ANIMAL BREEDING NOTES

CHAPTER 3

EIGENVALUES, EIGENVECTORS, AND DIAGONALIZATION OF A SQUARE MATRIX

Eigenvalues and eigenvectors

The **eigenvalues** of a **square** matrix $A_{n \times n}$ are the n roots of its characteristic polynomial:

$$|A - \lambda I| = 0$$

The **set of eigenvalues** (or latent roots) is called the **spectrum** and can be denoted as:

$$\lambda(A) = \{\lambda_1, \lambda_2, \ldots, \lambda_n\}$$

Associated with these n eigenvalues are n **eigenvectors**. The eigenvectors **must** satisfy the equation:

$$Au_i = \lambda_i u_i, \ i = 1, \ldots, n$$

where u_i = i^{th} eigenvector.

Diagonalizable matrix: a square matrix A is called diagonalizable if there is an invertible matrix P such that $P^{-1}AP$ is diagonal. The matrix P is said to diagonalize A (Anton, 1981, pg. 269).

Theorem: If A is an n × n matrix the following statements are equivalent:

(a) A is diagonalizable, and

(b) A has n linearly independent eigenvectors.

Proof:

(a) ⇒ (b) Assume A is diagonalizable, then there is an invertible matrix P such that:

$$P^{-1}AP = D$$

$$AP = PD$$

$$[Ap_1 \; Ap_2 \; \dots \; Ap_n] = [\lambda_1 p_1 \; \lambda_2 p_2 \; \dots \; \lambda_n p_n]$$

$$\Rightarrow Ap_i = \lambda_i p_i, \quad i = 1, \dots, n$$

where

$\{\lambda_i\}$ are the eigenvalues of A, and

$\{p_i\}$ are the corresponding eigenvectors, which are independent because P^{-1} exists.

(b) ⇒ (a) Assume A has n linearly independent eigenvectors (p_i) with associated eigenvalues $\{\lambda_i\}$. Consider the product,

$$A[\, p_1 \, p_2 \, \dots \, p_n \,].$$

But $Ap_i = \lambda_i p_i, \quad i = 1, \dots, n$. Thus, AP = PD, where D = diag $\{\lambda_i\}$.

Since the columns of P are linearly independent, P^{-1} exists, so

$$P^{-1}AP = D \Rightarrow A \text{ is diagonalizable.}$$

This theorem indicates that if a square matrix A is diagonalizable, we can find a set of n linearly independent eigenvectors. The matrix A can be singular or non-singular.

Placing the latent vectors of matrix A together to form a matrix we obtain:

$$U = [\, u_1 \, u_2 \, \dots \, u_n \,]$$

where U is an $n \times n$ square matrix, rank (U) = n, hence U is non-singular and U^{-1} exists. Thus, forming the equations $Au_i = \lambda u_i$, for i = 1, ..., n we get:

$$A[\, u_1 \, u_2 \, \dots \, u_n \,] = [\, u_1 \, u_2 \, \dots \, u_n \,] \text{diag} \{\lambda_i\}$$

or

$$AU = UD$$

$$\Rightarrow \quad A = UDU^{-1} \quad \text{and} \quad D = U^{-1}AU$$

where

$$D = \text{canonical form of } A \text{ under similarity.}$$

Furthermore, if **A is symmetric**, i.e., $A' = A$, there exist an **orthogonal U** (i.e., $UU' = U'U = I$) such that

$$A = UDU' \quad \text{and} \quad D = U'AU$$

Spectral Decomposition

For a **non-symmetric matrix** $A_{n\times n}$, the spectral decomposition is:

$$A = \lambda_1 u_1 v^1 + \lambda_2 u_2 v^2 + \cdots + \lambda_n u_n v^n$$

where $\{\lambda_i\}$ = eigenvalues of A

$\{u_i\}$ = eigenvectors of A

$\{v^i\}$ = rows of the matrix U^{-1}

Proof:

$$A = UDU^{-1}$$

Let $\begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$ be the row vectors of matrix U^{-1}.

Then, $$A = \begin{bmatrix} u_1 & u_2 & \cdots & u_n \end{bmatrix} \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix} \begin{bmatrix} v^1 \\ v^2 \\ \vdots \\ v^n \end{bmatrix}$$

$$A = \lambda_1 u_1 v^1 + \lambda_2 u_2 v^2 + \cdots + \lambda_n u_n v^n$$

For a **symmetric matrix** $A_{n \times n}$, the spectral decomposition is:

$$A = \lambda_1 u_1 u^{1\prime} + \lambda_2 u_2 u^{2\prime} + \cdots + \lambda_n u_n u^{n\prime}$$

where $\{u^{i\prime}\}$ = rows of the **symmetric** matrix U^{-1}.

