DEMONSTRATIO MATHEMATICA
Vol. XXVIII No 3 1995

Witold Prószyński, Andrzej Sosnowski

MATRIX EXPRESSIONS FOR MINIMUM SEMINORM LEAST SQUARES SOLUTIONS OF LINEAR SYSTEMS AND THE RELATED INVERSE

1. Introduction

The expressions presented in this paper have initially been derived for the needs of data analysis in geodetic monitoring of relative movements in engineering structures. The mathematical models used there are linearized rank-defficient systems solved by the method of least squares. The solutions of those systems are often required to be of minimum l^2 norm for a specified subvector of unknowns.

The initial approach has later been expanded into a more general problem of seeking the minimum N-seminorm M-least squares solutions (with matrices N and M being at least positive semi-definite) and determining a suitable inverse generating these solutions (see [1]).

The derivations are based on a Moore-Penrose inverse, providing a clearly interpretable structure of the obtained matrix expressions as well as enabling an apriori analysis of the problem itself to detect areas of non-unique solutions. With an auxiliary assumption of minimum l^2 norm for a full solution vector, being applied in the case of non-uniqueness, a unique minimum N-seminorm M-least-squares inverse is defined.

The derived expressions are complementary to those given in ([1]).

2. Formulation of the problem

Let us consider a matrix equation

$$Ax = b \tag{1}$$

where: $A \in R^{m\times n}$, $b \in R^m$, $x \in R^n$, and matrices $M \in R^{m\times m}$, $N \in R^{n\times n}$, each symmetric, at least positive semi-definite.

The problem is to determine $\widehat{x}$ such that it minimizes the seminorm

$$\|Ax - b\|_M = [(Ax - b)^T M(Ax - b)]^{1/2}$$

and has a minimum seminorm

$$\|x\|_N = (x^T N x)^{1/2}.$$

The $\widehat{x}$ as above is a minimun N-seminorm M-least-squares solution of (1).

The problem can be formulated in an equivalent way defined as: detemine $\widehat{x}$ such that it minimizes the norm $\|H(Ax - b)\|$ and has a minimum norm $\|Kx\|$, where $H \in R^{p\times m}$ and $K \in R^{q\times n}$ such that $H^T H = M$, $K^T K = N$. Hence $\widehat{x}$ is a minimun N-seminorm least-squares solution for

(2) $$A_* x = b_*$$

where $A_* = HA$, $b_* = Hb$.

The assumption of symmetry of M and N does not diminish the level of generality in formulation of the problem, since

$$\|x\|_M = \|x\|_{(M+M^T)^{1/2}} = \|x\|_{M'}$$

where M' is symmetric.

The same applies to the second seminorm based on the matrix N.

3. Derivation of the formula for $\widehat{x}$

Any least squares solution of (2) (see [2]) can be written as

(3) $$x_g = A_*^+ b_* + Uz$$

where: A_*^+ - a Moore–Penrose inverse of A (i.e. A_*^+ satisfies conditions:

$$A_* A_*^+ A_* = A_*, \quad A_*^+ A_* A_*^+ = A_*^+, \quad (A_* A_*^+)^T = A_* A_*^+,$$
$$(A_*^+ A_*)^T = A_*^+ A_*), \quad U = I - A_*^+ A_*, \quad z \in R^n \text{ - an arbitrary vector.}$$

We shall now find z such that it minimizes $\|x_g\|_N = \|Kx_g\|$. Denoting $D = KU$ and $c = -KA_*^+ b_*$ we have $\|Kx_g\| = \|Dz - c\|$, hence we are looking for a least-squares solution of the equation

(4) $$Dz = c.$$

The class of least-squares solutions of (4) is given by the formula

$$z_g = D^+ c + (I - D^+ D)g, \quad g \in R^n \quad \text{an arbitrary vector.}$$

Hence, substituting for **z** in (3) $\mathbf{z}_g$ as above, we get final expression for $\widehat{x}$

(5) $$\widehat{x} \equiv x_g = A_*^+ b_* + UD^+ c + U(I - D^+ D)g.$$

This expression describes any least-squares solution of (2) which has a minimum norm $\|Kx\|$. x_g is also a minimun N-seminorm M-least-squares solution of (1). Using well known properties of Moore-Penrose inverse, we have $A_*^+ = (HA)^+ = (A^T H^T HA)^+ A^T H^T = (A^T MA)^+ A^T H^T$,

$U = I - (A^TMA)^+(A^TMA)$ and is symetric, and $D^+ = (KU)^+ = (U^TK^TKU)^+U^TK^T = (UNU)^+UK^T$. Finally, the expression (5) can be rewritten in terms of martices M and N as follows

(6) $$\hat{x} \equiv x_g = \{I - C(CNC)^+CN\}B^+A^TMb + C\{I - (CNC)^+CNC\}g$$

where $B = A^TMA$, $C = I - B^+B$.

