VEC AND VECH OPERATORS FOR MATRICES, WITH SOME USES IN JACOBIANS AND MULTIVARIATE STATISTICS

BU-643-M

by

April, 1978

H. V. Henderson* and S. R. Searle
Biometrics Unit, Cornell University, Ithaca, New York

Abstract

The vec of a matrix $\mathbf{X}$ stacks columns of $\mathbf{X}$ one under another in a single column; the vech of a square matrix $\mathbf{X}$ does the same thing but starting each column at its diagonal element. The Jacobian of a one-to-one transformation $\mathbf{X} \to \mathbf{Y}$ is then $\|\partial(\mathrm{vec}\mathbf{X})/\partial(\mathrm{vec}\mathbf{Y})\|$ when $\mathbf{X}$ and $\mathbf{Y}$ are non-symmetric and it is $\|\partial(\mathrm{vech}\mathbf{X})/\partial(\mathrm{vech}\mathbf{Y})\|$ when $\mathbf{X}$ and $\mathbf{Y}$ are symmetric. Kronecker product properties of $\mathrm{vec}(\mathbf{ABC}) = (\mathbf{C}' \otimes \mathbf{A})\mathrm{vec}\mathbf{B}$ permit easy evaluation of this determinant in many cases. They also provide succinct descriptions in multivariate statistics.

1. Introduction

An operator on matrices in which there has recently been resurgent interest in statistics is that of stacking the columns of a matrix one underneath the other to form a single vector. The operator is coming to be known as vec. Thus, for $\mathbf{x}_i$, $i = 1, \cdots, c$, being the c columns of the $r \times c$ matrix $\mathbf{X}$,

$$\mathbf{X} = [\mathbf{x}_1\ \mathbf{x}_2\ \cdots\ \mathbf{x}_c] \quad \text{and} \quad \mathrm{vec}\mathbf{X} = \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \\ \vdots \\ \mathbf{x}_c \end{bmatrix}. \tag{1}$$

Paper No. BU-643-M in the Biometrics Unit Mimeo Series.
* Supported by a New Zealand National Research Advisory Council Fellowship.

An early reference to this idea is Sylvester [1884], who used it in connection with linear equations. Roth [1934] develops results for using the operation on a product matrix, Aitken [1949] mentions the idea in connection with Jacobians and, more recently, Neudecker [1968, 1969] has exploited the concept in a variety of ways for statistics. The operation has been referred to variously as the column string of $\underset{\sim}{X}$ and the pack of $\underset{\sim}{X}$, with vec$\underset{\sim}{X}$ (for "vector form of $\underset{\sim}{X}$") being the description currently in vogue. The equivalent notations vec$\underset{\sim}{X}$ and vec$(\underset{\sim}{X})$ are used interchangeably, the parentheses being employed only when deemed necessary for clarity.

Variations on vec$\underset{\sim}{X}$ are also available. For $\underset{\sim}{X}$ square, Searle [1978] defines vech$\underset{\sim}{X}$ in the same way that vec$\underset{\sim}{X}$ is defined, except that for each column of $\underset{\sim}{X}$ only that part of it which is on or below the diagonal of $\underset{\sim}{X}$ is put into vech$\underset{\sim}{X}$ ("vector-half" of $\underset{\sim}{X}$). In this way vech$\underset{\sim}{X}$, for $\underset{\sim}{X}$ symmetric, contains only the distinctly different elements* of $\underset{\sim}{X}$, a feature that is useful in deriving Jacobians for transformations from one symmetric array of variables to another. For example, with

$$\underset{\sim}{X} = \begin{bmatrix} a & b \\ b & c \end{bmatrix}, \quad (\text{vec}\underset{\sim}{X})' = [a\ b\ b\ c] \quad \text{and} \quad (\text{vech}\underset{\sim}{X})' = [a\ b\ c]. \tag{2}$$

Ideas similar to vech$\underset{\sim}{X}$ for symmetric $\underset{\sim}{X}$, but with different notations, have been used by Tracy and Singh [1972] and Vetter [1975], and also by Aitken [1949] and by Browne [1974], who confine attention to elements of $\underset{\sim}{X}$ on and above the diagonal.

A generalization of vech$\underset{\sim}{X}$ for $\underset{\sim}{X}$ symmetric is that of putting into a single vector just the distinctly different elements of any patterned matrix in which, solely on account of the pattern, some elements occur more than once. We then define vecp$(\underset{\sim}{X})$ as the vector of distinctly different elements of a patterned matrix $\underset{\sim}{X}$, the exact elements in vecp$(\underset{\sim}{X})$ being determined by the nature of the pattern in $\underset{\sim}{X}$.

---

* Distinctly different means functionally independent, so that if a and b are distinctly different then $\partial a/\partial b = 0$.

For example, when $\underset{\sim}{X}$ is skew-symmetric we define veck($\underset{\sim}{X}$) (for "vector skew" of $\underset{\sim}{X}$) as the vector made up from the elements below the diagonal of $\underset{\sim}{X}$; e.g.,

$$\underset{\sim}{X} = \begin{bmatrix} 0 & -a & -b \\ a & 0 & -c \\ b & c & 0 \end{bmatrix}, \quad \text{veck}(\underset{\sim}{X}) = \begin{bmatrix} a \\ b \\ c \end{bmatrix}. \tag{3}$$

Again, this turns out to be a useful concept in deriving Jacobians. Other examples could be illustrated, such as vecp($\underset{\sim}{X}$) for triangular, circulant, centrosymmetric, or diagonal matrices, and so on. Details and extensions of these ideas (extensions to orthogonal matrices, for example), will be reported by one of us (H.V.H.) in a dissertation.

2. Kronecker Products

The vec of a product of matrices gives rise to the Kronecker product of matrices, defined for $\underset{\sim}{A}_{p\times q} = \{a_{ij}\}$ and $\underset{\sim}{B}_{r\times s}$, of orders $p \times q$ and $r \times s$ respectively, as $\underset{\sim}{A} \otimes \underset{\sim}{B} = \{a_{ij}\underset{\sim}{B}\}_{pr\times qs}$ with $i = 1, \cdots, p$ and $j = 1, \cdots, q$. Well-known properties of the Kronecker product include

$$\begin{aligned} (\underset{\sim}{A} \otimes \underset{\sim}{B})(\underset{\sim}{X} \otimes \underset{\sim}{Y}) &= \underset{\sim}{AX} \otimes \underset{\sim}{BY}, & (\underset{\sim}{A} \otimes \underset{\sim}{B})^- &= \underset{\sim}{A}^- \otimes \underset{\sim}{B}^-, \\ |\underset{\sim}{A}_{p\times p} \otimes \underset{\sim}{B}_{r\times r}| &= |\underset{\sim}{A}|^r\,|\underset{\sim}{B}|^p \quad \text{and} & (\underset{\sim}{A} \otimes \underset{\sim}{B})' &= \underset{\sim}{A}' \otimes \underset{\sim}{B}', \end{aligned} \tag{4}$$

where, in these expressions, the necessary rank and conformability conditions for their existence are assumed to be satisfied, and where $\underset{\sim}{A}^-$ is a generalized inverse of $\underset{\sim}{A}$ such that $\underset{\sim}{AA}^-\underset{\sim}{A} = \underset{\sim}{A}$.

The relationship between $\underset{\sim}{A} \otimes \underset{\sim}{B}$ and $\underset{\sim}{B} \otimes \underset{\sim}{A}$ is clearly one of sequencing of rows and columns, since both have the same order and contain the same terms as elements, but arranged in different sequences. Pre- and post-multiplication of either

Kronecker product by appropriate permutation matrices therefore yields the other, the particular form of permutation matrix in this instance being denoted by $\underset{\sim}{I}_{(p,q)}$. Tracy and Singh [1972] (using the symbol $\underset{\sim}{I}_{(q)}$) define $\underset{\sim}{I}_{(p,q)}$ as a rearrangement of rows of $\underset{\sim}{I}_{pq}$, obtained by taking every q'th row starting at the first, then every q'th row starting at the second, and so on. For example, the rows of

$$\underset{\sim}{I}_{(2,3)} = \left[\begin{array}{ccc|ccc} 1 & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & 1 & \cdot & \cdot \\ \hline \cdot & 1 & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & 1 & \cdot \\ \hline \cdot & \cdot & 1 & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & 1 \end{array}\right]$$

