A NOTE ON SOLVING MATRIX EQUATIONS USING THE VEC OPERATOR*

by

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

Abstract

Solving linear matrix equations in an unknown matrix $\underset{\sim}{X}$ is discussed in terms of Kronecker products and the vec operator.

1. Introduction

Many special cases of the general matrix equation

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

are considered in the literature, especially cases with $\ell = 0$; e.g., Lancaster [1970], Rao and Mitra [1971] and Wimmer and Ziebur [1972], and the special case, $\underset{\sim}{A}\underset{\sim}{X} + \underset{\sim}{X}\underset{\sim}{B} = \underset{\sim}{C}$, reviewed by Hartwig [1975]. Other special cases with $\ell \neq 0$ are $\underset{\sim}{A}'\underset{\sim}{X} \pm \underset{\sim}{X}'\underset{\sim}{A} = \underset{\sim}{C}$, discussed by Hodges [1957]. In this note we derive an explicit solution to (1) and indicate its application to Hodges' special cases, suggesting, in so doing, that some of his limitations can be relaxed.

---

* Paper No. BU-626-M in the Biometrics Unit Mimeo Series, Cornell University.

[1] Department of Statistics, Carnegie-Mellon University, Pittsburgh, Pennsylvania.

[2] Supported by a New Zealand National Research Advisory Council Post-Graduate Fellowship.

Our approach is a generalization of Hartwig [1975] to solve (1) for $\underset{\sim}{X}$ by using Kronecker products and the vec operation of matrices so as to rewrite (1) in vector form $\underset{\sim}{A}\underset{\sim}{x} = \underset{\sim}{b}$ 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 $\underset{\sim}{A}^{-}$ being a generalized inverse of $\underset{\sim}{A}$ such that $\underset{\sim}{A}\underset{\sim}{A}^{-}\underset{\sim}{A} = \underset{\sim}{A}$. To do this, recall the definition of the Kronecker product of two matrices $\underset{\sim}{A}_{m\times n}$ and $\underset{\sim}{B}_{p\times q}$ as being the $mp \times nq$ matrix $\underset{\sim}{A} \otimes \underset{\sim}{B} = \{a_{ij}\underset{\sim}{B}\}$ for $i = 1 \cdots m$ and $j = 1 \cdots n$. And, for $\underset{\sim}{X}_{m\times n}$ partitioned as $\underset{\sim}{X} = [\underset{\sim}{x}_1 \cdots \underset{\sim}{x}_n]$ with $\underset{\sim}{x}_j$ being the j^{th} column, the vec operator as defined, for example, by Neudecker [1969] gives

$$\text{vec}\underset{\sim}{X} = [\underset{\sim}{x}_1' \; \underset{\sim}{x}_2' \cdots \underset{\sim}{x}_n']' \, , \tag{2}$$

so that $\text{vec}\underset{\sim}{X}$ is the column vector consisting of the columns of $\underset{\sim}{X}$ stacked as a single column. Then, as shown by MacCrae [1974],

$$\text{vec}\underset{\sim}{X}' = \underset{\sim}{I}_{(n,m)}\text{vec}\underset{\sim}{X} \tag{3}$$

where $\underset{\sim}{I}_{(m,n)}$ is the $mn \times mn$ permuted identity matrix with its $(i,j)^{th}$ $m \times n$ sub-matrix having 1 in position (j,i) and 0 elsewhere. Furthermore,

$$\text{vec}\underset{\sim}{A}\underset{\sim}{B}\underset{\sim}{C} = (\underset{\sim}{C}' \otimes \underset{\sim}{A})\text{vec}\underset{\sim}{B} \tag{4}$$

and

$$(\underset{\sim}{B} \otimes \underset{\sim}{A}) = \underset{\sim}{I}_{(m,p)}(\underset{\sim}{A} \otimes \underset{\sim}{B})\underset{\sim}{I}_{(q,n)} \, , \tag{5}$$

as given in Neudecker [1969] and MacCrae [1974], respectively.