Thus, if the set of linearly independent eigenvectors $\{u_i\}$ of a non-singular matrix A is orthogonalized, the spectral decomposition of A becomes:

$$A = \lambda_1 u_1 u^{1\prime} + \lambda_2 u_2 u^{2\prime} + \cdots + \lambda_n u_n u^{n\prime}$$

where $\{u_i\}$ = orthogonal eigenvectors of A $(u_i' \, u_{i'}) = 0$.

Furthermore, if the $\{u_i\}$ are normalized such that $u_i' \, u_i = 1$, the spectral decomposition is:

$$A = \lambda_1 e_1 e_1' + \lambda_2 e_2 e_2' + \cdots + \lambda_n e_n e_n'$$

where $\{e_i\}$ = orthonormal eigenvectors of A $(e_i' \, e_{i'} = 0$ and $e_i' \, e_i = 1)$.

Results:

1) If A is an $n \times n$ matrix, U is an $n \times n$ non-singular matrix and $D = PAP^{-1}$, then the characteristic polynomial, characteristic roots, trace, determinant and rank of D are identical to those of A.

Proof:

(a) $|D - \lambda I| = |PAP^{-1} - \lambda I|$

$$= \left| PAP^{-1} - \lambda PP^{-1} \right|$$

$$= \left| P(A - \lambda I)P^{-1} \right|$$

$$= \left| P \right| \left| A - \lambda I \right| \left| P^{-1} \right|$$

$$= \left| A - \lambda I \right| \left| PP^{-1} \right|$$

$$= \left| A - \lambda I \right| \left| I \right|$$

$$= \left| A - \lambda I \right|$$

$\Rightarrow$ the characteristic polynomial and the characteristic roots of D and A are the same.

(b) $\text{trace}(D) = \text{trace}(PAP^{-1})$

$$= \text{trace}(P^{-1}PA)$$

$$= \text{trace}(A)$$

Also, note that

$$\text{trace}(D) = \sum_{i=1}^{n} \lambda_i = \text{trace}(A)$$

(c) $\left| D \right| = \left| PAP^{-1} \right|$

$$= \left| P \right| \left| A \right| \left| P^{-1} \right|$$

$$= \left| P \right| \left| P^{-1} \right| \left| A \right|$$

But the product of the determinants of two square matrices of the same order is equal to the determinant of the product of these matrices (Searle, 1966, pg. 76), thus

$$\left| D \right| = \left| PP^{-1} \right| \left| A \right|$$

$$\left| D \right| = \left| A \right|$$

where

$$|D| = \prod_{i\_1}^{n} \lambda_i = |A|$$

(d) $\text{rank}(D) = \text{rank}(PAP^{-1})$

$\text{rank}(PA) \leq \text{rank}(A)$

Let $B = PA \Rightarrow A = P^{-1}B$.

$\text{rank}(A = P^{-1}B) \leq \text{rank}(B) = \text{rank}(PA)$

$\Rightarrow \quad \text{rank}(PA) = \text{rank}(A)$

Similarly,

$\text{rank}(BP^{-1}) \leq \text{rank}(B)$

Let $C = BP^{-1} \Rightarrow B = CP$.

$\text{rank}(B = CP) \leq \text{rank}(C) = \text{rank}(BP^{-1})$

$\Rightarrow \quad \text{rank}(BP^{-1}) = \text{rank}(B)$

Thus,

$$\begin{aligned} \text{rank}(PAP^{-1}) &= \text{rank}(BP^{-1}) \\ &= \text{rank}(B) \\ &= \text{rank}(PA) \\ &= \text{rank}(A) \end{aligned}$$

$\Rightarrow \quad \text{rank}(D) = \text{rank}(PAP^{-1}) = \text{rank}(A)$

(Proof taken from Goldberger, 1964, pg. 25 & 29)

2) If D is a **diagonal matrix** its latent roots are its diagonal elements.

Proof:

$$|D - \lambda I| = 0$$

$$|\{d_{ii} - \lambda\}| = 0$$

$$|\{d_{ii} - \lambda\}| = (d_{11} - \lambda)(d_{22} - \lambda) \cdots (d_{nn} - \lambda)$$

$$\Rightarrow \quad \{\lambda_i = d_{ii}\}$$

3) For a **symmetric matrix** A, if

$$(A - \lambda_1 I)\, u_1 = 0 \qquad \text{and} \qquad (A - \lambda_2 I)\, u_2 = 0$$

where

$$\lambda_1 \neq 0, \;\; \lambda_2 \neq 0 \;\; \text{and} \;\; \lambda_1 \neq \lambda_2$$

then

$$u_1' u_2 = 0.$$

Proof:

$$\begin{aligned} u_2'(A - \lambda_1 I)\, u_1 &= u_2' A u_1 - \lambda_1 u_2' u_1 \\ &= 0 \end{aligned}$$