(where dots represent zeros), are rows 1, 4, 2, 5, 3 and 6, respectively, of $\underset{\sim}{I}_6$. From this it is easily seen that

$$\underset{\sim}{I}_{(p,1)} = \underset{\sim}{I}_{(1,p)} = \underset{\sim}{I}_p, \quad [\underset{\sim}{I}_{(p,q)}]' = \underset{\sim}{I}_{(q,p)} \quad \text{and} \quad \underset{\sim}{I}_{(p,q)}\underset{\sim}{I}_{(q,p)} = \underset{\sim}{I}_{pq}, \tag{5}$$

and that

$$\underset{\sim}{B}_{r\times s} \otimes \underset{\sim}{A}_{p\times q} = \underset{\sim}{I}_{(p,r)}(\underset{\sim}{A}_{p\times q} \otimes \underset{\sim}{B}_{r\times s})\underset{\sim}{I}_{(s,q)} \,. \tag{6}$$

A particular example of (6), derived from using (5), is

$$\underset{\sim}{I}_{(n,n)}(\underset{\sim}{B}_{n\times n} \otimes \underset{\sim}{A}_{n\times n}) = (\underset{\sim}{A}_{n\times n} \otimes \underset{\sim}{B}_{n\times n})\underset{\sim}{I}_{(n,n)} \tag{7}$$

for $\underset{\sim}{A}$ and $\underset{\sim}{B}$ square, of order n. Another example is

$$\underset{\sim}{A}_{p\times p} \otimes \underset{\sim}{B}_{r\times r} = (\underset{\sim}{A}_{p\times p} \otimes \underset{\sim}{I}_r)(\underset{\sim}{I}_p \otimes \underset{\sim}{B}_{r\times r}) = \underset{\sim}{I}_{(r,p)}(\underset{\sim}{I}_r \otimes \underset{\sim}{A}_{p\times p})\underset{\sim}{I}_{(p,r)}(\underset{\sim}{I}_p \otimes \underset{\sim}{B}_{r\times r})$$

which leads to the determinantal result in (4) because $\underset{\sim}{I}_r \otimes \underset{\sim}{A}_{p\times p}$ is block diagonal with r matrices $\underset{\sim}{A}_{p\times p}$ on its diagonal.

Although the definition of $\underset{\sim}{I}_{(p,q)}$ is given by Tracy and Singh [1972], it was MacRae [1974] who introduced the notation $\underset{\sim}{I}_{(p,q)}$, to emphasize its order, namely pq. She called it a <u>permuted</u> <u>identity</u> matrix (a name which actually describes any permutation matrix), and defined $\underset{\sim}{I}_{(p,q)}$ as square, of order pq, partitioned into q rows and p columns of submatrices of order p $\times$ q, such that the (i,j)'th such submatrix has unity as its (j,i)'th element and zeros elsewhere. (Actually, MacRae's [1974] definition is a mite vague, in regard to specifying the partitioning and the order of the submatrices, but her uses of $\underset{\sim}{I}_{(p,q)}$ leave no doubt that the preceding definition is what is meant.)

3. Properties of the Vec and Vech Operators

3.1. Products, traces and transposes

The definitions of vec and of Kronecker product show that $\text{vec}(\underset{\sim}{x}\underset{\sim}{y}') = \underset{\sim}{y} \otimes \underset{\sim}{x}$, an extension of which is

$$\text{vec}(\underset{\sim}{A}\underset{\sim}{B}\underset{\sim}{C}) = (\underset{\sim}{C}' \otimes \underset{\sim}{A})\text{vec}\underset{\sim}{B} \ , \tag{8}$$

derived by Roth [1934] and rediscovered by Neudecker [1969]. From this it is easy to get

$$\text{vec}(\underset{\sim}{A}\underset{\sim}{B}) = (\underset{\sim}{I} \otimes \underset{\sim}{A})\text{vec}\underset{\sim}{B} = (\underset{\sim}{B}' \otimes \underset{\sim}{A})\text{vec}\underset{\sim}{I} = (\underset{\sim}{B}' \otimes \underset{\sim}{I})\text{vec}\underset{\sim}{A} \tag{9}$$

and for $\underset{\sim}{A}$ non-singular this leads to

$$\text{vec}\underset{\sim}{A}^{-1} = (\underset{\sim}{A}^{-1'} \otimes \underset{\sim}{A}^{-1})\text{vec}\underset{\sim}{A} \ . \tag{10}$$

A second useful result from Neudecker [1969] is

$$\text{tr}(\underset{\sim}{A}\underset{\sim}{B}) = (\text{vec}\underset{\sim}{A}')'\text{vec}\underset{\sim}{B} \tag{11}$$

which, together with (8), gives

$$\text{tr}(\underset{\sim}{A}\underset{\sim}{X}'\underset{\sim}{B}\underset{\sim}{X}\underset{\sim}{C}) = (\text{vec}\underset{\sim}{X})'(\underset{\sim}{A}'\underset{\sim}{C}' \otimes \underset{\sim}{B})\text{vec}\underset{\sim}{X} = (\text{vec}\underset{\sim}{X})'(\underset{\sim}{C}\underset{\sim}{A} \otimes \underset{\sim}{B}')\text{vec}\underset{\sim}{X} \ . \tag{12}$$

And for the vec of a transposed matrix $\underset{\sim}{A}'$, we have

$$\mathrm{vec}[(\underset{\sim}{A}')_{q\times p}] = \underset{\sim}{I}_{(q,p)}\mathrm{vec}(\underset{\sim}{A}_{p\times q}) \tag{13}$$

for $\underset{\sim}{I}_{(q,p)}$ defined in Section 2.

3.2. Relationships for symmetric matrices

The number of elements in $\mathrm{vech}(\underset{\sim}{X}_{n\times n})$ is, by the definition of the vech operator, $\frac{1}{2}n(n+1)$. When $\underset{\sim}{X}$ is symmetric, the elements of $\mathrm{vec}\underset{\sim}{X}$ are those of $\mathrm{vech}\underset{\sim}{X}$ with some repetitions. Therefore $\mathrm{vec}\underset{\sim}{X}$ and $\mathrm{vech}\underset{\sim}{X}$ for symmetric $\underset{\sim}{X}$ are linear transformations of one another. We represent these transformations by the matrices $\underset{\sim}{H}$ and $\underset{\sim}{G}$ defined, respectively, by the equalities

$$\mathrm{vech}\underset{\sim}{X} = \underset{\sim}{H}\mathrm{vec}\underset{\sim}{X}, \quad \mathrm{vec}\underset{\sim}{X} = \underset{\sim}{G}\mathrm{vech}\underset{\sim}{X} \quad \text{and} \quad \mathrm{vec}\underset{\sim}{X} = \underset{\sim}{GH}\mathrm{vec}\underset{\sim}{X} \tag{14}$$

where, for $\underset{\sim}{X} = \underset{\sim}{X}'$ of order $n \times n$,

$$\underset{\sim}{H} \text{ is } \tfrac{1}{2}n(n+1) \times n^2 \quad \text{and} \quad \underset{\sim}{G} \text{ is } n^2 \times \tfrac{1}{2}n(n+1)\ . \tag{15}$$

Examples of $\underset{\sim}{H}$ and $\underset{\sim}{G}$ for $n = 3$, shown alongside $\underset{\sim}{I}_{(3,3)}$, with which they have several relationships, are

