JACOBIANS FOR TRANSFORMATIONS OF PATTERNED MATRICES

by

Harold V. Henderson*

Biometrics Unit, Cornell University, Ithaca, New York

BU-687-M August 1979

ABSTRACT

The notion of vec and vech operators is extended to vecp operators for patterned matrices in which elements occur repetitively. When $\underset{\sim}{X}$ has pattern p, vecp$\underset{\sim}{X}$ is defined to be the vector of the distinct elements of $\underset{\sim}{X}$, formed by retaining only the first occurrence of each element in vec$\underset{\sim}{X}$. In the context of transformations of $\underset{\sim}{X}$, vecp$\underset{\sim}{X}$ will then be the vector of the functionally independent variables in $\underset{\sim}{X}$. This feature is useful in deriving Jacobians of transformations of patterned matrices, including cases where the patterns are not necessarily the same.

Paper No. BU-687-M in the Biometrics Unit.

* Supported by a New Zealand National Research Advisory Council Fellowship.

1. INTRODUCTION

The vec operator stacks the columns of a matrix one under the other to form a single vector. Thus, for

$\underset{\sim}{x}_i$, $i = 1, \ldots, n$, being the n columns of the $m \times n$ matrix $\underset{\sim}{X}$,

$$\underset{\sim}{X} = [\underset{\sim}{x}_1 \; \underset{\sim}{x}_2 \; \cdots \; \underset{\sim}{x}_n] \quad \text{and} \quad \text{vec}\underset{\sim}{X} = \begin{bmatrix} \underset{\sim}{x}_1 \\ \underset{\sim}{x}_2 \\ \vdots \\ \underset{\sim}{x}_n \end{bmatrix} . \tag{1}$$

The operation has been referred to variously as the column string or stack of $\underset{\sim}{X}$ and the pack of $\underset{\sim}{X}$, with vec$\underset{\sim}{X}$ for "vector of columns of $\underset{\sim}{X}$" being the notation currently in vogue. The vec operator is useful in deriving Jacobians involving transformations of $\underset{\sim}{X}$ having functionally independent elements or, as we say, "no pattern".

For square $\underset{\sim}{X}$, Searle [1978] and subsequently Henderson and Searle [1979], define vech$\underset{\sim}{X}$ in the same way as vec$\underset{\sim}{X}$, except that for each column of $\underset{\sim}{X}$ only that part of it which is on or below the diagonal is put into vech$\underset{\sim}{X}$ (for "vector-half" of $\underset{\sim}{X}$) . In this way, vech$\underset{\sim}{X}$ for $\underset{\sim}{X}$ symmetric, with no other equality relationships among its elements, contains only the functionally independent elements of $\underset{\sim}{X}$ or variables in $\underset{\sim}{X}$, a feature that is useful in deriving Jacobians for transformations from one symmetric array of variables to another. For example, with

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

Symmetric matrices are a subset of a broader class of matrices in which elements occur repetitively in some manner as to give the matrix an appearance of having a pattern. The concept of the vech operator is generalized to be applicable to patterned matrices of this nature. Any particular class of patterned matrices is characterized by a mnemonic letter which is incorporated in the vec notation, for example, we also use vect for triangular matrices and vecd for diagonal matrices. The formulation is for $\underset{\sim}{X}$ of pattern p (with no other equality relationships), in which case vecp$\underset{\sim}{X}$ is defined to be the vector of the distinct elements of $\underset{\sim}{X}$, formed by retaining only the first occurrence of each element in vec$\underset{\sim}{X}$. In this way vecp$\underset{\sim}{X}$ and vec$\underset{\sim}{X}$ are linear transformations of each other. When $\underset{\sim}{X}$ is not of pattern p, vecp$\underset{\sim}{X}$ may still be defined as a certain linear transformation of vec$\underset{\sim}{X}$, but in this case vec$\underset{\sim}{X}$ need not be a linear transformation of vecp$\underset{\sim}{X}$, e.g., vech$\underset{\sim}{X}$ for nonsymmetric square $\underset{\sim}{X}$. In the context of transformations of $\underset{\sim}{X}$ of pattern p, vecp$\underset{\sim}{X}$ will then be the vector of the functionally independent variables in $\underset{\sim}{X}$. The usefulness of vech in deriving Jacobians extends at once to vecp .

The general problem is to find the Jacobian of a transformation of one patterned matrix to another. The development is quite straightforward, and leads at once to detailed consideration of symmetric, (lower) triangular, and diagonal matrices; other patterned matrices that could be handled just as easily include circulant, centrosymmetric, Jacobi, tri-diagonal matrices and so on. The appeal of this method of deriving Jacobians is that it is applicable to *any* patterned matrices whose elements are repetitions of functionally independent variables.

An extension of these ideas, which we have not explored in connection with Jacobians (except for the degenerate case of skew-symmetry), is for a matrix $\underset{\sim}{X}$, having not only pattern arising from repetitions of elements but also structure, in that these elements are also linear combinations of distinct components. 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.,

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

The approach of Henderson and Searle [1979] to derive Jacobians for transformations involving no-pattern and symmetric matrices is here extended to transformations $\underset{\sim}{X}_1 \rightarrow \underset{\sim}{X}_2$ of other patterned matrices, including cases where $\underset{\sim}{X}_1$ and $\underset{\sim}{X}_2$ have patterns that are not necessarily the same. In preparation for this relationships of vecp$\underset{\sim}{X}$ to vec$\underset{\sim}{X}$ are first noted.