$$\Rightarrow \quad \lambda_1 u_2' u_1 = u_2' A u_1$$

Similarly,

$$\lambda_2 u_1' u_2 = u_1' A u_2$$

But

$$u_2' u_1 = u_1' u_2$$

and

$$\begin{aligned} u_2' A u_1 &= (u_2' A u_1)' \\ &= u_1' A' u_2 \\ &= u_1' A u_2 \end{aligned}$$

$$\Rightarrow \quad \lambda_2 u_1' u_2 = \lambda_1 u_2' u_1$$

or

$$u_1'u_2 = \frac{\lambda_1}{\lambda_2} u_2'u_1$$

But $\lambda_1 \neq \lambda_2 \neq 0$

$\Rightarrow$ $u_1'u_2 = 0$

4) For a **symmetric matrix** A there is an orthogonal matrix P that diagonalizes A. Then, the latent roots of A are the diagonal elements of D = P′AP and the rank (A) = number of diagonal elements of D.

Proof:

(a) $|D - \lambda I| = |P'AP - \lambda I|$

$= |P'(A - \lambda I)P|$

$= |A - \lambda I| \; |P'P|$

$= |A - \lambda I|$

b) $\text{rank}(D) = \text{rank}(P'AP)$

$= \text{rank}(A)$, by 1)(d) above

= number of diagonal elements of D

Latent roots all different

If the **latent roots** of a matrix $A_{n\times n}$ are **all different**, then the corresponding **latent vectors** are **linearly independent**. Furthermore, if the matrix A is **symmetric**, the latent vectors are **mutually orthogonal**.

Multiple latent roots

If various latent roots are the same, then a **linearly independent set of vectors** should be found for **each set of repeated latent roots**.

For a **symmetric** matrix $A_{n\times n}$ with multiple latent roots, a procedure to obtain pairwise orthogonal sets of eigenvectors for each set of repeated latent roots is the following:

(a) Given that the rank $(A - \lambda_i I) = n - m_i$, for i = 1, ..., n, where m_i = multiplicity of λ_i (i.e., the number of time λ_i appears), the equation

$$(A - \lambda_i I)\, u_i \quad = \quad 0$$

has m_i linearly independent non-null (LINN) solutions u_i. Denote one solution by v_{i1}. Now consider solving the system

$$\begin{aligned} (A - \lambda_i I)\, u_i &= 0 \\ v_{i1}'\, u_i &= 0 \end{aligned}$$

simultaneously for u_i. This set has m_{i-1} LINN. Any one of them, e.g., u_{i2} is a latent vector of A and it is orthogonal to u_{i1} and to any latent vector corresponding to $\lambda_{i'} \neq \lambda_i$ because of the orthogonality property of latent vectors from different latent roots (see result [3] above).

If $m_i = 3$, solve for v_{i3} using the set of equations:

$$\begin{aligned} (A - \lambda_i I)\, u_i &= 0 \\ v_{i1}'\, u_i &= 0 \\ v_{i2}'\, u_i &= 0 \end{aligned}$$

This equation system yields m_{i-2} LINN solutions for u_i.

(b) Continue this process until all the m solutions are obtained. The set of eigenvalues, $\{u_i\}$, obtained this way are pairwise orthogonal within **and** across sets of repeated eigenvalues. The matrix formed by these eigenvectors, however, is **not** orthogonal, i.e., $UU' \neq U'U \neq I$. To orthogonalize the matrix U, simply **divide each eigenvector by its length**, i.e., by $[u_i'u_i]^{1/2}$. The resulting matrix is orthogonal because $UU' = I \Rightarrow U'U = I$.

For a **non-symmetric matrix** A there is no guarantee that the set of linearly independent eigenvectors $\{u_i\}$ will be pairwise orthogonal, i.e., $u_i'u_i$ may **not** be zero for some pairs of vectors. Hence, to orthogonalize a matrix of eigenvectors U for **non-symmetric matrix A** use the **Gram-Schmidt process** or other suitable orthogonalization procedure (e.g., see Golub and Van Loan, 1983).

Gram-Schmidt process

Given a **linearly independent** set of vectors $\{x_1, x_2, \ldots, x_n\}$ (i.e., an arbitrary basis), there exists a set of mutually perpendicular (i.e., orthogonal) vectors $\{u_1, u_2, \ldots, u_n\}$ that have the same linear span.

To construct a mutually orthogonal set of vectors $\{u_i\}$ from the set of vectors $(x_i\}$ proceed as follows (for proof see Scheffé, 1959, pg. 382).

(a) Set $u_1 = x_1$.