$$\underset{\sim}{I}_{(3,3)} = \left[\begin{array}{ccc|ccc|ccc} 1&\cdot&\cdot&\cdot&\cdot&\cdot&\cdot&\cdot&\cdot\\ \cdot&\cdot&\cdot&1&\cdot&\cdot&\cdot&\cdot&\cdot\\ \cdot&\cdot&\cdot&\cdot&\cdot&\cdot&1&\cdot&\cdot\\ \hline \cdot&1&\cdot&\cdot&\cdot&\cdot&\cdot&\cdot&\cdot\\ \cdot&\cdot&\cdot&\cdot&1&\cdot&\cdot&\cdot&\cdot\\ \cdot&\cdot&\cdot&\cdot&\cdot&\cdot&\cdot&1&\cdot\\ \hline \cdot&\cdot&1&\cdot&\cdot&\cdot&\cdot&\cdot&\cdot\\ \cdot&\cdot&\cdot&\cdot&\cdot&1&\cdot&\cdot&\cdot\\ \cdot&\cdot&\cdot&\cdot&\cdot&\cdot&\cdot&\cdot&1 \end{array}\right], \quad \underset{\sim}{H}_{6\times 9} = \left[\begin{array}{ccc|ccc|ccc} 1&\cdot&\cdot&\cdot&\cdot&\cdot&\cdot&\cdot&\cdot\\ \cdot&\frac{1}{2}&\cdot&\frac{1}{2}&\cdot&\cdot&\cdot&\cdot&\cdot\\ \cdot&\cdot&\frac{1}{2}&\cdot&\cdot&\cdot&\frac{1}{2}&\cdot&\cdot\\ \hline \cdot&\cdot&\cdot&\cdot&1&\cdot&\cdot&\cdot&\cdot\\ \cdot&\cdot&\cdot&\cdot&\cdot&\frac{1}{2}&\cdot&\frac{1}{2}&\cdot\\ \hline \cdot&\cdot&\cdot&\cdot&\cdot&\cdot&\cdot&\cdot&1 \end{array}\right] \quad \text{and} \quad \underset{\sim}{G}_{9\times 6} = \left[\begin{array}{ccc|cc|c} 1&\cdot&\cdot&\cdot&\cdot&\cdot\\ \cdot&1&\cdot&\cdot&\cdot&\cdot\\ \cdot&\cdot&1&\cdot&\cdot&\cdot\\ \hline \cdot&1&\cdot&\cdot&\cdot&\cdot\\ \cdot&\cdot&\cdot&1&\cdot&\cdot\\ \cdot&\cdot&\cdot&\cdot&1&\cdot\\ \hline \cdot&\cdot&1&\cdot&\cdot&\cdot\\ \cdot&\cdot&\cdot&\cdot&1&\cdot\\ \cdot&\cdot&\cdot&\cdot&\cdot&1 \end{array}\right]. \tag{16}$$

The equation $\mathrm{vec}\underset{\sim}{X} = \underset{\sim}{G}\mathrm{vech}\underset{\sim}{X}$ of (14), for symmetric $\underset{\sim}{X}$ of order n, is true for all such $\underset{\sim}{X}$ for a unique $\underset{\sim}{G}$, of the form shown in (16) for n = 3. But this uniqueness does not apply to $\underset{\sim}{H}$; the form in (16) can be modified, for example, such that in each row that contains two elements $\frac{1}{2}$, one of them can be α and the other $1-\alpha$ for any α, with a different α for each such row, if desired. Furthermore, $\underset{\sim}{G}$ is not only unique in the manner described, but it has full column rank $\frac{1}{2}n(n+1)$; and $\underset{\sim}{H}$ always has full row rank.

Two useful properties of $\underset{\sim}{G}$ and $\underset{\sim}{H}$ that hold for all forms of $\underset{\sim}{H}$ are

$$\underset{\sim}{H}\underset{\sim}{G} = \underset{\sim}{I}_{\frac{1}{2}n(n+1)} \quad \text{and} \quad \underset{\sim}{I}_{(n,n)}\underset{\sim}{G} = \underset{\sim}{G} . \tag{17}$$

This means that $\underset{\sim}{H}$ is a left inverse of $\underset{\sim}{G}$, one form of which is

$$\underset{\sim}{H} = (\underset{\sim}{G}'\underset{\sim}{G})^{-1}\underset{\sim}{G}' . \tag{18}$$

This is the form illustrated in (16), for which we also have

$$\underset{\sim}{G}\underset{\sim}{H} = \tfrac{1}{2}[\underset{\sim}{I}_{n^2} + \underset{\sim}{I}_{(n,n)}] \quad \text{and} \quad \underset{\sim}{H}\underset{\sim}{I}_{(n,n)} = \underset{\sim}{H} . \tag{19}$$

3.3. Relationships for square matrices

Suppose $\underset{\sim}{A}$ is square of order n, but not necessarily symmetric. Then pre-multiplying (13) by $\underset{\sim}{G}'$ and using (5) and (17) gives $\underset{\sim}{G}'\text{vec}\underset{\sim}{A} = \underset{\sim}{G}'\text{vec}\underset{\sim}{A}'$. Define

$$\begin{aligned} \text{diag}\underset{\sim}{A} &\equiv \text{diagonal matrix of diagonal elements of } \underset{\sim}{A} , \\ &\equiv \{\delta_{ij}a_{ij}\} \end{aligned} \tag{20}$$

for $i,j = 1, \cdots, n$ and where δ_{ij} is the Kronecker delta. Then it is readily shown that

$$\underset{\sim}{G}'\text{vec}\underset{\sim}{A} = \underset{\sim}{G}'\text{vec}\underset{\sim}{A}' = \text{vech}(\underset{\sim}{A} + \underset{\sim}{A}' - \text{diag}\underset{\sim}{A}) \tag{21}$$

where

$$\underset{\sim}{A} + \underset{\sim}{A}' - \text{diag}\underset{\sim}{A} = \{a_{ij} + a_{ji} - \delta_{ij}a_{ij}\} \tag{22}$$

for $i, j = 1, \cdots, n$. It might be noted in passing that

$$\text{for symmetric } \underset{\sim}{A}, \qquad \underset{\sim}{A} + \underset{\sim}{A}' - \text{diag}\underset{\sim}{A} = \underset{\sim}{0} \Rightarrow \underset{\sim}{A} = \underset{\sim}{0}$$

and (23)

$$\text{otherwise} \qquad \underset{\sim}{A} + \underset{\sim}{A}' - \text{diag}\underset{\sim}{A} = \underset{\sim}{0} \Rightarrow \underset{\sim}{A} = -\underset{\sim}{A}' \, .$$

Suppose that with symmetric $\underset{\sim}{X}$ we use $\underset{\sim}{A}\underset{\sim}{X}\underset{\sim}{A}'$ in place of $\underset{\sim}{X}$, in the last equality of (14). Then, on also using (8) we get

$$(\underset{\sim}{A} \otimes \underset{\sim}{A})\text{vec}\underset{\sim}{X} = \underset{\sim}{G}\underset{\sim}{H}(\underset{\sim}{A} \otimes \underset{\sim}{A})\text{vec}\underset{\sim}{X} \, . \tag{24}$$

And so, on using the second equality in (14), we have

$$(\underset{\sim}{A} \otimes \underset{\sim}{A})\underset{\sim}{G}\text{vech}\underset{\sim}{X} = \underset{\sim}{G}\underset{\sim}{H}(\underset{\sim}{A} \otimes \underset{\sim}{A})\underset{\sim}{G}\text{vech}\underset{\sim}{X} \, . \tag{25}$$

In (24) and (25) $\underset{\sim}{X}$ is symmetric. Therefore, in (24) vec$\underset{\sim}{X}$ always has some elements repeated, whereas this is not so for vech$\underset{\sim}{X}$ in (25). There, vech$\underset{\sim}{X}$ can take any value we please. We therefore let vech$\underset{\sim}{X}$ take, in turn, the values of the columns of $\underset{\sim}{I}$ of order $\frac{1}{2}n(n+1)$ and so derive

$$(\underset{\sim}{A} \otimes \underset{\sim}{A})\underset{\sim}{G} = \underset{\sim}{G}\underset{\sim}{H}(\underset{\sim}{A} \otimes \underset{\sim}{A})\underset{\sim}{G} \, . \tag{26}$$

From this, for $\underset{\sim}{H}$ of (18) and (19), it is then easy to derive

$$\underset{\sim}{H}(\underset{\sim}{A} \otimes \underset{\sim}{A}) = \underset{\sim}{H}(\underset{\sim}{A} \otimes \underset{\sim}{A})\underset{\sim}{G}\underset{\sim}{H} \, . \tag{27}$$

Determinants that occur in certain Jacobians can be evaluated with the aid of (26), as is now shown.

3.4. Determinants

We prove the following result:

$$|\underset{\sim}{H}(\underset{\sim}{A} \otimes \underset{\sim}{A})\underset{\sim}{G}| = |\underset{\sim}{A}|^{n+1} \tag{28}$$

for $\underset{\sim}{A}$ of order n.