2. PROPERTIES OF VEC AND VECP OPERATORS

Let M_p denote a set of $m \times n$ matrices of pattern p, each of which have the same pattern of repetitions of $r \leq mn$ distinct elements. For example, the set M_h of $n \times n$ symmetric matrices, with no other equality relationships, has $r = \frac{1}{2}n(n+1)$. Then for $\underset{\sim}{X} \in M_p$ define

vecp$\underset{\sim}{X}$ as the vector of the r distinct elements of $\underset{\sim}{X}$, formed by retaining only the first occurrence of each element in vec$\underset{\sim}{X}$. (4)

Therefore, vec$\underset{\sim}{X}$ is a linear transformation of vecp$\underset{\sim}{X}$:

$$\text{vec}\mathbf{X} = \mathbf{Q}\text{vecp}\mathbf{X} \quad , \tag{5}$$

where $\mathbf{Q}$ is unique (for each M_p) of order $mn \times r$, with full column rank. Hence $\mathbf{Q}$ has a left inverse, $\mathbf{P}$ say, of order $r \times mn$ satisfying

$$\mathbf{P}\mathbf{Q} = \mathbf{I}_r \quad ; \tag{6}$$

$\mathbf{P}$ has full row rank but is not unique (unless $r = mn$, in which case $\mathbf{X}$ has no pattern so $\mathbf{P} = \mathbf{I}$ and $\text{vecp}\mathbf{X} = \text{vec}\mathbf{X}$) with

$$\mathbf{P} = (\mathbf{Q}'\mathbf{Q})^{-1}\mathbf{Q}' \tag{7}$$

the Moore-Penrose inverse of $\mathbf{Q}$ being one possible value. Then

$$\text{vecp}\mathbf{X} = \mathbf{P}\text{vec}\mathbf{X} \tag{8}$$

and

$$\text{vec}\mathbf{X} = \mathbf{Q}\mathbf{P}\text{vec}\mathbf{X} \quad . \tag{9}$$

When $\mathbf{X} \notin M_p$, $\text{vecp}\mathbf{X}$ may still be well defined as a linear transformation of $\text{vec}\mathbf{X}$. For example, for $\mathbf{X}$ square, $\text{vech}\mathbf{X}$ is the vector of elements on or below the diagonal of $\mathbf{X}$, and $\text{vecd}\mathbf{X}$ is the vector of diagonal elements of $\mathbf{X}$. When $\mathbf{X} \notin M_p$, $\text{vecp}\mathbf{X}$ is not necessarily the vector of distinct elements of $\mathbf{X}$ as defined in (4); i.e., for $\mathbf{X} \notin M_p$ (5) may not be true, but (8) always is for some $\mathbf{P}$.

In the context of transformations of $\mathbf{X}$ of pattern p (with no other equality relationships), $\text{vecp}\mathbf{X}$ will then be the vector of the functionally independent variables in $\mathbf{X}$. This feature of vecp operators is useful in deriving Jacobians for transformations of patterned matrices and is now explored.

3. JACOBIANS

For the transformation $\underset{\sim}{X}_1 \rightarrow \underset{\sim}{X}_2$, linear or non-linear, the differentials of $\underset{\sim}{X}_1$ are linear in those of $\underset{\sim}{X}_2$ (Deemer and Olkin [1951]); we define this linearity in terms of the matrix $\underset{\sim}{C}_{12}$ by

$$\text{vec}(d\underset{\sim}{X}_1) \equiv \underset{\sim}{C}_{12}\text{vec}(d\underset{\sim}{X}_2) \quad . \tag{10}$$

Now when $\underset{\sim}{X}_i \in M_{p_i}$ for $i = 1$ and 2, from (5) and (8),

$$\text{vecp}_1(d\underset{\sim}{X}_1) = \underset{\sim}{P}_1\text{vec}(d\underset{\sim}{X}_1) = \underset{\sim}{P}_1\underset{\sim}{C}_{12}\text{vec}(d\underset{\sim}{X}_2) = \underset{\sim}{P}_1\underset{\sim}{C}_{12}\underset{\sim}{Q}_2\text{vecp}_2(d\underset{\sim}{X}_2) \tag{11}$$

and so, identifying partial derivatives from relations between differentials (Neudecker [1969]), we have

$$\frac{\partial\,\text{vecp}_1(\underset{\sim}{X}_1)}{\partial\,\text{vecp}_2(\underset{\sim}{X}_2)} = (\underset{\sim}{P}_1\underset{\sim}{C}_{12}\underset{\sim}{Q}_2)' \quad . \tag{12}$$

When the transformation is one-to-one, $\underset{\sim}{X}_1$ and $\underset{\sim}{X}_2$ both have the same number of functionally independent variables so that (12) is square and the Jacobian, $J^{\#}_{\underset{\sim}{X}_1 \rightarrow \underset{\sim}{X}_2}$, is the absolute value of the determinant

$$J_{\underset{\sim}{X}_1 \rightarrow \underset{\sim}{X}_2} = |\underset{\sim}{P}_1\underset{\sim}{C}_{12}\underset{\sim}{Q}_2| \quad . \tag{13}$$

This determinant may be developed further either directly or by using singular-value decomposition of $\underset{\sim}{C}_{12}$ and then the canonical form. Henderson and Searle [1979] outline the latter, for symmetric matrices and those with no pattern; and we now develop this further for general patterned matrices.

