THE VEC-PERMUTATION MATRIX, THE VEC OPERATOR AND KRONECKER PRODUCTS: A REVIEW

by

Harold V. Henderson

Biometrics Section, Ruakura Agricultural Research Centre, Hamilton, New Zealand

and

S. R. Searle

Biometrics Unit, Cornell University, Ithaca, New York

ABSTRACT

The vec-permutation matrix $\text{vec}_m^n$ is defined by the equation

$$\text{vec}_m^n A_{mn} = I_m \text{vec}_n A',$$

where vec is the vec operator such that vec$A$ is the vector of columns of $A$ stacked one under the other. The variety of definitions, names and notations for $I_m^n$ are discussed, and its properties are developed by simple proofs in contrast to certain lengthy proofs in the literature that are based on descriptive definitions. For example, the role of $I_m^n$ in reversing the order of Kronecker products is succinctly derived using the vec operator. The matrix $M_{mn}$ is introduced as $M_{mn} = I_m \text{vec}_n M$; it is the matrix having for rows, every $n$'th row of $M$, of order $mn \times c$, starting with the first, then every $n$'th row starting with the second, and so on. Special cases of $M_{mn}$ are discussed.

---

Paper No. BU-645-M in the Biometrics Unit, Cornell University.

* Supported by a New Zealand National Research Advisory Council Research Fellowship during his sojourn at Cornell.
1. NOTATION

Let $A$ and $B$ be matrices of order $m \times n$ and $p \times q$, respectively. Rows of a matrix will be denoted by Greek letters transposed and columns by Roman letters:

$$A = \{a_{ij}\} = [a_{11} \ldots a_{1n} \ldots a_{m1} \ldots a_{mn}]$$
$$B = \{b_{rs}\} = [b_{11} \ldots b_{1q} \ldots b_{p1} \ldots b_{pq}]$$

$I_n$ is the identity matrix of order $n$, with $e_i$ denoting its $i$'th column and then

$$I_n = \sum_{i=1}^{n} e_i e_i' .$$

2. THE VEC OPERATOR AND KRONECKER PRODUCTS

2.1. The vec operator and its many names and notations

The vec operator stacks the columns of a matrix one underneath the other to form a single vector. Thus for $A$ of (1)

$$\text{vec}A = \begin{bmatrix} a_{11} \\ \vdots \\ a_{mn} \end{bmatrix} ,$$

with vecA (for "vector of columns of $A$") being the notation currently in vogue. The equivalent notations vecA and vec(A) are used interchangeably, the parentheses being employed only when deemed necessary for clarity.
An early reference to this idea of stacking elements of a matrix in a vector is Sylvester \cite{1884a,b,c} who used it in connection with linear equations. Roth \cite{1934}, using the notation $A^c$, develops results for using the operation on a product matrix, Aitken \cite{1949} mentions the idea in connection with Jacobians, and Koopmans et al. \cite{1950} introduce the notation vec. More recently, the concept has been exploited in a variety of ways, for example, in solving linear matrix equations and in matrix differentiation, from which it is seen to be useful in deriving Jacobians of matrix transformations. The paper by Henderson and Searle \cite{1979} outlines these applications and highlights several uses in statistics; e.g., in rewriting multivariate linear models in a univariate form, in developing the dispersion matrix of elements of a matrix such as a Wishart matrix and, from this, in deriving fourth moments in a general linear model.

As a result of these applications, the concept has, in recent years, been used by numerous writers. There is vec$A$ or vec$(A)$ used by Neudecker \cite{1968, 1969a,b}, Browne \cite{1974} (who also uses $a$), Swain \cite{1975} (who also uses $A$), Conlisk \cite{1976}, Balestra \cite{1976}, Anderson et al. \cite{1977}, Anderson \cite{1978}, Searle \cite{1978, 1979}, Brewer \cite{1978}, Magnus and Neudecker \cite{1979} and Henderson and Searle \cite{1979}. Equivalent notations are $S(A)$, for stacking columns of $A$, used by Nissen \cite{1968} who also uses $a$; and $L(A)$ used by Conlisk \cite{1969}; also, $\tilde{A}$ for the column-rolled-out form of $A$, as it is described by Cole \cite{1969}; and cs$A$ as the column string of $A$ used by Vetter \cite{1970, 1973, 1975}, Kucera \cite{1974} and Mitra \cite{1977}. There is also
Variations on what we have defined as vecA are also available: for example, having it be the column vector derived from writing each transposed row of A one under the other. This stacks the elements of A in lexicon order and is, in our notation, equivalent to vecA', where A' is the transpose of A. Notations for this include a used by Lancaster [1969, 1970], A_r by Tracy and Dwyer [1969], Singh [1972] and Tracy and Singh [1972a,b], v(A) by Barnett [1973], vrA as the vector of transposed rows of A by Legault-Giguère [1974] and Giguère and Styan [1974] and a_r by Nel [1978]; and even vecA by McDonald and Swaminathan [1973], Bentler and Lee [1975], McDonald [1976] and Pukelsheim [1977]. Other variations include the row vector of rows of A, (vecA')' in our notation, which Roth [1934] denotes by A_R and Vetter [1970, 1973, 1975] by rsA, as the row string of A. A final form, the row vector of transposed columns of A, (vecA)' has not yet been seen in the literature!