Proof: Consider first the case where the canonical form of A under similarity $U^{-1}AU = D$ exists, with D being a diagonal matrix of the n latent roots λ_i of A for $i = 1, \cdots, n$. Then

$$|H(A \otimes A)G| = |H(UDU^{-1} \otimes UDU^{-1})G|$$

$$= |H(U \otimes U)(D \otimes D)(U^{-1} \otimes U^{-1})G|$$

$$= |H(U \otimes U)GH(D \otimes D)GH(U^{-1} \otimes U^{-1})G|, \quad \text{from (26)}$$

$$= |H(U \otimes U)G| \; |H(U^{-1} \otimes U^{-1})G| \; |H(D \otimes D)G|$$

$$= |H(U \otimes U)GH(U^{-1} \otimes U^{-1})G| \; |H(D \otimes D)G|$$

$$= |H(U \otimes U)(U^{-1} \otimes U^{-1})G| \; |H(D \otimes D)G|, \quad \text{from (26)}$$

$$= |HG| \; |H(D \otimes D)G| \tag{29}$$

$$= |H(D \otimes D)G|, \quad \text{from (17).} \tag{30}$$

The nature of D means that $D \otimes D$ is a diagonal matrix with elements $\lambda_i\lambda_j$ for $i,j = 1, \cdots, n$ in lexicon order and so,

$$|H(A \otimes A)G| = |H\{G \text{ with } (i,j)\text{'th row multiplied by } \lambda_i\lambda_j\}|$$

$$= \prod_{i=1}^{n} \lambda_i\lambda_j \prod_{j>i}^{n} \tfrac{1}{2}(\lambda_i\lambda_j + \lambda_j\lambda_i) \text{ by the nature of } H \text{ as illustrated in (16)} \tag{31}$$

$$= \prod_{i=1}^{n} \prod_{j\geq i}^{n} \lambda_i\lambda_j = \prod_{i=1}^{n} \lambda_i^{n+1} = |A|^{n+1}, \tag{32}$$

which is (28). It is easily shown that this result is invariant to whatever left inverse of G is used, for H.

If $\underset{\sim}{A}$ is such that $\underset{\sim}{U}^{-1}$ does not exist in $\underset{\sim}{A}\underset{\sim}{U} = \underset{\sim}{U}\underset{\sim}{D}$, we say that $\underset{\sim}{A}$ is defective, in which case the singular-valued decomposition $\underset{\sim}{L}\underset{\sim}{A}\underset{\sim}{M}' = \underset{\sim}{\Delta}$ can be used in the development of (30) in place of $\underset{\sim}{U}^{-1}\underset{\sim}{A}\underset{\sim}{U} = \underset{\sim}{D}$. In $\underset{\sim}{L}\underset{\sim}{A}\underset{\sim}{M}' = \underset{\sim}{\Delta}$, $\underset{\sim}{L}$ and $\underset{\sim}{M}$ are orthogonal matrices and $\underset{\sim}{\Delta}^2$ is the diagonal matrix of eigenvalues of $\underset{\sim}{A}\underset{\sim}{A}'$. ($\underset{\sim}{A}$ here is square and diagonal elements of $\underset{\sim}{\Delta}$ are the positive square roots of those of $\underset{\sim}{\Delta}^2$.) It will then be found that (29) becomes

$$\begin{aligned}
|\underset{\sim}{H}(\underset{\sim}{A} \otimes \underset{\sim}{A})\underset{\sim}{G}| &= |\underset{\sim}{H}(\underset{\sim}{L}'\underset{\sim}{M} \otimes \underset{\sim}{L}'\underset{\sim}{M})\underset{\sim}{G}| \; |\underset{\sim}{H}(\underset{\sim}{\Delta} \otimes \underset{\sim}{\Delta})\underset{\sim}{G}| \\
&= |\underset{\sim}{L}'\underset{\sim}{M}|^{n+1} \, |\underset{\sim}{\Delta}|^{n+1}, \quad \text{by (30)} \\
&= |\underset{\sim}{A}|^{n+1}, \qquad (33)
\end{aligned}$$

because $\underset{\sim}{L}\underset{\sim}{A}\underset{\sim}{M}' = \underset{\sim}{\Delta}$ gives $\underset{\sim}{A} = \underset{\sim}{L}'\underset{\sim}{\Delta}\underset{\sim}{M}$ and hence $|\underset{\sim}{A}| = |\underset{\sim}{L}'\underset{\sim}{M}| \; |\underset{\sim}{\Delta}|$. Thus (30) applies whether $\underset{\sim}{A}$ is defective or not. Q.E.D.

4. Applications

4.1. Linear matrix equations

The equation

$$\sum_{i=1}^{k} \underset{\sim}{A}_i \underset{\sim}{X} \underset{\sim}{B}_i + \sum_{j=1}^{\ell} \underset{\sim}{D}_j \underset{\sim}{X}' \underset{\sim}{E}_j = \underset{\sim}{C} \qquad (34)$$

in $\underset{\sim}{X}$, of order $m \times n$, is a broad generalization of linear matrix equations discussed in the literature, one particular form being the Procrustes' problem (e.g., Schöneman [1966]). Other special cases are those of $\ell = 0$ considered in Lancaster [1970], Rao and Mitra [1971], and Wimmer and Ziebur [1972], and, the particular case $\underset{\sim}{A}\underset{\sim}{X} + \underset{\sim}{X}\underset{\sim}{B} = \underset{\sim}{C}$ in Hartwig [1975]; and $\underset{\sim}{A}'\underset{\sim}{X} \pm \underset{\sim}{X}'\underset{\sim}{A} = \underset{\sim}{C}$ in Hodges [1957]. All of these and more are embodied in (34), which can be solved directly, just as they stand, by using the vec operator, as in Vetter [1975].

Taking the vec of both sides of (34) and using (8) and (13) gives

$$\left\{\sum_{i=1}^{k}(\underset{\sim}{B}_i' \otimes \underset{\sim}{A}_i) + \left[\sum_{j=1}^{\ell}(\underset{\sim}{E}_j' \otimes \underset{\sim}{D}_j)\right]\underset{\sim}{I}_{(n,m)}\right\} \mathrm{vec}\underset{\sim}{X} = \mathrm{vec}\underset{\sim}{C} \; . \tag{35}$$

Solving this is easy: if the matrix multiplying $\mathrm{vec}\underset{\sim}{X}$ is non-singular, there is but a single solution for $\mathrm{vec}\underset{\sim}{X}$, otherwise there are many solutions, based on any generalized inverse of that matrix (see, for example, Searle [1971, p. 9, Theorem 2]). For any solution $\mathrm{vec}\underset{\sim}{X}$, with given m and n, $\underset{\sim}{X}_{m\times n}$ is then uniquely determined. Simplifications for certain special cases are available in Kadane et al. [1977].

4.2. Jacobians

If $\underset{\sim}{y}(\underset{\sim}{x})$ is a vector of n differentiable functions of the elements of $\underset{\sim}{x}$, of order n, of such a nature that the transformation $\underset{\sim}{x} \to \underset{\sim}{y}$ is 1-to-1, then the Jacobian of the transformation is defined as the absolute value of the determinant

$$J_{\underset{\sim}{x}\to\underset{\sim}{y}} = \left|\left\{\frac{\partial x_i}{\partial y_j}\right\}\right| = 1\Big/\left|\left\{\frac{\partial y_i}{\partial x_j}\right\}\right| \tag{36}$$

for $i, j = 1, \cdots, n$. Vector representation of (36) is

$$J_{\underset{\sim}{x}\to\underset{\sim}{y}} = \left|\frac{\partial \underset{\sim}{x}}{\partial \underset{\sim}{y}}\right| = 1\Big/\left|\frac{\partial \underset{\sim}{y}}{\partial \underset{\sim}{x}}\right| \; . \tag{37}$$

In the distribution theory of multivariate statistics we often need the Jacobian of a 1-to-1 transformation from variables represented by $\underset{\sim}{X}$ to those represented by $\underset{\sim}{Y}$. We first consider the case of $\underset{\sim}{X}$ and $\underset{\sim}{Y}$ non-symmetric or, to put it more carefully, the case of $\underset{\sim}{X}$ and $\underset{\sim}{Y}$ each consisting of distinctly different elements.

(a) Non-symmetric transformations

The Jacobian matrix of the transformation $\underset{\sim}{X} \to \underset{\sim}{Y}$, when $\underset{\sim}{X}$ and $\underset{\sim}{Y}$ each consist of distinctly different elements, is the matrix of every element of $\underset{\sim}{X}$ differentiated with respect to every element of $\underset{\sim}{Y}$. By the definition of the vec operator and analogous to (37) the Jacobian is therefore the absolute value of

$$J_{\underset{\sim}{X} \to \underset{\sim}{Y}} = \left|\frac{\partial \operatorname{vec}\underset{\approx}{X}}{\partial \operatorname{vec}\underset{\sim}{Y}}\right| = 1 \Big/ \left|\frac{\partial \operatorname{vec}\underset{\approx}{Y}}{\partial \operatorname{vec}\underset{\sim}{X}}\right| = 1/J_{\underset{\sim}{Y} \to \underset{\sim}{X}} \tag{38}$$

as given by Neudecker [1969]. We give three examples.