The form of $\underset{\sim}{C}_{12}$ depends on the patterns of $\underset{\sim}{X}_1$ and $\underset{\sim}{X}_2$ and on the transformation involved. Define $C_{p_1,p_2}(\underset{\sim}{X}_1 \rightarrow \underset{\sim}{X}_2)$ to be the set of matrices $\underset{\sim}{C}_{12}$ satisfying (10):

$$C_{p_1,p_2}(\underset{\sim}{X}_1 \rightarrow \underset{\sim}{X}_2) = \{\underset{\sim}{C}_{12} : \text{vec}(d\underset{\sim}{X}_1) = \underset{\sim}{C}_{12}\text{vec}(d\underset{\sim}{X}_2) \quad \text{for} \quad \underset{\sim}{X}_1 \in M_{p_1} \rightarrow \underset{\sim}{X}_2 \in M_{p_2}\}. \tag{14}$$

A useful property of the matrices of this set is stated in the following lemma.

<u>Lemma 1</u>: When $\underset{\sim}{C}_{12} \in C_{p_1,p_2}(\underset{\sim}{X}_1 \rightarrow \underset{\sim}{X}_2)$

$$\underset{\sim}{C}_{12}\underset{\sim}{Q}_2 = \underset{\sim}{Q}_1\underset{\sim}{P}_1\underset{\sim}{C}_{12}\underset{\sim}{Q}_2 \quad . \tag{15}$$

<u>Proof</u>: Substituting (10) in (9) gives

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

and applying (5) yields

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

which is true for all $\underset{\sim}{X}_2 \in M_{p_2}$, in which case $\text{vecp}_2(d\underset{\sim}{X}_2)$ can take any value whatever since it contains the r_2 functionally independent elements of $d\underset{\sim}{X}_2$. Therefore $\text{vecp}_2(d\underset{\sim}{X}_2)$ can be given, in turn, values that are the columns of $\underset{\sim}{I}$ of order r_2, so establishing (15).

This result may be used to show that:

<u>Lemma 2</u>: $\underset{\sim}{P}_1\underset{\sim}{C}_{12}\underset{\sim}{Q}_2$ is invariant to whatever left inverse of $\underset{\sim}{Q}_1$ is used for $\underset{\sim}{P}_1$.

<u>Proof</u>: Let $\underset{\sim}{P} \neq \underset{\sim}{P}_1$ be some other left inverse. Then pre-multiply (15) by $\underset{\sim}{P}$ and use $\underset{\sim}{P}\underset{\sim}{Q}_1 = \underset{\sim}{I}$ to give

$$\underset{\sim}{P}\underset{\sim}{C}_{12}\underset{\sim}{Q}_2 = \underset{\sim}{P}_1\underset{\sim}{C}_{12}\underset{\sim}{Q}_2 \quad . \tag{16}$$

In view of this, we usually use $\underset{\sim}{P}_1 = (\underset{\sim}{Q}_1'\underset{\sim}{Q}_1)^{-1}\underset{\sim}{Q}_1'$ of (7), the Moore-Penrose inverse of $\underset{\sim}{Q}_1$, for which $\underset{\sim}{Q}_1\underset{\sim}{P}_1$ is symmetric.

The determinant (13) that occurs in certain Jacobians can be evaluated with the aid of these lemmas and decompositions of $\underset{\sim}{C}_{12}$, as will be shown in Theorem 1.

We first consider $\underset{\sim}{C}_{12}$ being rectangular, and then square but defective, utilizing the singular-value decomposition (SVD) of $\underset{\sim}{C}_{12}$. This is stated as lemma 3; a proof is available, for example, in Stewart [1973, p. 318].

Lemma 3: The singular-value decomposition of $\underset{\sim}{C}_{12}$ of order $m_1n_1 \times m_2n_2$, for orthogonal matrices $\underset{\sim}{U}$ and $\underset{\sim}{V}$, is

$$\begin{aligned} \underset{\sim}{V}'\underset{\sim}{C}_{12}\underset{\sim}{U} = \underset{\sim}{S} &= [\underset{\sim}{\Sigma} \quad \underset{\sim}{0}] \text{ when } m_1n_1 \le m_2n_2 \\ &= \begin{bmatrix} \underset{\sim}{\Sigma} \\ \underset{\sim}{0} \end{bmatrix} \text{ when } m_1n_1 > m_2n_2 \quad . \end{aligned} \tag{17}$$

Here $\underset{\sim}{\Sigma}^2 = \text{diag}(\sigma_1^2, \ldots, \sigma_{m_1n_1}^2)$ is the diagonal matrix of the eigen-values of $\underset{\sim}{C}_{12}\underset{\sim}{C}_{12}'$ and $\underset{\sim}{\Sigma} = \text{diag}(\sigma_1, \ldots, \sigma_{m_1n_1})$ is the diagonal matrix of the singular values of $\underset{\sim}{C}_{12}$. Then

$$\underset{\sim}{C}_{12} = \underset{\sim}{V}\underset{\sim}{S}\underset{\sim}{U}' \quad . \tag{18}$$

3.1. A general theorem

An expression for the determinant $|\underset{\sim}{P}_1\underset{\sim}{C}_{12}\underset{\sim}{Q}_2|$ of (13), incorporating the SVD of $\underset{\sim}{C}_{12}$, is now derived.

<u>Theorem 1</u>: Consider the transformation $X_1 \in M_{p_1} \to X_2 \in M_{p_2}$ for which $\mathrm{vec}(dX_1) = C_{12}\mathrm{vec}(dX_2)$. When:

(1) X_1 and X_2 have the same number of functionally independent elements, i.e., $r_1 = r_2 = r$,

(2) the SVD of C_{12} is $V'C_{12}U = S$ and

(3) for $i = 1$ and $j = 2$,

$$Q_2 = Q_j P_j U' Q_2 \quad \text{and} \quad Q_j = Q_i P_i S Q_j \ , \tag{19}$$

then

$$J_{X_1 \to X_2} = |P_1 C_{12} Q_2| = |P_1 V Q_1|\,|P_1 S Q_2|\,|P_2 U' Q_2| \ . \tag{20}$$

<u>Proof</u>: Substitute $C_{12} = VSU'$ from (18) in $P_1 C_{12} Q_2$ and use (19) to give $P_1 C_{12} Q_2 = P_1 V Q_1 P_1 S Q_2 P_2 U' Q_2$. Apply $|AB| = |A|\,|B|$ for square matrices to yield (20).

