J. Wiley, New York.

Bouillon, T.L. and P.L. Odell (1971). *Generalized Inverse Matrices*. Wiley-Interscience, New York.

Rao, C.R. and S.K. Mitra (1971). *Generalized Inverse of Matrices and its Applications*. J. Wiley, New York.

The book by Rao and Mitra is probably the most cited reference in the geodetic literature dealing with generalized inverses.

Books written by and for geodesists which treat the theory of generalized inverses and give geodetic applications are:

Bjerhammar, A. (1973). *Theory of Errors and Generalized Matrix Inverses*. Elsevier, Amsterdam.

Bjerhammar was one of the first who started a systematic study of the problem of inverting matrices of arbitrary order and rank. (see e.g. his paper [Bjerhammar, 1951].).

Grafarend, E., H. Heister, R. Kelm, H. Kropff, and B. Schaffrin (1979). *Optimierung Geodätischer Messoperationen*. Herbert Wichmann Verlag Karlsruhe. Band II.

Koch, K.R. (1980). *Parameterschätzung und Hypothesentests in linearen Modellen*. Dümmler.

Meissl, P. (1982). *Least Squares Adjustment: A Modern Approach*. Mitteilungen der Geodätischen Institute der Technischen Universität Graz. Folge 43.

Where scientific papers are concerned, there is an overwhelming list of geodetic papers which deal with the theory of generalized inverses and S-transformations. Here are some typical examples:

Baarda, W. (1973). *S-transformations and Criterion Matrices*, Netherlands Geodetic Commission, Publications on Geodesy, New Series Vol. 5, No. 1, 1973 Delft.

Baarda was the first who made a systematic study of the consequences of datum definitions on the precision description of geodetic networks.

(For an application of the theory of S-transformations to the problem of precision testing in geodetic networks, see also my lecture notes [Teunissen, *Quality Control in Geodetic Networks*]).

Bjerhammar, A. (1951). *Rectangular Reciprocal Matrices, with Special Reference to Geodetic Calculations,* No. 20, 188-220.

Blaha, G. (1971). *Inner Adjustment Constraints with Emphasis on Range Observations.* Report No. 148, Department of Geodetic Science of the Ohio State University, Columbus.

Grafarend, E. and B. Schaffrin (1974). *Unbiased Free Net Adjustment,* Survey Review 22 (1974), 200-218.

Grafarend, E. and B. Schaffrin (1976). *Equivalence of Estimable Quantities and Invariants in Geodetic Networks,* ZfV 101, S. 485-491.

Meissl, P. (1962). *Die Innere Genauigkeit eines Punkthaufens,* Öz Vermessungswesen 50 (1962), 159-165, 186-194.

Mierlo, J. van (1979). *Free Network Adjustment and S-transformations,* DGK B, Nr. 252, S. 41-54. München 1980.

Mittermayer, E. (1971). *Zur Ausgleichung Freier Netze.* ZfV 97, S. 481-489.

Pelzer, H. (1974). *Zur Behandlung Singulärer Ausgleichungsproblemen,* ZfV 99, S. 181-194.

Pope, A. (1971). *Transformation of Covariance Matrices Due to Changes in Minimal Control,* Paper pres. American Geophysical Union, San Francisco 1971.