L-structured matrices and linear matrix equations

Magnus, J.R.

Published in:
Linear and Multilinear Algebra

Publication date:
1983

L-structured Matrices and Linear Matrix Equations*

JAN R. MAGNUS

The London School of Economics, London WC2A 2AE, UK

Conditions for the existence of solutions, and the general solution of linear matrix equations are given, when it is known *a priori* that the solution matrix has a given structure (e.g. symmetric, triangular, diagonal). This theory is subsequently extended to matrix equations that are linear in several unknown 'structured' matrices, and to partitioned matrix equations.

1. INTRODUCTION AND NOTATION

Linear matrix equations can take a variety of forms, simple examples of which are

$$AXB = C, \qquad AX + XB = C,$$
$$A'X + X'A = C, \qquad AX_1 - X_2B = C.$$

Conditions for the existence of solutions and the general solution of these equations are well-known [22, 2], either directly from the matrix equation, or indirectly from the equivalent vector equation.

The purpose of this paper is to solve linear matrix equations, when it is *a priori* known that the solution matrix X (or the solution matrices $X_1 \ldots X_r$) has a given structure (symmetric, triangular, diagonal, or otherwise).

Patterned matrices (with equality relationships only among their elements) were studied by Tracy and Singh [29] with the purpose of finding matrix derivatives of certain matrix transformations, and to use these in evaluating their Jacobians. Vetter [31] discussed linear matrix equations where the solution matrix is known to be symmetric, while Henderson and Searle [9] and Magnus and Neudecker [15] recently derived many new results that are relevant for transformations involving symmetric or lower triangular matrices.

In all these papers, only equality relationships among the elements of the solution matrix X are assumed. In the present paper any set of linear relationships between the elements of X is permitted. The totality of real matrices of a given order that satisfy a given set of linear restrictions forms an L-structure ("L" stands for linear). This concept is defined and discussed in section 2. In section 3 the equation $Q \operatorname{vec} X = \operatorname{vec} C$ is solved, where X is L-structured. A more general class of matrices, the "extended" L-structure, is introduced and the solution of $Q \operatorname{vec} X = \operatorname{vec} C$ provided for that case. Also, the solution of a system of equations in one L-structured X is discussed. Section 4 solves the linear equation in several unknown L-structured matrices and section 5 deals with partitioned matrices. In section 6 some well-known L-structures (e.g., symmetry and triangularity) are characterized. An appendix, listing some properties of the Moore–Penrose inverse, concludes the paper.

This paper deals with L-structured matrices. Since, however, the class of L-structured matrices embodies 'unstructured' matrices (where no restrictions are placed on its elements) as a special case, the theorems derived in this paper apply to unstructured matrices in particular.

All matrices are real; capital letters represent matrices; lowercase letters denote vectors or scalars. An (m, n) matrix is one having m rows and n columns; A^{+} denotes the Moore–Penrose (MP) inverse of A; the identity matrix of order s is denoted I_s. If A is an (m, n) matrix, then $\operatorname{vec} A$ is the $(mn, 1)$ vector that stacks the columns of A one underneath the other. Thus, for a_i, $i = 1 \ldots n$, being the n columns of A,

$$A = (a_1 a_2 \ldots a_n) \quad \text{and} \quad \operatorname{vec} A = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}.$$

The Kronecker product of an (m, n) matrix $A = (a_{ij})$ and an (s, t) matrix B is the (ms, nt) matrix

$$A \otimes B = (a_{ij}B).$$

Well-known properties of the Kronecker product, see e.g. [3, p. 235], [16, p. 8] or [20], include

$$(A \otimes B)(C \otimes D) = AC \otimes BD, \qquad (A \otimes B)' = A' \otimes B',$$

$$(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}, \qquad (A \otimes B)^{+} = A^{+} \otimes B^{+},$$

if the necessary rank and conformability conditions in these expressions are satisfied. A standard result on vecs, due to Roth [24] and rediscovered by Neudecker [20], is

$$\text{vec } ABC = (C' \otimes A)\text{vec } B, \tag{1.1}$$

if the matrix product ABC exists.

The commutation matrix K_{mn}, a row-permutation of the identity matrix I_{mn}, was introduced by Tracy and Dwyer [28] and studied in detail by Magnus and Neudecker [14]. It is an orthogonal (mn, mn) matrix, defined implicitly by the relationship

$$K_{mn}\text{vec } A = \text{vec } A', \tag{1.2}$$

where A is an arbitrary (m, n) matrix.

In the following, orders of matrices are deleted when clear from the context or irrelevant.

2. L-STRUCTURED MATRICES

The object of this section is to define the concept of an L-structure, and to discuss some of its properties. The discussion will be based on the idea of a subspace. Throughout the present section I denote by $\mathbb{R}^p$ the real vector space of finite dimension $p > 0$, and by $\mathscr{D}$ a given subspace (or linear manifold) of $\mathbb{R}^p$ of dimension $s \leq p$. Let $d_1, \ldots, d_s$ be a set of basis vectors for $\mathscr{D}$, then the (p, s) matrix

$$\Delta = (d_1, \ldots, d_s)$$

will be called a *basis matrix* for $\mathscr{D}$. Such a basis matrix is, of course, not unique. The following lemma gives necessary and sufficient conditions that two matrices Δ_1 and Δ_2 are basis matrices for the same subspace $\mathscr{D}$.

LEMMA 2.1 *Let $\mathscr{D}_1$ and $\mathscr{D}_2$ be s-dimensional subspaces of $\mathbb{R}^p$. Let Δ_1 and Δ_2 be basis matrices for $\mathscr{D}_1$ and $\mathscr{D}_2$ respectively. Then any of the following four conditions is necessary and sufficient for $\mathscr{D}_1 = \mathscr{D}_2$:*

(i) $\Delta_1 = \Delta_2 E$ *for some nonsingular* (s, s) *matrix* E;

(ii) $\Delta_2 \Delta_2^+ \Delta_1 = \Delta_1$;

(iii) $\Delta_1 \Delta_1^+ = \Delta_2 \Delta_2^+$;

(iv) $(Q\Delta_1)(Q\Delta_1)^+ = (Q\Delta_2)(Q\Delta_2)^+$ *for all matrices* Q *with* p *columns*.

