7 : APPENDIX

Vectors and Matrices

An n-tuple vector x is defined as an ordered set of n numbers. Usually we write these numbers $x_1, \ldots, x_n$ in a *column* in the order indicated by their subscripts. The transpose of x is the *row* vector $x' = [x_1, \ldots, x_n]$ and the transpose of x' is x again. That is $(x')' = x$. We can display x within the line by writing $x = [x_1, \ldots, x_n]'$, which saves space.

An $m \times n$ real matrix X is an array of numbers which are set in m rows and n columns :

$$(1) \qquad \begin{aligned} X &= \begin{bmatrix} x_{11} & x_{12} & \ldots & x_{1n} \\ x_{21} & x_{22} & \ldots & x_{2n} \\ \vdots & \vdots & & \vdots \\ x_{m1} & x_{m2} & \ldots & x_{mn} \end{bmatrix} = \begin{bmatrix} x_{1.} \\ x_{2.} \\ \vdots \\ x_{m.} \end{bmatrix} \\ &= \begin{bmatrix} x_{.1} & x_{.2} & \ldots & x_{.n} \end{bmatrix}. \end{aligned}$$

The subscript associated with the generic element x_{ij} is to indicate that it is located in the ith row and the jth column. It often makes sense to regard an n-tuple column vector as a matrix of order $n \times 1$.

In (1), the matrix X has been represented also as an ordered set of n column vectors $x_{.1}, \ldots, x_{.n}$ and as an ordered set of m row vectors $x_{1.}, \ldots, x_{m.}$. In the notation $x_{.j}$, which stands for the jth column, the dot in the subscript is a placeholder which indicates that the row subscript ranges over all the values $1, \ldots, m$. The notation $x_{i.}$, which stands for the ith row, is to be interpreted likewise.

A matrix is often represented in a summary fashion by writing $X = [x_{ij}]$, which is appropriate when we need to display the generic element.

Although we commonly talk of matrix algebra, it may be more helpful, in the first instance, to regard a matrix as a mathematical notation rather than as a mathematical object in its own right. In this way, we can avoid attributing a specific algebra to matrices; which leaves them free to represent whatever abstract objects we care to define. Bearing this in mind, one should not be too confused by the many different roles assumed by matrices. Nevertheless, we may define three basic operations for matrices which are applied in many algebras.

(2) The sum of the $m \times n$ matrices $A = [a_{ij}]$ and $B = [b_{ij}]$ is the $m \times n$ matrix $A + B = C = [c_{ij}]$ wherein the generic element is $c_{ij} = a_{ij} + b_{ij}$.

(3) The product of the the matrix $A = [a_{ij}]$ and the scalar λ is the matrix $\lambda A = [\lambda a_{ij}]$.

(4) The product of the $m \times n$ matrix $A = [a_{ij}]$ and $n \times p$ matrix $B = [b_{jk}]$ is the $m \times p$ matrix $AB = C = [c_{ik}]$ wherein the generic element is $c_{ik} = \sum_j a_{ij}b_{jk} = a_{i1}b_{1k} + a_{i2}b_{2k} + \cdots + a_{in}b_{nk}$.

We shall begin with an account of the algebra of real n-dimensional vector spaces. In this context, we shall regard an $n \times m$ matrix simply as an ordered set of m column vectors, each of n elements; and, to begin, we shall consider only the operations under (2) and (3).

Real Coordinate Vector Spaces

The set of all n-tuple vectors with real-valued elements constitutes a real coordinate space $\mathcal{R}^n$ which is closed in respect of the operations of the addition of vectors and their multiplication by arbitrary scalars. Thus, if $x, y \in \mathcal{R}^n$, then $(x + y) \in \mathcal{R}^n$; and, if $\lambda, \mu \in \mathcal{R}$ are scalars, then $\lambda x, \mu y \in \mathcal{R}^n$. Combining these two results, we have

(5) $$(\lambda x + \mu y) \in \mathcal{R}^n \quad \text{if} \quad \lambda, \mu \in R \quad \text{and} \quad x, y \in \mathcal{R}^n.$$

A linear subspace $\mathcal{S} \subset \mathcal{R}^n$ is a subset of the vectors of $\mathcal{R}^n$ which is closed in respect of the two operations. Thus

(6) If $(\lambda x + \mu y) \in \mathcal{S}$ for all $x, y \in \mathcal{S}$ and $\lambda, \mu \in \mathcal{R}$, then $\mathcal{S}$ is a vector subspace.