Further simplification is possible when C_{12} is square. In this case condition (19) can be generalized by allowing i and j to be 1 or 2; thus giving (possibly) further routes for evaluating (13). This follows immediately from Theorem 1 and is stated as a corollary.

<u>Corollary 1</u>: When conditions 1 and 2 of Theorem 1 and the following hold:

(4) X_1 and X_2 have the same number of elements, i.e., $m_1 n_1 = m_2 n_2 = mn$, so that C_{12} is square, and

(3*) condition 3 of Theorem 1 is modified for i and j being 1 and/or 2,

then, for i and j of 3*,

$$J_{X_1 \to X_2} = |P_1 C_{12} Q_2| = |P_1 V Q_i|\,|P_i S Q_j|\,|P_j U' Q_2| \ . \tag{21}$$

3.2. Reduction to simple matrices

Corollary 1 reduces $|\tilde{P}_1\tilde{C}_{12}\tilde{Q}_2|$ to evaluating $|\tilde{P}_i\tilde{M}\tilde{Q}_j|$ where $\tilde{M}$ is an orthogonal or diagonal matrix. But every orthogonal matrix is normal (Lancaster [1969, p. 81]) and hence simple (diagonable) and, of course, so is any diagonal matrix. A generalization is to evaluate $|\tilde{P}_1\tilde{M}\tilde{Q}_2|$ for $\tilde{M}$ being simple, which is given in the following theorem. When $\tilde{C}_{12}$ is itself simple we can proceed directly, so we state the result in terms of $\tilde{C}_{12}$ rather than $\tilde{M}$.

Theorem 2: When the conditions of corollary 1 are satisfied, with $i = j = 1$ or 2 in 3* and $\tilde{C}_{12}$ is simple with canonical form

$$\tilde{U}^{-1}\tilde{C}_{12}\tilde{U} = \tilde{D}_{12} = \text{diag}(\lambda_1, \cdots, \lambda_{mn}) \quad , \tag{22}$$

where $\lambda_1, \cdots, \lambda_{mn}$ are the eigenvalues of $\tilde{C}_{12}$,

$$J_{\tilde{X}_1 \to \tilde{X}_2} = |\tilde{P}_1\tilde{C}_{12}\tilde{Q}_2| = |\tilde{P}_1\tilde{Q}_2||\tilde{P}_i\tilde{D}_{12}\tilde{Q}_i| \quad . \tag{23}$$

Proof: Let $\tilde{C}_{12} = \tilde{U}\tilde{D}_{12}\tilde{U}^{-1}$ take the place of $\tilde{V}\tilde{S}\tilde{U}'$ in corollary 1. Then (21) becomes

$$|\tilde{P}_1\tilde{C}_{12}\tilde{Q}_2| = |\tilde{P}_1\tilde{U}\tilde{Q}_i\tilde{P}_i\tilde{U}^{-1}\tilde{Q}_2||\tilde{P}_i\tilde{D}_{12}\tilde{Q}_i|$$

and using (19) together with $\tilde{U}\tilde{U}^{-1} = \tilde{I}$ leads to (23).

The important case of $\tilde{X}_1$ and $\tilde{X}_2$ having the same pattern follows immediately from (23) on dropping subscripts and using $\tilde{P}\tilde{Q} = \tilde{I}$, as is now stated.

Corollary 2: Further, when $\tilde{X}_1$ and $\tilde{X}_2$ have the same pattern, M_p

$$J_{\tilde{X}_1 \to \tilde{X}_2} = |\tilde{P}\tilde{C}_{12}\tilde{Q}| = |\tilde{P}\tilde{D}_{12}\tilde{Q}| \quad . \tag{24}$$

4. APPLICATIONS

Henderson and Searle [1979] developed aspects of this approach for deriving Jacobians of transformations involving no-pattern and symmetric matrices, dealing with the transformations $\underset{\sim}{Y} = \underset{\sim}{A}\underset{\sim}{X}\underset{\sim}{B}$, $\underset{\sim}{Y} = \underset{\sim}{A}\underset{\sim}{X}\underset{\sim}{A}'$, $\underset{\sim}{Y} = \underset{\sim}{X}^{-1}$ and $\underset{\sim}{Y} = \underset{\sim}{X}^{p}$. Additional transformations and patterns are considered here and results developed for arbitrary patterned matrices. Applications are then made in turn, and by way of illustration, for no-pattern, symmetric, (lower) triangular, skew-symmetric and diagonal matrices. This method of deriving Jacobians is applicable to transformations of any patterned matrices whose elements are repetitions of functionally independent variables.

The relationship

$$\text{vec}(\underset{\sim}{A}\underset{\sim}{B}\underset{\sim}{C}) = (\underset{\sim}{C}' \otimes \underset{\sim}{A})\text{vec}\underset{\sim}{B} \quad , \tag{25}$$

where $\underset{\sim}{A} \otimes \underset{\sim}{B} = \{a_{ij}\underset{\sim}{B}\}$ is the Kronecker product of $\underset{\sim}{A}$ and $\underset{\sim}{B}$, permits easy evaluation of the Jacobian in many cases.