2.2. Origins of the Kronecker product

The Kronecker product of two matrices is defined for A and B of (1) as the mp x nq matrix

\[ A \otimes B = \{ a_{ij} B \} , \]  

(3)
which is a particular case of the tensor product for transformations as discussed, for example, by Halmos [1958, p. 97]. In its original setting it seems to have been first studied by Zehfuss [1858], and although he dealt only with its determinant, Rutherford [1933] appropriately calls \( A \otimes B \) the Zehfuss matrix of \( A \) and \( B \). Loewy [1910, pp. 149-150] refers to Zehfuss's determinantal result as Kronecker's theorem which, according to Hensel [1891, p. 319], Kronecker (1823-91) had for sometime given in his algebra lectures in Berlin, which Bell [1937, p. 478] notes, he presented regularly from 1861 "principally on his personal researches, after the necessary introductions". Thus, although the name of Kronecker is now generally associated with the \( A \otimes B \) operation, with an early use as pointed out by Dr. George Styan being Murnaghan [1938, p. 68], the exact origin of this association is hard to find.

2.3. The vec and Kronecker product operators are connected

The definitions of vec and of Kronecker product show that

\[ \text{vec}(ab') = b \otimes a, \]  

so that \( \text{vec}[(Aa)(b'C)] = (C' \otimes A)\text{vec}(ab'). \) This, together with \( B = \{b_{ij}\} = \sum_{i,j} e_i e_j' \), leads to

\[ \text{vec}(ABC) = (C' \otimes A)\text{vec}B, \]  

(4)

a result derived by Roth [1934], and hence called Roth's column lemma by Hartwig [1975], and rediscovered by Aitken [1949], Koopmans et al. [1950], Nissen [1968] and Neudecker [1969b].

3. VEC-PERMUTATION MATRICES

The vectors vec\(A\) and vec\(A'\) contain the same elements, in different sequences. We define \( I_{m,n} \) as the permutation matrix such that, for \( A \) of order \( m \) by \( n \),

\[ \text{vec}A = I_{m,n} \text{vec}A', \]  

(5)

and give it the name vec-permutation matrix.

There has been increasing interest over the last decade in vec-permutation matrices in matrix algebra, mathematics, statistics, econometrics and psychometrics. This wide interest in what we call
\(I_{m,n}\) of (5) is partly responsible for the diversity of its literature, for duplication of published results, for varying definitions and for a wide range of notation and nomenclature. For example, (5) defines what \(I_{m,n}\) does, in contrast to definitions discussed later that describe what \(I_{m,n}\) looks like. This is an important distinction because operational definitions of \(I_{m,n}\), like (5), lead to succinct derivations of its properties, in contrast to some recent and rather tedious proofs based on descriptive definitions. An example of this is the important role of \(I_{m,n}\) in reversing the order of Kronecker products; e.g., for \(A_{m \times n}\) and \(B_{p \times q}\)

\[
B \otimes A = I_{m,p} (A \otimes B) I_{n,q}.
\]

(6)

In these circumstances, and because of the topical nature of the subject, a cohesive account and succinct development of properties of vec-permutation matrices is needed. This is attempted in Section 4. But first we discuss the variety of definitions, names and notations to be found in the literature and summarized in Table 1, for a vec-permutation matrix, using our notation \(I_{m,n}\).

(Show Table 1 here)

3.1. Definitions from reversing the order of Kronecker products of matrices and vectors

Ledermann [1936], in material from his Ph.D. thesis supervised by A. C. Aitken, gives an elegant proof of (6), although without explicit development of the vec-permutation matrices. Using column vectors \(r\), \(x\), \(s\) and \(y\) of order \(m\), \(n\), \(p\) and \(q\), respectively, with \(r = Ax\) and \(s = By\), he proceeds as follows.
"The two products $A \otimes B$ and $B \otimes A$ are related to each other by an identity

$$Q(A \otimes B)P^{-1} = (B \otimes A),$$

where $P$ and $Q$ are permutation matrices which depend only on the types of $A$ and $B$ and not on their elements.

Proof: Apart from the order, the vectors $x \otimes y$ and $y \otimes x$ contain the same elements $\ldots$. We can therefore find a permutation matrix $P$ of degree $nq$ such that

$$(y \otimes x) = P(x \otimes y)$$

and similarly

$$(s \otimes r) = Q(r \otimes s),$$

where $Q$ is a permutation matrix of degree $mp$. Evidently $P$ and $Q$ do not depend on the elements of $x, y, r, s$ but only on the numbers $m, n, p, q$. By $[(A \otimes B)(x \otimes y) = Ax \otimes By]$ we have

$$(r \otimes s) = (A \otimes B)(x \otimes y),$$

$$(s \otimes r) = (B \otimes A)(y \otimes x).$$

On premultiplying the first equation by $Q$ and substituting $\ldots$ we get

$$(s \otimes r) = Q(A \otimes B)(x \otimes y) = (B \otimes A)P(x \otimes y).$$

Since there is obviously no linear relation between the elements of $x \otimes y$, we obtain

$$Q(A \otimes B) = (B \otimes A)P$$

or

$$Q(A \otimes B)P^{-1} = (B \otimes A),$$

as was to be proved."
[For consistent notation we have replaced \( \times \) in the original form of this quotation by \( \otimes \), and changed subscripts to conform with (1); and we do this in all quotations, without further comment.]

Ledermann [1936] then, in our notation, defines \( I_{m,n} \) as the permutation matrix such that

\[
I_{m,n} (a \otimes b) = b \otimes a,
\]

for any column vectors \( a \) and \( b \) of order \( m \) and \( n \), respectively.