Notice that the final formula for $\hat{x}$ does not depend on the H and K decompositions of matrices M and N respectively.

We shall conclude the above derivations by formulating the following theorems:

THEOREM 1. *A minimun N-seminorm M-least-squares solution of* (1), *not necessarily unique, always exists.* ∎

THEOREM 2. *Any minimun N-seminorm M-least-squares solution of* (1) *is given by the formula* (5) *or* (6). ∎

4. Uniqueness of solution

In general a minimun N-seminorm M-least-squares solution of (1) is not unique. This solution becomes unique in the case when the last component of formula (5) is zero for any g, i.e.

(7) $$U(I - D^+D) = 0.$$

Since $U = (I - A_*^+A_*)$ is an orthogonal projection on $\mathrm{Ker}(A_*)$, and $(I - D^+D)$ is an orthogonal projection on $\mathrm{Ker}(D)$, the condition (7) is equivalent to

$$\mathrm{Ker}(A_*) \perp \mathrm{Ker}(D).$$

But $\mathrm{Ker}(D) = \mathrm{Ker}(K(I - A_*^+A_*)) = \mathrm{Ker}(A_*)^\perp \cup (\mathrm{Ker}(A_*) \cap \mathrm{Ker}(K))$, where $A_* = HA$. Hence we can formulate the following proposition:

PROPOSITION. *Each of the following conditions is necessary and sufficient for the uniqueness of minimum N-seminorm M-least-squares solution of equation* (1):

(i) $\mathrm{Ker}(HA) \cap \mathrm{Ker}(K) = \{0\}$, *for any* H *and* K *such that* $M = H^TH$, $N = K^TK$;

(ii) $\mathrm{Ker}\begin{bmatrix} HA \\ K \end{bmatrix} = \{0\}$, *for any* H *and* K *as above*;

(iii) $\mathrm{rank}\begin{bmatrix} HA \\ K \end{bmatrix} = n$, *for any* H *and* K *as above*;

(iv) $\mathrm{rank}[A^TMA + N] = n$. ∎

The unique solution of (1) is given by

(8) $$x_u = A_*^+b_* + UD^+c$$

or equivalently

$$x_u = (I - UD^+K)A_*^+ b_*. \tag{9}$$

In the case of non-uniqueness (the last component in (5) is not zero) in order to choose a specific solution from among all possible solutions we shall put an auxiliary condition of minimum l^2 norm for a full vector of unknowns. We shall thus find g in (5) such that $\|x_g\|$ is a minimum, which leads to the least squares solution of the system

$$Wg = -x_u \tag{10}$$

where $W = U(I - D^+D)$.

Hence

$$g = -W^+ x_u \tag{11}$$

and finally

$$\widehat{x}_m = x_{gm} = (I - WW^+)x_u = (I - WW^+)(I - UD^+K)A_*^+ b_*. \tag{12}$$

We shall denote by x_{go} another particular choice in the class of non- unique solutions (5) obtained with $g = 0$. The expression is identical with that for a unique solution, i.e. x_u in (8) or (9). In the special case when $D = 0$ we obtain $x_u = A_*^+ b_*$ and $W = U = (I - A_*^+ A_*)$. Hence $x_{go} = x_u$. Finally we have

$$D = 0 \Rightarrow x_{gm} = x_{go} = A_*^+ b_*.$$

5. A specific minimum N-seminorm M-least-squares inverse and its properties

The expression (12) enables us to define a specific inverse of A which can be classified as a particular choice of the minimum N-seminorm M-least-squares inverse A_{MN}^+ (see [1]).

DEFINITION. The matrix $A_{MN}^\oplus$ is said to be a specific minimum N-seminorm M-least-squares inverse of A, if

$$A_{MN}^\oplus = (I - WW^+)(I - UD^+K)A_*^+ H \tag{13}$$

where: $W = U(I - D^+D)$; $D = KU$; $U = I - A_*^+ A_*$; $A_* = HA$.

The formula (13) can be rewritten in terms of martices M and N, showing its independence of a particular choice of H and K. The vector $A_{MN}^\oplus b$ is such a minimum N-seminorm M-least-squares solution of (1) which has a minimum l^2 norm in R^n.

Another possibility for proposing a specific choice of A_{MN}^+ could be taking as a basis the solution $x_{go} = x_u$, which would simplify the expression defining the inverse. However, the properties of solutions generated by such an inverse would in the case of non-uniqueness be not so explicite as with x_{gm}.