Examples

(i) The transformation $\underset{\sim}{Y}_{p\times q} = \underset{\sim}{A}\underset{\sim}{X}_{p\times q}\underset{\sim}{B}$ has Jacobian

$$J_{\underset{\sim}{X} \to \underset{\sim}{Y}} = |\underset{\sim}{A}|^{-q}\, |\underset{\sim}{B}|^{-p} \,. \tag{39}$$

This is so because

$$1/J_{\underset{\sim}{X} \to \underset{\sim}{Y}} = \left|\frac{\partial \operatorname{vec}(\underset{\sim\sim\sim}{AXB})}{\partial \operatorname{vec}\underset{\sim}{X}}\right| = \left|\frac{\partial(\underset{\sim}{B}' \otimes \underset{\sim}{A})\operatorname{vec}\underset{\sim}{X}}{\partial \operatorname{vec}\underset{\sim}{X}}\right|, \quad \text{from (8)}$$

$$= |\underset{\sim}{B}' \otimes \underset{\sim}{A}| = |\underset{\sim}{A}|^{q}\, |\underset{\sim}{B}|^{p}, \quad \text{from (4)}.$$

Result (39) is well-known and is to be found, for example, in Deemer and Olkin [1951, Theorem 3.6] and Neudecker [1969, equation 7.1.1].

The preceding example is a linear transformation from $\underset{\sim}{X}$ to $\underset{\sim}{Y}$. For non-linear transformations, (38) can be used only through invoking the property that even then the total differentials are linear functions of each other; i.e., for vector variables $\underset{\sim}{x}$ and $\underset{\sim}{y}$ (with $\underset{\sim}{x}$ having distinct elements) there exists a matrix $\underset{\sim}{M}$ such that

$$d\underset{\sim}{y} = \underset{\sim}{M}d\underset{\sim}{x}; \quad \text{and also,} \quad \frac{\partial \underset{\approx}{y}}{\partial \underset{\sim}{x}} = \underset{\sim}{M}' \,, \tag{40}$$

by a theorem from Neudecker [1969].

This result is, of course, also true for linear transformations, although its use is redundant in such cases. Nevertheless, it is useful to observe for the linear case $\underset{\sim}{Y} = \underset{\sim}{A}\underset{\sim}{X}\underset{\sim}{B}$ just considered, that

$$d\underset{\sim}{Y} = \underset{\sim}{A}(d\underset{\sim}{X})\underset{\sim}{B} \tag{41}$$

and so

$$d(\mathrm{vec}\underset{\sim}{Y}) = \mathrm{vec}(d\underset{\sim}{Y}) = (\underset{\sim}{B}' \otimes \underset{\sim}{A})\mathrm{vec}(d\underset{\sim}{X}) = (\underset{\sim}{B}' \otimes \underset{\sim}{A})d(\mathrm{vec}\underset{\sim}{X})$$

leads by way of (40) to

$$\frac{\partial \mathrm{vec}\underset{\sim}{Y}}{\partial \mathrm{vec}\underset{\sim}{X}} = (\underset{\sim}{B}' \otimes \underset{\sim}{A})'$$

and hence to (39).

(ii) For $\underset{\sim}{Y}_{n\times n} = \underset{\sim}{X}^{-1}$, the Jacobian is $J_{\underset{\sim}{X} \to \underset{\sim}{Y}} = |\underset{\sim}{X}|^{2n}$. This transformation is non-linear and so we take differentials of $\underset{\sim}{X}\underset{\sim}{Y} = \underset{\sim}{I}$, giving $\underset{\sim}{0} = d\underset{\sim}{I} = d(\underset{\sim}{X}\underset{\sim}{Y}) = \underset{\sim}{X}(d\underset{\sim}{Y}) + (d\underset{\sim}{X})\underset{\sim}{Y}$, and so

$$d\underset{\sim}{Y} = -\underset{\sim}{X}^{-1}(d\underset{\sim}{X})\underset{\sim}{X}^{-1} . \tag{42}$$

Since this is akin to (41) we can immediately use (40), and through (39) get, on ignoring sign,

$$J_{\underset{\sim}{X} \to \underset{\sim}{Y}} = |-\underset{\sim}{X}^{-1}|^{-n} \, |\underset{\sim}{X}^{-1}|^{-n} = |\underset{\sim}{X}|^{2n} . \tag{43}$$

This result is, of course, available in many places; e.g., Anderson [1958, p. 349], Dwyer [1967, equation (15.9)], and Kshirsagar [1972, p. 525, but with failure to acknowledge non-symmetry].

(iii) For $\underset{\sim}{Y} = \underset{\sim}{X}^p$ we have $d\underset{\sim}{Y} = \sum_{i=1}^{p} \underset{\sim}{X}^{i-1}(d\underset{\sim}{X})\underset{\sim}{X}^{p-i}$, so that taking vec of both sides and using (8) gives

$$\mathrm{vec}(d\underset{\sim}{Y}) = \sum_{i=1}^{p} (\underset{\sim}{X}'^{\,p-i} \otimes \underset{\sim}{X}^{i-1})\mathrm{vec}(d\underset{\sim}{X}) \tag{44}$$

from which (38) yields

$$1/J_{\underset{\sim}{X}\to\underset{\sim}{Y}} = \left|\sum_{i=1}^{p} (\underset{\sim}{X}'^{p-i} \otimes \underset{\sim}{X}^{i-1})\right| .$$

For non-defective $\underset{\sim}{X}$, as defined following equation (32), with $\underset{\sim}{D}$ being the diagonal matrix of the eigenvalues, this becomes

$$\begin{aligned} 1/J_{\underset{\sim}{X}\to\underset{\sim}{Y}} &= \left|\sum_{i=1}^{p} (\underset{\sim}{D}^{p-i} \otimes \underset{\sim}{D}^{i-1})\right| \\ &= \left| \text{ dg}\left\{ \sum_{i=1}^{p} \lambda_1^{p-i}\lambda_1^{i-1} \;\; \sum_{i=1}^{p} \lambda_1^{p-i}\lambda_2^{i-1} \cdots \sum_{i=1}^{p} \lambda_{n-1}^{p-i}\lambda_n^{i-1} \;\; \sum_{i=1}^{p} \lambda_n^{p-i}\lambda_n^{i-1} \right\}\right| \end{aligned} \tag{45}$$

where n is the order of $\underset{\sim}{X}$ and dg{ } is a diagonal matrix having the terms in the braces along its diagonal. Therefore

$$1/J_{\underset{\sim}{X}\to\underset{\sim}{Y}} = \prod_{i=1}^{n} \prod_{j=1}^{n} \sum_{r=1}^{p} \lambda_i^{p-r}\lambda_j^{r-1} . \tag{46}$$

In the literature we have found only a special case of this, namely for p = 2,

$$1/J_{\underset{\sim}{X}\to\underset{\sim}{Y}} = \prod_{i=1}^{n} \prod_{j=1}^{n} (\lambda_i + \lambda_j) = \left(\prod_{i=1}^{n} \lambda_i\right)^n \prod_{i=1}^{n} \prod_{j=1}^{n} (1 + \lambda_j/\lambda_i) , \tag{47}$$

as in Neudecker [1969].

(b) Symmetric transformations

Comparable to (38), when $\underset{\sim}{X} = \underset{\sim}{X}'$ and $\underset{\sim}{Y} = \underset{\sim}{Y}'$, we have

$$J_{\underset{\sim}{X}\to\underset{\sim}{Y}} = \left|\frac{\partial \text{vech}\underset{\sim}{X}}{\partial \text{vech}\underset{\sim}{Y}}\right| = 1\Big/\left|\frac{\partial \text{vech}\underset{\sim}{Y}}{\partial \text{vech}\underset{\sim}{X}}\right| = 1/J_{\underset{\sim}{Y}\to\underset{\sim}{X}} . \tag{48}$$

Examples

(i) For $\underset{\sim}{Y} = \underset{\sim}{AXA}'$, the Jacobian is

$$J_{\underset{\sim}{X}\to\underset{\sim}{Y}} = |\underset{\sim}{A}|^{-(n+1)} . \tag{49}$$

This is so because

$$1/J_{\underset{\sim}{X} \to \underset{\sim}{Y}} = \left|\frac{\partial \text{vech}(\underset{\sim}{A}\underset{\sim}{X}\underset{\sim}{A}')}{\partial \text{vech}\underset{\sim}{X}}\right| = \left|\frac{\partial \underset{\sim}{H}(\underset{\sim}{A} \otimes \underset{\sim}{A})\text{vec}\underset{\sim}{X}}{\partial \text{vech}\underset{\sim}{X}}\right|, \quad \text{from (8) and (14)}$$

$$= \left|\frac{\partial \underset{\sim}{H}(\underset{\sim}{A} \otimes \underset{\sim}{A})\underset{\sim}{G}\text{vech}\underset{\sim}{X}}{\partial \text{vech}\underset{\sim}{X}}\right|, \quad \text{from (14) again}$$

$$= |\underset{\sim}{H}(\underset{\sim}{A} \otimes \underset{\sim}{A})\underset{\sim}{G}| = |\underset{\sim}{A}|^{n+1}, \quad \text{from (28).} \tag{50}$$

(ii) The transformation $\underset{\sim}{Y} = \underset{\sim}{X}^{-1}$ for $\underset{\sim}{X}$ symmetric, has $J_{\underset{\sim}{X} \to \underset{\sim}{Y}} = |\underset{\sim}{X}|^{n+1}$. In this case (42) still applies, but because $\underset{\sim}{X}$ is symmetric we now use (49), rather than (39), and so, after again ignoring sign, $J_{\underset{\sim}{X} \to \underset{\sim}{Y}} = |\underset{\sim}{X}|^{n+1}$, as in Dwyer [1967]. Note the difference between this and $|\underset{\sim}{X}|^{2n}$ for the non-symmetric case in (43).

