ON SOLVING MATRIX EQUATIONS USING THE VEC OPERATOR*

BU-620-M | by | July, 1977

H. V. Henderson[1], S. R. Searle and J. B. Kadane[2]
Biometrics Unit, Cornell University, Ithaca, New York

1. Introduction

A general form of matrix equation in the unknown matrix $\underline{X}$ is

$$\sum_{i=1}^{k} \underline{A}_i \underline{X} \underline{B}_i + \sum_{j=1}^{\ell} \underline{D}_j \underline{X}' \underline{E}_j = \underline{C} \tag{1}$$

where $\underline{X}$ can be rectangular and where all matrices can have elements in the complex field. Many special cases of (1) are considered in the literature, particularly the form

$$\sum_{i=1}^{k} \underline{A}_i \underline{X} \underline{B}_i = \underline{C} \tag{2}$$

which is (1) without terms in $\underline{X}'$. Lancaster [1970] reviews numerous cases of (2), introducing a lemma of Krein's, involving contour integrals, to do so. He particularly presents techniques for solving the special form

$$\sum_{i,j=0}^{k} \alpha_{ij} \underline{A}^i \underline{X} \underline{B}^j = \underline{C} \tag{3}$$

of (2) and also the much discussed but simpler form

---

* Paper No. BU-620-M in the Biometrics Unit Mimeo Series, Cornell University.

1 Supported by a New Zealand National Research Advisory Council Post-Graduate Fellowship.

2 Department of Statistics, Carnegie-Mellon University, Pittsburgh, Pennsylvania.

$$\underset{\sim}{A}\underset{\sim}{X} + \underset{\sim}{X}\underset{\sim}{B} = \underset{\sim}{C} \tag{4}$$

where $\underset{\sim}{A}$ and $\underset{\sim}{B}$ are necessarily square, which is also reviewed and extended by Hartwig [1975]. He does not, however, consider the special case of (4) which arises in Schönemann's [1966] solution to the Orthogonal Procrustes problem, in which $\underset{\sim}{X}$ is taken as being square and orthogonal and a solution to (4) is then sought when $\underset{\sim}{A}$ and $\underset{\sim}{B}$ are symmetric. Hearon [1977] discusses the existence of non-singular solutions to (4). Another special case of (2) is $\underset{\sim}{AXB} = \underset{\sim}{C}$ dealt with by Rao and Mitra [1971], with solution $\underset{\sim}{X} = \underset{\sim}{A}^{-}\underset{\sim}{CB}^{-} + \underset{\sim}{Z} - \underset{\sim}{A}^{-}\underset{\sim}{AZBB}^{-}$ provided the consistency condition $\underset{\sim}{AA}^{-}\underset{\sim}{CB}^{-}\underset{\sim}{B} = \underset{\sim}{C}$ holds, where $\underset{\sim}{A}^{-}$ is a generalized inverse of $\underset{\sim}{A}$ satisfying $\underset{\sim}{AA}^{-}\underset{\sim}{A} = \underset{\sim}{A}$, and where $\underset{\sim}{Z}$ is arbitrary of appropriate order. A generalization of (3) considered by Wimmer and Ziebur [1972] is $\sum_{i=1}^{k} f_i(\underset{\sim}{A})\underset{\sim}{X}g_i(\underset{\sim}{B}) = \underset{\sim}{C}$, where f_i and g_i are analytic in the neighborhood of the eigenvalues of $\underset{\sim}{A}$ and $\underset{\sim}{B}$ respectively.

Two more special cases of (1) are the pair of equations

$$\underset{\sim}{A}'\underset{\sim}{X} \pm \underset{\sim}{X}'\underset{\sim}{A} = \underset{\sim}{C} . \tag{5}$$

Although not mentioned by Lancaster [1970], they are discussed by Hodges [1957] as two distinctly different equations, a distinction which we find to be unnecessary. He gives explicit solutions for $\underset{\sim}{X}$ when all matrices are square and $\underset{\sim}{A}$ is non-singular. In section 3 we extend this to rectangular $\underset{\sim}{X}$ and $\underset{\sim}{A}$ of full column rank.

Although Lancaster suggests that equation (2) "is far from tractable", the purpose of this note is to suggest that even its more general form (1) does admit an explicit solution for $\underset{\sim}{X}$. It is achieved by using Kronecker (direct) products and the vec operator to rewrite (1) in familiar vector form typified by $\underset{\sim}{Ax} = \underset{\sim}{b}$ with solutions $\underset{\sim}{x} = \underset{\sim}{A}^{-}\underset{\sim}{b} + (\underset{\sim}{I} - \underset{\sim}{A}^{-}\underset{\sim}{A})\underset{\sim}{z}$ for arbitrary $\underset{\sim}{z}$ and $\underset{\sim}{A}$ and $\underset{\sim}{b}$ given.

