Canad. Math. Bull. Vol. **23** (1), 1980

ON LINEAR MATRIX EQUATIONS

BY
P. SCOBEY AND D. G. KABE

ABSTRACT. Some results from the theory of minimization of vector quadratic forms (subjected to linear restrictions) are used to obtain particular solutions to the usual types of linear matrix equations. An answer to a question raised by Greville [1] is supplied.

1. **Introduction** In linear parametric estimation theory of Normal Multivariate Statistical Analysis we are required to solve several types of normal equations. Among these equations are linear matrix equations of the type (i) $DX = V$, (ii) $FX = W_1$, $XH = W_2$, (iii) $AXB = C$, (iv) $AXB + CXD = E$, or their combinations, where X may be a symmetric matrix. Such equations are often solved in statistical literature by using weighted least squares theory, and elsewhere by using unweighted least squares theory. In statistical literature only, obtaining of particular solutions is sufficient, because the scalar test criteria of statistics are unique, whatever be the particular solutions. We use some simple results on traces of vector quadratic forms (subjected to linear restrictions) minimization theory and obtain particular solutions to the above types of matrix equations. Following the practice in mathematical literature we obtain unweighted least squares solutions; however, we also indicate procedures for obtaining weighted least squares solutions.

Some results found useful in the sequel are stated in the next section. Section 3 is devoted to solving linear matrix equations, and in Section 4 we answer a question raised by Greville [1].

Sometimes the same symbol denotes different quantities; however, its meaning is made explicit in the context. All matrices in the paper are assumed to be full rank matrices, although all our results are valid for less than full rank matrices provided the pseudoinverses are properly handled.

2. **Some useful results.** Let X be a $p \times N$ matrix, μ a $p \times N$ matrix, Δ an $N \times N$ symmetric matrix, Σ a $p \times p$ symmetric matrix, D a given $q \times N$ matrix, and V a given $p \times q$ matrix, $N \geq \max(p, q)$. Then Kabe [2] shows that

$$\text{(1)} \qquad \min \operatorname{tr} \Sigma^{-}(X-\mu)\Delta(X-\mu)', \quad \text{subject to} \quad DX' = V',$$

is given by

$$\text{(2)} \qquad \operatorname{tr} \Sigma^{-}(V-\mu D')(D\Delta^{-1}D')^{-1}(V-\mu D')',$$

Received by the editors September 16, 1978.

https://doi.org/10.4153/CMB-1980-006-4 Published online by Cambridge University Press

and the value $\hat{X}$ of X which yields this minimum is

$$\hat{X}=\mu+(V-\mu D')(D\Delta^{-1}D')^{-1}D\Delta^{-1}+\theta[I-D'(D\Delta^{-1}D')^{-1}D\Delta^{-1}], \tag{3}$$

where θ is an arbitrary $p\times N$ matrix. The solution (3) is called a weighted least squares solution, with weight matrix Δ, to the (consistent) system of linear equations

$$DX'=V'. \tag{4}$$

If in (3) we set $\mu=0$, and $\Delta=I$, then the resulting solution is an unweighted least squares solution to (4).

If in (1) X and μ are $p\times p$ symmetric matrices and Δ is a $p\times p$ symmetric matrix, then a particular solution $\hat{X}$ to X is

$$\begin{aligned}\hat{X}&=\mu+(V-\mu D')(D\Delta^{-1}D')^{-1}\\&\quad+\Delta^{-1}D'(D\Delta^{-1}D')^{-1}(V-\mu D')'[I-D'(D\Delta^{-1}D')^{-1}D\Delta^{-1}],\end{aligned} \tag{5}$$

where $V'D'=DV$. In this case also the minimum of (1) is given by (2). Note that the rank of $\hat{X}$ is twice the rank of V.

In case D and V in (1) are partitioned into a number of parts, say $D=(D_1'D_2')'$, $V=(V_1V_2)$, then (2) may be written as

$$\begin{aligned}&\operatorname{tr}\Sigma^{-1}[(V_1-\mu D_1')\phi(V_1-\mu D_1')'\\&\quad+(V_2-\mu D_2'-(V_1-\mu D_1')\phi\rho)\psi(V_2-\mu D_2'-(V_1-\mu D_1')\phi\rho)'],\end{aligned} \tag{6}$$

where

$$\phi=(D_1\Delta^{-1}D_1')^{-1},\qquad \rho=D_1\Delta^{-1}D_2',\qquad \psi=(D_2\Delta^{-1}D_2'-\rho'\phi\rho)^{-1}. \tag{7}$$