Proof $\mathscr{D}_1 = \mathscr{D}_2 \leftrightarrow \Delta_1$ and Δ_2 span the same subspace $\leftrightarrow \Delta_1 = \Delta_2 E$ for some nonsingular E. Further, (iv) → (iii) by choosing $Q = I_p$, (iii) → (ii) by postmultiplying (iii) with Δ_1, (ii) → (i) by letting $E = \Delta_2^+ \Delta_1$, and (i) → (iv) since $(AB)(AB)^+ = AA^+$ holds for any A and nonsingular B (Lemma A.1 in the Appendix), and therefore in particular for $A = Q\Delta_2$ and $B = E$. ∎

Let X be a real (m, n) matrix. Suppose that among its mn elements x_{ij} there exist $mn - s$ linear relationships. If these restrictions are linearly independent, then X has only s "free" variables, the other $mn - s$ being determined by the linear restrictions. The collection of all real (m, n) matrices that satisfy a given set of linear restrictions constitutes an L-structure. The following definition formalizes this concept.

DEFINITION 2.1 Let $\mathscr{D}$ be an s-dimensional subspace of $\mathbb{R}^{mn}$, and Δ any (mn, s) basis matrix for $\mathscr{D}$. The subset of real (m, n) matrices given by

$$L(\Delta) = \{ X \mid \text{vec } X \in \mathscr{D} \}$$

will be called an L-*structure*, and s will be called the *dimension* of the L-structure.

Thus s, the dimension of the subspace $\mathscr{D}$, is equal to the number of "free" variables in X. It is, of course, obvious from Lemma 2.1(i) that if Δ is a basis matrix for $\mathscr{D}$, then so is ΔE for any nonsingular E. Hence $L(\Delta) = L(\Delta E)$ for each nonsingular E. This fact suggests that it might be more convenient to regard L as function of $\mathscr{D}$ rather than Δ. However, the basis matrix Δ is relevant in any discussion on matrix equations, and I prefer, therefore, to retain Definition 2.1 as it stands.

One example of an L-structure is the class of real *symmetric* (n, n) matrices. The linear restrictions are simply the $n(n-1)/2$ equalities $x_{ij} = x_{ji}$, and so the dimension of the L-structure is $n(n+1)/2$. For

$n = 2$, one choice for Δ would be

$$\Delta = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

By saying that the class of real symmetric matrices constitutes an L-structure, it is understood, of course, that symmetry is the *only* structure imposed. Consider, for example, the following three classes of symmetric matrices:

$$X_1 = \begin{pmatrix} x & y \\ y & z \end{pmatrix}, \qquad X_2 = \begin{pmatrix} x & y \\ y & y^2 \end{pmatrix}, \qquad X_3 = \begin{pmatrix} 1 & y \\ y & 1 \end{pmatrix},$$

with $x, y, z \in \mathbb{R}$. Clearly X_1 is an L-structure, but X_2 and X_3 are not since, in view of Definition 2.1, only *linear* combinations of the "free" parameters are permitted. Thus, no powers (as in X_2) or constants other than zeros (as in X_3). In section 3 "extended" L-structures are introduced which allow constants such as in X_3.

Other examples of L-structures are (strictly) triangular, skew-symmetric, diagonal, circulant, and Toeplitz matrices. Also L-structured (of dimension 4) is

$$\begin{pmatrix} x & z & x+z \\ v & y & v+y \\ x+v & z+y & x+v+z+y \end{pmatrix}.$$

If s, the dimension of the L-structure, is zero, then the L-structure is the null matrix. If $s = mn$, then no restrictions are placed on the elements of X. In this paper I will assume that $1 \leqslant s \leqslant mn$, thus excluding the null matrix but including the unstructured case.

Each basis matrix Δ defines an L-structure, and several of these Δ-matrices have been studied in the literature. A Δ-matrix associated with symmetry was first studied in [29] and, more extensively, in [9] an [15]. The latter paper also deals with a Δ associated with the class of lower triangular matrices. Δ-matrices for strictly lower triangular, skew-symmetric, and diagonal matrices were studied in [21].

The main result of this section is

THEOREM 2.1 *Consider the class of real (m, n) matrices defined by the L-structure $L(\Delta)$ of dimension s. Then the following three statements are equivalent:*

(i) $X \in L(\Delta)$;

(ii) *there exists an $(s, 1)$ vector $\psi(X)$ such that* vec $X = \Delta\psi(X)$; *the*

vector $\psi(X)$ *is uniquely determined and* $\psi(X) = \Delta^+ \text{vec } X$;
  (iii) $(I - \Delta\Delta^+)\text{vec } X = 0$.

Further, let A *and* B *be square matrices of orders* (n, n) *and* (m, m) *respectively. Then the following two statements are equivalent:*

  (iv) $BXA' \in L(\Delta)$ *for all* $X \in L(\Delta)$;
  (v) $\Delta\Delta^+ (A \otimes B)\Delta = (A \otimes B)\Delta$.

Finally, if A *and* B *are nonsingular, either* (iv) *or* (v) *implies*

  (vi) $[\Delta^+ (A \otimes B)\Delta]^{-1} = \Delta^+ (A^{-1} \otimes B^{-1})\Delta$ *and* $[\Delta'(A \otimes B)\Delta]^{-1}$
$= \Delta^+ (A^{-1} \otimes B^{-1})\Delta^{+\prime}$.

Proof (i)$\leftrightarrow$(ii): $X \in L(\Delta) \leftrightarrow \text{vec } X$ lies in the space $\mathscr{D}$ spanned by $\Delta \leftrightarrow$ a vector ψ exists such that $\text{vec } X = \Delta\psi$. Since Δ has full column rank s, $\Delta^+ \Delta = I_s$, and ψ is uniquely given by $\Delta^+ \text{vec } X$.
  (ii) $\leftrightarrow$ (iii): $\Delta\psi(X) = \text{vec } X$ for $\psi(X) = \Delta^+ \text{vec } X \leftrightarrow \Delta\Delta^+ \text{vec } X = \text{vec } X$.
  (iv) $\leftrightarrow$ (v): $\Delta\Delta^+ (A \otimes B)\Delta = (A \otimes B)\Delta \leftrightarrow \Delta\Delta^+ (A \otimes B)\Delta\psi(X)$ $= (A \otimes B)\Delta\psi(X)$ for all values of $\psi(X) \leftrightarrow \Delta\Delta^+ (A \otimes B)\text{vec } X = (A \otimes B)\text{vec } X$ for all $X \in L(\Delta) \leftrightarrow \Delta\Delta^+ \text{vec } BXA' = \text{vec } BXA'$ for all $X \in L(\Delta) \leftrightarrow BXA' \in L(\Delta)$ for all $X \in L(\Delta)$.
  (v) $\rightarrow$ (vi): $\Delta^+ (A^{-1} \otimes B^{-1})\Delta\Delta^+ (A \otimes B)\Delta = \Delta^+ (A^{-1} \otimes B^{-1})(A \otimes B)\Delta = \Delta^+ \Delta = I_s$. The second assertion follows from the symmetry of $\Delta\Delta^+$. ∎

Two comments are in order. First, the converse of (vi) does not hold, that is nonsingularity of A and B, and $[\Delta^+ (A \otimes B)\Delta]^{-1}$ $= \Delta^+ (A^{-1} \otimes B^{-1})\Delta$ together do not imply that BXA' and X belong to the same $L(\Delta)$-structure. To see this, let

$$A = B' = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad \Delta' = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Then,

$$\Delta^+ (A \otimes B)\Delta\Delta^+ (A^{-1} \otimes B^{-1})\Delta = I_2,$$

but

$$(I - \Delta\Delta^+)(A \otimes B)\Delta \neq 0.$$

Secondly, matrices A and B where BXA' has the same L-structure as X are often easy to find. If, for example, X is (skew-)symmetric, then

so is AXA'. Also, if X is (strictly) lower triangular, and P and Q are lower triangular, then PXQ is (strictly) lower triangular.