Conlisk [1976] has a similar idea in mind for \( I_{n,n'} \), using the notation \( M \). [Actually, he writes (p. 760) \( (a \otimes b)M = b \otimes a \), which is clearly incorrect, from dimension considerations alone. However, he correctly claims (p. 763) that \( M(a \otimes B_{n \times n'}) = B \otimes a \).] Defining \( I_{m,n'} \) as in (7), in terms of reversing the order of Kronecker products of vectors, was also suggested by a referee of Henderson and Searle [1979]. The suggestion has merit, but it emphasizes a feature of \( I_{m,n} \) that is second to the more fundamental and more general definition given in (5) in terms of the vec operator. In fact, (7) is the special case of (5) with \( A \) being \( b \otimes a' \), of rank 1.

Murnaghan [1938] was also interested in reversing the order of Kronecker products and, in this connection, gives an explicit formulation of what we call a vec-permutation matrix. He denotes the position of \( a_{i,j} \) in \( A \otimes B \) by the row index-pair \((i,r)\) and column index-pair \((j,s)\), and notes (p. 68-69) that for \( A \) and \( B \) square of order \( m \) and \( n \), respectively,
"B ⊗ A is obtainable from A ⊗ B by applying the same permutation to the rows and columns of the latter. Let us denote by \((i,r)^*\) the position of \((r,i)\) when the ordering is dictionary-like, the label with the range \(n\) coming first; e.g., if \(m = 2, n = 3\) the dictionary order where the label with the range 2 comes first is \((1,1), (1,2), (1,3), (2,1), (2,2), (2,3)\) whilst the dictionary order when the label with the range 3 comes first is \((1,1), (1,2), (2,1), (2,2), (3,1), (3,2)\). Hence \((1,1)^* = (1,1), (1,2)^* = (1,3), (1,3)^* = (2,2), (2,1)^* = (1,2), (2,2)^* = (2,1), (2,3)^* = (2,3)\). Then if the \((j,s)\) column of \(A \otimes B\) is transferred to the \((j,s)^*\) position and the \((i,r)\) row to the \((i,r)^*\) position we obtain \(B \otimes A\). In other words

\[
(B \otimes A) = P(A \otimes B)P^{-1}
\]

where \(P\) is the permutation matrix associated with \(\{(i,r)^*\}\). Notice that \(P\) is quite independent of the elements of \(A\) and \(B\) being completely determined by their dimensions."

A clearer statement of Murnaghan's [1938] development of \(P\), which is our \(I_{m,n}\), is that
row \( (i,j) \) of \( I_{m,n} \) is row \( (j,i) \) of \( I_{mn} \), \hspace{1cm} (9) 

for \( i = 1, \cdots, n \) and \( j = 1, \cdots, m \). Note that \( I_{mn} \) is the identity matrix of order \( mn \), to be distinguished from \( I_{m,n} \), a vec-permutation matrix of the same order. Searle [1966, p. 216] makes a more explicit statement of this situation:

"It is apparent from (3) that for \( A_{m \times n} = \{ a_{ij} \} \) and \( B_{p \times q} = \{ b_{rs} \} \) the elements of both \( A \otimes B \) and \( B \otimes A \) consist of all possible products \( a_{ij} b_{rs} \). In fact, \( B \otimes A \) is simply \( A \otimes B \) with the rows and columns each in a different order. 

Thus [in general]

\[
B \otimes A = P(A \otimes B)Q 
\]

(10)

where \( P \) and \( Q \) are each a product of \( E \)-type elementary operators [permutation matrices].

For any values of \( i, j, r \) and \( s \) the element \( a_{ij} b_{rs} \) is located in \( A \otimes B \) in the \( r \)'th row and \( s \)'th column of the \( i,j \)'th sub-matrix \( [a_{ij}, B] \). It is therefore in row \( [p(i-1) + r] \) and column \( [q(j-1) + s] \) of \( A \otimes B \). 

In \( B \otimes A \), however, \( b_{rs} a_{ij} \) is in row \( i \) and column \( j \) of the \( r,s \)'th sub-matrix \( [b_{rs}, A] \). It is therefore in row \( [m(r-1) + i] \) and column \( [n(s-1) + j] \) of \( B \otimes A \). Consequently the interchanging of rows and columns implied in (10), to obtain \( B \otimes A \) from \( A \otimes B \), can be specified as follows. The \( [i + (j-1)m] \)'th row of \( P \) is the \( [(i-1)p + j] \)'th row of \( I_{mp} \), for \( i = 1, 2, \cdots, m \) and \( j = 1, 2, \cdots, p \), and the \( [i + (j-1)n] \)'th column of \( Q \).
is the \([(i - 1)q + j]\)'th column of \(I_{nq}\), for \(i = 1, 2, \ldots, n\) and \(j = 1, 2, \ldots, q\)."

Comparison of (6) with (10) shows that \(P = I_{m,p}\) and \(Q = I_{q,n}\); so that Searle's [1966] expression for \(P\) may be restated, in our notation, as

\[
\text{row } (i - 1)m + j \text{ of } I_{m,n} \text{ is row } (j - 1)n + i \text{ of } I_{mn}, \quad (11)
\]

for \(i = 1, \ldots, n\) and \(j = 1, \ldots, m\). This is, of course, equivalent to (9) because row \((i,j)\) of \(I_{m,n}\) is its \([(i - 1)m + j]\)'th row. Expressions (9) and (11) show consecutive rows of \(I_{m,n}\) to be every \(n\)'th row of \(I_{mn}\) starting with the first, and so on; a fact which Tracy and Dwyer [1969] rediscovered and used as its definition, as is now discussed.
3.2. Defining the row permutations

Tracy and Dwyer [1969, p. 1579] introduce

\[
\tilde{I}(n) \text{ as the matrix obtained by rearranging}
\]
the rows of \( \tilde{I} \) by taking \\
every \( n \)’th row starting with the first, then \\
every \( n \)’th row starting with the second, and so on.

\[
(I2)
\]