Notice that a vector subspace must include the zero vector which is the origin of any coordinate system.

Linear Combinations

If $x_1, \ldots, x_p$ is a set of n-tuple vectors and if $\lambda_1, \ldots, \lambda_p$ is a set of scalars, then the sum $\lambda_1 x_1 + \cdots + \lambda_p x_p$ is said to be a *linear combination* of these vectors.

The set of all linear combinations of the column vectors in the $n \times m$ matrix $X = [x_{.1}, \ldots, x_{.m}]$ is said to be the *manifold* of that matrix, and it is denoted by $\mathcal{M}(X)$. Clearly, this manifold is a linear subspace of $\mathcal{R}^n$. The subspace is said to be spanned or generated by the column vectors.

A set of vectors $x_1, \ldots, x_p$ is said to be *linearly dependent* if there exists a set of scalars $\lambda_1, \ldots, \lambda_p$, not all of which are zeros, such that $\lambda_1 x_1 + \cdots + \lambda_p x_p = 0$. If there are no such scalars, then the vectors are said to be *linearly independent*.

A set of linearly independent vectors spanning a subspace $\mathcal{S}$ is said to constitute a *basis* of $\mathcal{S}$. It can be shown that any two bases of $\mathcal{S}$ comprise the same number of vectors which is the maximum number of linearly independent vectors. This number is called the *dimension* of $\mathcal{S}$, written $\text{Dim}\{\mathcal{S}\}$.

Any vector $x \in \mathcal{S}$ can be expressed as a unique linear combination of the vectors of a basis of $\mathcal{S}$. That is to say, if $v_1, \ldots, v_p$ with $p = \text{Dim}\{\mathcal{S}\}$ is a basis of $\mathcal{S}$ and if $x = \lambda_1 v_1 + \cdots + \lambda_p v_p$, then the scalars $\lambda_1, \ldots, \lambda_p$ are unique.

The *natural basis* of $\mathcal{R}^n$ is the set of vectors $e_1, \ldots, e_n$ which constitute the columns of the $n \times n$ identity matrix $I_n = [e_1, \ldots, e_n]$. If $[x_1, \ldots, x_n]' = x \in \mathcal{R}^n$ is any vector, then it can be expressed uniquely in terms of the natural basis as follows:

$$x = x_1 e_1 + x_2 e_2 + \cdots + x_n e_n. \tag{7}$$

The scalars x_i are the coordinates of x relative to the natural basis and the vectors $x_i e_i$ are the projections of x onto the one-dimensional subspaces or lines spanned by the individual basis vectors. More generally, the vector

$$[x_1, \ldots, x_p, 0, \ldots, 0]' = x_1 e_1 + \cdots + e_p x_p \tag{8}$$

is the projection of x onto the subspace spanned jointly by the vectors $e_1, \ldots, e_p$ which are a subset of the basis set.

Inner Products and Orthogonality

The inner product of two vectors $x, y \in \mathcal{R}^n$ is defined by

$$\langle x, y \rangle = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n \tag{9}$$

The norm of the vector x is

$$\|x\| = \sqrt{\langle x, x \rangle}. \tag{10}$$

We define the angle θ between the vectors x and y by

$$\cos\theta = \frac{\langle x, y \rangle}{\|x\|.\|y\|}. \tag{11}$$

Two vectors x and y are said to be orthogonal or perpendicular if $\langle x, y \rangle = 0$. A vector x is said to be orthogonal to a subspace $\mathcal{Y}$ if $\langle x, y \rangle = 0$ for all $y \in \mathcal{Y}$. A subspace $\mathcal{X}$ is said to be orthogonal to a subspace $\mathcal{Y}$ if $\langle x, y\rangle = 0$ for all $x \in \mathcal{X}$ and $y \in \mathcal{Y}$. In that case, the only vector which $\mathcal{X}$ and $\mathcal{Y}$ have in common is the zero vector. Thus $\mathcal{X} \cap \mathcal{Y} = 0$.

If $\mathcal{A}$ and $\mathcal{B}$ are two subspaces, not necessarily orthogonal, such that $\mathcal{A} \cap \mathcal{B} = 0$, then the set of all vectors $z = a + b$ with $a \in \mathcal{A}$ and $b \in \mathcal{B}$ is called a direct sum of $\mathcal{A}$ and $\mathcal{B}$ written $\mathcal{A} \oplus \mathcal{B}$.