(b) Compute vectors u_2 to u_n using the formula:

$$u_i = x_i - \sum_{j=1}^{i-1} \frac{(x_i{}'u_j)}{u_j{}'u_j} u_j , \quad 2 \leq j \leq n$$

For example:

$$u_2 = x_2 - \frac{(x_2{}'u_1)}{u_1{}'u_1} u_1$$

$$u_3 = x_3 - \frac{(x_3{'}u_1)}{u_1{'}u_1} u_1 - \frac{(x_3{'}u_2)}{u_2{'}u_2} u_2$$

We can also build a **mutually orthogonal set of vectors** $\{z_i\}$ from the set $\{x_i\}$ which, in addition, have **length equal to 1**. This process, called **normalizing a set of vectors**, can be easily accomplished by dividing each vector u_i by its length, i.e.,

$$Z_i = \frac{u_i}{[u_i{'}u_i]^{1/2}} \quad \text{for } i = 1, \ldots, n$$

where $[uu_i]^{1/2}$ is the length of vector u_i.

Remarks:

(1) $[z_i{'}z_i] = \frac{[u_i{'}u_i]^{1/2}}{[u_i{'}u_i]^{1/2}} = 1$

(2) $(xz_j)\, z_j$ is the projection of x_i on z_j.

(3) $\sum_{j=1}^{i-1} (xz_j)z_j$ is the projection of x_i on the linear span of $z_1, z_2, \ldots, z_{i-1}$ (or the span of $x_1, x_2, \ldots, x_{i-1}$ because both $\{z_i\}$ and $\{x_i\}$ have the same span).

(4) The Gram-Schmidt (GS) procedure has poor numerical properties. There is typically a severe loss of orthogonality among the computed z_i (Golub and Van Loan, 1983, pg. 151). Better methods are:

(a) Modified Gram-Schmidt (Golub & Van Loan, 1983, pg. 152).

(b) Householder Orthogonalization (Golub & Van Loan, 1983, pg. 148).

Computation of eigenvalues and eigenvectors

Consider a matrix $A_{3\times3}$,

$$A = \begin{bmatrix} 13 & -4 & 2 \\ -4 & 13 & -2 \\ 2 & -2 & 10 \end{bmatrix}$$

a) **Eigenvalues**

The characteristic equations of A is:

$$\begin{vmatrix} 13-\lambda & -4 & 2 \\ -4 & 13-\lambda & -2 \\ 2 & -2 & 10-\lambda \end{vmatrix} = 0$$

To find the latent roots we need to expand the above determinant. The **diagonal expansion** (appropriate for (A + D) matrices, D = diagonal, Searle, 1966, pg. 71-73), when the diagonal elements of D are equal, i.e., $\{d_{ii}\} = \{-\lambda\}$. Then,

$$(-\lambda)^n + (-\lambda)^{n-1}\left(a_{11} + a_{22} + \ldots + a_{nn}\right) + (-\lambda)^{n-2}\left\{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} + \ldots + \begin{vmatrix} a_{11} & a_{1n} \\ a_{n1} & a_{nn} \end{vmatrix}\right\} + \ldots + |A| = 0$$

or

$$|A - \lambda I| = \sum_{i=0}^{n} (-\lambda)^{n-i}\,\mathrm{trace}_i\,(A) = 0$$

where

trace_0 (A) = 1

trace_i (A) = sum of the principal minors (i.e., minors with diagonals that coincide with the diagonal of $|A|$) of order i of $|A|$

trace_n (A) = $|A|$

Thus, the diagonal expansion for $A_{3\times3}$ above is:

$$-\lambda^3+\lambda^2(13+13+10)-\lambda\left\{\begin{vmatrix}13 & -4\\ -4 & 13\end{vmatrix}+\begin{vmatrix}13 & 2\\ 2 & 10\end{vmatrix}+\begin{vmatrix}13 & -2\\ -2 & 10\end{vmatrix}\right\}+\begin{vmatrix}13 & -4 & 2\\ -4 & 13 & -2\\ 2 & -2 & 10\end{vmatrix} = 0$$

$$-\lambda^3 + 36\lambda^2 - \lambda\,(153 + 126 + 126) + 1458 = 0$$

$$-\lambda^3 + 36\lambda^2 - 405\,\lambda + 1458 = 0$$

$$\lambda^3 - 36\lambda^2 + 405\,\lambda - 1458 = 0$$

To solve for this third degree polynomial equation:

(1) Find an integer number λ_1 such that:

(a) Is an exact division of the constant 1458

(b) It satisfies the equation $\lambda^3 - 36\lambda^2 + 405\lambda - 1458 = 0$

(2) Divide $\lambda^3 - 36\lambda + 405\lambda - 1458$ by $\lambda - \lambda_1$. The result of this division is a second degree polynomial, which can be solved by the quadratic formula. The quadratic formula is:

$$\lambda = \frac{-b \pm \left[(b^2) - 4(a)(c)\right]^{1/2}}{2\,a}$$

whose solutions satisfy the quadratic equation

$$a\,\lambda^2 + b\,\lambda + c = 0$$

In our example, the integers ±1, ±2, ±3, ±6 and ±9 are exact divisions of 1458, but only 9 satisfies the above cubic equation, i.e., $(9)^3 - 36(9)^2 + 405(9) - 1458 = 0$. Thus, the first root of the cubic polynomial above is 9, because this polynomial can be written as:

$$(\lambda - 9)(\lambda^2 + b\lambda + c) = 0$$

for some integers b and c. To find them, divide the cubic polynomial by $\lambda - 9$, i.e.,

$$\lambda^3 - 36\lambda^2 + 405\lambda - 1458 : \lambda - 9 = \lambda^2 - 27\lambda + 162$$

$$\begin{array}{r} -\lambda^3 \overset{(+)}{\underset{-}{}} 9\lambda^2 \\ \hline 0 - 27\lambda^2 + 405\lambda \end{array}$$

$$\begin{array}{r} \overset{(+)}{\underset{-}{}} 27\lambda^2 \overset{(-)}{\underset{+}{}} 243\lambda \\ \hline 0 + 162\lambda - 1458 \end{array}$$

$$\begin{array}{r} (-)162\lambda \overset{(+)}{\underset{-}{}} 1458 \\ \hline 0 \quad + \quad 0 \end{array}$$

$$\Rightarrow \qquad (\lambda - 9)\,(\lambda^2 - 27\lambda + 162) \;=\; 0$$

$$\Rightarrow \qquad \lambda_1 \;=\; 9$$

Solving for the quadratic equation $\lambda^2 - 27\lambda + 162 = 0$ yields,

$$\Rightarrow \qquad \lambda \;=\; \frac{-27 \pm \left[(27)^2 \pm 4(1)(162)\right]^{1/2}}{2(1)}$$

$$\Rightarrow \qquad \lambda \;=\; \frac{27 \pm 9}{2}$$

$$\Rightarrow \qquad \lambda_2 \;=\; \frac{18}{2} \;=\; 9$$

and

$$\lambda_3 \;=\; \frac{36}{2} \;=\; 18$$

which satisfy

$$(\lambda - 9)(\lambda - 18) \quad = \quad \lambda^2 - 27\lambda + 162 \quad = \quad 0$$

Thus, the set of eigenvalues is:

$$\{\lambda_1, \lambda_2, \lambda_3\} \quad = \quad \{9, 9, 18\}$$

b) **Eigenvectors**

We need to solve the equations:

$$Au_i \;=\; \lambda_i u_i, \;\text{ for } i = 1, 2, 3.$$

Eigenvector for $\lambda_1 = 9$

The set of equations to solve is:

$$(A - \lambda_i I)u_i \;=\; 0$$

$$\begin{bmatrix} (13-9) & -4 & 2 \\ -4 & (13-9) & -2 \\ 2 & -2 & (10-9) \end{bmatrix} \begin{bmatrix} u_{11} \\ u_{21} \\ u_{31} \end{bmatrix} = 0$$

$$\begin{bmatrix} 4 & -4 & 2 \\ -4 & 4 & -2 \\ 2 & -2 & 1 \end{bmatrix} \begin{bmatrix} u_{11} \\ u_{21} \\ u_{31} \end{bmatrix} = 0$$

By theorem 4) of **Linear Equations** (Chapter 2), a vector of solutions to the above system is given by:

$$u_i \;=\; (H - I)z$$

for an arbitrary z, where H = GA. To obtain G, look for dependencies in A, zero out the dependent rows and columns, and invert the resulting non-singular matrix. The first row of A is equal to (−1)row 2 and (2)row 3. The same dependencies exist among columns. Thus,

$$H \;=\; GA \quad = \begin{bmatrix} ¼ & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 4 & -4 & 2 \\ -4 & 4 & -2 \\ 2 & -2 & 1 \end{bmatrix} = \begin{bmatrix} 1 & -1 & ½ \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

and

$$u_1 \quad = \begin{bmatrix} (1-1) & -1 & \tfrac{1}{2} \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix} = \begin{bmatrix} -z_2 + \tfrac{1}{2} z_3 \\ -z_2 \\ -z_3 \end{bmatrix}$$

Arbitrarily choose $z_2 = -1$ and $z_3 = 0$, then, an eigenvector for $\lambda_1 = 9$ is:

$$u_1 \quad = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$$

Eigenvector for $\lambda_2 = 9$

The system of equations to be solved for $\lambda_2 = 9$ is the same as for $\lambda_1 = 9$. Here we choose a different set of values for z_2 and z_3 such that u_2 and u_1 are linearly independent, i.e., let

$z_2 = 1$ and $z_3 = 4$.

Thus, the second eigenvector is:

$$u_2 \quad = \begin{bmatrix} 1 \\ -1 \\ -4 \end{bmatrix}$$

Note: because A is symmetric, $u_2'u_1 = 0$.