We are now ready to deal with linear matrix equations where the solution matrix X is known to be L-structured.

3. THE EQUATION $Q \operatorname{vec} X = \operatorname{vec} C$

The linear matrix equation

$$A_1 X B_1 + A_2 X B_2 + \cdots + A_s X B_s = C \tag{3.1}$$

in X, of order (m, n), is a generalization of many matrix equations discussed in the literature, the most famous being the Lyapunov equation

$$AX + XA' = C.$$

Equation (3.1) was first studied almost a century ago by Sylvester [27], who considered it as a system of mn equations for the mn elements x_{ij} of X, but did not recognize the matrix of this system of equations as a sum of Kronecker products. A review of the literature before 1932 is provided in [13, section 46]. Roth [24, Theorem 4] gives necessary and sufficient conditions for consistency. A general solution of (3.1) is implicit in [22, p. 409], and explicit in [12], where much of what is known about the solution of (3.1) is brought together, using mainly methods of contour integration. In [32], equation (3.1) is attacked via Taylor's formula for matrix functions, see also [33].

More general still is the equation

$$\sum_{i=1}^{s} A_i X B_i + \sum_{j=1}^{t} D_j X' E_j = C, \tag{3.2}$$

special cases of which include

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

as studied in [10]. Using the vec-operator and the commutation matrix K as defined in (1.2), the matrix equation (3.2) can be transformed into one vector equation, as in [31]:

$$\left[\sum_i (B_i' \otimes A_i) + \sum_j (E_j' \otimes D_j) K_{mn} \right] \operatorname{vec} X = \operatorname{vec} C.$$

All these cases and more are embodied in the vector equation

$$Q \operatorname{vec} X = \operatorname{vec} C. \tag{3.3}$$

The conditions for consistency and the general solution of (3.3) are well-known [22], also if Q is singular.

In case X is L-structured, the consistency and solution of (3.3) are more complicated. For symmetric X, this problem was raised in [31], but not solved satisfactorily. The first complete solution for symmetric (and lower triangular) X was provided in [15, Lemma 7.1]. This solution is now generalized to arbitrary L-structures.

THEOREM 3.1   *Let $L(\Delta)$ be a given L-structure. The equation*

$$Q \operatorname{vec} X = \operatorname{vec} C$$

has a solution in $L(\Delta)$ if and only if

$$Q\Delta(Q\Delta)^{+} \operatorname{vec} C = \operatorname{vec} C,$$

in which case the general solution is

$$\operatorname{vec} X = \Delta(Q\Delta)^{+} \operatorname{vec} C + \Delta\left[I - (Q\Delta)^{+} Q\Delta \right] \operatorname{vec} P,$$

where P is an arbitrary matrix of appropriate order. In particular:

(i)   *The solution—if it exists—is unique iff $Q\Delta$ has full column rank, in which case*

$$\operatorname{vec} X = \Delta(\Delta'Q'Q\Delta)^{-1}\Delta'Q'\operatorname{vec} C;$$

(ii)   *A solution exists for all C iff $Q\Delta$ has full row rank, in which case*

$$\operatorname{vec} X = \Delta\Delta'Q'(Q\Delta\Delta'Q')^{-1}\operatorname{vec} C + \Delta\left[I - \Delta'Q'(Q\Delta\Delta'Q')^{-1}Q\Delta \right]\operatorname{vec} P.$$

Proof   Since X is L-structured, we have $\operatorname{vec} X = \Delta\psi(X)$ and thus

$$Q\Delta\psi(X) = \operatorname{vec} C.$$

The consistency condition and solution of this vector equation follows from Lemma A.2 in the Appendix. Thus, if a solution exists, it takes the form

$$\psi(X) = (Q\Delta)^{+} \operatorname{vec} C + \left[I - (Q\Delta)^{+} Q\Delta \right]\operatorname{vec} P,$$

for arbitrary P. Premultiplication by Δ gives the solution for $\operatorname{vec} X$. If the equation is consistent, the solution is unique iff $(Q\Delta)^{+} Q\Delta = I$, that is iff $Q\Delta$ has full column rank. In that case

$$(Q\Delta)^{+} = (\Delta'Q'Q\Delta)^{-1}\Delta'Q'$$

and (i) follows. A solution exists for all C iff $(Q\Delta)(Q\Delta)^+ = I$, that is iff $Q\Delta$ has full row rank. In this case

$$(Q\Delta)^+ = \Delta'Q'(Q\Delta\Delta'Q')^{-1}$$

and (ii) follows. ∎

An alternative way of solving the vector equation $Q\,\text{vec}\,X = \text{vec}\,C$ under the constraint $X \in L(\Delta)$ would be to use the fact, established in Theorem 2.1, that $X \in L(\Delta) \leftrightarrow (I - \Delta\Delta^+)\text{vec}\,X = 0$. The constrained problem can thus be written as

$$\begin{pmatrix} Q \\ I - \Delta\Delta^+ \end{pmatrix}\text{vec}\,X = \begin{pmatrix} \text{vec}\,C \\ 0 \end{pmatrix},$$

where $I - \Delta\Delta^+$ is an idempotent matrix which in view of Lemma 2.1(iii) is uniquely determined by $\mathscr{D}$. It turns out, however, that this representation of the problem leads to more cumbersome expressions than in Theorem 3.1. Accordingly, I do not explore this avenue further.

It was noted before that the only constants that an L-structured matrix may contain are zeros. This restriction, however, can easily be removed by introducing "extended" L-structures as follows:

DEFINITION 3.1   Let $L(\Delta)$ be a given L-structure, and A a given (m, n) matrix of real constants. The set of real (m, n) matrices defined by

$$L(\Delta, A) = \{X \mid X - A \in L(\Delta)\}$$

will be called an *extended L-structure*.