(iii) For $\underset{\sim}{Y} = \underset{\sim}{X}^p$ for $\underset{\sim}{X} = \underset{\sim}{X}'$ and $\underset{\sim}{Y} = \underset{\sim}{Y}'$ we use (14) to change (44) from vec to vech vectors, which gives

$$1/J_{\underset{\sim}{X} \to \underset{\sim}{Y}} = \left|\underset{\sim}{H} \sum_{i=1}^{p} (\underset{\sim}{D}^{p-i} \otimes \underset{\sim}{D}^{i-1})\underset{\sim}{G}\right|, \tag{51}$$

and this leads to

$$1/J_{\underset{\sim}{X} \to \underset{\sim}{Y}} = \prod_{i=1}^{n} \prod_{j=i}^{n} \sum_{r=1}^{p} \lambda_i^{p-r}\lambda_j^{r-1}, \tag{52}$$

which differs from (46) for the non-symmetric case in that j starts at 1 in (46) but at i in (52).

(c) Other patterned matrices

For patterned matrices of any sort, in which $\underset{\sim}{X}_1$ and $\underset{\sim}{X}_2$ may be of different patterns but each having the same number of distinctly different elements, we can consider the transformation $\underset{\sim}{X}_1 \to \underset{\sim}{X}_2$. Let $\underset{\sim}{M}_{p_1}$ denote a class of matrices of pattern

p_i with $\underset{\sim}{X}_i \in \underset{\sim}{M}_{p_i}$ for i = 1, 2. Define $\text{vecp}_i(\underset{\sim}{X}_i)$ as the vector of all distinct elements of $\underset{\sim}{X}_i$. Then, corresponding to (14), we have

$$\text{vecp}_i(\underset{\sim}{X}_i) = \underset{\sim}{P}_i\text{vec}(\underset{\sim}{X}_i) \quad \text{and} \quad \text{vec}(\underset{\sim}{X}_i) = \underset{\sim}{Q}_i\text{vecp}_i(\underset{\sim}{X}_i) = \underset{\sim}{Q}_i\underset{\sim}{P}_i\text{vec}(\underset{\sim}{X}_i)$$

where $\underset{\sim}{P}_i$ and $\underset{\sim}{Q}_i$ correspond exactly to $\underset{\sim}{H}$ and $\underset{\sim}{G}$ of (14), and have similar properties: $\underset{\sim}{P}_i\underset{\sim}{Q}_i = \underset{\sim}{I}$, $\underset{\sim}{Q}_i$ has full column rank and is unique, in the sense that $\underset{\sim}{G}$ is, and $\underset{\sim}{P}_i = (\underset{\sim}{Q}_i'\underset{\sim}{Q}_i)^{-1}\underset{\sim}{Q}_i'$ is one possible value of $\underset{\sim}{P}_i$. Then for the transformation $\underset{\sim}{X}_1 \to \underset{\sim}{X}_2$, linear or non-linear, the differentials of $\underset{\sim}{X}_2$ are linear in those of $\underset{\sim}{X}_1$ (Deemer and Olkin [1951]); we define this linearity in terms of the matrix $\underset{\sim}{C}_{12}$ by

$$\text{vec}(d\underset{\sim}{X}_1) \equiv \underset{\sim}{C}_{12}\text{vec}(d\underset{\sim}{X}_2) \; . \tag{53}$$

Now

$$\text{vecp}_1(d\underset{\sim}{X}_1) = \underset{\sim}{P}_1\text{vec}(d\underset{\sim}{X}_1) = \underset{\sim}{P}_1\underset{\sim}{C}_{12}\text{vec}(d\underset{\sim}{X}_2) = \underset{\sim}{P}_1\underset{\sim}{C}_{12}\underset{\sim}{Q}_2\text{vecp}_2(d\underset{\sim}{X}_2)$$

and so, on applying (40),

$$J_{\underset{\sim}{X}_1 \to \underset{\sim}{X}_2} = |\underset{\sim}{P}_1\underset{\sim}{C}_{12}\underset{\sim}{Q}_2| \; . \tag{54}$$

The absolute value of this is the Jacobian of a transformation from $\underset{\sim}{X}_1$ of pattern p_1 to $\underset{\sim}{X}_2$ of pattern p_2.

Example

$\underset{\sim}{X}_1 = \underset{\sim}{X}_2\underset{\sim}{X}_2'$ for $\underset{\sim}{X}_2$ lower triangular.

Define $\text{vecp}_2\underset{\sim}{X}_2$ as $\text{vect}\underset{\sim}{X}_2$, just like $\text{vech}\underset{\sim}{X}_2$ but with $\text{vec}\underset{\sim}{X}_2 = \underset{\sim}{Q}_2\text{vect}\underset{\sim}{X}_2$ where $\underset{\sim}{Q}_2$ has the form of $\underset{\sim}{G}$ in (16) but with all off-diagonal submatrices being null, and then $\underset{\sim}{Q}_2'$ is a possible $\underset{\sim}{P}_2$. And since $\underset{\sim}{X}_1$ is symmetric, with $\text{vecp}_1\underset{\sim}{X}_1 \equiv \text{vech}\underset{\sim}{X}_1$, we have $\underset{\sim}{P}_1 = \underset{\sim}{H}$ as in (14). To derive $\underset{\sim}{C}_{12}$, note that

$$d\underset{\sim}{X}_1 = (d\underset{\sim}{X}_2)\underset{\sim}{X}_2' + \underset{\sim}{X}_2(d\underset{\sim}{X}_2')$$

so that from (8)

$$\mathrm{vec}(d\underset{\sim}{X}_1) = (\underset{\sim}{X}_2 \otimes \underset{\sim}{I})\mathrm{vec}(d\underset{\sim}{X}_2) + (\underset{\sim}{I} \otimes \underset{\sim}{X}_2)\mathrm{vec}(d\underset{\sim}{X}_2') ,$$

which, on using (7) and (13) becomes

$$\mathrm{vec}(d\underset{\sim}{X}_1) = [\underset{\sim}{I}_{n^2} + \underset{\sim}{I}_{(n,n)}](\underset{\sim}{X}_2 \otimes \underset{\sim}{I})\mathrm{vec}(d\underset{\sim}{X}_2) .$$

Therefore, from (53)

$$\underset{\sim}{C}_{12} = [\underset{\sim}{I}_{n^2} + \underset{\sim}{I}_{(n,n)}](\underset{\sim}{X}_2 \otimes \underset{\sim}{I})$$

and so in (54)

$$J_{\underset{\sim}{X}_1 \to \underset{\sim}{X}_2} = |\underset{\sim}{H}(\underset{\sim}{I}_{n^2} + \underset{\sim}{I}_{(n,n)})(\underset{\sim}{X}_2 \otimes \underset{\sim}{I})\underset{\sim}{Q}_2| .$$

It can then be shown that this reduces to

$$J_{\underset{\sim}{X}_1 \to \underset{\sim}{X}_2} = 2^n \prod_{i=1}^{n} [x_{2(ii)}]^{n-i+1} \tag{55}$$

as in Deemer and Olkin [1951, Theorem 4].

4.3. Uses in multivariate statistics

There are many uses for vec and vech operators in the distribution theory of multivariate statistics, in addition to their use in differentiation and deriving Jacobians. We give but four such examples.