Tracy and Singh [1972a,b] and Singh [1972], restricting \( \tilde{I}(n) \) to be 
of order \( mn \), establish (5) and (6) and call \( \tilde{I}(n) \) a permuted identity 
matrix, a name which actually applies to any permutation matrix. We 
call \( \tilde{I}(n) \) of (12) a generalized vec-permutation matrix because it is 
defined for any order, not necessarily a multiple of \( n \), whereas 

\[
\tilde{I}_{m,n} \equiv \tilde{I}(n) \text{ of order } mn .
\]

(13)

For example, the rows of 

\[
I_{2,3} = \begin{bmatrix}
1 & . & . & . & . & . \\
. & . & . & 1 & . & . \\
. & . & . & . & 1 & . \\
. & . & 1 & . & . & . \\
. & . & . & . & 1 & . \\
\end{bmatrix}
\]

(14)

(where dots represent zeros), are every third row, namely rows 1, 4, 
2, 5, 3 and 6, respectively, of \( \tilde{I}_{6} \).
3.3. Descriptive definitions of $I_{m,n}$

In contrast to operational definitions like (5) and (7) through (13), there is a number of what can be called descriptive definitions based on describing what $I_{m,n}$ looks like. They amount to defining

$$I_{m,n}$$

as a square matrix of order $mn$, partitioned into an $n$ by $m$ array of submatrices each of order $m$ by $n$, such that the $(i,j)$'th submatrix $S_{ij}$, say, has unity in its $(j,i)$'th position and zeros elsewhere.

For example, $I_{2,3}$ of (14) is shown partitioned in accord with (15).

Variations of (15) include Vetter's [1970] $I_{ij}^{ij}$, Hartwig's [1972, p. 540] $P$ (which should be $P^{-1}$) for his use in $P^{-1}_{m,n}(A_{m\times m} \otimes B_{n\times n})P = B \otimes A$; Vetter's [1973, 1975] $E_{m\times n}^{mxn}$; the "permutated identity matrix" $I_{(m,n)}$ of MacRae [1974], who notes (p. 338) that it "is identical to the matrix $I_{(m,n)}$ [of order $mn$] defined by Tracy and Dwyer [1969]" and the universal flip matrix $P_{m,n}$ of Hartwig and Morris [1975]. Balestra [1976, p. 21], using $P_{m,n}$, clarifies the partitioning in MacRae's [1974] definition. He also gives a symbolic version of (15):

$$I_{m,n} = P_{m,n} = \{S_{ij}\}_{i=1}^{m} \otimes_{j=1}^{n}$$

where

$$S_{ij} = e_{i} \otimes e_{j}^{'} = e_{j} \otimes e_{i}^{'} = e_{i}^{'} \otimes e_{j}^{'}$$

which is a generalization of $I_{(n,n)}$ in Anderson et al. [1977], with $S_{ij} = e_{i} \otimes e_{j}^{'}$ presented in Anderson [1978, p. 53].
A further formulation of (15) and of (16), where $H_{ij}$ is $S_{ij}$, is

$$I_{m,n} = K_{nm} = \sum_{i=1}^{n} \sum_{j=1}^{m} (H_{ij} \otimes H'_{ij})$$

where $H_{ij}$ is the $n \times m$ matrix with a 1 in position $(i,j)$ and zeros elsewhere. \hfill (17)

Vetter [1973, 1975] presents (17) as does Brewer [1977, 1978], who uses $U_{nm}$ for $I_{m,n}$. Magnus and Neudecker [1979] also use (17), which they derive from (5). They call it a 'commutation matrix' because of its role in reversing ('commuting') the order of Kronecker products, and denote it by $K_{nm}$.

Definition (15) and its variations impose an artificial partitioning on $I_{m,n}$, which disguises its inherent permutation features. Only the partitioning into $n$ blocks of $m$ rows or $m$ blocks of $n$ columns, but not both, is natural. This natural partitioning is implicit in the operational definitions and is explicit in a further formulation given by Balestra [1976]:

$$I_{m,n} = \{I \otimes e_i^{1}\}_{i=1}^{n} = \{I_{n} \otimes e_j^{j}\}_{j=1}^{m}, \hfill (18)$$

the second of which is also given by Swain [1975] who uses $I_{m,n}$ for $I_{m,n}$. Denoting the direct sum $A_1 \oplus \cdots \oplus A_n$ by $\bigoplus_{i=1}^{n} A_i$, we see that (18) may be written as

$$I_{m,n} = \{\bigoplus_{j=1}^{m} e_i^{1}\}_{i=1}^{n} = \{\bigoplus_{i=1}^{n} e_j^{j}\}_{j=1}^{m}, \hfill (19)$$

where in the first partitioning in (19) the $e_{i}^{1}$ have order $n$, and in the second the $e_{j}^{j}$ have order $m$. 

Results (18) and (19) are nicely illustrated by (14). Two further formulations of \( I_{m,n} \) developed by Magnus and Neudecker [1979] are

\[
I_{m,n} = \sum_{i=1}^{m} (e_i' \otimes I_{n} \otimes e_i) \quad \text{and} \quad I_{m,n} = \sum_{j=1}^{n} (e_j' \otimes I_{m} \otimes e_j').
\] (20)

Although Hartwig and Morris [1975] give a block formulation definition of \( I_{m,n} \), they give an ingenious card shuffling interpretation which is in fact based on (5), our definition. They identify the permutation implicit in \( I_{m,n} \) with the "generalized out faro-shuffle" for a one dimensional deck of \( mn \) cards. Their description, on p. 451, with further details in Morris and Hartwig [1976], is as follows.