Thus, $L(\Delta, A)$ contains all matrices X for which $\text{vec}\,X \in \mathscr{D} + \text{vec}\,A$. The space $\mathscr{D} + \text{vec}\,A$ is a linear (affine) variety of $\mathbb{R}^{mn}$, that is, a translation of the subspace $\mathscr{D}$ by $\text{vec}\,A$. Note that $L(\Delta) = L(\Delta, 0)$.

Simple examples of extended L-structures are

$$\begin{pmatrix} 1 & x \\ x & 1 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} x & y \\ x+1 & y-1 \end{pmatrix},$$

and the corresponding "A"-matrices are respectively

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 0 & 0 \\ 1 & -1 \end{pmatrix}.$$

The following result can now be proved as a corollary to Theorem 3.1.

COROLLARY 3.1 *Let $L(\Delta, A)$ be a given extended L-structure. The equation $Q \operatorname{vec} X = \operatorname{vec} C$ has a solution in $L(\Delta, A)$ if and only if*

$$Q\Delta(Q\Delta)^+ (\operatorname{vec} C - Q \operatorname{vec} A) = \operatorname{vec} C - Q \operatorname{vec} A,$$

in which case the general solution is

$$\operatorname{vec} X = \left[I - \Delta(Q\Delta)^+ Q \right] \operatorname{vec} A + \Delta(Q\Delta)^+ \operatorname{vec} C$$
$$+ \Delta\left[I - (Q\Delta)^+ Q\Delta \right] \operatorname{vec} P,$$

where P is arbitrary.

Proof Let $\overline{X} = X - A$, then $\overline{X} \in L(\Delta) \leftrightarrow X \in L(\Delta, A)$, and

$$Q \operatorname{vec} X = \operatorname{vec} C, \qquad X \in L(\Delta, A)$$

is equivalent to

$$Q \operatorname{vec} \overline{X} = \operatorname{vec} C - Q \operatorname{vec} A, \qquad \overline{X} \in L(\Delta).$$

The result now follows from Theorem 3.1. ∎

Theorem 3.1 also provides the framework to solve simultaneously several equations in one L-structured matrix X. Suppose we are given r equations in one unknown L-structured matrix X,

$$Q_i \operatorname{vec} X = \operatorname{vec} C_i, \qquad i = 1, \ldots, r, \quad X \in L(\Delta). \qquad (3.4)$$

Let

$$Q' = (Q_1', Q_2', \ldots, Q_r') \quad \text{and} \quad C = (C_1, C_2, \ldots, C_r),$$

then (3.4) reduces to $Q \operatorname{vec} X = \operatorname{vec} C$, which is the equation solved in Theorem 3.1. The special case $r = 2$ has received relatively little attention in the literature, with the exception of the equations $AX = C$, $XB = D$, see [5] and [23, Theorem 2.3.3]. More recently [4, pp. 208–209], the pair of consistent linear equations $Ax = a$, $Bx = b$ was considered, and their common solutions, if any, studied. Using Cline's [6] results on the Moore–Penrose inverse of a partitioned matrix, it is possible to study the case $r = 2$ in more detail. The following theorem deals with this situation where two equations are given in one common X. Notice that, since the theorem applies to L-structured X, it applies in particular to *unstructured* X, where no restrictions are placed on the elements of X and $\Delta = I$.

THEOREM 3.2 *Let* $L(\Delta)$ *be a given L-structure. A necessary and sufficient condition for the equations*

$$Q_1 \text{vec } X = \text{vec } C_1 \quad and \quad Q_2 \text{vec } X = \text{vec } C_2,$$

each of which is consistent in $L(\Delta)$, *to have a common solution in* $L(\Delta)$, *is that*

$$(I - SS^+)P_2 P_1^+ \text{vec } C_1 = (I - SS^+)\text{vec } C_2,$$

where

$$P_1 = Q_1 \Delta,$$
$$P_2 = Q_2 \Delta,$$
$$S = P_2(I - P_1^+ P_1),$$

in which case the general solution is

$$\text{vec } X = \Delta(I - S^+ P_2)P_1^+ \text{vec } C_1 + \Delta S^+ \text{vec } C_2$$
$$+ \Delta(I - P_1^+ P_1 - S^+ S)\text{vec } P,$$

where P is arbitrary.

Proof We write the equations $Q_i \text{vec } X = \text{vec } C_i$ $(i = 1, 2)$ as one equation:

$$\begin{pmatrix} Q_1 \Delta \\ Q_2 \Delta \end{pmatrix} \psi(X) = \begin{pmatrix} \text{vec } C_1 \\ \text{vec } C_2 \end{pmatrix} \tag{3.5}$$

or, for short,

$$Qx = c.$$

Let

$$U = (Q_1 \Delta)' \quad and \quad V = (Q_2 \Delta)'.$$

Then

$$Q' = (U, V), \qquad\qquad Q^+ = (U, V)^{+\prime},$$

$$QQ^+ = (U, V)^+(U, V), \qquad Q^+ Q = (U, V)(U, V)^+.$$

Using the results and notation of Lemma A.3, we find

$$I - QQ^+ = I - (U, V)^+(U, V)$$

$$= \begin{pmatrix} I - U^+ U + U^+ VZ(I - R^+ R)V'U^{+\prime} & -U^+ V(I - R^+ R)Z \\ -Z(I - R^+ R)V'U^{+\prime} & (I - R^+ R)Z \end{pmatrix}.$$

Consistency requires $(I - QQ^+)c = 0$, that is

$$(I - U^+U)\text{vec } C_1 + U^+VZ(I - R^+R)V'U^{+\prime}\text{vec } C_1$$
$$- U^+V(I - R^+R)Z \text{ vec } C_2 = 0$$

and

$$- Z(I - R^+R)V'U^{+\prime}\text{vec } C_1 + (I - R^+R)Z \text{ vec } C_2 = 0.$$

Now, Z is nonsingular, and $(I - R^+R)Z = Z(I - R^+R)$. Also, consistency of the equation $Q_1\text{vec } X = \text{vec } C_1$ implies $(I - U^+U)\text{vec } C_1 = 0$. The two equations thus have a common solution iff

$$(I - R^+R)V'U^{+\prime}\text{vec } C_1 = (I - R^+R)\text{vec } C_2. \qquad (3.6)$$