4.1. Transformations of matrices of the same pattern

The first set of four examples are for transformations $\underset{\sim}{X} \rightarrow \underset{\sim}{Y}$ when $\underset{\sim}{X}$ and $\underset{\sim}{Y}$ have the same pattern, M_p of order $m \times n$, with r functionally independent elements. The Jacobian $J^*_{\underset{\sim}{Y}\rightarrow\underset{\sim}{X}}$ is the absolute value of $J_{\underset{\sim}{Y}\rightarrow\underset{\sim}{X}}$ which is derived; the more usual $J^*_{\underset{\sim}{X}\rightarrow\underset{\sim}{Y}}$ is obtainable as $1/J^*_{\underset{\sim}{Y}\rightarrow\underset{\sim}{X}}$.

Examples

(i) The transformation $\underset{\sim}{Y} = \underset{\sim}{A}\underset{\sim}{X}\underset{\sim}{B}$ has Jacobian based on

$$J_{\underset{\sim}{Y}\rightarrow\underset{\sim}{X}} = |\underset{\sim}{P}(\underset{\sim}{B}' \otimes \underset{\sim}{A})\underset{\sim}{Q}| \quad . \tag{26}$$

This is so because applying (25) to $d\underset{\sim}{Y} = \underset{\sim}{A}(d\underset{\sim}{X})\underset{\sim}{B}$ gives

$$\text{vec}(d\underset{\sim}{Y}) = (\underset{\sim}{B}' \otimes \underset{\sim}{A})\text{vec}(d\underset{\sim}{X}) \tag{27}$$

so that $\underset{\sim}{C}_{12} = \underset{\sim}{B}' \otimes \underset{\sim}{A}$ in (10) and (13).

Although use of differentials is redundant for $\underset{\sim}{Y} = \underset{\sim}{A}\underset{\sim}{X}\underset{\sim}{B}$, a linear transformation, their use unifies the treatment as being also applicable to nonlinear transformations, like the next example.

(ii) For $\underset{\sim}{Y}_{n\times n} = \underset{\sim}{X}^{-1}_{n\times n}$ the Jacobian is based on

$$J_{\underset{\sim}{Y}\to\underset{\sim}{X}} = |-\underset{\sim}{P}(\underset{\sim}{X}^{-1'} \otimes \underset{\sim}{X}^{-1})\underset{\sim}{Q}| \quad . \tag{28}$$

This follows from (26) because taking differentials of $\underset{\sim}{X}\underset{\sim}{Y} = \underset{\sim}{I}$ leads to $d\underset{\sim}{Y} = -\underset{\sim}{X}^{-1}(d\underset{\sim}{X})\underset{\sim}{X}^{-1}$, which has the same form as $d\underset{\sim}{Y} = \underset{\sim}{A}(d\underset{\sim}{X})\underset{\sim}{B}$ used in deriving (26).

(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 (25) gives

$$\text{vec}(d\underset{\sim}{Y}) = \sum_{i=1}^{p} (\underset{\sim}{X}'^{p-i} \otimes \underset{\sim}{X}^{i-1})\text{vec}(d\underset{\sim}{X}) \tag{29}$$

from which (13) yields

$$J_{\underset{\sim}{Y}\to\underset{\sim}{X}} = \left|\underset{\sim}{P}\sum_{i=1}^{p} (\underset{\sim}{X}'^{p-i} \otimes \underset{\sim}{X}^{i-1})\underset{\sim}{Q}\right| \quad . \tag{30}$$

(iv) The transformation $\underset{\sim}{Y} = \underset{\sim}{X}^{-p}$ can be expressed as $\underset{\sim}{Y} = \underset{\sim}{V}^{-1}$ and $\underset{\sim}{V} = \underset{\sim}{X}^p$. The Jacobian may be derived via the chain rule $J_{\underset{\sim}{Y}\to\underset{\sim}{X}} = J_{\underset{\sim}{Y}\to\underset{\sim}{V}}J_{\underset{\sim}{V}\to\underset{\sim}{X}}$, which on using (28), (30) and (15) yields

$$J_{\underset{\sim}{Y}\to\underset{\sim}{X}} = \left|-\underset{\sim}{P}\sum_{i=1}^{p} (\underset{\sim}{X}^{-i'} \otimes \underset{\sim}{X}^{i-p-1})\underset{\sim}{Q}\right| \quad . \tag{31}$$

These examples for transformations of arbitrary patterned matrices are now applied to matrices having particular patterns.

(a) No-pattern matrices

When all the elements of $\underset{\sim}{X}$ are functionally independent, $\text{vecp}\underset{\sim}{X} \equiv \text{vec}\underset{\sim}{X}$ and $\underset{\sim}{P} \equiv \underset{\sim}{Q} \equiv \underset{\sim}{I}$. The Jacobian (13) reduces to $J_{\underset{\sim}{Y}\to\underset{\sim}{X}} = |\underset{\sim}{C}_{12}|$ and is available for these and other examples in Table 2. Henderson and Searle [1979] detail the derivation for the first three examples, and the final example is an easy consequence of these.

(b) Symmetric matrices

The vech operator is the particular vecp operator for symmetric matrices, with $\underset{\sim}{H}$ and $\underset{\sim}{G}$, as defined by Henderson and Searle [1979], taking the place of $\underset{\sim}{P}$ and $\underset{\sim}{Q}$, respectively. For symmetric $\underset{\sim}{X}$ and $\underset{\sim}{Y}$ of order n (13) becomes

$$J_{\underset{\sim}{Y}\to\underset{\sim}{X}} = \left|\frac{\partial\,\text{vech}\underset{\sim}{Y}}{\partial\,\text{vech}\underset{\sim}{X}}\right| = |\underset{\sim}{H}\underset{\sim}{C}_{12}\underset{\sim}{G}| \quad . \tag{32}$$

Although detailed derivations of (32) for the examples except (iv) are given in that paper, it is instructive to outline that for $\underset{\sim}{Y} = \underset{\sim}{A}\underset{\sim}{X}\underset{\sim}{A}'$ in this more general setting, to illustrate the conditions in the lemmas and the theorems.