Let $\mathcal{X}$ and $\mathcal{Y}$ be orthogonal subspaces of $\mathcal{R}^n$ and let their direct sum be $\mathcal{X} \oplus \mathcal{Y} = \mathcal{R}^n$. That is to say, any vector $z \in \mathcal{R}^n$ can be written uniquely as $z = x + y$ with $x \in \mathcal{X}$ and $y \in \mathcal{Y}$ and with $\langle x, y\rangle = 0$. Then $\mathcal{X}$ and $\mathcal{Y}$ are orthogonal complements and we may write $\mathcal{X} = \mathcal{O}(\mathcal{Y})$ and $\mathcal{Y} = \mathcal{O}(\mathcal{X})$. We may indicate that $\mathcal{X}$ and $\mathcal{Y}$ are orthogonal by writing $\mathcal{X} \perp \mathcal{Y}$. Likewise, if $\langle x, y\rangle = 0$, then we can write $x \perp y$.

Parametric Representations of Subspaces

Let $\mathcal{M}(B)$ be the manifold of the matrix $B = [b_{.1}, \ldots, b_{.m}]$ comprising m linearly independent vectors. That is to say, $\mathcal{M}(B)$ is the linear subspace comprising all linear combinations of the column vectors of B. Then any vector $x \in \mathcal{M}(B)$ can be expressed as $x = B\lambda = \lambda_1 b_{.1} + \cdots + \lambda_m b_{.m}$ where $\lambda = [\lambda_1, \ldots, \lambda_m]'$ is an m-tuple vector. Let $p \notin \mathcal{M}(B)$ be an n-tuple vector. Then $\mathcal{A} = \mathcal{M}(B) + p$ is called an *affine subspace* of $\mathcal{R}^n$ or a *translated vector subspace*, and the vector p is termed the translation. Any vector $x \in \mathcal{A}$ can be expressed as $x = B\lambda + p$ for some λ. It makes sense to define the dimension of $\mathcal{A}$ to be equal to that of $\mathcal{M}(B)$

Now consider an $n \times (n - m)$ matrix A comprising a set of $n - m$ linearly independent vectors which are orthogonal to those of B such that $A'B = 0$ and $\mathcal{M}(A) \oplus \mathcal{M}(B) = \mathcal{R}^n$. Then, on multiplying our equation by A', we get

$$\begin{aligned} A'x &= A'B\lambda + A'p \\ &= A'p = c. \end{aligned} \tag{12}$$

Thus we have an alternative definition of the affine subspace as the set $\mathcal{A} = \{x; A'x = c\}$.

There are two limiting cases. In the first case, the matrix A is replaced by an $n \times 1$ vector a, and we have the equation $a'x = c$ where c is a scalar. Then the set $\mathcal{A} = \{x; a'x = c\}$ represents a hyperplane in $\mathcal{R}^n$. A hyperplane is a subspace—in this case an affine subspace—whose dimension is one less than that of the space in which it resides.

In the second case, A becomes a matrix of order $n \times (n - 1)$. Then the set $\mathcal{A} = \{x; A'x = c\}$ defines a line in $\mathcal{R}^n$ which is an one-dimensional affine subspace.

Matrices as Linear Transformations

An $m \times n$ matrix A can be regarded as a linear transformation mapping from the space $\mathcal{R}^n$ to the space $\mathcal{R}^m$. Thus the equation $y = Ax$ describes how the n-tuple vector $x = [x_1, \ldots, x_n]'$ is mapped into the m-tuple vector $y = [y_1, \ldots, y_m]'$.

If $x, z \in \mathcal{R}^n$ are vectors and if $\lambda, \mu \in \mathcal{R}$ are scalars, then

$$A(\lambda x + \mu z) = \lambda Ax + \mu Az. \tag{13}$$

This equation represents the linearity of the transformation. Thus, for a linear transformation, the operations of vector addition and scalar multiplication may be performed equally in $\mathcal{R}^n$, prior to the transformation of the product, or in $\mathcal{R}^m$, after the separate transformations of the two vectors.