Recalling that $Q' = (U, V)$, the general solution of (3.5) is, for arbitrary P,

$$\psi(X) = (U, V)^{+\prime}\begin{pmatrix} \text{vec } C_1 \\ \text{vec } C_2 \end{pmatrix} + \left[I - (U, V)(U, V)^+ \right]\text{vec } P$$

$$= \begin{pmatrix} U^+ - U^+VH \\ H \end{pmatrix}'\begin{pmatrix} \text{vec } C_1 \\ \text{vec } C_2 \end{pmatrix} + (I - UU^+ - RR^+)\text{vec } P$$

$$= U^{+\prime}\text{vec } C_1 - H'(V'U^{+\prime}\text{vec } C_1 - \text{vec } C_2)$$

$$+ (I - UU^+ - RR^+)\text{vec } P.$$

Using the condition for consistency (3.6) and the definition of H, we find

$$\psi(X) = U^{+\prime}\text{vec } C_1 - R^{+\prime}(V'U^{+\prime}\text{vec } C_1 - \text{vec } C_2)$$

$$+ (I - UU^+ - RR^+)\text{vec } P$$

$$= (I - R^{+\prime}V')U^{+\prime}\text{vec } C_1 + R^{+\prime}\text{vec } C_2$$

$$+ (I - UU^+ - RR^+)\text{vec } P.$$

Premultiply with Δ, and define

$$P_1 = U', \qquad P_2 = V', \qquad S = R',$$

and the results follows.   ∎

4. THE CASE OF SEVERAL UNKNOWN MATRICES

Roth [24] studied the consistency of the linear matrix equation

$$A_1X_1B_1 + A_2X_2B_2 + \cdots + A_rX_rB_r = C, \qquad (4.1)$$

but he did not provide a general solution for the X_i ($i = 1 \ldots r$). In a later paper [25], he studied a special case of (4.1), viz. the equation

$$AX - YB = C, \qquad (4.2)$$

and he provided an elegant and simple condition for consistency. New proofs of the consistency condition of (4.2) were derived in [7] and [1]. The latter paper also gives the general solution of (4.2).

Both (4.1) and (4.2) are special cases of the linear equation

$$Q_1 \text{vec } X_1 + Q_2 \text{vec } X_2 + \cdots + Q_r \text{vec } X_r = \text{vec } C, \qquad (4.3)$$

which can be written as a simple vector equation by defining

$$Q = (Q_1, Q_2, \ldots, Q_r) \quad \text{and} \quad X = (X_1, X_2, \ldots, X_r).$$

The equation (4.3) then reduces to $Q \text{vec } X = \text{vec } C$.

In case the X_i are L-structured (rather than unstructured), there exist vectors $\psi_i(X_i)$ and matrices Δ_i ($i = 1 \ldots r$) such that

$$\Delta_i \psi_i(X_i) = \text{vec } X_i, \qquad i = 1 \ldots r.$$

Thus, defining

$$\Delta = \begin{bmatrix} \Delta_1 & & & 0 \\ & \Delta_2 & & \\ & & \ddots & \\ 0 & & & \Delta_r \end{bmatrix} \quad \text{and} \quad \psi(X) = \begin{bmatrix} \psi_1(X_1) \\ \psi_2(X_2) \\ \vdots \\ \psi_r(X_r) \end{bmatrix},$$

(4.3) further reduces to $Q\Delta\psi(X) = \text{vec } C$. Note that X is an L-structured matrix with $\Delta\psi(X) = \text{vec } X$. Hence, Theorem 3.1 applies straightforwardly.

I shall now take up the linear equation

$$Q_1 \text{vec } X_1 + Q_2 \text{vec } X_2 = \text{vec } C,$$

where only two matrices are unknown. Although this equation can be solved using Theorem 3.1, as outlined above, it seems potentially useful to provide a more detailed solution. This is done in the following theorem, which embodies the solution of (4.2) as a special case.

THEOREM 4.1 *Let* $L(\Delta_1)$ *and* $L(\Delta_2)$ *be two given L-structures. A necessary and sufficient condition for the equation*

$$Q_1 \text{vec } X_1 + Q_2 \text{vec } X_2 = \text{vec } C \qquad (4.4)$$

to have a solution for $X_1 \in L(\Delta_1)$ and $X_2 \in L(\Delta_2)$ is that

$$(P_1 P_1^+ + RR^+)\text{vec}\, C = \text{vec}\, C,$$

where

$$P_1 = Q_1 \Delta_1,$$
$$P_2 = Q_2 \Delta_2,$$
$$R = (I - P_1 P_1^+)P_2,$$

in which case the general solution is

$$\begin{pmatrix} \text{vec}\, X_1 \\ \text{vec}\, X_2 \end{pmatrix} = \begin{pmatrix} \Delta_1 & 0 \\ 0 & \Delta_2 \end{pmatrix} \begin{bmatrix} P_1^+ - P_1^+ P_2 H & \vdots & S_{11} & \vdots & S_{12} \\ H & \vdots & S_{12}' & \vdots & S_{22} \end{bmatrix} \begin{bmatrix} \text{vec}\, C \\ \text{vec}\, W_1 \\ \text{vec}\, W_2 \end{bmatrix},$$

where

$$H = R^+ + (I - R^+ R)ZP_2'P_1^{+\prime}P_1^+(I - P_2 R^+),$$

$$Z = \left[I + (I - R^+ R)P_2'P_1^{+\prime}P_1^+ P_2(I - R^+ R) \right]^{-1},$$

$$S_{11} = I - P_1^+ P_1 + P_1^+ P_2 Z(I - R^+ R)P_2'P_1^{+\prime},$$

$$S_{12} = -P_1^+ P_2(I - R^+ R)Z,$$

$$S_{22} = (I - R^+ R)Z,$$

and W_1 and W_2 are arbitrary matrices.

Proof We can write equation (4.4) as

$$\left(P_1 \vdots P_2 \right) \begin{pmatrix} \psi_1(X_1) \\ \psi_2(X_2) \end{pmatrix} = \text{vec}\, C.$$

From Lemma A.2 we know that consistency requires

$$\left(P_1 \vdots P_2 \right)\left(P_1 \vdots P_2 \right)^+ \text{vec}\, C = \text{vec}\, C.$$