Eigenvector for $\lambda_3 = 18$

The set of equations for u_3 is:

$$\begin{bmatrix} -5 & -4 & 2 \\ -4 & -5 & -2 \\ 2 & -2 & -8 \end{bmatrix} \begin{bmatrix} u_{13} \\ u_{23} \\ u_{33} \end{bmatrix} = 0$$

Here, (2) [column 2 – column 1] = column 3. Thus, delete the first row and column, replace them with zeroes and compute G, i.e.,

$$G = \begin{bmatrix} 0 & 0 \quad 0 \\ 0 & \begin{bmatrix} -5 & -2 \\ -2 & -8 \end{bmatrix}^{-1} \\ 0 & \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & -\frac{2}{9} & \frac{1}{18} \\ 0 & \frac{1}{18} & -\frac{5}{36} \end{bmatrix}$$

Thus, u_3 is:

$$(GA - I)\,z = \left\{ \begin{bmatrix} 0 & 0 & 0 \\ 0 & -\frac{2}{9} & \frac{1}{18} \\ 0 & \frac{1}{18} & -\frac{5}{36} \end{bmatrix} \begin{bmatrix} -5 & -4 & 2 \\ -4 & -5 & -2 \\ 2 & -2 & -8 \end{bmatrix} - \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \right\} \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix}$$

$$(GA - I)\,z = \begin{bmatrix} (0-1) & 0 & 0 \\ 1 & (1-1) & 0 \\ -\frac{1}{2} & 0 & (1-1) \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix} = \begin{bmatrix} -z_1 \\ z_1 \\ -\frac{1}{2} z_1 \end{bmatrix}$$

Arbitrarily set $z_1 = -2$. The resulting eigenvector is:

$$u_3 = \begin{bmatrix} 2 \\ -2 \\ 1 \end{bmatrix}$$

The matrix U is:

$$U = \begin{bmatrix} 1 & 1 & 2 \\ 1 & -1 & -2 \\ 0 & -4 & 1 \end{bmatrix}$$

We can verify that,

$$D = U^{-1}AU = \begin{bmatrix} 1 & 1 & 2 \\ 1 & -1 & -2 \\ 0 & -4 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 13 & -4 & 2 \\ -4 & 13 & -2 \\ 2 & -2 & 10 \end{bmatrix} \begin{bmatrix} 1 & 1 & 2 \\ 1 & 1 & 2 \\ 0 & -4 & 1 \end{bmatrix} = \begin{bmatrix} 9 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & 18 \end{bmatrix}$$

and

$$A = UDU^{-1} = \begin{bmatrix} 1 & 1 & 2 \\ 1 & -1 & -2 \\ 0 & -4 & 1 \end{bmatrix} \begin{bmatrix} 9 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & 18 \end{bmatrix} \begin{bmatrix} 1 & 1 & 2 \\ 1 & -1 & -2 \\ 0 & -4 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} 13 & -4 & 2 \\ -4 & 13 & -2 \\ 2 & -2 & 10 \end{bmatrix}$$

The matrix D is the canonical form of A under similarity

Because A is symmetric in this example, an orthogonal matrix E can be obtained from matrix U by dividing the elements of each vector u_i by its length, i.e., by $[u_i'u_i]^{1/2}$.

The lengths of the eigenvectors of A, i.e., the $[u_i'u_i]^{1/2}$ for i = 1, 2, 3, are:

$$[u_1' u_1]^{1/2} = \left[(1)^2 + (1)^2 + (0)^2\right]^{1/2} = \sqrt{2}$$

$$[u_2' u_2]^{1/2} = \left[(1)^2 + (-1)^2 + (-4)^2\right]^{1/2} = \sqrt{18}$$

$$[u_3' u_3]^{1/2} = \left[(2)^2 + (-2)^2 + (1)^2\right]^{1/2} = 3$$

Thus, the orthogonal matrix of eigenvectors, E, is:

$$E = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{18}} & \frac{2}{3} \\ \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{18}} & \frac{-2}{3} \\ 0 & \frac{-4}{\sqrt{18}} & \frac{1}{3} \end{bmatrix}$$

We can verify that EE′ = E′E = I, and

$$A = EDE' = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{18}} & \frac{2}{3} \\ \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{18}} & \frac{-2}{3} \\ 0 & \frac{-4}{\sqrt{18}} & \frac{1}{3} \end{bmatrix} \begin{bmatrix} 9 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & 18 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{18}} & \frac{-1}{\sqrt{18}} & \frac{-4}{\sqrt{18}} \\ \frac{2}{3} & \frac{-2}{3} & \frac{1}{3} \end{bmatrix} = \begin{bmatrix} 13 & -4 & 2 \\ -4 & 13 & -2 \\ 2 & -2 & 10 \end{bmatrix}$$

and the **canonical form of A under orthogonal similarity** (only for symmetric A), i.e., D, is:

$$D = E'AE = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{18}} & \frac{-1}{\sqrt{18}} & \frac{-4}{\sqrt{18}} \\ \frac{2}{3} & \frac{-2}{3} & \frac{1}{3} \end{bmatrix} \begin{bmatrix} 13 & -4 & 2 \\ -4 & 13 & -2 \\ 2 & -2 & 10 \end{bmatrix} \begin{bmatrix} \frac{1}{2} & \frac{1}{\sqrt{18}} & \frac{2}{3} \\ \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{18}} & \frac{-2}{3} \\ 0 & \frac{-4}{\sqrt{18}} & \frac{1}{3} \end{bmatrix} = \begin{bmatrix} 9 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & 18 \end{bmatrix}$$

Numerical methods to compute eigenvalues and eigenvectors

Eigenvalues: Schur decomposition (Golub and Van Loan, 1983, pg. 192).