There are various linear spaces associated with a transformation A. The *null space* of A, denoted $\mathcal{N}(A) = \{x; Ax = 0\}$, is the set of all $x \in \mathcal{R}^n$ which are mapped into the zero vector $0 \in \mathcal{R}^m$. The dimension of the null space is called the nullity of A, denoted $\text{Null}(A)$.

The *range space* of A, denoted $\mathcal{R}(A) = \{y; y = Ax, x \in \mathcal{R}^n\}$ is the set of all $y \in \mathcal{R}^m$ which represent the mapping under A of some $x \in \mathcal{R}^n$. We can also write $\mathcal{R}(A) = A\mathcal{R}^n$, which makes it explicit that the domain of A is the space $\mathcal{R}^n$. The dimension of $\mathcal{R}(A)$ is the *rank* of A denoted $\text{Rank}(A)$. There is the following relationship between the rank of A, its nullity and the dimension of its domain $\mathcal{R}^n$:

$$\text{Null}(A) + \text{Rank}(A) = \text{Dim}\{\mathcal{R}^n\}. \tag{14}$$

Notice that $\mathcal{R}(A)$ is simply the set of all linear combinations of the columns of the matrix A which we have previously described as the manifold of $\mathcal{M}(A)$ of the matrix.

The Composition of Linear Transformations

Let the $p \times n$ matrix $A = [a_{kj}]$ represent a transformation from $\mathcal{R}^n$ to $\mathcal{R}^p$ and let the $m \times p$ matrix $B = [b_{ik}]$ represent a transformation from $\mathcal{R}^p$ to $\mathcal{R}^m$. Then their composition is the matrix $C = BA = [c_{ij}]$ of order $m \times n$ representing a transformation from $\mathcal{R}^n$ to $\mathcal{R}^m$. Sylvester's theorem states that

$$\text{Rank}(BA) \leq \min\big\{\text{Rank}(B), \text{Rank}(A)\big\}. \tag{15}$$

Proof. We have $\mathcal{R}(B) = B\mathcal{R}^p$ and $\mathcal{R}(BA) = B\mathcal{R}(A)$ with $\mathcal{R}(A) \subset \mathcal{R}^p$; so it must be true that $\text{Dim}\{B\mathcal{R}(A)\} = \text{Rank}(BA) \leq \text{Rank}(B) = \text{Dim}\{B\mathcal{R}^p\}$.

To prove that $\text{Rank}(BA) \leq \text{Rank}(A)$, consider a basis $a_1, \ldots, a_g$ of $\mathcal{R}(A)$ consisting of a set of $g = \text{Rank}(A)$ linearly independent vectors. Then the range space $\mathcal{R}(BA) = B\mathcal{R}(A)$ of the composition is spanned by the transformed vectors $Ba_1, \ldots, Ba_g$. For, if $y = \lambda_1 a_1 + \cdots + \lambda_g a_g$ is any vector in $\mathcal{R}(A)$, then its image is $z = By = \lambda_1 Ba_1 + \cdots + \lambda_g Ba_g$. It follows that, in any basis of $\mathcal{R}(BA)$, the number of linearly independent vectors does not exceed $g = \text{Rank}(A)$ and may be less. Thus $\text{Rank}(BA) \leq \text{Rank}(A)$.

Transformations on $\mathcal{R}^n$

We may regard an $n \times n$ matrix A as a transformation on $\mathcal{R}^n$. There are a few such transformation which are of prime importance. We shall consider, in turn, the inverse matrix, the class of orthonormal matrices, and the class of orthogonal projectors.

The inverse matrix A^{-1} is defined by the conditions $AA^{-1} = A^{-1}A = I_n$. It follows from Sylvester's Theorem that

$$n = \text{Rank}(I_n) \leq \min\bigl\{\text{Rank}(A), \text{Rank}(A^{-1})\bigr\},$$

which implies that $\text{Rank}(A) \geq n$. But we also have the result that $n = \text{Rank}(A) + \text{Null}(A)$, which implies that $\text{Rank}(A) \leq n$. Therefore $\text{Rank}(A) = n$; and this is the necessary and sufficient condition for the existence of the inverse matrix.

A matrix C is said to be orthonormal if it satisfies the conditions $C'C = CC' = I$.

(16) If C is orthonormal and if $x, y \in \mathcal{R}^n$ are any two vectors, then $\|x\| = \|Cx\|$ and $\|y\| = \|Cy\|$. Also the angle θ between Cx and Cy is the same as the angle between x and y.