For a consistent equation, the general solution is

$$\begin{pmatrix} \psi_1(X_1) \\ \psi_2(X_2) \end{pmatrix} = \left(P_1 \vdots P_2 \right)^+ \text{vec}\, C + \left[I - \left(P_1 \vdots P_2 \right)^+ \left(P_1 \vdots P_2 \right) \right]\text{vec}\, W,$$

where W is arbitrary. The theorem then follows from Lemma A.3. ∎

COROLLARY 4.1   (*Continuation of Theorem* 4.1). *If* $(P_1 \vdots P_2)$ *has full column rank, equation* (4.4) *has a solution if and only if*

$$M_2\left[I - P_1(P_1'M_2P_1)^{-1}P_1' \right]M_2\operatorname{vec} C = 0$$

or, equivalently,

$$M_1\left[I - P_2(P_2'M_1P_2)^{-1}P_2' \right]M_1\operatorname{vec} C = 0,$$

where

$$M_1 = I - P_1(P_1'P_1)^{-1}P_1',$$

and

$$M_2 = I - P_2(P_2'P_2)^{-1}P_2'.$$

If this is the case, the solution is unique and takes the form

$$\operatorname{vec} X_1 = \Delta_1(P_1'M_2P_1)^{-1}P_1'M_2\operatorname{vec} C$$

$$= \Delta_1(P_1'P_1)^{-1}P_1'\left[I - P_2(P_2'M_1P_2)^{-1}P_2'M_1 \right]\operatorname{vec} C,$$

$$\operatorname{vec} X_2 = \Delta_2(P_2'P_2)^{-1}P_2'\left[I - P_1(P_1'M_2P_1)^{-1}P_1'M_2 \right]\operatorname{vec} C$$

$$= \Delta_2(P_2'M_1P_2)^{-1}P_2'M_1\operatorname{vec} C,$$

Proof   Immediate from Lemma A.4.   ∎

COROLLARY 4.2   (*Continuation of Theorem* 4.1). *A solution of* (4.4) *exists for all C iff* $(P_1 \vdots P_2)$ *has full row rank, in which case the general solution is*

$$\begin{pmatrix} \operatorname{vec} X_1 \\ \operatorname{vec} X_2 \end{pmatrix} = \begin{pmatrix} \Delta_1 & 0 \\ 0 & \Delta_2 \end{pmatrix}\left[\begin{array}{c|c|c} P_1'T & I - P_1'TP_1 & -P_1'TP_2 \\ P_2'T & -P_2'TP_1 & I - P_2'TP_2 \end{array} \right]\begin{bmatrix} \operatorname{vec} C \\ \operatorname{vec} W_1 \\ \operatorname{vec} W_2 \end{bmatrix},$$

where

$$T = (P_1P_1' + P_2P_2')^{-1}$$

and W_1 *and* W_2 *are arbitrary matrices.*

Proof   Immediate from Lemma A.5.   ∎

5. PARTITIONED MATRICES

Consider the matrix equation $AXB = C$, where X is partitioned as

$$X = \begin{pmatrix} X_{11} & \cdots & X_{1t} \\ \vdots & & \vdots \\ X_{r1} & \cdots & X_{rt} \end{pmatrix}. \tag{5.1}$$

The equation $AXB = C$ may of course be solved by Theorem 3.1. It seems, however, that a more fruitful approach is available, one that preserves the structure of the submatrices X_{ij}. For example, suppose X_{11} is symmetric, while the other submatrices are unstructured. One would wish for an operator which stacks the elements of X in such a way that vec X_{11} would appear as its first subvector. The vec operator, then, is inappropriate. At this point a new matrix product must be introduced.

A new matrix product (Tracy–Singh)

This new product was defined by Tracy and Singh [30] and, independently, by McDonald and Swaminathan [17]. Both papers develop this product in the context of matrix differentiation. For my purpose a special case of their product suffices. Let

$$A = (A_1 : \ldots : A_r) \quad \text{and} \quad E = (E_1 : \ldots : E_t)$$

be two partitioned matrices. Let the order of A_i be (p, m_i) and the order of E_j (q, n_j), $i = 1 \ldots r$, $j = 1 \ldots t$. Further,

$$m = \sum_{i=1}^{r} m_i \quad \text{and} \quad n = \sum_{j=1}^{t} n_j.$$

Then A and E have orders (p, m) and (q, n) respectively. The new product is

$$A \boxtimes E = (A_1 \otimes E_1 : A_2 \otimes E_1 : \ldots : A_r \otimes E_1 : A_1 \otimes E_2 :$$
$$\ldots : A_r \otimes E_2 : \ldots : A_r \otimes E_t), \tag{5.2}$$

which is a (pq, mn) matrix, just as $A \otimes E$. In [26, p. 95] it is wrongly alleged that this product is the same as the matrix product introduced in [11], and generalized in [23, p. 12]. Many properties of the new product $A \boxtimes E$ are similar to those of the Kronecker product, see [30]. In particular, let X be as in (5.1), where the order of X_{ij} is (m_i, n_j) so

that the order of X is (m, n). Define the mn-vector

$$w(X) = \left[(\text{vec } X_{11})', (\text{vec } X_{12})', \dots, (\text{vec } X_{1t})', (\text{vec } X_{21})', \right.$$
$$\left. \dots, (\text{vec } X_{2t})', \dots, (\text{vec } X_{rt})' \right]' \tag{5.3}$$

The matrix equation $AXB = C$, where

$$A = (A_1 : \dots : A_r) \quad \text{and} \quad B' = (B'_1 : \dots : B'_t),$$

and X is partitioned as in (5.1), can now be written as

$$(B' \boxtimes A)w(X) = \text{vec } C, \tag{5.4}$$

which has the advantage over the equivalent expression $(B' \otimes A)$ $\text{vec } X = \text{vec } C$ that the ordering of the elements of X is more sensible. If the submatrices of X are L-structured, there exist basis matrices Δ_{ij} and vectors $\psi_{ij}(X_{ij})$ such that

$$\Delta_{ij}\psi_{ij}(X_{ij}) = \text{vec } X_{ij}, \qquad i = 1 \dots r, \quad j = 1 \dots t.$$

Define

$$\Delta = \begin{bmatrix} \Delta_{11} & & & \\ & \Delta_{12} & & \\ & & \ddots & \\ & & & \Delta_{rt} \end{bmatrix} \quad \text{and} \quad \psi(X) = \begin{bmatrix} \psi_{11}(X_{11}) \\ \psi_{12}(X_{12}) \\ \vdots \\ \psi_{rt}(X_{rt}) \end{bmatrix},$$

then the vector equation (5.4) reduces to

$$Q\Delta\psi(X) = \text{vec } C, \tag{5.5}$$

where $Q = B' \boxtimes A$. Theorem 3.1 can be applied to (5.5) to yield the solution of the partitioned matrix equation $AXB = C$.

It is clear that more complicated partitioned linear matrix equations can be solved by analogy with the procedure of this section.

