DIAGONALISATION OF A SYMMETRIC MATRIX

Characteristic Roots and Characteristic Vectors. 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.

(1) 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 $x_2' A x_1 = \lambda_1 x_2' x_1$ and $x_1' A x_2 = \lambda_2 x_1' x_2$. But $A = A'$ implies that $x_2' A x_1 = x_1' A x_2$, 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.

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

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 $\Lambda = \mathrm{Diag}\{\lambda_1, \ldots, \lambda_n\}$ is the matrix with λ_i as its ith diagonal elements and with zeros in the non-diagonal positions. Postmultiplying 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

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

Principal Components of a Symmetric Matrix. The characteristic vectors of a symmetric matrix may be ordered according to the declining values of their associated characteristic roots. If $A = A'$ is a symmetric positive-definite matrix such that $x'Ax > 0$ for all x, then the characteristic roots are all positive, and the ordered set of characteristic vectors is apt to be called the principal components of the matrix.

The first principal component is the vector c_1 which fulfils the following criterion

$$\text{(4)} \qquad \text{Maximise} \quad c'Ac \quad \text{subject to} \quad c'c = 1.$$

To fulfil this condition, we should evaluate the following Lagrangean function:

$$\text{(5)} \qquad L(c, \lambda) = c'Ac - \lambda c'c.$$

Differentiating L with respect to c and setting the result to zero gives a first-order condition $Ac_1 - \lambda c_1 = 0$, which is in accordance with the defining condition of the characteristic vectors and characteristic roots of A. By premultiplying the equation by c_1' and by observing the condition that $c_1'c_1 = 1$, we find that $\lambda_1 = c_1'Ac_1$, which is also in accordance with the previous results.

The second principal component is the vector c_2 which fulfils the following criterion:

$$\text{(6)} \qquad \text{Maximise} \quad c'Ac \quad \text{subject to} \quad c'c = 1 \quad \text{and} \quad c'c_1 = 0.$$

It is straightforward to show that the vector c_2 is the characteristic vector of A which is associated with the second largest characteristic root.

The characteristic vectors of a symmetric positive-definite matrix are aligned with the principal axes of the ellipsoid defined by the equation $x'Ax = r$, where r is a squared radius, which can be set to unity for convenience. Consider, for example, the criterion

$$\text{(7)} \qquad \text{Maximise} \quad c'c \quad \text{subject to} \quad c'Ac = 1.$$

This relates to the problem of finding the axis spanned by the vector c where the extent of ellipsoid is greatest. To fulfil the condition, we should evaluate the following Lagrangean function:

$$\text{(8)} \qquad L(c, \lambda) = c'c - \mu c'Ac.$$

Differentiating L with respect to c and setting the result to zero gives the first-order condition $c - \mu Ac = 0$ or, equivalently, $Ac - \mu^{-1}c = 0$, which indicates that c is a characteristic vector of A. Since $c'c = \mu$, we see that the principal axis of the ellipsoid is aligned with the characteristic vector which corresponds to the largest value of μ, which is the smallest characteristic root of A.