2. The Vector Solution

We solve (1) by applying to it the vec operator defined in (2), together with (3) and (4), to get

$$\left[\sum_{i=1}^{k}(\underset{\sim}{B}_i' \otimes \underset{\sim}{A}_i) + \sum_{j=1}^{\ell}(\underset{\sim}{E}_j' \otimes \underset{\sim}{D}_j)\underset{\sim}{I}_{(n,m)}\right]\text{vec}\underset{\sim}{X} = \text{vec}\underset{\sim}{C} \; . \qquad (6)$$

Then defining $\underset{\sim}{G}$ as the matrix pre-multiplying vec$\underset{\sim}{X}$, gives (6) as

$$\underset{\sim}{G}\text{vec}\underset{\sim}{X} = \text{vec}\underset{\sim}{C} \; . \qquad (7)$$

The existence and uniqueness of a solution vec$\underset{\sim}{X}$ to (7) now determines the existence and uniqueness of a solution $\underset{\sim}{X}$, of given dimensions, to (1). It is clear that (7) has a solution if and only if $\underset{\sim}{G}$ and the augmented matrix $[\underset{\sim}{G} \vdots \text{vec}\underset{\sim}{C}]$ have the same rank. When this condition holds, both (1) and (7) are consistent and solutions to (7) are, for arbitrary $\underset{\sim}{Z}$ of the same order as $\underset{\sim}{X}$,

$$\text{vec}\underset{\sim}{X} = \underset{\sim}{G}^{-}\text{vec}\underset{\sim}{C} + [\underset{\sim}{I} - \underset{\sim}{G}^{-}\underset{\sim}{G}]\text{vec}\underset{\sim}{Z} \; . \qquad (8)$$

Defining vec^{-1} as the inverse vec operator which re-forms $\underset{\sim}{X}$ from vec$\underset{\sim}{X}$ so that $\text{vec}^{-1}(\text{vec}\underset{\sim}{X}) = \underset{\sim}{X}$, we have explicit solutions for $\underset{\sim}{X}$ from (8)

$$\begin{aligned} \underset{\sim}{X} &= \text{vec}^{-1}(\underset{\sim}{G}^{-}\text{vec}\underset{\sim}{C} + [\underset{\sim}{I} - \underset{\sim}{G}^{-}\underset{\sim}{G}]\text{vec}\underset{\sim}{Z}) \; , \\ &= \text{vec}^{-1}(\underset{\sim}{G}^{-}\text{vec}\underset{\sim}{C} + [\underset{\sim}{I} - \underset{\sim}{G}^{-}\underset{\sim}{G}]\underset{\sim}{z}) \; , \qquad (9) \end{aligned}$$

where $\underset{\sim}{z}$ is an arbitrary vector. When $\underset{\sim}{G}$ is non-singular the solution is unique:

$$\underset{\sim}{X} = \text{vec}^{-1}(\underset{\sim}{G}^{-1}\text{vec}\underset{\sim}{C}) \; .$$

3. The equations $\underset{\sim}{A}'\underset{\sim}{X} \pm \underset{\sim}{X}'\underset{\sim}{A} = \underset{\sim}{C}$

Hodges [1957] develops solutions for the pair of equations

$$\underset{\sim}{A}'\underset{\sim}{X} \pm \underset{\sim}{X}'\underset{\sim}{A} = \underset{\sim}{C} \qquad (10)$$

treating them as two distinct equations, a distinction we find to be unnecessary.

Also, his solutions are confined to having $\underset{\sim}{A}$ non-singular and all matrices square. We extend this to $\underset{\sim}{X}$ and $\underset{\sim}{A}$ being rectangular of order m × n with $\underset{\sim}{A}$ having full column rank, whereupon $\underset{\sim}{A}$ has a left inverse, $\underset{\sim}{L} = (\underset{\sim}{A}'\underset{\sim}{A})^{-1}\underset{\sim}{A}'$, say (which is also a generalized inverse of $\underset{\sim}{A}$). Then, on rewriting (10) in the form of (6) and using (5) and the identity $\underset{\sim}{I}_{(m,n)}\underset{\sim}{I}_{(n,m)} = \underset{\sim}{I}_{mn}$ given by MacCrae [1974], (10) reduces to (7) with

$$\underset{\sim}{G} = [\underset{\sim}{I}_{n^2} \pm \underset{\sim}{I}_{(n,n)}](\underset{\sim}{I}_n \otimes \underset{\sim}{A}') \ . \qquad (11)$$