6. EXAMPLES OF Δ

In order to simplify implementation of the theory developed in this paper, I present in this section Δ-matrices for eight L-structures that are most likely to appear in practical situations. Each defines a class of square matrices, say of order (n, n). The L-structures are (with their dimensions in brackets): (1) symmetric $[n(n + 1)/2]$, (2) lower triangular $[n(n + 1)/2]$, (3) skew-symmetric $[n(n - 1)/2]$, (4) strictly lower triangular $[n(n - 1)/2]$, (5) diagonal $[n]$, (6) circulant $[n]$, (7) Toeplitz

J. R. MAGNUS

$[2n - 1]$, and (8) symmetric Toeplitz matrices $[n]$. For the case $n = 3$ sensible choices for Δ are (with dots representing zeros):

$$
\Delta_1 = \begin{pmatrix}
1 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & 1 & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & 1 & \cdot & \cdot & \cdot & \cdot \\
\cdot & 1 & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & 1 & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & 1 & \cdot & \cdot \\
\cdot & \cdot & 1 & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & 1 & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 1
\end{pmatrix}
\qquad
\Delta_2 = \begin{pmatrix}
1 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & 1 & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & 1 & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & 1 & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & 1 & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 1
\end{pmatrix}
$$

$$
\Delta_3 = \begin{pmatrix}
1 & \cdot & \cdot \\
\cdot & 1 & \cdot \\
-1 & \cdot & \cdot \\
\cdot & \cdot & \cdot \\
\cdot & \cdot & 1 \\
\cdot & -1 & \cdot \\
\cdot & \cdot & -1
\end{pmatrix}
\qquad
\Delta_4 = \begin{pmatrix}
\cdot & \cdot & \cdot \\
1 & \cdot & \cdot \\
\cdot & 1 & \cdot \\
\cdot & \cdot & \cdot \\
\cdot & \cdot & 1 \\
\cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot
\end{pmatrix}
\qquad
\Delta_5 = \begin{pmatrix}
1 & \cdot & \cdot \\
\cdot & \cdot & \cdot \\
\cdot & 1 & \cdot \\
\cdot & \cdot & \cdot \\
\cdot & \cdot & 1
\end{pmatrix}
$$

$$
\Delta_6 = \begin{pmatrix}
1 & \cdot & \cdot \\
\cdot & 1 & \cdot \\
\cdot & \cdot & 1 \\
\cdot & \cdot & 1 \\
1 & \cdot & \cdot \\
\cdot & 1 & \cdot \\
\cdot & 1 & \cdot \\
\cdot & \cdot & 1 \\
1 & \cdot & \cdot
\end{pmatrix}
\qquad
\Delta_7 = \begin{pmatrix}
1 & \cdot & \cdot & \cdot & \cdot \\
\cdot & 1 & \cdot & \cdot & \cdot \\
\cdot & \cdot & 1 & \cdot & \cdot \\
\cdot & \cdot & \cdot & 1 & \cdot \\
1 & \cdot & \cdot & \cdot & \cdot \\
\cdot & 1 & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & 1 \\
\cdot & \cdot & \cdot & 1 & \cdot \\
1 & \cdot & \cdot & \cdot & \cdot
\end{pmatrix}
\qquad
\Delta_8 = \begin{pmatrix}
1 & \cdot & \cdot \\
\cdot & 1 & \cdot \\
\cdot & \cdot & 1 \\
\cdot & 1 & \cdot \\
1 & \cdot & \cdot \\
\cdot & 1 & \cdot \\
\cdot & \cdot & 1 \\
\cdot & 1 & \cdot \\
1 & \cdot & \cdot
\end{pmatrix}
$$

APPENDIX: SOME PROPERTIES OF THE MOORE-PENROSE INVERSE

The Moore–Penrose (MP) inverse of a matrix was introduced by Moore [18, 19] and rediscovered by Penrose [22].

DEFINITION   An (m, n) matrix X is the MP inverse of an (n, m) matrix A if

$$AXA = A, \quad XAX = X, \quad (AX)' = AX, \quad (XA)' = XA.$$

The MP inverse of A is denoted as A^+.

The MP inverse exists and is unique. In this appendix, some properties of the MP inverse which are used in the text, are presented. The reader who is not familiar with general properties of MP inverses is referred to [4] and [23]. All matrices are taken to be real.

LEMMA A.1   *For any A and nonsingular B,*

$$AB(AB)^+ = AA^+.$$

Proof   Since $AB = AB(AB)^+AB$, and B is nonsingular, we have $A = AB(AB)^+A$, and hence $AA^+ = AB(AB)^+AA^+ = (AB)^{+\prime}(AB)'$ $AA^+ = (AB)^{+\prime}B'A'AA^+ = (AB)^{+\prime}B'A' = (AB)^{+\prime}(AB)' = AB(AB)^+$. ∎

LEMMA A.2 [22]   *A necessary and sufficient condition for the vector equation $Ax = b$ to have a solution is that*

$$AA^+b = b,$$

in which case the general solution is

$$x = A^+b + (I - A^+A)q,$$

where q is an arbitrary vector of appropriate order.

LEMMA A.3 [6]   *The MP inverse of the partitioned matrix (U, V) is*

$$(U, V)^+ = \begin{pmatrix} U^+ - U^+VH \\ H \end{pmatrix}$$

with

$$H = R^+ + (I - R^+R)ZV'U^{+\prime}U^+(I - VR^+)$$
$$R = (I - UU^+)V$$
$$Z = \left[I + (I - R^+R)V'U^{+\prime}U^+V(I - R^+R) \right]^{-1}.$$

Further,

$$(U, V)(U, V)^+ = UU^+ + RR^+$$

and

$$(U, V)^+ (U, V)$$

$$= \begin{pmatrix} U^+ U - U^+ VZ(I - R^+ R)V'U^{+\prime} & U^+ V(I - R^+ R)Z \\ Z(I - R^+ R)V'U^{+\prime} & I - (I - R^+ R)Z \end{pmatrix}.$$

LEMMA A.4   *If the partitioned matrix* (U, V) *has full column rank, its MP inverse is*

$$(U, V)^+ = \begin{bmatrix} (U'M_v U)^{-1} U'M_v \\ (V'V)^{-1} V'[I - U(U'M_v U)^{-1} U'M_v] \end{bmatrix}$$

$$= \begin{bmatrix} (U'U)^{-1} U'[I - V(V'M_u V)^{-1} V'M_u] \\ (V'M_u V)^{-1} V'M_u \end{bmatrix},$$

where

$$M_u = I - U(U'U)^{-1} U'$$

$$M_v = I - V(V'V)^{-1} V'.$$

Further,

$$(U, V)(U, V)^+ = I - M_v[I - U(U'M_v U)^{-1} U']M_v$$

$$= I - M_u[I - V(V'M_u V)^{-1} V']M_u.$$

Proof   If (U, V) has full column rank, then

$$(U, V)^+ = [(U, V)'(U, V)]^{-1}(U, V)' = \begin{pmatrix} U'U & U'V \\ V'U & V'V \end{pmatrix}^{-1} \begin{pmatrix} U' \\ V' \end{pmatrix}.$$

The inverse of $(U, V)'(U, V)$ can be expressed in two equivalent ways, viz.