(i) For $\underset{\sim}{Y} = \underset{\sim}{A}\underset{\sim}{X}\underset{\sim}{A}'$ with $\underset{\sim}{X}$ and $\underset{\sim}{Y}$ being symmetric of order n

$$J_{\underset{\sim}{Y}\to\underset{\sim}{X}} = |\underset{\sim}{A}|^{n+1} \quad . \tag{33}$$

From (26), with $\underset{\sim}{B}' = \underset{\sim}{A}$, we have

$$J_{\underset{\sim}{Y} \to \underset{\sim}{X}} = |\underset{\sim}{H}(\underset{\sim}{A} \otimes \underset{\sim}{A})\underset{\sim}{G}| \quad . \tag{34}$$

Lemma 1, with $\underset{\sim}{C}_{12} = \underset{\sim}{A} \otimes \underset{\sim}{A}$, shows $(\underset{\sim}{A} \otimes \underset{\sim}{A})\underset{\sim}{G} = \underset{\sim}{G}\underset{\sim}{H}(\underset{\sim}{A} \otimes \underset{\sim}{A})\underset{\sim}{G}$, for <u>any</u> square $\underset{\sim}{A}$ (because $\underset{\sim}{Y} = \underset{\sim}{A}\underset{\sim}{X}\underset{\sim}{A}'$ is symmetric when $\underset{\sim}{X}$ is). This result with $\underset{\sim}{A} \otimes \underset{\sim}{A}$ replaced by $\underset{\sim}{U}' \otimes \underset{\sim}{U}'$ and $\underset{\sim}{S} \otimes \underset{\sim}{S}$ of the SVD of $\underset{\sim}{A} \otimes \underset{\sim}{A} = (\underset{\sim}{V} \otimes \underset{\sim}{V})(\underset{\sim}{S} \otimes \underset{\sim}{S})(\underset{\sim}{U}' \otimes \underset{\sim}{U}')$ satisfies (19), and so applying corollary 1 reduces (34) to

$$|\underset{\sim}{H}(\underset{\sim}{A} \otimes \underset{\sim}{A})\underset{\sim}{G}| = |\underset{\sim}{H}(\underset{\sim}{V} \otimes \underset{\sim}{V})\underset{\sim}{G}|\,|\underset{\sim}{H}(\underset{\sim}{S} \otimes \underset{\sim}{S})\underset{\sim}{G}|\,|\underset{\sim}{H}(\underset{\sim}{U}' \otimes \underset{\sim}{U}')\underset{\sim}{G}| \quad . \tag{35}$$

Each of the determinants in (35) can be evaluated using corollary 2 for simple matrices (because $\underset{\sim}{V}$ and $\underset{\sim}{U}'$ are orthogonal and $\underset{\sim}{S}$ is diagonal and thus simple). Henderson and Searle [1979] show that when $\underset{\sim}{M}$ is simple (of order n), it may be replaced by its diagonal matrix of eigenvalues giving

$$|\underset{\sim}{H}(\underset{\sim}{M} \otimes \underset{\sim}{M})\underset{\sim}{G}| = |\underset{\sim}{M}|^{n+1} \quad , \tag{36}$$

which substituted into (35) yields

$$\begin{aligned} J_{\underset{\sim}{Y} \to \underset{\sim}{X}} = |\underset{\sim}{H}(\underset{\sim}{A} \otimes \underset{\sim}{A})\underset{\sim}{G}| &= |\underset{\sim}{V}|^{n+1}|\underset{\sim}{S}|^{n+1}|\underset{\sim}{U}'|^{n+1} \\ &= |\underset{\sim}{V}\underset{\sim}{S}\underset{\sim}{U}'|^{n+1} = |\underset{\sim}{A}|^{n+1} \quad , \end{aligned} \tag{37}$$

as claimed in (33).

Of the remaining transformations (ii) and (iii) are discussed by Henderson and Searle [1979] and (iv) is obtained directly from these, as given in Table 2.

Table 2: Jacobians for transformations of matrices having the same pattern

Transformation $Y_{m\times n} = F(X_{m\times n})$	$X, Y \in M_p$ with r functionally independent elements $\left\|\frac{\partial \text{vecp} Y}{\partial \text{vecp} X}\right\|$	X and Y have no pattern $\left\|\frac{\partial \text{vec} Y}{\partial \text{vec} X}\right\|$	X and Y symmetric of order n $\left\|\frac{\partial \text{vech} Y}{\partial \text{vech} X}\right\|$
AXB	$\|P(B' \otimes A)Q\|$	$\|B\|^m \|A\|^n$	$B = sA'$, $s^{\frac{1}{2}n(n+1)}\|A\|^{n+1}$
AXA'	$\|P(A \otimes A)Q\|$	$\|A\|^{2n}$	$\|A\|^{n+1}$
AX	$\|P(I \otimes A)Q\|$	$\|A\|^n$	$A = sI$, $s^{\frac{1}{2}n(n+1)}$
XB	$\|P(B' \otimes I)Q\|$	$\|B\|^m$	$B = sI$, $s^{\frac{1}{2}n(n+1)}$
sX	s^r	s^{mn}	$s^{\frac{1}{2}n(n+1)}$
X^{-1}	$\|-P(X' \otimes X)^{-1}Q\|$	$(-1)^{n^2}\|X\|^{-2n}$	$(-1)^{\frac{1}{2}n(n+1)}\|X\|^{-(n+1)}$
X^p	$\left\|P\left\{\sum_{i=1}^{p}(D^{p-i} \otimes D^{i-1})\right\}Q\right\|$	$\prod_{k=1}^{n}\prod_{j=1}^{n}\left(\sum_{i=1}^{p}\lambda_k^{p-i}\lambda_j^{i-1}\right)$	$\prod_{k=1}^{n}\prod_{j=k}^{n}\left(\sum_{i=1}^{p}\lambda_k^{p-i}\lambda_j^{i-1}\right)$
X^{-p}	$\left\|-P\left\{\sum_{i=1}^{p}D^{-i} \otimes D^{i-p-1}\right\}Q\right\|$	$(-1)^{n^2}\prod_{k=1}^{n}\prod_{j=1}^{n}\left(\sum_{i=1}^{p}\lambda_k^{-i}\lambda_j^{i-p-1}\right)$	$(-1)^{\frac{1}{2}n(n+1)}\prod_{k=1}^{n}\prod_{j=k}^{n}\left(\sum_{i=1}^{p}\lambda_k^{-i}\lambda_j^{i-p-1}\right)$