(a) Distributions of transformed variables

Suppose $\underset{\sim}{X} = [\underset{\sim}{x}_1 \cdots \underset{\sim}{x}_c]$ is a matrix of random variables having mean value $\underset{\sim}{M} = [\underset{\sim}{m}_1 \cdots \underset{\sim}{m}_c]$, with columns $\underset{\sim}{x}_j$ each having variance-covariance matrix $\underset{\sim}{V}$, and all pairs of columns uncorrelated. We summarize this as

$$\underset{\sim}{X} \text{ has uncorrelated columns } \underset{\sim}{x}_j \sim (\underset{\sim}{m}_j, \underset{\sim}{V}), \text{ for } j = 1, \cdots, c. \tag{56}$$

A question arises as to the distribution of transformed variables $\underset{\sim}{B}\underset{\sim}{X}$ and $\underset{\sim}{X}\underset{\sim}{C}$ for some matrices $\underset{\sim}{B}$ and $\underset{\sim}{C}$. The case of $\underset{\sim}{B}\underset{\sim}{X}$ is easy:

$$\underset{\sim}{B}\underset{\sim}{X} \text{ has uncorrelated columns } \underset{\sim}{B}\underset{\sim}{x}_j \sim (\underset{\sim}{B}\underset{\sim}{m}_j,\ \underset{\sim}{B}\underset{\sim}{V}\underset{\sim}{B}'),\ \text{for } j = 1, \cdots, c. \tag{57}$$

For $\underset{\sim}{X}\underset{\sim}{C}$, first observe that from the uncorrelated and variance properties of the columns of $\underset{\sim}{X}$ in (56)

$$\operatorname{var}(\operatorname{vec}\underset{\sim}{X}) = \begin{bmatrix} \underset{\sim}{V} & \underset{\sim}{0} \cdots & \underset{\sim}{0} \\ \underset{\sim}{0} & \underset{\sim}{V} \cdots & \underset{\sim}{0} \\ \vdots & & \vdots \\ \underset{\sim}{0} & \underset{\sim}{0} \cdots & \underset{\sim}{V} \end{bmatrix} = (\underset{\sim}{I} \otimes \underset{\sim}{V})\ , \tag{58}$$

and so (57) is $\operatorname{vec}(\underset{\sim}{B}\underset{\sim}{X}) \sim [\operatorname{vec}(\underset{\sim}{B}\underset{\sim}{M}),\ \underset{\sim}{I} \otimes \underset{\sim}{B}\underset{\sim}{V}\underset{\sim}{B}']$. Then with $E(\underset{\sim}{X}\underset{\sim}{C}) = \underset{\sim}{M}\underset{\sim}{C}$ we also have

$$\operatorname{var}[\operatorname{vec}(\underset{\sim}{X}\underset{\sim}{C})] = \operatorname{var}[(\underset{\sim}{C}' \otimes \underset{\sim}{I})\operatorname{vec}\underset{\sim}{X}] = (\underset{\sim}{C}' \otimes \underset{\sim}{I})(\underset{\sim}{I} \otimes \underset{\sim}{V})(\underset{\sim}{C} \otimes \underset{\sim}{I}) = \underset{\sim}{C}'\underset{\sim}{C} \otimes \underset{\sim}{V}$$

and so

$$\operatorname{vec}(\underset{\sim}{X}\underset{\sim}{C}) \sim [\operatorname{vec}(\underset{\sim}{M}\underset{\sim}{C}),\ \underset{\sim}{C}'\underset{\sim}{C} \otimes \underset{\sim}{V}]\ . \tag{59}$$

The dispersion matrix here, $\underset{\sim}{C}'\underset{\sim}{C} \otimes \underset{\sim}{V}$, summarizes the result given in Anderson [1958, penultimate equation following (2) on p. 52]. We also have the derivation, for $\underset{\sim}{X}_{p\times q}$

$$\begin{aligned} \operatorname{var}[\operatorname{vec}(\underset{\sim}{X}'\underset{\sim}{A})] &= \operatorname{var}[(\underset{\sim}{A}' \otimes \underset{\sim}{I})\operatorname{vec}\underset{\sim}{X}'] = \operatorname{var}[(\underset{\sim}{A}' \otimes \underset{\sim}{I})\underset{\sim}{I}_{(q,p)}\operatorname{vec}(\underset{\sim}{X})],\ \text{from (13)} \\ &= (\underset{\sim}{A}' \otimes \underset{\sim}{I})\underset{\sim}{I}_{(q,p)}(\underset{\sim}{I} \otimes \underset{\sim}{V})\underset{\sim}{I}_{(p,q)}(\underset{\sim}{A} \otimes \underset{\sim}{I}) \\ &= (\underset{\sim}{A}'\underset{\sim}{V}\underset{\sim}{A} \otimes \underset{\sim}{I})\ , \end{aligned} \tag{60}$$

after using (7), (5) and (4). Many useful special cases can be derived from these results, when features such as $\underset{\sim}{M} = \underset{\sim}{\mu}\underset{\sim}{1}'$, $\underset{\sim}{M} = \mu\underset{\sim}{1}\underset{\sim}{1}'$, $\underset{\sim}{B}\underset{\sim}{V}\underset{\sim}{B}' = \underset{\sim}{I}$, $\underset{\sim}{C}'\underset{\sim}{C} = \underset{\sim}{I}$ and $\underset{\sim}{A}'\underset{\sim}{V}\underset{\sim}{A} = \underset{\sim}{I}$ arise, either with or without normality.

(b) The multivariate linear model

Under normality assumptions, the traditional univariate linear model $\underset{\sim}{y} = \underset{\sim}{X}\underset{\sim}{b} + \underset{\sim}{e}$ can be denoted as

$$\underset{\sim}{y} \sim N(\underset{\sim}{X}\underset{\sim}{b},\ \sigma^2\underset{\sim}{I}) \tag{61}$$

where $\underset{\sim}{X}\underset{\sim}{b}$ is the expected value of $\underset{\sim}{y}$ and $\sigma^2\underset{\sim}{I}$ is its variance. It is shown in Searle [1978] how the customary multivariate linear model $\underset{\sim}{Y}_{n\times p} = \underset{\sim}{X}_{n\times q}\underset{\sim}{B}_{q\times p} + \underset{\sim}{E}_{n\times p}$ can easily be formulated in a similar manner as the univariate model

$$\text{vec}\underset{\sim}{Y} \sim N[(\underset{\sim}{I}_p \otimes \underset{\sim}{X})\text{vec}\underset{\sim}{B},\ \underset{\sim}{\Sigma} \otimes \underset{\sim}{I}_n] \tag{62}$$

where the rows of $\underset{\sim}{E}$ are n, independently and identically, normally distributed random vectors having zero mean and variance-covariance matrix $\underset{\sim}{\Sigma}$.

The formulation (62) provides ready opportunity for considering the multivariate linear model in a univariate framework. For example, testing the hypothesis $H: \underset{\sim}{K}'\underset{\sim}{b} = \underset{\sim}{m}$ in the model (61) can be done by way of the statistic $F = Q/s\hat{\sigma}^2$ where $\underset{\sim}{K}'$ has full row rank s, where $\hat{\sigma}^2 = \underset{\sim}{y}'\underset{\sim}{P}\underset{\sim}{y}/f$ for $\underset{\sim}{P} = \underset{\sim}{I} - \underset{\sim}{X}(\underset{\sim}{X}'\underset{\sim}{X})^-\underset{\sim}{X}'$ and $f = n - r(\underset{\sim}{X})$, and where

$$Q = (\underset{\sim}{K}'\underset{\sim}{b}^o - \underset{\sim}{m})'[\underset{\sim}{K}'(\underset{\sim}{X}'\underset{\sim}{X})^-\underset{\sim}{K}]^{-1}(\underset{\sim}{K}'\underset{\sim}{b}^o - \underset{\sim}{m}) \tag{63}$$

for $\underset{\sim}{b}^o = (\underset{\sim}{X}'\underset{\sim}{X})^-\underset{\sim}{X}'\underset{\sim}{y}$. Through the use of (62), Searle [1978] shows how adaptation of (63) leads to Hotelling's T_0^2-statistic for testing, in the multivariate model $\underset{\sim}{Y} = \underset{\sim}{X}\underset{\sim}{B} + \underset{\sim}{E}$, the hypothesis that some rows of $\underset{\sim}{B}$ have pre-assigned values.