The inverse of the partitioned matrix

$$\begin{bmatrix}D_1\Delta^{-1}D_1' & D_1\Delta^{-1}D_2'\\ D_2\Delta^{-1}D_1' & D_2\Delta^{-1}D_2'\end{bmatrix} \tag{8}$$

is given by

$$\begin{bmatrix}\phi+\phi\rho\psi\rho'\phi & -\phi\rho\psi\\ -\psi\rho'\phi & \psi\end{bmatrix} \tag{9}$$

Now consider (1) with double linear restrictions, i.e.,

$$\min \operatorname{tr}\Sigma^{-}(X-\mu)\Delta(X-\mu)',\quad \text{subject to}\quad DX'B'=C'. \tag{10}$$

To solve (10), we first set $DX'=V'$, and write (2) as

$$\operatorname{tr}\phi(V-\mu D')'\Sigma^{-1}(V-\mu D'), \tag{11}$$

and then minimize (11) under the restrictions

$$DX'B'=C',\quad \text{i.e.,}\quad BV=C. \tag{12}$$

https://doi.org/10.4153/CMB-1980-006-4 Published online by Cambridge University Press

The minimum of (11) subject to (12) is given by

$$\text{tr}\ \phi(C-B\mu D')'(B\Sigma B')^{-1}(C-B\mu D'), \tag{13}$$

and the value $\hat{V}$ of V which yields this minimum is

$$\hat{V}=\mu D'+\Sigma B'(B\Sigma B')^{-1}(C-B\mu D'). \tag{14}$$

On substituting $\hat{V}$ for V in (3), a particular solution $\bar{X}$ that minimizes (10) is

$$\bar{X}=\nu+\Sigma B'(B\Sigma B')^{-1}(C-B\mu D')(D\Delta^{-1}D')^{-1}D\Delta^{-1}, \tag{15}$$

and forms a weighted least squares solution to the system (12).

If Δ is deficient in rank, and D, B, C are of specified rank, then Δ can be chosen in such a way that the solution $\hat{X}$ is of a specified rank, see Mitra [4]. However, this is not always possible if X is a symmetric matrix.

We proceed to consider solutions to linear matrix equations. We simply show a procedure to compute a solution and do not discuss necessary and sufficient conditions for the existence of such solutions.

3. **Linear matrix equations.** Khatri and Mitra [3], and Mitra [4, 5, 6], in a number of papers have utilized several available results from statistical literature, especially from multivariate linear regression theory, experimental designs theory, and Rao's [7] MINQUE theory to solve linear matrix equations and supply the required analysis for the existence and other properties of the solutions of such equations. Thus, e.g., Mitra [6] considers the equation

$$D_1X'B_1'+D_2X'B_2'=E', \tag{16}$$

and solves it by a very complicated procedure and further remarks ([6], p. 825, Section 4) that when one or more terms of the same type are added on the lefthand side of (16), a solution to (16) appears difficult to compute. It is true that a unique solution to (16), or even a particular solution to (16), is very difficult to compute by using weighted least squares theory. However, a particular solution by using unweighted squares theory can be easily obtained to the system (16). It is also possible to compute a fixed rank solution to (16). We split up E into parts $E=E_1+E_2$ such that rank $E=$ rank E_1+ rank E_2, where E_1 may be a null matrix, and set

$$D_1X'B_1'=E_1', \qquad D_2X'B_2'=E_2'. \tag{17}$$

We further set

$$D_1X'=V_1', \qquad D_2X'=V_2', \tag{18}$$

and by assuming $\Sigma=I$, $\Delta=I$, $\mu=0$, in (1), minimize (1) subject to the restrictions (18). The minimum, from (6), is

$$\text{tr}\,[V_1\phi V_1'+(V_2-V_1\phi\rho)\psi(V_2-V_1\phi\rho)'], \tag{19}$$

where now $D=(D_1'D_2')'$ and

(20) $$\phi=(D_1D_1')^{-1}, \qquad \rho=D_1D_2', \qquad \psi=(D_2D_2'-\rho'\phi\rho)^{-1}.$$

The value $\hat{X}$ of X which solves this minimum value problem from (3) is

(21) $$\hat{X}=(V_1V_2)(DD')^{-1}D=V_1[\phi D_1+\phi\rho\psi\rho'\phi D_1-\phi\rho\psi D_2] + V_2[-\psi\rho'\phi D_1+\psi D_2].$$

Now we write (16) as

(22) $$B_1V_1+B_2V_2=E,$$

and minimize (19), under the restrictions (22). However, to avoid this difficult minimization procedure, we shall solve

(23) $$B_1V_1=E_1, \qquad B_2V_2=E_2,$$

and substitute the solutions in (21). The unweighted least squares solutions $\hat{V}_1$ and $\hat{V}_2$ to (23) are

(24) $$\hat{V}_1=B_1'(B_1B_1')^{-1}E_1, \qquad \hat{V}_2=B_2'(B_2B_2')^{-1}E_2.$$

On substituting these solutions into (21), a solution $\bar{X}$ that satisfies (16) is

(25) $$\bar{X}=B_1'(B_1B_1')^{-1}E_1[\phi D_1+\phi\rho\psi\rho'\phi D_1-\phi\rho\psi D_2] + B_2'(B_2B_2')^{-1}E_2[-\psi\rho'\phi D_1+\psi D_2].$$

Actually in practice any E_1 and E_2, such that $E=E_1+E_2$, may be chosen. Obviously, this procedure can be applied to any number of terms on the lefthand side of (16).