$D = \text{diag}(\lambda_1, \cdots, \lambda_n)$ is the diagonal matrix of eigenvalues of X.

Table 2 (continued)

Transformation $Y = F(X)$	X and Y lower triangular of order n $\left\lvert\frac{\partial \operatorname{vect} Y}{\partial \operatorname{vect} X}\right\rvert$	X and Y skew-symmetric of order n $\left\lvert\frac{\partial \operatorname{veck} Y}{\partial \operatorname{veck} X}\right\rvert$	X and Y diagonal of order n $\left\lvert\frac{\partial \operatorname{vecd} Y}{\partial \operatorname{vecd} X}\right\rvert$
AXB	$\prod_{i=1}^{n} b_{ii}^{n-i+1} \prod_{j=i}^{n} a_{jj}$	$B = sA'$, $s^{\frac{1}{2}n(n+1)}\lvert A\rvert^{n-1}$	$\prod_{i=1}^{n} a_{ii}b_{ii}$
AXA'	$\prod_{i=1}^{n} a_{ii}^{n+1}$	$\lvert A\rvert^{n-1}$	$\prod_{i=1}^{n} a_{ii}^{2}$
AX	$\prod_{i=1}^{n} a_{ii}^{i}$	$A = sI$, $s^{\frac{1}{2}n(n-1)}$	$\prod_{i=1}^{n} a_{ii}$
XB	$\prod_{i=1}^{n} b_{ii}^{n-i+1}$	$B = sI$, $s^{\frac{1}{2}n(n-1)}$	$\prod_{i=1}^{n} b_{ii}$
sX	$s^{\frac{1}{2}n(n+1)}$	$s^{\frac{1}{2}n(n-1)}$	s^{n}
X^{-1}	$(-1)^{\frac{1}{2}n(n+1)} \prod_{i=1}^{n} x_{ii}^{-(n+1)}$	$(-1)^{\frac{1}{2}n(n-1)}\lvert X\rvert^{-(n-1)}$, n even	$(-1)^{n} \prod_{i=1}^{n} x_{ii}^{-2}$
X^{p}	$\prod_{k=1}^{n} \prod_{j=k}^{n} \left(\sum_{i=1}^{p} x_{kk}^{p-i} x_{jj}^{i-1} \right)$	$\prod_{k=1}^{n-1} \prod_{j=k+1}^{n} \left(\sum_{i=1}^{p} \lambda_{k}^{p-i} \lambda_{j}^{i-1} \right)$, n even, p odd	$p^{n} \prod_{k=1}^{n} x_{kk}^{p-1}$
X^{-p}	$(-1)^{\frac{1}{2}n(n+1)} \prod_{k=1}^{n} \prod_{j=k}^{n} \left(\sum_{i=1}^{p} x_{kk}^{-i} x_{jj}^{i-p-1} \right)$	$(-1)^{\frac{1}{2}n(n-1)} \prod_{k=1}^{n-1} \prod_{j=k+1}^{n} \left(\sum_{i=1}^{p} \lambda_{k}^{-i} \lambda_{j}^{i-p-1} \right)$, n even, p odd	$(-1)^{n} p^{n} \prod_{k=1}^{n} x_{kk}^{-p-1}$

(c) Skew-symmetric matrices

When $\mathbf{X}$ is square, with particular use for skew-symmetric $\mathbf{X} = -\mathbf{X}'$, we define

veck$\mathbf{X}$ as the vector formed by stacking the columns of $\mathbf{X}$ starting below the diagonal , (38)

as illustrated in (3). Then $\mathbf{K}$ and $\mathbf{L}$ are the particular forms of $\mathbf{P}$ and $\mathbf{Q}$, in (5)-(9), corresponding to veck. For skew-symmetric $\mathbf{X}$ of order $n \times n$ vec$\mathbf{X} = \mathbf{L}$veck$\mathbf{X}$, where $\mathbf{L}$ is of order $n^2 \times \frac{1}{2}n(n-1)$, for example, when $n=3$

$$\mathbf{L} = \left[\begin{array}{cc:c} \cdot & \cdot & \cdot \\ 1 & \cdot & \cdot \\ \cdot & 1 & \cdot \\ \hdashline -1 & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & 1 \\ \hdashline \cdot & -1 & \cdot \\ \cdot & \cdot & -1 \\ \cdot & \cdot & \cdot \end{array}\right] . \qquad (39)$$

The $\mathbf{K}$ corresponding to the Moore-Penrose inverse of $\mathbf{L}$, $(\mathbf{L}'\mathbf{L})^{-1}\mathbf{L}'$ is $\frac{1}{2}\mathbf{L}'$, for which $\mathbf{LK} = \frac{1}{2}\mathbf{LL}' = \frac{1}{2}(\mathbf{I}_{n^2} - \mathbf{I}_{n,n})$, where $\mathbf{I}_{n,n}$ is the vec-permutation of order n^2, defined by Henderson and Searle [1979].