"In this shuffle a deck of \( mn \) cards, labeled from top to bottom, is cut into \( m \) portions of \( n \) cards, and each portion is given in clockwise fashion to one of \( m \) players seated at a circular table, starting with the dealer. If, starting with the dealer, in clockwise fashion, each player plays his top card when it is his turn, until all cards have been played, we obtain the permutation \( \Pi \), labeled from the bottom cards up.

When \( m = 2 \) this reduces to the classical out faro-shuffle (in which an even deck of cards is cut in halves and then ruffled such that the first and the last cards remain in fixed positions) which is the basis to several remarkable card tricks."

A further application of these permutations, kindly brought to our attention by Dr. Stephen Barnett, is that of Whelchel and Guinn [1970] who are concerned with shuffling data in computer storage. In this context, they refer to \( I_{n,n} \) as the "shuffle matrix", and denote it by \( S_{n,n} \).
3.4. Defining $I_{m,n}$ as a matrix derivative operator

McDonald and Swaminathan [1973, p. 39] denote $I_{m,n}$ by $E'$, and subsequently in McDonald [1976] by $E_{n,m}$ to eliminate ambiguity, defined as the derivative operator

$$I_{m,n} = E_{n,m} = \frac{\partial X'}{\partial X}, \text{ where } X_{m \times n} \text{ is m.i.v.} \tag{21}$$

[A matrix is said to be mathematically independent and variable (m.i.v.) if no elements are functionally dependent or constant.] They recognize $E_{n,m}$ as a permutation matrix of $I_{m,n}$, and give (1973, p. 39) its form:

"the general element of $E[I_{n,m}]$, $e_{gh}$, is equal to unity if $j = k'$ and $k = j'$ with

$$g = n(j - 1) + k, \quad 0 < j \leq m, \quad 0 < k \leq n,$$

$$h = m(j' - 1) + k', \quad 0 < j' \leq n, \quad 0 < k' \leq m,$$

and is zero otherwise."

This is a less concise statement than each of (9), (11) and (13).

Rather than define $I_{m,n}$ as the outcome of differentiation, as in (21) and which demands the m.i.v. property, and which fails to highlight the permutation properties of $I_{m,n}$, we prefer to derive (21) from the more fundamental definition (5) using the standard result $\frac{\partial X}{\partial x} = A'$ for $x$ being m.i.v.; and then for $X$ being m.i.v.

$$\frac{\partial X}{\partial \text{vec} X} = \frac{\partial \text{vec} X}{\partial \text{vec} X} = \frac{\partial I_{n,m}}{\partial \text{vec} X} = (I_{n,m})' = I_{m,n}.$$
4. PROPERTIES OF VEC-PERMUTATION MATRICES

Except for (6) which we consider separately, well-known properties of \( I_{m,n} \) are given without proof by Tracy and Singh [1972a,b], Singh [1972], Vetter [1973, p. 354] who suggests verification "by construction", MacRae [1974, p. 339] who comments they "can be verified by direct examination", Balestra [1976] and Swain [1975], and very recently by Brewer [1978] and Magnus and Neudecker [1979] whose proofs seem somewhat lengthy. Using \( I_{m,n} \) as defined in (5), we develop considerably shorter and simpler proofs.

Applying (5) to vecA' in (5) itself gives vecA = \( I_{m,n} I_{m,n} \) vecA so that

\[
I_{m,n} I_{m,n} = I_{m,n} .
\]  
(22)

Then, because \( I_{m,n} \) is a permutation matrix and so is orthogonal, we have

\[
(I_{m,n})^{-1} = (I_{m,n})' = I_{n,m} .
\]  
(23)

Also because vecA = vecA'

\[
I_{m,n} = I_{m,n} = I_{m} .
\]  
(24)

[We use the notation \( (I_{m,n})' \) and \( I_{m,n}' \) interchangeably.]

4.1. Reversing the order of Kronecker products of matrices

Vec-permutation matrices are related very directly to Kronecker products through the identity (6):

\[
B_{p \times q} \otimes A_{m \times n} = I_{m, p} (A \otimes B) I_{n, q} .
\]  
(25)
Derivation of (25) comes from using the two fundamental properties of the vec operator, (4) and (5), as follows:

\[(B \otimes A)\text{vec}X = \text{vec}AXB' = I_{m,p} \text{vec}X'A'\]

\[= I_{m,p}(A \otimes B)\text{vec}X'\]

\[= I_{m,p}(A \otimes B)I_{n,q} \text{vec}X',\]

for any \(X\) of order \(n\) by \(q\). Letting vec\(X\) be in turn the columns of an identity matrix of order \(nq\) yields (25).

This concise proof was motivated by Barnett's [1973] use of the vec operator to establish (25) but without explicit development of the permutation matrices. Similar developments are also given by Hartwig and Morris [1975] and more recently by Magnus and Neudecker [1979], even though the latter do not use vec to prove results like (22) and (23).