Note that (25) satisfies the simultaneous equations in (17), see Mitra [5]. We note that B_1 and B_2 may be identity matrices.

We illustrate the procedure for finding a fixed rank solution to a given matrix equation by a numerical example.

EXAMPLE 1. We wish to find a 2×3 matrix X that satisfied

(26) $$DX'=d'X'=[1,2,1]X'=V'=v'=[4,3],$$

and is of rank 2. We choose Δ in (1) such that D does not belong to its range space; such a Δ of rank 2 is

(27) $$\Delta=\begin{bmatrix}1 & -1 & 2\\ -1 & 1 & -2\\ 2 & -2 & 9\end{bmatrix}$$

The matrix Δ is of rank 2, the two nonzero roots of Δ are 1 and 10 and the respective latent vectors are

$$\left[\frac{2}{3},\frac{-2}{3},\frac{-1}{3}\right],\qquad [(18)^{-1/2},-(18)^{-1/2},4(18)^{-1/2}], \tag{28}$$

and the latent vector corresponding to the zero root is proportional to

$$t'=[1,1,0]. \tag{29}$$

The Moore-Penrose Δ^- of Δ is given by

$$20\Delta^-=\begin{bmatrix} 9 & -9 & 4\\ -9 & 9 & 4\\ -4 & 4 & 4\end{bmatrix} \tag{30}$$

We note that D does not belong to the range space of Δ, and hence, following the practice of the maximum likelihood estimation of the mean vector of a singular normal distribution, we restrict the solution $\hat{X}$ to the range space of Δ. By assuming $\mu=0$ in (1), we find that

$$\hat{X}=v(d'\Delta^-d)^{-1}d'\Delta'=\begin{bmatrix}-52 & 52 & 32\\ -39 & 39 & 24\end{bmatrix}/21. \tag{31}$$

We note that $t'\hat{X}'=0$, and that $\hat{X}$ is of rank 2. The theory behind this example is very complicated, see Mitra [4]. This procedure does not always suceed if X is symmetric, although it succeeds in finding fixed rank solutions to systems (16) and (17).

We note that (5) is a more general solution to (4), if X is symmetric, than the one obtained by Khatri and Mitra ([3, p. 579, Theorem 2.1).

It is possible to extend Khatri and Mitra's [3] theory to obtain symmetric solutions (although not of fixed ranks) to all types of linear matrix equations given in Section 1. Thus, e.g., we give a symmetric solution to the system

$$D_1X'=V_1',\qquad D_2X'B_2'=E_2'. \tag{32}$$

We first set

$$D_1X'=V_1',\qquad D_2X'=V_2',\qquad D=(D_1'D_2')', \tag{33}$$

and then, setting $\mu=0$, $\Delta=I$, $\Sigma=I$ in (5), we find that

$$\hat{X}=(V_1V_2)(DD')^{-1}D \tag{34}$$

satisfies (33). Further, by setting $B_2V_2=E_2$, we find that $\hat{V}_2=B_2'(B_2B_2')^{-1}E_2$ satisfies $B_2V_2=E_2$. Thus a symmetric $\bar{X}$ that satisfies (32) is

$$\begin{aligned} X&=(V_1,B_2'(B_2B_2')^{-1}E_2)(DD')^{-1}D\\ &\quad+D'(DD')^{-1}(V_1,B_2'(B_2B_2')^{-1}E_2)'[I-D'(DD')^{-1}D], \end{aligned} \tag{35}$$

provided that $V_1'D_1'=D_1V_1$ and $B_2'D_2'=D_2B_2$.

https://doi.org/10.4153/CMB-1980-006-4 Published online by Cambridge University Press

Similarly, we may obtain a symmetric solution to $D_1X'B_1'+D_2X'B_2'=E$. It is difficult to obtain a weighted least squares solution, expressible in the standard solutions format, to the equations $FX=W$, $XH=T$.

Now by a numerical example we illustrate our procedure for obtaining a nonsymmetric solution X to the system $D_1X'B_1'+D_2X'B_2'=E'$.