Eigenvectors: (Golub and Van Loan, 1983, pg. 238).

Note: To compute eigenvectors when the matrix [A – λI] is full rank:

(a) Substitute zeroes for the elements of the i^{th} dependent row and column of matrix A, and compute G,

(b) Compute $u_i = [GA - I]z$, and

(c) Assign arbitrary values to the necessary z_i's and compute u_i.

Iterative procedures to find eigenvectors and eigenvalues

Dominant eigenvalue: if a symmetric matrix $A_{n\times n}$ of rank r has eigenvalues $\{\lambda_1\ \lambda_2\ \ldots\ \lambda_n\}$, where

$$|\lambda_1| > |\lambda_2| \geq \ldots \geq |\lambda_{n-1}| \geq |\lambda_n|$$

then, **λ_1 is called the dominant eigenvalue of A**.

1) **Power method (or iterative method) to approximate the dominant eigenvector (v_1) and the dominate eigenvalue (λ_1) of a diagonalizable matrix A.**

(a) Arbitrarily select a vector, e.g., x_o, to be the first approximation to the dominant eigenvector v_1.

(b) Compute $\tilde{v}_{1(1)} = A\,\tilde{v}_{1(0)}$ and scale it down, i.e., divide each element of $\tilde{v}_{1(1)}$ by its largest element.

(c) Compute $\tilde{v}_{1(2)} = A\,\tilde{v}_{1(1)} (= A^2\ \tilde{v}_{1(0)})$ and scale it down.

(d) Compute the first approximation to the dominant eigenvalue as follows:

$$\lambda_1 \approx \frac{\tilde{v}_{1(1)}'\tilde{v}_{1(2)}}{\tilde{v}_{1(1)}'\tilde{v}_{1(1)}} = \tilde{\lambda}_{1(1)}$$

This computation is based on the equality

$$\frac{v_1' A v_1}{v_1' v_1} = \frac{v_1' \lambda v_1}{v_1' v_1} = \frac{\lambda v_1' v_1}{v_1' v_1} = \lambda$$

Thus,

$$\lambda_1 \approx \frac{\tilde{v}_{1(1)}{}' A \tilde{v}_{1(1)}}{\tilde{v}_{1(1)} \tilde{v}_{1(1)}} = \frac{\tilde{v}_{1(1)}{}' \lambda_1 \tilde{v}_{1(1)}}{\tilde{v}_{1(1)}{}' \tilde{v}_{1(1)}} = \lambda_1 v_{1(1)}{}' \tilde{v}_{1(1)} = \tilde{\lambda}_{1(1)}$$

(e) Repeat the computations

$$\tilde{\lambda}_{1(i)} = \frac{\tilde{v}_{1(i-1)}{}' \tilde{v}_{1(i)}}{\tilde{v}_{1(i-1)}{}' \tilde{v}_{1(i-1)}}$$

until the convergence criterion is met. Such convergence criterion could be:

(i) The estimated relative error:

$$\left| \frac{\tilde{\lambda}_{1(i)} - \tilde{\lambda}_{1(i-1)}}{\tilde{\lambda}_{1(i)}} \right| < E$$

i.e., the absolute value of the difference between the estimates of λ_1 in the i^{th} and $(i-1)^{th}$ iterations relative to the estimate of λ_1 in the i^{th} iteration is less than a chosen number E. For instance, E = 0.00001.

(ii) The estimated percentage error:

$$\left| \frac{\tilde{\lambda}_{1(i)} - \tilde{\lambda}_{1(i-1)}}{\tilde{\lambda}_{1(i)}} \right| * 100 \leq 100\,E$$

(f) At convergence after p iterations, the vector x_p is a good approximation to the dominant eigenvector v_1, and the scalar $\lambda_{1(p)}$ is a good approximation to the dominant eigenvalue λ_1.

Proof:

Let A be an n×n diagonalizable matrix of rank r. If A is diagonalizable, A has a set of n independent eigenvectors $\{v_1, v_2, \ldots, v_n\}$. To see this, note that AV = VD, where D = diag $\{\lambda_i\}$ and V = $[v_1\ v_2\ \ldots\ v_n]$ (for proof see Searle, 1966, pg. 168, or Anton, 1981, pg. 269).