Result (25) shows the exact form of the relationship between \(A \otimes B\) and \(B \otimes A\), which simply entails a resequencing of rows and of columns of \(A \otimes B\) to obtain \(B \otimes A\). Postmultiplying (25) by \(I_{n,q}\) and using (22) gives

\[I_{m,p}(A \otimes B) = (B \otimes A)I_{n,q}.\]  

(26)

An early reference to reversing the order of Kronecker products of rectangular matrices is Ledermann [1936] although without explicit development of permutation matrices. But Murnaghan [1938, pp. 68-69] deals with square matrices, noting for square \(A\) and \(B\), the validity of \(B \otimes A = P(A \otimes B)P^{-1}\) as in (8). Vartak [1955] generalizes this to (9), namely \(B \otimes A = P(A \otimes B)Q\), for rectangular \(A\) and \(B\), indicating that a generalization of Murnaghan's proof can be con-
structed. Searle [1966, p. 216] shows that $a_{ij}b_{rs}$ is in row $i + (r - 1)m$ of $B \otimes A$ and row $(i - 1)p + r$ of $A \otimes B$, with a similar analysis for columns, and is thus able to give the explicit form of $P$ and $Q$, as detailed after equation (10). Hartwig [1972, p. 540] gives Murnaghan's result but misprints the form for $P^{-1}$ as being that for $P$. Singh [1972, p. 22] identifies columns in (26) and thence establishes (25). Tracy and Singh [1972b], MacRae [1974] and Swain [1975] give (25) without proof; so does Vetter [1973, p. 354] who comments that it is "well-known ... though the explicit transposition relationship is not usually given". In similar vein is the remark that "the proof [not given] ... is detailed and rather tedious" by Bentler and Lee [1975, p. 148], to whom McDonald [1976, p. 90] attributes (25). Balestra [1976, p. 23] uses (18), and Brewer [1978] suggests substituting (17). These developments are tedious in comparison to our proof using the vec operator.

4.2. Trace and determinant of $\text{I}_{m,n}$

The case of $m = n$ is easy. Because $\text{I}_{m,n}$ is real, orthogonal and symmetric, it has eigenvalues $\pm 1$ with multiplicities $\frac{1}{2}n(n \pm 1)$, respectively. The determinant and trace, being the product and sum of the eigenvalues, respectively, are

$$|\text{I}_{n,n}| = (-1)^{\frac{1}{2}n(n-1)} \quad \text{and} \quad \text{tr}(\text{I}_{n,n}) = n.$$ 

For the more general situation, when $m$ and $n$ are not necessarily equal, the recurrence relation $|\text{I}_{m,n}| = (-1)^{\frac{1}{2}m(m-1)(n-1)}|\text{I}_{m,n-1}|$
yields
\[ |I_{m,n}| = (-1)^{\frac{1}{2}m(m-1)n(n-1)}. \]

This derivation is given by Hartwig and Morris [1975, p. 450] and Magnus and Neudecker [1979, p. 383] with an inductive proof presented by Swain [1975, Appendix A].

An expression for the trace is more difficult. Magnus and Neudecker [1979] prove that
\[ \text{tr}(I_{m,n}) = 1 + \gcd(m-1, n-1), \]
where \( \gcd(m,n) \) is the greatest common divisor of \( m \) and \( n \). A less compact form of this result is available in Hartwig and Morris [1975] who additionally derive expressions for the characteristic and minimal polynomials.

### 4.3. A generalization: \( M(n) \) and \( M_{m,n} \)

Tracy and Dwyer [1969] introduce

\[ M_{\sim}(n) \]

as the \( r \times c \) matrix formed by rearranging the rows of \( M \), of order \( r \times c \), by taking every \( n' \)th row starting with the first, then every \( n' \)th row starting with the second, and so on,

\[ M_{\sim}(n) = I(\sim)M. \]

For example, with dots denoting zeros, for
\[
\begin{align*}
M = \begin{bmatrix} a' \\ b' \\ c' \end{bmatrix}, \quad M(2) = \begin{bmatrix} a' \\ c' \\ b' \end{bmatrix} = \begin{bmatrix} 1 & . & . \\ . & 1 & . \\ . & . & 1 \end{bmatrix} = I(2)M.
\end{align*}
\]

Papers subsequent to Tracy and Dwyer [1969] have focused on \( I(n) \), rather than on the more general notion of \( \sim(n) \) that has wider applicability. In this context, we now define \( M_{m,n} \) for \( M \) with \( mn \) rows as

\[
M_{m,n} = I_{m,n}M
\]

\[
= \sim(n) \quad \text{when } M \text{ has } mn \text{ rows}.
\]

Our symbol \( I_{m,n} \) for the vec-permutation matrix conforms with \( M_{m,n} \) of (29) because it is \( M_{m,n} \) with \( M = I \); i.e.,

\[
I_{m,n} = I_{m,n}I,
\]

so motivating use of the letter \( I \) for the matrix \( I_{m,n} \) in preference to any other letter. In the special case of \( m = n \) there might sometimes be convenience in alternative symbols; e.g., \( P = I_{n,n} \) or \( K_n \) as introduced by Magnus and Neudecker [1979].

Within the framework of the definition (29) of \( M_{m,n} \) it can be noted that the defining property, (5), and results (25) and (26) can be rewritten:

\[
\text{vec}A = (\text{vec}A')_m,n,
\]

\[
B \otimes A = [(A \otimes B')_m,p]_n,q = [(A' \otimes B')_n,q]_m,p,
\]

and

\[
(A \otimes B)_{m,p} = (B' \otimes A')_{q,n}.
\]
4.4. Equalities from reversing the order of Kronecker products

Alternative forms of (26) are now developed. In (34), 
\( (A \otimes B)_{\sim_m, p} \) is by definition a reordering of the rows of \( A \otimes B \) by taking every \( p \)'th row starting with the first, and so on. Since \( B \) has \( p \) rows, the \( i \)'th block of rows of \( (A \otimes B)_{\sim_m, p} \) is \( A \otimes \beta'_i \), for \( \beta'_i \) of (1) so that

\[
\begin{align*}
\mathbb{I}_{m, p} (A \otimes B) = (A \otimes B)_{\sim_m, p} = \{ A \otimes \beta'_1 \} & \quad \text{for } \sim_i \sim l \sim 1 \sim i = 1 \ldots p \\
= \begin{bmatrix} A \otimes \beta'_1 \\
\vdots \\
A \otimes \beta'_p \end{bmatrix} = \{ a_j \otimes \beta'_1 \} \sim i = 1 \ldots p = \{ a_j \beta'_1 \} \sim i = 1 \ldots p \\
= \begin{bmatrix} A \otimes \beta'_1 \\
A \otimes \beta'_p \end{bmatrix}
\end{align*}
\]

(35)

Notice that \( (A \otimes B)_{\sim_m, p} \) affords a compact notation for \( \begin{bmatrix} A \otimes \beta'_1 \\
A \otimes \beta'_p \end{bmatrix} \).