Here are some properties of $A^{\oplus}_{MN}$:

1° It can be easily verified that $A^{\oplus}_{MN}$ satisfies the following relationships

$$MAA^{\oplus}_{MN}A = MA, \quad NA^{\oplus}_{MN}AA^{\oplus}_{MN} = NA^{\oplus}_{MN},$$
$$(MAA^{\oplus}_{MN})^T = MAA^{\oplus}_{MN}, \quad (NA^{\oplus}_{MN}A)^T = NA^{\oplus}_{MN}A$$

being identical with the n.s. conditions for N-seminorm M-least-squares g-inverse of A, i.e. A^{+}_{MN} (see [1]).

2° $A^{\oplus}_{MN}$ generates the solution $\widehat{x}$ of (2) having the error vector identical with that for A^{+}_{*}. We shall introduce auxiliary notation $x^{+} = A^{+}_{*}b_{*}$

$$A_{*}\widehat{x} - b_{*} = A_{*}A^{+}_{*}b_{*} + A_{*}(I - A^{+}_{*}A_{*})z - b_{*} = A_{*}A^{+}_{*}b_{*} - b_{*} = A_{*}x^{+} - b_{*}.$$

3° With $D = K(I - A^{+}_{*}A_{*}) = \mathbf{0}$, $A^{\oplus}_{MN}$ becomes equal to $A^{+}_{*}H$ or equivalently $(A^TMA)^{+}A^TM$.

4° With $K = I$, $A^{\oplus}_{MN}$ gets reduced to $A^{+}_{*}H$, and if additionaly $H = I$, we get $A^{\oplus}_{MN} = A^{+}$.

5° If A_{*} is of full column rank, then obviously $A^{\oplus}_{MN}$ does not depend on K and is equal to $A^{+}_{*}H$, or equivalently $(A^TMA)^{+}A^TM$.

6. Introductory analysis of a given system

Prior to solving a particular problem one should examine whether the solution to the system (1) will be unique or there will be a set of possible solutions. This can be done with the use of the uniqueness conditions given in section 4. The third and the fourth condition seem to be most effective for practical computations.

The a priori examination of the structure of matrix A may provide some guidance on specification of the matrix K to obtain a unique solution. Such an examination can be based on a property presented below.

Let the system (1) be given in the following partitioned form

$$[A_1 \vdots A_2]\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = b \tag{18}$$

where: $A_1(m \times n_1)$; $A_2(m \times n_2)$.

If $\operatorname{rank} HA_1 = n_1$, $\operatorname{rank} HA_2 < n_2$ and $\operatorname{rank} HA = \operatorname{rank} HA_1 + \operatorname{rank} HA_2$, then a necessary and sufficient structure of K (denoted by K_r) to ensure the uniqueness of solution of (1) is as follows

$$K_r = [\,\cdot\; K_{r2}]$$

where $\operatorname{rank}[(HA_2)^T\, K^T_{r2}] = n_2$; the dot denotes an arbitrary block.

In particular, we can not obtain the uniqueness of solution of (1) by specifying $K = [K_1\ 0]$. Furtheremore, the structure of $U = I - A_*^+ A_*$ is

$$U = \begin{bmatrix} 0 & 0 \\ 0 & I - A_{*2}^+ A_{*2} \end{bmatrix} \tag{19}$$

and hence $D = KU = 0$. The specific minimum N-seminorm M-least-squares inverse $A_{MN}^{\oplus}$ takes the form $A_{MN}^{\oplus} = A_*^+ H = (A^T M A)^+ A^T M$ for any K_1. The solution generated by $A_{MN}^{\oplus}$, i.e. $\widehat{x}_m = A_{MN}^{\oplus} b = A_*^+ b_*$, is a Moore–Penrose solution of (2).

7. Discussion of characteristic examples

EXAMPLE 1. For given matrices A, M and N

$$A = \begin{bmatrix} 1 & 2 & 1 & 2 \\ 2 & 4 & 2 & 4 \\ 3 & 5 & 4 & 7 \end{bmatrix} \qquad M = \begin{bmatrix} 1 & 3 & 6 \\ 3 & 9 & 18 \\ 6 & 18 & 54 \end{bmatrix} \qquad N = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

carry out an a priori check on the uniqueness of solution of (1) and then compute $A_{MN}^{\oplus}$.