$$\begin{bmatrix} (U'M_v U)^{-1} & -(U'M_v U)^{-1} U'V(V'V)^{-1} \\ -(V'V)^{-1} V'U(U'M_v U)^{-1} & (V'V)^{-1} + (V'V)^{-1} V'U(U'M_v U)^{-1} U'V(V'V)^{-1} \end{bmatrix}$$

or

$$\begin{bmatrix} (U'U)^{-1} + (U'U)^{-1} U'V(V'M_u V)^{-1} V'U(U'U)^{-1} & -(U'U)^{-1} U'V(V'M_u V)^{-1} \\ -(V'M_u V)^{-1} V'U(U'U)^{-1} & (V'M_u V)^{-1} \end{bmatrix}$$

The result follows.   ∎

LEMMA A.5 *If the partitioned matrix (U, V) has full row rank, its MP inverse is*

$$(U, V)^+ = \begin{pmatrix} U'P \\ V'P \end{pmatrix},$$

where

$$P = (UU' + VV')^{-1}.$$

Further,

$$(U, V)^+ (U, V) = \begin{pmatrix} U'PU & U'PV \\ V'PU & V'PV \end{pmatrix}.$$

Proof If (U, V) has full row rank, then

$$(U, V)^+ = (U, V)' \left[(U, V)(U, V)' \right]^{-1}$$

$$= \begin{pmatrix} U' \\ V' \end{pmatrix} (UU' + VV')^{-1}.$$

The result follows.

References

[1] J. K. Baksalary and R. Kala, The matrix equation $AX - YB = C$, *Linear Algebra Appl.* **25** (1979), 41–43.

[2] J. K. Baksalary and R. Kala, The matrix equation $AXB + CYD = E$, *Linear Algebra Appl.* **30** (1980), 141–147.

[3] R. Bellman, *Introduction to matrix analysis*, 2nd edition, McGraw Hill, New York (1970).

[4] A. Ben-Israel and T. N. E. Greville, *Generalized inverses: theory and applications*, Wiley, New York (1974).

[5] F. Cecioni, *Sopra alcune operazioni algebriche sulle matrici*, Ann. Scu. Norm. Sup. Pisa, Sc. Fis. e Mat., Vol. 11, Pisa (1911), especially Chapter I, section 4.

[6] R. E. Cline, Representations for the generalized inverse of a partitioned matrix, *SIAM J. Appl. Math.*, **12** (1964), 588–600.

[7] H. Flanders and H. K. Wimmer, On the matrix equations $AX - XB = C$ and $AX - YB = C$, *SIAM J. Appl. Math.*, **32** (1977), 707–710.

[8] G. Hadley, *Linear Algebra*, Addison-Wesley, Reading, Massachusetts (1961).

[9] H. V. Henderson and S. R. Searle, Vec and vech operators for matrices, with some uses in Jacobians and multivariate statistics, *Canadian J. Statist.*, **7** (1979), 65–81.

[10] J. H. Hodges, Some matrix equations over a finite field, *Ann. Mat. Pura Appl. Ser. 4*, **44** (1957), 245–250.

[11] C. G. Khatri and C. R. Rao, Solutions to some functional equations and their applications to characterization of probability distributions, *Sankhyā Ser. A* **30** (1968), 167–180.

[12] P. Lancaster, Explicit solutions of linear matrix equations, *SIAM Rev.* **12** (1970), 544–566.

[13] C. C. MacDuffee, *The theory of matrices*, Berlin (1933), reprinted by Chelsea, New York.

[14] J. R. Magnus and H. Neudecker, The commutation matrix: some properties and applications, *Ann. Statist.* 7 (1979), 381–394.

[15] J. R. Magnus and H. Neudecker, The elimination matrix: some lemmas and applications, *SIAM J. Alg. Discr. Meth.*, 1 (1980), 422–449.

[16] M. Marcus and H. Minc, *A survey of matrix theory and matrix inequalities*, Allyn and Bacon, Inc., Boston (1964).

[17] R. P. McDonald and H. Swaminathan, A simple matrix calculus with application to multivariate analysis, *General Systems* 18 (1973), 37–54.

[18] E. H. Moore, On the reciprocal of the general algebraic matrix (abstract), *Bull. Amer. Math. Soc.* 26 (1920), 394–395.

[19] E. H. Moore, *General Analysis*, Memoirs of the American Philosophical Society, vol. 1, Philadelphia (1935), especially Part I, section 29.

[20] H. Neudecker, Some theorems on matrix differentiation with special reference to Kronecker matrix products, *J. Amer. Statist. Assoc.* 64 (1969), 953–963.

[21] H. Neudecker, *On Jacobians of transformations with skew-symmetric, strictly (lower) triangular or diagonal matrix arguments*, Report AE 8/80, Institute for Actuarial Science and Econometrics, University of Amsterdam (1980).

[22] R. Penrose, A generalized inverse for matrices, *Proc. Cambridge Philos. Soc.* 51 (1955), 406–413.

[23] C. R. Rao and S. K. Mitra, *Generalized inverse of matrices and its applications*, Wiley, New York (1971).

[24] W. E. Roth, On direct product matrices, *Bull. Amer. Math. Soc.* 40 (1934), 461–468.

[25] W. E. Roth, The equations $AX - YB = C$ and $AX - XB = C$ in matrices, *Proc. Amer. Math. Soc.* 3 (1952), 392–396.

[26] H. Swaminathan, Matrix calculus for functions of partitioned matrices, *General Systems* 21 (1976), 95–99.

[27] M. Sylvester, *Comptes Rendus des Séances de l'Académie des Sciences*, Paris (1884), 117–118, 409–412, 432–436, and 527–529.

[28] D. S. Tracy and P. S. Dwyer, Mulivariate maxima and minima with matrix derivatives, *J. Amer. Statist. Assoc.* 64 (1969), 1576–1594.

[29] D. S. Tracy and R. P. Singh, Some modifications of matrix differentiation for evaluating Jacobians of symmetric matrix transformations, *Symmetric Functions in Statistics, Proceedings of a Symposium in Honor of Paul S. Dwyer* (Derrick S. Tracy, ed.), Univ. Windsor, Windsor, Ontario (1972), 203–224.

[30] D. S. Tracy and R. P. Singh, A new matrix product and its applications in partitioned matrix differentiation, *Statistica Neerlandica* 26 (1972), 143–157.

[31] W. J. Vetter, Linear structures and solutions of linear matrix equations, *Linear Algebra Appl.* 10 (1975), 181–188.

[32] H. Wimmer and A. D. Ziebur, Solving the matrix equation $\sum_{p-1}^{r} f_p(A) X g_p(B) = C$, *SIAM Rev.* 14 (1972), 318–323.

[33] H. Wimmer and A. D. Ziebur, Blockmatrizen und lineare Matrizengleichungen, *Math. Nachr.* 59 (1974), 213–219.