Applying (35) and (36) to the right-hand side of (34) yields the further equalities:

\[
\begin{align*}
\mathbb{I}_{m, p} (A \otimes B) = (B \otimes A)_{\sim m, q} & = (B \otimes A)_{\sim n, q} \\
= (B' \otimes A')_{\sim q, n} & = \{ B' \otimes a'_j \} \sim j = 1 \ldots n = \{ B \otimes a_j \} \sim j = 1 \ldots n \\
= [B \otimes a_1 \ldots B \otimes a_n] & = \{ \beta'_1 \otimes a_j \} \sim i = 1 \ldots p = \{ a_j \beta'_1 \} \sim j = 1 \ldots n .
\end{align*}
\]

(37)

(38)

These equalities lead to the equivalent descriptive formulations of \( \mathbb{I}_{m, n} \), (15), (16), (17) and (18), on substituting \( A = \mathbb{I}_m \) and \( B = \mathbb{I}_n \) in (35) - (38) to give \( \mathbb{I}_{m, n} = (\mathbb{I}_m \otimes \mathbb{I}_n)_{m, n} \).
4.5. Reversing the order of Kronecker products of vectors

The special cases of \( A \) and/or \( B \) being vectors are worthy of note. For column vectors \( a \) and \( b \) of order \( m \) and \( p \), respectively, putting \( n = 1 \), then \( q = 1 \) and then \( n = q = 1 \), in (26) gives

\[
(a \otimes B)_{m, p} = I_{m, p} (a \otimes B) = B \otimes a ,
\]

(39)

\[
(A \otimes b)_{m, p} = I_{m, p} (A \otimes b) = b \otimes A ,
\]

(40)

and

\[
(a \otimes b)_{m, p} = I_{m, p} (a \otimes b) = b \otimes a .
\]

(41)

Transposing (39) - (41), or putting \( p = 1 \) and \( m = 1 \) in (26), gives the corresponding results for row vectors. Using \( n = 1 \) and \( p = 1 \) in (26), (36) and (38) reveals the familiar commutativity property of Kronecker products of vectors:

\[
\text{a} \otimes b' = ab' = b' \otimes a .
\]

4.6. The vec-permutation matrix and Kronecker products of 3 or more matrices

Extending the vec-permutation matrix to cyclically permute order in Kronecker products of three or more matrices is straightforward, as indicated by Magnus and Neudecker [1979]. Balestra [1976, p. 24] introduced \( I_{mp,s} \), for use in this connection but perhaps surprisingly did not explicitly give its use to cyclically permute a Kronecker product of three matrices. We present this result by immediate application of (25) to \( (A \otimes B) \otimes C = A \otimes (B \otimes C) \) for \( A, B \) and \( C \) of order \( m \times n, p \times q \) and \( s \times t \), respectively:
Properties of vec-permutation matrices with the same set of
three indices, developed by Balestra [1976, pp. 24-25] and Magnus
and Neudecker [1979] with lengthy algebra, are now shown to be easy
consequences of (42) - (44): put \( n = q = t = 1 \) in (42) - (44), so
that \( A \otimes B \otimes C \) becomes \( a \otimes b \otimes c \) and the final vec-permutation
matrices on the right-hand sides of (42) - (44) reduce to 1, giving

\[
I_{\sim m, s} (a \otimes b \otimes c) = I_{\sim m, s} (a \otimes b \otimes c)
\]

\[
= (I_{\sim m, s} \otimes I_{\sim p}) (I_{\sim m} \otimes I_{\sim p, s}) (a \otimes b \otimes c).
\]  

Let \( a \otimes b \otimes c \) take in turn the columns of \( I_{\sim m, s} \), and hence

\[
I_{\sim m, s} = I_{\sim m, s} (I_{\sim m, s} \otimes I_{\sim p}) (a \otimes b \otimes c).
\]  

Since \( I_{\sim m, s} = I_{\sim m, s} \), interchanging \( m \) and \( p \) in (46) yields

\[
I_{\sim m, s} = I_{\sim m, s} = I_{\sim m, s} (I_{\sim m, s} \otimes I_{\sim m, s}).
\]  

Postmultiplying the middle two equalities in (47) by \( I_{\sim s, mp} \) gives,
using (22),

\[
I_{\sim m, ps} I_{\sim p, ms} I_{\sim s, mp} = I_{\sim m, ps}.
\]
This development, based on (45), is in contrast to the lengthy manipulations of Balestra [1976] who uses (18) to establish (46) - (48), and Magnus and Neudecker [1979] who give (42) and use (20) to develop (48) and all of (47) except its final equality.

A special case of (42) is

\[(a \otimes b \otimes C)_{m,s} = b' \otimes C \otimes a.\]  

(49)

Using \(a \otimes b' = ab'\), (49) becomes

\[(ab' \otimes C)_{m,s} = b' \otimes C \otimes a\]  

(50)

which, applied on

\[I_{m,n} = (I_{m} \otimes I_{n})_{m,n} = \left(\sum_{i=1}^{m} e_{i} e_{i}' \otimes I_{n}\right)_{m,n},\]  

(51)

with \(C = I_{n}\) and \(a = b = e_{i}\), yields the Magnus and Neudecker [1979] formulations for \(I_{m,n}\) in (20).