(c) Fourth moments in a general linear model

A more general representation of the linear model is $\underset{\sim}{y} = \underset{\sim}{X}\underset{\sim}{\beta} + \underset{\sim}{Z}\underset{\sim}{u}$ where $\underset{\sim}{\beta}$ represents fixed effects and $\underset{\sim}{u}$ represents random effects having zero mean and variance-covariance matrix $\underset{\sim}{D}$. Then $\text{var}(\underset{\sim}{y}) = \underset{\sim}{V} = \underset{\sim}{Z}\underset{\sim}{D}\underset{\sim}{Z}'$. A common form of $\underset{\sim}{D}$ is the following function of the elements of $\underset{\sim}{\sigma}^2 = \{\sigma_i^2\}$, $i = 1, \cdots, k$,

$$\underset{\sim}{D} = \underset{\sim}{D}\{\sigma^2\} = \text{dg}\,\{\sigma_1^2 \underset{\sim}{I}_{c_1} \cdots \sigma_k^2 \underset{\sim}{I}_{c_k}\}\ . \qquad (64)$$

Under these conditions Anderson et al. [1977] show, using the tools discussed here, that a matrix involving fourth moments of $\underset{\sim}{y}$, namely

$$\underset{\sim}{F} = \text{var}[(\underset{\sim}{y} - \underset{\sim}{X}\underset{\sim}{\beta}) \otimes (\underset{\sim}{y} - \underset{\sim}{X}\underset{\sim}{\beta})] = \text{var}[\text{vec}(\underset{\sim}{y} - \underset{\sim}{X}\underset{\sim}{\beta})(\underset{\sim}{y} - \underset{\sim}{X}\underset{\sim}{\beta})'] \qquad (65)$$

is

$$\underset{\sim}{F} = (\underset{\sim}{V} \otimes \underset{\sim}{V})[\underset{\sim}{I}_{n^2} + \underset{\sim}{I}_{(n,n)}] + (\underset{\sim}{Z}\underset{\sim}{D}^{\frac{1}{2}} \otimes \underset{\sim}{Z}\underset{\sim}{D}^{\frac{1}{2}})\Big(\text{dg}[\text{vec}(\underset{\sim}{D}\{\underset{\sim}{\gamma}\})]\Big)(\underset{\sim}{D}^{\frac{1}{2}}\underset{\sim}{Z}' \otimes \underset{\sim}{D}^{\frac{1}{2}}\underset{\sim}{Z}'), \qquad (66)$$

where $\text{dg}[\underset{\sim}{x}]$ is a diagonal matrix of the elements of $\underset{\sim}{x}$ and $\underset{\sim}{D}\{\underset{\sim}{\gamma}\}$ has the same form as (64), using the kurtosis parameters γ_i in place of the variances σ_i^2. In the case of normality, (66) reduces to $\underset{\sim}{F} = (\underset{\sim}{V} \otimes \underset{\sim}{V})[\underset{\sim}{I}_{n^2} + \underset{\sim}{I}_{(n,n)}]$. Analogous results are given by Pukelsheim [1977, p. 327] and Rao [1971, p. 447].

(d) The Wishart Distribution

Suppose the columns of $\underset{\sim}{X}_{p\times n}$ are independently, identically normally distributed with mean $\underset{\sim}{0}$ and dispersion matrix $\underset{\sim}{V}$. We write, similar to (56),

$$\underset{\sim}{X} \text{ has uncorrelated columns } \underset{\sim}{x}_j \sim \text{i.i.d. } N(\underset{\sim}{0}, \underset{\sim}{V}), \text{ for } j = 1, \cdots, n. \qquad (67)$$

Then $\underset{\sim}{S} = \underset{\sim}{X}\underset{\sim}{X}'$ follows the p-dimensional Wishart distribution with scale matrix $\underset{\sim}{V}$ and n degrees of freedom. Because $\underset{\sim}{X}\underset{\sim}{X}' = \sum_{j=1}^{n} \underset{\sim}{x}_j\underset{\sim}{x}_j'$, the mean value of $\underset{\sim}{S}$ is readily obtainable from (67) as

$$E(\underset{\sim}{S}) = \sum_{j=1}^{n} E(\underset{\sim}{x}_j\underset{\sim}{x}_j') = n\underset{\sim}{V}\ , \qquad (68)$$

a well-known result. Second moments of elements of $\underset{\sim}{S}$ come from

$$\text{var}(\text{vec}\underset{\sim}{S}) = \text{var}\Big[\sum_{j=1}^{n} \text{vec}(\underset{\sim}{x}_j\underset{\sim}{x}_j')\Big] = \sum_{j=1}^{n} \text{var}[\text{vec}(\underset{\sim}{x}_j\underset{\sim}{x}_j')]\ ,$$

using the independence of the $\underset{\sim}{x}_j$'s given in (67). Then, by having $\underset{\sim}{x}_j$ play the part of $\underset{\sim}{y} - \underset{\sim}{X}\underset{\sim}{\beta}$ in (65), the consequent normality applied to (66) gives

$$\mathrm{var}(\mathrm{vec}\underset{\sim}{S}) = n(\underset{\sim}{V} \otimes \underset{\sim}{V})[\underset{\sim}{I}_{n^2} + \underset{\sim}{I}_{(n,n)}] \; . \tag{69}$$

These succinct matrix representations, (68) and (69), correspond to the scalar expressions of, for example, Anderson [1958, p. 161].

Acknowledgment

Grateful thanks go to Friedrich Pukelsheim for suggested improvements to an early draft of this paper.

References

Aitken, A. C. [1949]. On the Wishart distribution in statistics. Biometrika 36, 59-62.

Anderson, R. D., Quaas, R. L., and Searle, S. R. [1977]. Fourth moments in the general linear model; and the variance of translation invariant quadratic forms. Paper No. BU-630-M in the Biometrics Unit Mimeo Series, Cornell University. 11 pp.

Anderson, T. W. [1958]. An Introduction to Multivariate Statistical Analysis. John Wiley and Sons, New York.

Browne, M. W. [1974]. Generalized least squares estimators in the analysis of covariance structures. S. Afr. Statist. J. 8, 1-24.

Deemer, Walter L. and Olkin, Ingram [1951]. The Jacobians of certain matrix transformations useful in multivariate analysis. Based on lectures of P. L. Hsu at the University of North Carolina, 1947. Biometrika 38, 345-367.

Dwyer, P. S. [1967]. Some applications of matrix derivatives in multivariate analysis. J. Amer. Statist. Assoc. 62, 607-625.

Hartwig, R. E. [1975]. $\underset{\sim}{A}\underset{\sim}{X} - \underset{\sim}{X}\underset{\sim}{B} = \underset{\sim}{C}$, resultants and generalized inverses. SIAM J. Appl. Math. 28, 154-183.

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

Kadane, J. B., Henderson, H. V., and Searle, S. R. [1977]. A note on solving matrix equations using the vec operator. Paper No. BU-626-M in the Biometrics Unit Mimeo Series, Cornell University.

Kshirsagar, A. M. [1972]. Multivariate Analysis. Marcell Dekker, New York.

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

MacRae, E. C, [1974]. Matrix derivatives with an application to an adaptive linear decision problem. Ann. Statist. 2, 337-346.

Neudecker, H. [1968]. The Kronecker matrix product and some of its applications in econometrics. Stat. Neerlandica 22, 69-81.

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

Pukelsheim, Friedrich [1977]. On Hsu's model in regression analysis. Math. Operationsforsch. Statist., Ser. Statistics 8, 323-331.

Rao, C. R. [1971]. Minimum variance quadratic unbiased estimation of variance components. J. Multivariate Anal. 1, 445-456.

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

Roth, W. E. [1934]. On direct product matrices. Bull. Amer. Math. Soc. 40, 461-468.

Schöneman, Peter H. [1966]. A generalized solution of the orthogonal Procrustes problem. Psychometrika 31, 1-10.

Searle, S. R. [1971]. Linear Models. John Wiley and Sons, New York.

Searle, S. R. [1978]. A univariate formulation of the multivariate linear model. Contributions to Survey Sampling and Applied Statistics, Papers in Honor of H. O. Hartley (Ed. H. A. David), Academic Press, 181-189.

Sylvester, M. [1884]. Sur la solution du cas le plus général des équations linéaires en quantités binaires, c'est-à-dire en quaternions ou en matrices du second ordre. Comptes Rendus de l'Académie des Sciences (Paris) 99, 117-118.

Tracy, D. S. and Singh, R. P. [1972]. 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 (Ed. Derrick S. Tracy), University of Windsor, Ontario, Canada, 203-224.

Vetter, W. J. [1975]. Linear structures and solutions of linear matrix equations. Linear Algebra Appl. 10, 181-188.

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.

Rev 7/78