Let $\{\lambda_1, \lambda_2, \ldots, \lambda_n\}$ be the set of eigenvalues of A and **assume** that

$$|\lambda_1| > |\lambda_2| \geq \ldots \geq |\lambda_n|.$$

By theorem 9(a) of Anton (1981, pg. 155), the set of linearly independent eigenvectors of matrix A, $\{v_1, v_2, \ldots, v_n\}$ form a basis for R^n. Thus, an arbitrary vector x_0 in R^n can be expressed as a linear combination of $v_1, v_2, \ldots, v_n$, i.e.,

$$x_0 = m_1v_1 + m_2v_2 + \ldots + m_nv_n.$$

But

$$\begin{aligned} Ax_0 &= m_1Av_1 + m_2Av_2 + \ldots + m_nAv_n \\ &= m_1\lambda_1v_1 + m_2\lambda_2v_2 + \ldots + m_n\lambda_nv_n \end{aligned}$$

and

$$\begin{aligned} A(Ax_0) &= m_1\lambda_1Av_1 + m_2\lambda_2Av_2 + \ldots + m_n\lambda_nAv_n \\ A^2x_0 &= m_1\lambda v_1 + m_2\lambda v_2 + \ldots + m_n\lambda v_n \\ &\vdots \\ A^px_0 &= m_1\lambda v_1 + m_2\lambda v_2 + \ldots + m_n\lambda v_n. \end{aligned}$$

Because $\lambda_1 \neq 0$,

$$A^p x_0 = \lambda_1^p \left[m_1 v_1 + m_2 \left(\frac{\lambda_2}{\lambda_1}\right)^p v_2 + \ldots + m_n \left(\frac{\lambda_2}{\lambda_1}\right)^p v_n \right]$$

Also, as $p \to \infty$, $\left\{ \left(\frac{\lambda_2}{\lambda_1}\right)^p, \ldots, \left(\frac{\lambda_n}{\lambda_1}\right)^p \right\} \to 0$ because $|\lambda_1| > |\lambda_2| \geq \ldots \geq |\lambda_n|$.

Thus,

$$A^px_0 \approx \lambda m_1v_1 \text{ as } p \to \infty.$$

If $m_1 \neq 0$, then $(\lambda m_1)v_1$ is a multiple of the dominant eigenvector v_1. Thus, λm_1v_1 is also a dominant eigenvector. Therefore, as p increases, $A^p x_0$ becomes an increasingly better estimate of **a** dominant eigenvector.

2) **Deflation method to approximate nondominant eigenvectors and eigenvalues**

The deflation method is based on the following theorem (Anton, 1981, pg. 336; Searle, 1966, pg. 187).

Theorem: Let A be a **symmetric** n×n matrix of rank r with eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$. If v_1 is **an** eigenvector of A corresponding to λ_1, and $[v_1' v_1]^{1/2} = 1$, i.e., v_1 has unit length, then:

(i) The matrix $A_1 = A - \lambda_1 v_1 v_1'$ has eigenvalues $0, \lambda_2, \lambda_3, \ldots, \lambda_n$.

(ii) The eigenvectors of A_1 corresponding to $\lambda_2, \lambda_3, \ldots, \lambda_n$ are also eigenvectors for $\lambda_2, \lambda_3, \ldots, \lambda_n$ in A.

Proof: see Searle, 1966, pg. 187.

The deflection method also rests on the assumption that

$$|\lambda_1| > |\lambda_2| > |\lambda_3| > \ldots \geq |\lambda_n|.$$

The procedure of deflation is as follows:

(a) Compute A_1 using the expression:

$$A_1 = A - \tilde{\lambda}_{1(p)} \tilde{z}_{1(p)} \tilde{z}_{1(p)}$$

where

$$\tilde{z}_{1(p)} = \frac{\tilde{v}_{1(p)}}{\left[\tilde{v}_{1(p)}' \tilde{v}_{1(p)}\right]^{1/2}}$$

(b) Estimate the dominant eigenvector (v_2) and the dominant eigenvalue (λ_2) of A_1 using the power method, i.e., obtain $\tilde{v}_{2(p)}$ and $\tilde{\lambda}_{2(p)}$.

(c) Repeat steps (a) and (b) for A_i, i = 3, ... , last dominant eigenvector and eigenvalue.

Warning: because λ_i and v_i are estimated iteratively, rounding errors accumulate quickly. Thus, the deflation method is recommended to estimate only two or three eigenvectors and eigenvalues (Anton, 1981, pg. 338).

References

Anton, H. 1981. Elementary Linear Algebra. John Wiley and Sons, NY.

Goldberger, A. S. 1964. Econometric Theory. John Wiley and Sons, NY.

Golub, G. H. and C. F. Van Loan. 1983. Matrix Computations. The John Hopkins University Press, MD.

Searle, S. R. 1966. Matrix Algebra for the Biological Sciences (including Applications in Statistics). John Wiley and Sons, NY.

Scheffé, H. 1959. The analysis of variance. John Wiley and Sons, Inc., NY.