Proof. In the first place, we have $\|x\| = \sqrt{\langle x, x\rangle} = \sqrt{x'x}$. Also, we have $\|Cx\| = \sqrt{\langle Cx, Cx\rangle} = \sqrt{xC'Cx} = \sqrt{x'x}$, so $\|x\| = \|Cx\|$. Likewise, we can show that $\|y\| = \|Cy\|$. Next, we have $\langle Cx, Cy\rangle = x'C'Cy = x'y$, so $\langle Cx, Cy\rangle = \langle x, y\rangle$. It follows that

$$\cos\theta = \frac{\langle Cx, Cy\rangle}{\|Cx\|.\|Cy\|} = \frac{\langle x, y\rangle}{\|x\|.\|y\|}, \tag{17}$$

where the last expression is for the cosine of the angle between x and y.

In view of the properties demonstrated above, the orthonormal matrix is said to represent an isometric transformation. In fact, in two and three dimensions, orthonormal matrices correspond either to reflections or to rotations.

Example. Consider the matrix

$$C = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}. \tag{18}$$

It is readily confirmed that $C'C = CC' = I$. Moreover, the effect of the transformation on any vector $x \in \mathcal{R}^2$ is to rotate it in an anticlockwise sense through an angle of θ relative to the natural coordinate system. Alternatively, we can conceive of the vector remaining fixed while the coordinate system rotates in an anticlockwise sense.

Let $C = C(\theta)$ and $Q = Q(\phi)$ be two orthonormal matrices defined in the manner of (18). Then their product CQ is also an orthonormal matrix:

$$(19) \qquad \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix} = \begin{bmatrix} \cos(\theta+\phi) & -\sin(\theta+\phi) \\ \sin(\theta+\phi) & \cos(\theta+\phi) \end{bmatrix}.$$

This result, which arises intuitively, may be established using the following trigonometrical identities which are known as the compound-angle formulae:

$$(20) \qquad \begin{aligned} \cos(\theta+\phi) &= \cos\theta\cos\phi - \sin\theta\sin\phi, \\ \sin(\theta+\phi) &= \sin\theta\cos\phi + \cos\theta\sin\phi. \end{aligned}$$

A symmetric idempotent matrix is one which satisfies the conditions $P = P' = P^2$ or, equivalently, the condition $P'(I-P) = 0$. Such a matrix is also described as an orthogonal or minimum-distance projector for reasons which will become apparent shortly. The matrix has the following characteristics:

(21) (i) If $x \in \mathcal{M}(P)$, then $Px = x$ and $(I-P)x = 0$,

(ii) The conditions $P(I-P) = 0$ and $(I-P)P = 0$ imply that $\mathcal{R}(I-P) = \mathcal{N}(P)$ and $\mathcal{R}(P) = \mathcal{N}(I-P)$,

(iii) The direct sum $\mathcal{M}(P) \oplus \mathcal{M}(I-P) = \mathcal{R}^n$ describes an orthogonal decomposition of $\mathcal{R}^n$ with $\mathcal{M}(P) \perp \mathcal{M}(I-P)$.

We may summarise by saying that P acts as an identity transformation on $\mathcal{M}(P)$ and as a zero transformation on $\mathcal{M}(I-P)$. Moreover, any vector $y \in \mathcal{R}^n$ is decomposed uniquely as $y = Py + (I-P)y$ where $Py \perp (I-P)y$.

The result under (20, i) is the consequence of the condition of idempotency, $P = P^2$ or $P(I-P) = 0$, which is the defining condition of any projector. The result under (20, iii) is the consequence of the additional condition of symmetry, $P = P'$, which induces the condition of orthogonality, $P'(I-P) = 0$. The latter ensures that the distance between a vector y and its image Py is minimised by the projector. Thus

(22) If $P = P' = P^2$ and if Px is any vector in $\mathcal{R}(P)$, then $\|y - Py\| \leq \|y - Px\|$.

Proof. We have

$$\begin{aligned}\|y - Px\|^2 &= \|(y - Py) - (Px - Py)\|^2 \\ &= (y - Py)'(y - Py) - 2(y - Py)'(Px - Py) \\ &\qquad + (Px - Py)'(Px - Py) \\ &= \|y - Py\|^2 + \|Px - Py\|^2 - 2y'(I - P)'P(x - y).\end{aligned}$$

It follows that $\|y - Px\|^2 = \|y - Py\|^2 + \|Px - Py\|^2$ for all $x, y \in \mathcal{R}^n$, or, equivalently, that $\|y - Px\| \geq \|y - Py\|$, if and only if $(I - P)'P = 0$, which is equivalent to the conditions $P = P' = P^2$.