2. The Vector Solution

2.1. The vec operator

For convenience we define the vec operator as given, for example, in Neudecker [1969]. For $\underset{\sim}{X}_{m\times n}$ partitioned as $\underset{\sim}{X} = [\underset{\sim}{x}_1 \cdots \underset{\sim}{x}_n]$, where $\underset{\sim}{x}_j$ is the j^{th} column of $\underset{\sim}{X}$, $\text{vec}\underset{\sim}{X}$ is defined as the $mn \times 1$ vector formed by stacking the columns of $\underset{\sim}{X}$ as a single column:

$$\text{vec}\underset{\sim}{X} = \begin{bmatrix} \underset{\sim}{x}_1 \\ \vdots \\ \underset{\sim}{x}_n \end{bmatrix} . \qquad (6)$$

With this definition it is clear that the vec of the transpose $\underset{\sim}{X}'$ is a vector containing the same elements as $\text{vec}\underset{\sim}{X}$. Indeed for $\underset{\sim}{X}_{m\times n}$

$$\text{vec}\underset{\sim}{X}' = \underset{\sim}{I}_{(n,m)}\text{vec}\underset{\sim}{X} \qquad (7)$$

where $\underset{\sim}{I}_{(m,n)}$ is the $mn \times mn$ permuted identity matrix defined in MacCrae [1974] with its ij^{th} $m \times n$ submatrix having 1 in position (j,i) and 0 elsewhere. She gives the identities: $\underset{\sim}{I}_{(m,1)} = \underset{\sim}{I}_{(1,m)} = \underset{\sim}{I}_m$, $\underset{\sim}{I}_{(m,n)} = \underset{\sim}{I}'_{(n,m)}$ and $\underset{\sim}{I}_{(m,n)}\underset{\sim}{I}_{(n,m)} = \underset{\sim}{I}_{mn}$, a particular case of the latter being $\underset{\sim}{I}^2_{(n,n)} = \underset{\sim}{I}_{n^2}$.

2.2. Kronecker products

For convenience we recall the definitions of the Kronecker (direct) product of two matrices $\underset{\sim}{A}_{m\times n}$ and $\underset{\sim}{B}_{p\times q}$. It is the $mp \times nq$ matrix $A \otimes B = \{a_{ij}\underset{\sim}{B}\}$ for $i = 1 \cdots m$ and $j = 1 \cdots n$ with a variety of well-known properties, for example,

$$(\underset{\sim}{A} \otimes \underset{\sim}{B})(\underset{\sim}{C} \otimes \underset{\sim}{D}) = \underset{\sim}{AC} \otimes \underset{\sim}{BD} \quad \text{and} \quad (\underset{\sim}{A} \otimes \underset{\sim}{B})^- = \underset{\sim}{A}^- \otimes \underset{\sim}{B}^- .$$

Our need for Kronecker products is their use in the vec of a matrix product:

$$\text{vec}\underline{ABC} = (\underline{C}' \otimes \underline{A})\text{vec}\underline{B} \tag{8}$$

a result given, for example, in Neudecker [1969]. MacCrae [1974] shows that permuted identity matrices may be used to reverse the order of a Kronecker product with $(\underline{B} \otimes \underline{A}) = \underline{I}_{(m,p)}(\underline{A} \otimes \underline{B})\underline{I}_{(q,n)}$.

2.3. The solution

Apply the vec operator (6) to equation (1) and get

$$\sum_{i=1}^{k} \text{vec}(\underline{A}_i\underline{X}\underline{B}_i) + \sum_{j=1}^{\ell} \text{vec}(\underline{D}_j\underline{X}'\underline{E}_j) = \text{vec}\underline{C} .$$

Using (8) gives

$$\sum_{i=1}^{k} (\underline{B}_i' \otimes \underline{A}_i)\text{vec}\underline{X} + \sum_{j=1}^{\ell} (\underline{E}_j' \otimes \underline{D}_j)\text{vec}\underline{X}' = \text{vec}\underline{C} .$$

Then by (7)

$$\left[\sum_{i=1}^{k} (\underline{B}_i' \otimes \underline{A}_i) + \sum_{j=1}^{\ell} (\underline{E}_j' \otimes \underline{D}_j)\underline{I}_{(n,m)}\right]\text{vec}\underline{X} = \text{vec}\underline{C} \tag{9}$$

which, on defining $\underline{G}$ as the matrix pre-multiplying vec$\underline{X}$ can be written as

$$\underline{G}\text{vec}\underline{X} = \text{vec}\underline{C} . \tag{10}$$

The existence and uniqueness of a solution vec$\underline{X}$ to (10) now determines the existence and uniqueness of a solution $\underline{X}$, of given dimensions, to (1). Since vec$\underline{X}$ is a vector, it is clear that (10) has a solution if and only if

$$r[\underline{G} \quad \text{vec}\underline{C}] = r(\underline{G}) \tag{11}$$

where $r(G)$ is the rank of G. When this condition holds, both (1) and (10) are consistent and solutions to (10) are, for arbitrary Z of the same order as X,

$$\text{vec}X = G^{-}\text{vec}C + [I - G^{-}G]\text{vec}Z \ . \tag{12}$$

This extends the approach of Hartwig [1975] for the special case (4) to the general case (1).