Solution. We shall use the condition (iv) (see Proposition in section 4). As $\operatorname{rank}[A^T M A + N] = 4$ the solution of (1) will be unique. To compute $A_{MN}^{\oplus}$ we shall use the expression (13) rewriting it in terms of M and N, i.e. putting

$$\begin{aligned} A_*^+ H &= (A^T M A)^+ A^T M = B^+ A^T M \\ U &= I - (A^T M A)^+ A^T M A = I - B^+ B = C \\ D^+ K &= (CNC)^+ CN \\ W &= C[I - (CNC)^+ CNC], \end{aligned}$$

where B and C are as those in (6).

As an intermediate step in computations we get $W = 0$, which numerically confirms the uniqueness of solution of (1) (see condition (7)). Finally, we obtain $A_{MN}^{\oplus}$ in the form

$$A_{MN}^{\oplus} = \begin{bmatrix} 0.05555 & 0.16667 & -0.11111 \\ 0.22222 & 0.66667 & -0.44444 \\ -0.05555 & -0.16667 & 0.11111 \\ -0.15079 & -0.45238 & 0.44444 \end{bmatrix}.$$

The above inverse satisfies each of the four relationships listed in section 5 (see property 1°).

EXAMPLE 2. For a given block matrix $A = [A_1 \vdots A_2]$

$$A = \begin{bmatrix} 3 & 2 & \vdots & 1 & 2 \\ 2 & 7 & \vdots & 2 & 4 \\ 3 & -5 & \vdots & 4 & 8 \end{bmatrix}$$

and $H = I$, determine a sufficient structure of K ensuring the uniqueness of solution of (1).

Compute $A^{\oplus}_{MN}$ assuming: a) $K = [3\,4\,2\,1]$; b) $K = [-1\,2\,0\,0]$.

Solution. As $\operatorname{rank} A_1 = 2$, $\operatorname{rank} A_2 = 1$ and $\operatorname{rank} A = 3$, i.e. $\operatorname{rank} A_1 + \operatorname{rank} A_2 = \operatorname{rank} A$, the required structure of K will be $K_r = [\,\cdot\; K_{r2}]$, with K_{r2} such that $\operatorname{rank}[A_2^T\, K_{r2}^T] = 2$.

The matrix K given in a) satisfies this requirement, and hence, the solution of (1) will be unique.

As an intermediate result of computations we get $W = 0$ which confirms the uniqueness of solution and finally

$$A^{\oplus}_{MN} = \begin{bmatrix} 0.48101 & -0.16456 & -0.03797 \\ -0.02532 & 0.11392 & -0.05063 \\ -0.76371 & -0.06329 & 0.13924 \\ 0.18565 & 0.16456 & 0.03797 \end{bmatrix}.$$

The matrix K given in b) does not posses a required structure, and since $D = 0$ we may use the property 3° (see section 5) with H=I. We get $A^{\oplus}_{MN} = A^+_* H = A^+$.

Computations carried out according to (13) confirm the above, as

$$A^{\oplus}_{MN} = \begin{bmatrix} 0.48101 & -0.16456 & -0.03797 \\ -0.02532 & 0.11392 & -0.05063 \\ -0.07848 & 0.05316 & 0.04304 \\ -0.15696 & 0.10633 & 0.08608 \end{bmatrix} = A^+.$$

Since in this case $W = U[I - D^+D] = U \neq 0$ (we get it also when executing the formula (13)) the solution of (1) is not unique and any solution of (1) can be written in the form $x_g = A^+b + Ug$, with g arbitrary.

8. Concluding remarks

The paper concentrates entirely on theoretical aspects of the problem, being the existance of solution and its uniqueness or non-uniqueness. The numerical aspects involved in using the derived expressions for practical computations are not covered in this paper. The question of constructing an effective numerical algorithm deserves a separate treatment.

References

[1] C. R. Rao and S. K. Mitra, *Generalized inverses of matrices and their applications*, John Wiley & Sons, Inc. 1971.

[2] C. R. Rao, *Linear models of mathematical statistics*, (trans. into Polish), PWN, Warszawa 1982.

[3] W. Prószyński, *Introduction into applications of the generalized Moore–Penrose inverse to chosen adjustment problems*, (in Polish), Geodezja i Kartografia, 2 (1981), 123–130.

[4] A. Olbrot and A. Sosnowski, *Worst case optimal observer design for linear discrete-time systems with unknown bounded square summable disturbances*, Int. J. Control, 48 (1988), 1937–1946.

Witold Prószyński—INSTITUTE OF APPLIED GEODESY
Andrzej Sosnowski—INSTITUTE OF MATHEMATICS
WARSAW UNIVERSITY OF TECHNOLOGY
00-660 WARSZAWA, POLAND

Received November 2nd, 1994.