Characteristic Roots and Vectors of a Symmetric Matrix

Let A be an $n \times n$ symmetric matrix such that $A = A'$, and imagine that the scalar λ and the vector x satisfy the equation $Ax = \lambda x$. Then λ is a characteristic root of A and x is a corresponding characteristic vector. We also refer to characteristic roots as latent roots or eigenvalues. The characteristic vectors are also called eigenvectors.

(23) The characteristic vectors corresponding to two distinct characteristic roots are orthogonal. Thus, if $Ax_1 = \lambda_1 x_1$ and $Ax_2 = \lambda_2 x_2$ with $\lambda_1 \neq \lambda_2$, then $x_1'x_2 = 0$.

Proof. Premultiplying the defining equations by x_2' and x_1' respectively, gives $x_2'Ax_1 = \lambda_1 x_2'x_1$ and $x_1'Ax_2 = \lambda_2 x_1'x_2$. But $A = A'$ implies that $x_2'Ax_1 = x_1'Ax_2$, whence $\lambda_1 x_2'x_1 = \lambda_2 x_1'x_2$. Since $\lambda_1 \neq \lambda_2$, it must be that $x_1'x_2 = 0$.

The characteristic vector corresponding to a particular root is defined only up to a factor of proportionality. For let x be a characteristic vector of A such that $Ax = \lambda x$. Then multiplying the equation by a scalar μ gives $A(\mu x) = \lambda(\mu x)$ or $Ay = \lambda y$; so $y = \mu x$ is another characteristic vector corresponding to λ.

(24) If $P = P' = P^2$ is a symmetric idempotent matrix, then its characteristic roots can take only the values of 0 and 1.

Proof. Since $P = P^2$, it follows that, if $Px = \lambda x$, then $P^2x = \lambda x$ or $P(Px) = P(\lambda x) = \lambda^2 x = \lambda x$, which implies that $\lambda = \lambda^2$. This is possible only when $\lambda = 0, 1$.

The Diagonalisation of a Symmetric Matrix

Let A be an $n \times n$ symmetric matrix, and let $x_1, \ldots, x_n$ be a set of n linearly independent characteristic vectors corresponding to its roots $\lambda_1, \ldots, \lambda_n$. Then we can form a set of normalised vectors

$$c_1 = \frac{x_1}{\sqrt{x_1' x_1}} \;,\; \ldots \;,\; c_n = \frac{x_n}{\sqrt{x_n' x_n}},$$

which have the property that

$$c_i' c_j = \begin{cases} 0, & \text{if } i \neq j; \\ 1, & \text{if } i = j. \end{cases}$$

The first of these reflects the condition that $x_i' x_j = 0$. It follows that $C = [c_1, \ldots, c_n]$ is an orthonormal matrix such that $C'C = CC' = I$.

Now consider the equation $A[c_1, \ldots, c_n] = [\lambda_1 c_1, \ldots, \lambda_n c_n]$ which can also be written as $AC = C\Lambda$ where $\Lambda = \text{Diag}\{\lambda_1, \ldots, \lambda_n\}$ is the matrix with λ_i as its ith diagonal elements and with zeros in the non-diagonal positions. Post-multiplying the equation by C' gives $ACC' = A = C\Lambda C'$; and premultiplying by C' gives $C'AC = C'C\Lambda = \Lambda$. Thus $A = C\Lambda C'$ and $C'AC = \Lambda$; and C is effective in diagonalising A.

Let D be a diagonal matrix whose ith diagonal element is $1/\sqrt{\lambda_i}$ so that $D'D = \Lambda^{-1}$ and $D'\Lambda D = I$. Premultiplying the equation $C'AC = \Lambda$ by D' and postmultiplying it by D gives $D'C'ACD = D'\Lambda D = I$ or $TAT' = I$, where $T = D'C'$. Also, $T'T = CDD'C' = C\Lambda^{-1}C' = A^{-1}$. Thus we have shown that

(25) For any symmetric matrix $A = A'$, there exists a matrix T such that $TAT' = I$ and $T'T = A^{-1}$.