For the solution (9) we need $\underset{\sim}{G}^-$, which is obtainable in this case by noticing that for k a positive integer $[\underset{\sim}{I}_{n^2} \pm \underset{\sim}{I}_{(n,n)}]^k = 2^{k-1}[\underset{\sim}{I}_{n^2} \pm \underset{\sim}{I}_{(n,n)}]$, because $\underset{\sim}{I}^2_{(n,n)} = \underset{\sim}{I}_{n^2}$. Then, since $\underset{\sim}{L}\underset{\sim}{A} = \underset{\sim}{I}$, a generalized inverse of (11) is

$$\underset{\sim}{G}^- = (\underset{\sim}{I}_n \otimes \underset{\sim}{A}')^-[\underset{\sim}{I}_{n^2} \pm \underset{\sim}{I}_{(n,n)}]^- = \tfrac{1}{4}(\underset{\sim}{I}_n \otimes \underset{\sim}{L}')[\underset{\sim}{I}_{n^2} \pm \underset{\sim}{I}_{(n,n)}] \ .$$

Furthermore, for (9)

$$\underset{\sim}{G}^-\underset{\sim}{G} = \tfrac{1}{4}(\underset{\sim}{I}_n \otimes \underset{\sim}{L}')[\underset{\sim}{I}_{n^2} \pm \underset{\sim}{I}_{(n,n)}]^2(\underset{\sim}{I}_n \otimes \underset{\sim}{A}') = \tfrac{1}{2}[(\underset{\sim}{I}_n \otimes \underset{\sim}{L}'\underset{\sim}{A}') \pm (\underset{\sim}{A}' \otimes \underset{\sim}{L}')\underset{\sim}{I}_{(n,m)}] \ ,$$

and substituting into (9) gives, after repeated use of (3), (4) and (5) and $\underset{\sim}{C} = \pm \underset{\sim}{C}'$,

$$\underset{\sim}{X} = \tfrac{1}{2}[\underset{\sim}{L}'\underset{\sim}{C} + 2\underset{\sim}{Z} - \underset{\sim}{L}'(\underset{\sim}{A}'\underset{\sim}{Z} \pm \underset{\sim}{Z}'\underset{\sim}{A})] \qquad (12)$$

as solutions to (10) where $\underset{\sim}{Z}$ is an arbitrary m × n matrix, provided the consistency condition, referred to earlier, holds true.

When $\underset{\sim}{X}$ and $\underset{\sim}{A}$ are square with $\underset{\sim}{A}$ non-singular (the case considered by Hodges), $\underset{\sim}{L} = \underset{\sim}{A}^{-1}$ and so (12) simplifies to

$$\underset{\sim}{X} = \tfrac{1}{2}(\underset{\sim}{A}^{-1'}\underset{\sim}{C} + \underset{\sim}{Z} \mp \underset{\sim}{A}^{-1'}\underset{\sim}{Z}'\underset{\sim}{A}) \ . \qquad (13)$$

Hodges writes his solutions as $\underset{\sim}{X} = \underset{\sim}{P}'(\pm \underset{\sim}{K} + \frac{1}{2}\underset{\sim}{Q}'\underset{\sim}{C}\underset{\sim}{Q})\underset{\sim}{Q}^{-1}$ with $\underset{\sim}{K} = \mp \underset{\sim}{K}'$ but otherwise arbitrary, and $\underset{\sim}{P}\underset{\sim}{A}\underset{\sim}{Q} = \underset{\sim}{I}$ with $\underset{\sim}{P}$ and $\underset{\sim}{Q}$ non-singular. This simplifies to $\underset{\sim}{X} = \frac{1}{2}\underset{\sim}{A}^{-1'}\underset{\sim}{C} \pm \underset{\sim}{P}'\underset{\sim}{K}\underset{\sim}{Q}^{-1}$, which generates the arbitrary part of the solution in a different manner to that in (13), the relationship between the two being $\pm 2\underset{\sim}{K} = \underset{\sim}{P}'^{-1}\underset{\sim}{Z}\underset{\sim}{Q} \mp \underset{\sim}{Q}'\underset{\sim}{Z}'\underset{\sim}{P}^{-1}$, with $\underset{\sim}{K} = \mp \underset{\sim}{K}'$ as required by Hodges.

References

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

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. JASA 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.

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.