(i) The transformation $\mathbf{Y} = \mathbf{AXA}'$ when $\mathbf{X}$ is skew-symmetric of order $n \times n$ has Jacobian based on

$$J_{\mathbf{Y}\to\mathbf{X}} = |\mathbf{A}|^{n-1} .$$

This is established in a similar manner to $|\mathbf{A}|^{n+1}$ for symmetric matrices in (33). Results for the further examples are similarly obtained, and are displayed in Table 2. Two comments on the transformation $\underset{\sim}{Y} = \underset{\sim}{X}^p$ for integer p need to be made. A skew-symmetric $\underset{\sim}{X}$ of odd order is singular, as it has both λ and $-\lambda$ as eigenvalues. When $\underset{\sim}{X} = -\underset{\sim}{X}'$, $\underset{\sim}{X}^p = (-1)^p \underset{\sim}{X}^{p'}$ is skew-symmetric if p is odd.

(d) (Lower) triangular matrices

For square $\underset{\sim}{X}$, with particular application for lower triangular $\underset{\sim}{X}$, vect$\underset{\sim}{X}$ for vector triangular is defined in the same way as vech$\underset{\sim}{X}$. Specifically, vect$\underset{\sim}{X}$ = $\underset{\sim}{T}$vec$\underset{\sim}{X}$ and when $\underset{\sim}{X}$ is lower triangular, of order $n \times n$, vec$\underset{\sim}{X}$ = $\underset{\sim}{T}'$vect$\underset{\sim}{X}$, where $\underset{\sim}{T} = \bigoplus_{i=1}^{n} \underset{\sim}{T}_i$ is the direct sum $\underset{\sim}{T}_1 \oplus \cdots \oplus \underset{\sim}{T}_n$ and $\underset{\sim}{T}_i$ is the last $n-i+1$ rows of $\underset{\sim}{I}_n$.

Table 2 presents results for transformations of triangular matrices, easily derived using this method.

(e) Diagonal matrices

For square $\underset{\sim}{X}$ we define

$$\text{vecd}\underset{\sim}{X} \text{ as the vector of diagonal elements of } \underset{\sim}{X} \qquad (40)$$

so that vecd$\underset{\sim}{X}$ = $\underset{\sim}{D}$vec$\underset{\sim}{X}$, where $\underset{\sim}{D} = \bigoplus_{i=1}^{n} \underset{\sim}{e}_i'$ and $\underset{\sim}{e}_i$ is the ith column of $\underset{\sim}{I}_n$. When $\underset{\sim}{X}$ is also diagonal vec$\underset{\sim}{X}$ = $\underset{\sim}{D}'$vecd$\underset{\sim}{X}$. Examples of Jacobians for transformations of diagonal matrices are given in Table 2.

4.2. Transformations involving matrices with different patterns

Theorems 1 and 2 give conditions under which the Jacobian for the transformation $\underset{\sim}{X}_1 \rightarrow \underset{\sim}{X}_2$, possibly of different patterns, may be simplified using the singular-value decomposition and canonical decomposition theorems. Simple examples of such transformations are to and from symmetric and triangular matrices of the same order, the Jacobians of which can be evaluated by straightforward application of these methods.

For example,

$$\underset{\sim}{Y} = \underset{\sim}{A}\underset{\sim}{X}' + \underset{\sim}{X}\underset{\sim}{A}' \text{ where } \underset{\sim}{A} \text{ and } \underset{\sim}{X} \text{ are lower triangular of order } n \qquad (41)$$

and $\underset{\sim}{Y}$ is, of course, symmetric of the same order, has Jacobian $J_{\underset{\sim}{Y}\to\underset{\sim}{X}} = 2^n \prod_{i=1}^{n} a_{ii}^{n-i+1}$, where a_{ii} is the ith diagonal element of $\underset{\sim}{A}$. This result leads immediately to the Jacobian for the Schur (Bartlett) decomposition $\underset{\sim}{Y} = \underset{\sim}{X}\underset{\sim}{X}'$, where $\underset{\sim}{X}$ is lower triangular and $\underset{\sim}{Y}$ symmetric, which has the same form as (41) in terms of the differentials $d\underset{\sim}{Y} = \underset{\sim}{X}(d\underset{\sim}{X}') + (d\underset{\sim}{X})\underset{\sim}{X}'$. Hence $J_{\underset{\sim}{Y}\to\underset{\sim}{X}} = 2^n \prod_{i=1}^{n} x_{ii}^{n-i+1}$, as discussed by Henderson and Searle [1979], using special cases of the methods formulated here.

The appeal of this method of deriving Jacobians is that it is applicable to one-to-one transformations of <u>any</u> patterned matrices whose elements are repetitions (or more generally linear combinations) of functionally independent variables.

ACKNOWLEDGMENTS

Grateful thanks go to S. R. Searle and D. L. Solomon for discussion of this topic.

REFERENCES

Deemer, W. L. and Olkin, I. [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.

Henderson, H. V. and Searle, S. R. [1979]. Vec and vech operators for matrices, with some uses in Jacobians and multivariate statistics. Canad. J. Statist., 7, No. 1 (in press).

Lancaster, P. [1969]. Theory of Matrices. Academic Press, New York.

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

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, New York, 181-189.

Stewart, G. W. [1973]. Introduction to Matrix Computations. Academic Press, New York.