Defining vec^{-1} as the inverse vec operator which re-forms X from $\text{vec}X$ so that $\text{vec}^{-1}(\text{vec}X) = X$, we have from (12)

$$X = \text{vec}^{-1}(G^{-}\text{vec}C + [I - G^{-}G]\text{vec}Z) \ . \tag{13}$$

The solution is unique when G is non-singular:

$$X = \text{vec}^{-1}(G^{-1}\text{vec}C) \ . \tag{14}$$

These are our general solutions to equation (1).

3. Special Cases

The general solution (13) is now applied to some special cases of equation (1) with extensions, not involving vec^{-1}, to well-known results.

3.1. $A'X \pm X'A = C$

Hodges' [1957] explicit solutions to (5), when all matrices are square and A is non-singular, can be extended to rectangular X and A when A has full column rank. In this case A has a left inverse L, say, for example, $(A'A)^{-1}A'$. Rewriting (5) leads to (10) with

$$G = [I_{n^2} \pm I_{(n,n)}](I_n \otimes A') \ . \tag{15}$$

For the solution (13) we need G^- which is obtainable in this case by noticing that for k a positive integer $[I_{n^2} \pm I_{(n,n)}]^k = 2^{k-1}[I_{n^2} \pm I_{(n,n)}]$, because $I^2_{(n,n)} = I_{n^2}$, then, since $LA = I$, a generalized inverse of (15) is

$$G^- = (I_n \otimes A')^-(I_{n^2} \pm I_{(n,n)})^- = \tfrac{1}{4}(I_n \otimes L')(I_{n^2} \pm I_{(n,n)}) .$$

Furthermore, for (13)

$$G^-G = \tfrac{1}{4}(I_n \otimes L')(I_{n^2} \pm I_{(n,n)})^2(I_n \otimes A') = \tfrac{1}{2}[(I_n \otimes L'A') \pm (A' \otimes L')I_{(n,m)}]. \quad (16)$$

Substituting into (13) gives

$$X = \tfrac{1}{2}[L'C + ZZ - L'(A'Z \pm Z'A)] \quad (17)$$

as the solutions to (5) when (11) holds where Z is an arbitrary m × n matrix.

When X and A are square with A non-singular (the case considered by Hodges), $L = A^{-1}$ and so (17) simplifies to

$$X = \tfrac{1}{2}(A^{-1'}C + Z \mp A^{-1'}Z'A) . \quad (18)$$

Hodges' solutions can be written as $X = P'(\pm K + \tfrac{1}{2}Q'CQ)Q^{-1}$ where $K = \mp K'$ but otherwise arbitrary and $PAQ = I$ with P and Q non-singular. This permits X to be rewritten as $X = \pm P'KQ^{-1} + \tfrac{1}{2}A^{-1'}C$, which generates the arbitrary part of the solution in a different manner to that in (18). The relationship between the two being $\pm 2K = P'^{-1}(Z \mp A^{-1'}Z'A)Q = P'^{-1}ZQ \mp Q'ZP^{-1} = -K'$. So $K = \mp K'$ as required by Hodges.

$$\underline{AX + XB = C,\ A = A',\ B = B',\ X'X = XX' = I} \quad (19)$$

This special case of (4) arises in Schönemann's [1966] solution to the Orthogonal Procrustes problem. Schönemann gives a solution to a consequence of (19)

$$XC'CX' = CC' \tag{20}$$

which may be obtained by taking the transpose of (4) together with the symmetry of A and B to give $X'A + BX' = C'$. Pre- and post-multiplying by X gives $AX + XB = XC'X$ which by (4) also equals C. So,

$$XC' = CX' \tag{21}$$

and then (20) follows by post-multiplying (21) by its transpose.

Note necessary but not sufficient conditions for a solution to (19) are given by (20), which Schönemann uses, or (21) a point that appears to have been overlooked by Schönemann. For example, if C is symmetric then $C'C = CC'$ so that $X = I$ is a solution to (20) but not a solution to (19) as $A + B$ need not equal C.

References

Hartwig, R. E. [1975]. AX - XB = C, resultants and generalized inverses. SIAM J. Applied Math. 28:154-183.

Hearon, J. Z. [1977]. Non-singular solutions of TA - BT = C. Linear Algebra and Its Applications 16:57-63.

Hodges, J. H. [1957]. Some matrix equations over a finite field. Annali di Matematica XLIV:245-250.

Lancaster, P. [1970]. Explicit solutions of linear matrix equations. SIAM Review 12:544-566.

MacCrae, E. C. [1974]. Matrix derivatives with an application to an adaptive linear decision problem. Annals of Statistics 2:337-346.

Neudecker, H. [1969]. Some theorems on matrix derivatives in multivariate analysis. J.A.S.A. 62:607-625.

Rao, C. R. and Mitra, S. K. [1971]. Further contributions to the theory of generalized inverse of matrices and its applications. Sankhyā 33:289-300.

Schönemann, P. H. [1966]. A generalized solution of the orthogonal Procrustes problem. Psychometrika 31:1-10.

Wimmer, H. and Ziebur, A. D. [1972]. Solving the matrix equation $\sum_{p=1}^{r} f_p(A)Xg_p(B) = C$. SIAM Review 14:318-323.