Generalization to Kronecker products of four or more matrices and vec-permutation matrices with four or more indices is straightforward.

ACKNOWLEDGMENTS

Grateful thanks go to Robert Anderson, Stephen Barnett, Paul Dwyer, Robert Hartwig and Friedrich Pukelsheim for suggested improvements to early drafts of this paper.
REFERENCES


Rogers, G. S. and Young, D. L. [1978]. On testing a multivariate linear hypothesis when the covariance matrix and its inverse have the same pattern. *J. Amer. Statist. Assoc.*, 73, 203-207.


Table 1: Definitions, notations and names for what, in this paper, is called the vec-permutation matrix, $\mathbf{I}_m,n$.  

<table>
<thead>
<tr>
<th>Definition</th>
<th>Writer</th>
<th>Notation</th>
<th>Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reversing (a) vectors</td>
<td>Ledermann [1936]</td>
<td>$\mathbf{P}$ and $\mathbf{Q}$</td>
<td>Permutation matrix</td>
</tr>
<tr>
<td></td>
<td>Conlisk [1976] (incorrect)</td>
<td>$\mathbf{M}$ for $\mathbf{I}_{n,n}$</td>
<td>Permutation matrix</td>
</tr>
<tr>
<td>Kronecker products (b) matrices</td>
<td>Murnaghan [1938]</td>
<td>$\mathbf{P}$</td>
<td>Permutation matrix</td>
</tr>
<tr>
<td></td>
<td>Vartak [1955]</td>
<td>$\mathbf{P}$ and $\mathbf{Q}$</td>
<td>Elementary operators</td>
</tr>
<tr>
<td></td>
<td>Searle [1966]</td>
<td>$\mathbf{P}$ and $\mathbf{Q}$</td>
<td>Elementary operators</td>
</tr>
<tr>
<td></td>
<td>Whelchel and Guinn [1970]</td>
<td>$\mathbf{S}<em>{n}^2$ for $\mathbf{I}</em>{n,n}$</td>
<td>Shuffle matrix</td>
</tr>
<tr>
<td>Defining row permutations</td>
<td>Tracy and Dwyer [1969]</td>
<td>$\mathbf{I}(n)$</td>
<td>Permutated identity matrix</td>
</tr>
<tr>
<td></td>
<td>Tracy and Singh [1972a,b]</td>
<td>$\mathbf{I}(n)$</td>
<td>Permutated identity matrix</td>
</tr>
<tr>
<td></td>
<td>Singh [1972]</td>
<td>$\mathbf{E}_{n,m}$</td>
<td></td>
</tr>
<tr>
<td>The derivative $\frac{\partial \mathbf{x}}{\partial \mathbf{X}}$</td>
<td>McDonald and Swaminathan [1973]</td>
<td>$\mathbf{E}$</td>
<td>Permutation matrix</td>
</tr>
<tr>
<td></td>
<td>Bentler and Lee [1975]</td>
<td>$\mathbf{E}$</td>
<td>Permutation matrix</td>
</tr>
<tr>
<td></td>
<td>McDonald [1976]</td>
<td>$\mathbf{E}_{n,m}$</td>
<td></td>
</tr>
<tr>
<td>Description of elements</td>
<td>Vetter [1970]</td>
<td>$\langle \mathbf{U}^{d_1}<em>{i,j} \mathbf{U}^{d_2}</em>{i,j} \rangle$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Hartwig [1972] (incorrect)</td>
<td>$\mathbf{P}$</td>
<td>Permutation matrix</td>
</tr>
<tr>
<td></td>
<td>Vetter [1973, 1975]</td>
<td>$\mathbf{P}^{m \times n}_{n \times m}$</td>
<td>Permutation matrix</td>
</tr>
<tr>
<td></td>
<td>MacRae [1974]</td>
<td>$\mathbf{I}_{(m,n)}$</td>
<td>Permutated identity matrix</td>
</tr>
<tr>
<td></td>
<td>Hartwig and Morris [1975]</td>
<td>$\mathbf{P}_{m,n}$</td>
<td>Universal flip matrix</td>
</tr>
<tr>
<td></td>
<td>Swain [1975]</td>
<td>$\mathbf{P}_{m,n}$</td>
<td>Permutation matrix</td>
</tr>
<tr>
<td></td>
<td>Balestra [1976]</td>
<td>$\mathbf{P}_{m,n}$</td>
<td>Permutated identity matrix</td>
</tr>
<tr>
<td></td>
<td>Brewer [1977, 1978]</td>
<td>$\mathbf{P}_{m,n}$</td>
<td>Permutated identity matrix</td>
</tr>
<tr>
<td></td>
<td>Magnus and Neudecker [1979]</td>
<td>$\mathbf{K}_{n,m}$</td>
<td>Commutation matrix</td>
</tr>
<tr>
<td>vecA related to vecA'</td>
<td>Barnett [1973]</td>
<td>$\mathbf{P}$ and $\mathbf{Q}$</td>
<td>Permutation matrix</td>
</tr>
<tr>
<td></td>
<td>Henderson and Searle [1979]</td>
<td>$\mathbf{I}_{(m,n)}$</td>
<td>vec-permutation matrix</td>
</tr>
<tr>
<td></td>
<td>Henderson and Searle (this paper)</td>
<td>$\mathbf{I}_{m,n}$</td>
<td>vec-permutation matrix</td>
</tr>
</tbody>
</table>