EXAMPLE 2. We consider the equation

(36) $d_1'X'b_1+d_2'X'b_2=1, \qquad d_1'=[1,2], \qquad b_1'=[0,1], \qquad d_2'=[2,1],$

and $b_2'=(4,3]$. We fitst set

$$d_1'X'=0, \qquad d_2'X'=v_2'=(v_{21},v_{22}), \tag{37}$$

and, by using (31), compute the solution $\hat{X}$ to (37) given by

$$\hat{X}=v_2[-\psi\rho'\phi d_1'+\psi d_2'], \tag{38}$$

where

(39) $\phi=(d_1'd_1)^{-1}=5^{-1}, \qquad \rho=d_1'd_2=4, \qquad \psi=(d_2'd_2-\rho'\phi\rho)^{-1}=5/9.$

By using (38) and (39) we have that

$$\hat{X}=\begin{bmatrix} v_{21} \\ v_{22} \end{bmatrix}\left[\frac{2}{3},\frac{-1}{3}\right], \tag{40}$$

where $v_2'b_2=4\ v_{21}+3\ v_{22}=1$, i.e., $v_{21}=4/25$, $v_{22}=3/25$, and hence we have that

$$\hat{X}=\begin{bmatrix} 8 & -4 \\ 6 & -3 \end{bmatrix}/75 \tag{41}$$

which satisfies (36). Thus, a particular solution to (16) with any number of members on the left-hand side of (16) can always be computed.

4. **The matrix equation TAT = T.** Let A be any $p\times p$ positive definite symmetric matrix and X and Y be any two $q\times p$, $q\leq p$, matrices of rank q. Then the Gramian matrix

$$H=A^{-1}-A^{-1}X'(YA^{-1}X')^{-1}YA^{-1}, \tag{42}$$

is exactly of rank $(p-q)$. If R, $(p-q)\times p$, is a semi-orthogonal matrix orthogonal to X, and P, $(p-q)\times p$, is a semi-orthogonal matrix orthogonal to Y, then by writing $H=P'GR$ for some $p\times pG$, we note that $RAP'GR=RAH=R$, i.e., $GR=(RAP')^{-1}R$ or $H=P'(RAP')^{-1}R$. On the other hand, it is known that

$$H=(EAF)^{-}, \tag{43}$$

https://doi.org/10.4153/CMB-1980-006-4 Published online by Cambridge University Press

where E and F are idempotent matrices and G^- denotes the Moore-Penrose inverse of G. Now set $X = Y$ in (42) and, using (43), find that

$$H = R'(RAR')^{-1}R. \tag{44}$$

In (44) we assume the symmetric matrix A to be of rank $(p-q)$, and R to be the $(p-q)\times p$ matrix of the first $(p-q)$ latent vectors of A corresponding to the first $(p-q)$ largest roots. In this case (44) gives

$$A^- = R'(RAR')^{-1}R = (EAE)^-. \tag{45}$$

On setting $T = Q^-$ in (43), we find that

$$TAT = Y = (EAE)^-, \tag{46}$$

has the solution $T = A^- = (EAE)^-$. The structure of (45) shows that no other inverse of A, except the Moore–Penrose inverse of A, can be written as $(EAE)^-$. If A is non-symmetric, then under certain conditions on P and R we may write $A^- = (EAF)^- = P'(RAP')^{-1}R$. Hence T must equal the Moore–Penrose inverse of A in (46). This answers a question raised by Greville ([1], p. 829, lines 14 and 15).

This research is supported by National Research Council of Canada Grant A-4018.

References

1. T. N. E. Greville, *Solution of the matrix equation $XAX = X$ and relations between oblique and orthogonal projectors*, SIAM J. Appl. Math., **26** (1974), 828–834.

2. D. G. Kabe, *Minima of vector quadratic forms with applications to statistics*, Metrika **12** (1968), 155–160.

3. C. G. Khatri and S. K. Mitra, *Hermitian and nonnegative definite solutions of linear matrix equations*, SIAM J. Appl. Math., **31** (1976), 579–585.

4. S. K. Mitra, *Fixed rank solutions to linear matrix equations*, Sankhya Ser. A, **35** (1972), 387–392.

5. ——, *Common solutions to a pair of matrix equations*, $A_1XB_1 = C_1$, $A_2XB_2 = C_2$, Proc. Cambridge Phil. Soc., **74,** (1973), 213–216.

6. ——, *The matrix equation $AXB + CXD = E$*, SIAM J. Appl. Math. **32** (1977), 823–825.

7. C. R. Rao, *Estimation of variance and covariance components in linear models*, J. Amer, Statisti. Assoc., **67** (1972), 112–115.

Department of Mathematics
St. Mary's University
Halifax, N.S., B3H 3C3