MTS 204: Linear Algebra II

Department of Mathematical Sciences
Federal University of Technology, Akure

1. Cayley-Hamilton Theorem

Before this theorem, we shall consider two other relevant theorems.

Theorem 1.1. *If $\lambda_1, \lambda_2, ..., \lambda_n$ are the characteristic roots, distinct or not, of a matrix A of order n, and if $g(A)$ is any polynomial function of A, then the characteristic roots of $g(A)$ are $g(\lambda_1), g(\lambda_2), ..., g(\lambda_n)$.*

Proof. Recall that $\det(A - I\lambda_n) = (\lambda_1 - \lambda)(\lambda_2 - \lambda)...(\lambda_n - \lambda)$. For any polynomial function of $g(A)$, it is to be shown that $\det(g(A) - I\lambda_n) = (g(\lambda_1) - \lambda)(g(\lambda_2) - \lambda)...(g(\lambda_n) - \lambda)$.

To show this, suppose that $g(x)$ is of degree r in x and that for a fixed value of λ, the roots of $g(x) - \lambda = 0$ are $x_1, x_2, ..., x_r$. Then we have $g(x) - \lambda = \alpha(x - x_1)(x - x_2)...(x - x_r)$, where α is the coefficient of x^r in $g(x)$.

Hence, $g(x) - \lambda I_n = \alpha(A - x_1 I_n)(A - x_2 I_n)...(A - x_r I_n)$, so that, if $\phi(\lambda)$ is the characteristic polynomial of A,

$$
\begin{aligned}
&\det(g(A) - \lambda I_n) \\
&= \alpha^n \det(A - x_1 I_n) \det(A - x_2 I_n)... \det(A - x_r I_n) \\
&= \alpha^n \phi(x_1)\phi(x_2)...\phi(x_r) \\
&= \alpha^n (\lambda_1 - x_1)(\lambda_2 - x_1)...(\lambda_n - x_1) \bullet (\lambda_1 - x_2)(\lambda_2 - x_2)...(\lambda_n - x_2) \bullet \cdots (\lambda_1 - x_r)(\lambda_2 - x_r)...(\lambda_n - x_r).
\end{aligned}
$$

By rearranging the orders of the factors, we obtain

$$
\begin{aligned}
&\det(g(A) - \lambda I_n) \\
&= \alpha(\lambda_1 - x_1)(\lambda_2 - x_1)...(\lambda_n - x_1) \bullet \alpha(\lambda_1 - x_2)(\lambda_2 - x_2)...(\lambda_n - x_2) \bullet \cdots \alpha(\lambda_1 - x_r)(\lambda_2 - x_r)...(\lambda_n - x_r) \\
&= (g(\lambda_1) - \lambda)(g(\lambda_2) - \lambda)...(g(\lambda_n) - \lambda).
\end{aligned}
$$

□

Note: If A^{-1} is non-singular and no characteristic root μ of A^{-1} is zero, then we have

$$\det(g(A^{-1} - \mu I)) = -\mu^n \det A^{-1} \det(A - \mu^{-1} I).$$

For this expression to vanish, the factor $\det(A - \mu^{-1} I)$ must be zero. The roots of $\det(A - \mu^{-1} I) = 0$ in μ^{-1} are $\lambda_1, \lambda_2, ..., \lambda_n$. Hence, the roots of μ of A^{-1} must be $\lambda_1^{-1}, \lambda_2^{-1}, ..., \lambda_n^{-1}$. This result can be stated generally as a theorem.

Theorem 1.2. *If the characteristic roots of A are $\lambda_1, \lambda_2, ..., \lambda_n$, then those of A^p, where p is any integer, are $\lambda_1^p, \lambda_2^p, ..., \lambda_n^p$, provided A^{-1} exists.*

Note: Suppose that A is a real symmetric matrix with characteristic roots $\lambda_1, \lambda_2, ..., \lambda_n$ and with characteristic equation $\phi(\lambda) = 0$. Then $\phi(A)$ is also a real symmetric matrix and by Theorem 1.1, its characteristic roots are $\phi(\lambda_1), \phi(\lambda_2), ..., \phi(\lambda_n)$ all of which are zero. Hence, $\phi(A) = 0$.

Theorem 1.3 (Cayley-Hamilton)**.** *If A is a square matrix of order n, then A satisfies its characteristic equation* $\det(A-\lambda I) = 0$, *i.e.* $\det(A-\lambda I) = A^n + \alpha_1 A^{n-1} + \alpha_2 A^{n-2} + ... + \alpha_n I = 0$, *where $\alpha_i' s$, $i = 1, 2, ..., n$ are constant.*

Example: Given the system of equations

$$\begin{aligned} 2x_1 + x_2 &= \lambda x_1 \\ x_1 + 2x_2 &= \lambda x_2, \end{aligned}$$

then

$$A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}.$$

The characteristic equation is $\det(A - \lambda I) = \lambda^2 - 4\lambda + 3 = 0$. The theorem states that A satisfies its characteristic equation. So, we set $\lambda = A$ to have $\det(A - AI) = A^2 - 4A + 3I = 0$, i.e.

$$\det(A - \lambda I) = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}^2 - 4\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} + 3\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 5 & 4 \\ 4 & 5 \end{bmatrix} - \begin{bmatrix} 8 & 4 \\ 4 & 8 \end{bmatrix} + \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

Example: Given the system of equations

1. $$\begin{aligned} 3x_1 + x_2 &= \lambda x_1 \\ 2x_1 + 2x_2 &= \lambda x_2, \end{aligned}$$

2. $$\begin{aligned} x_1 + 2x_2 &= \lambda x_1 \\ 2x_1 - x_2 &= \lambda x_2 \\ 4x_3 &= \lambda x_3, \end{aligned}$$

show that $\phi(A) = 0$, where A is the square matrix of each of the systems of equations.

1.1 The Nth Power of a Matrix

The Cayley-Hamilton theorem allows for a representation of high powers of a matrix. Suppose $\phi(A) = 0$ is expressed in the polynomial equation form

$$\phi(A) = A^n + \alpha_1 A^{n-1} + \alpha_2 A^{n-2} + ... + \alpha_n I = 0. \tag{1.1}$$

Having computed $A^2, A^3, ..., A^{n-1}$, then

$$A^n = -\left(\alpha_n I_n + \alpha_{n-1}A + \alpha_{n-2}A^2 + ... + \alpha_2 A^{n-2} + \alpha_1 A^{n-1}\right). \tag{1.2}$$

Multiplying (1.1) by A and substituting the expression for A^n, we have

$$A^{n+1} = \alpha_{n-2}\alpha_1 I_n + (\alpha_{n-2}\alpha_2 - \alpha_1)A + (\alpha_{n-2}\alpha_3 - \alpha_2)A^2 + ... + (\alpha_{n-2}^2 - \alpha_{n-3})A^{n-1}. \tag{1.3}$$

This means that any positive integral power of A is expressible as a linear combination of $I, A^2, A^3, ..., A^{n-1}$.

Example: Compute A^4 for $A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$ using the Cayley-Hamilton theory.

The characteristic equation is $det(A - \lambda I) = \lambda^2 - 2\lambda - 3 = 0$. Going by the Cayley-Hamilton theory, we have $A^2 - 2A - 3I$ such that

$$A^2 = 2A + 3I. \tag{1.4}$$

Then we have

$$A^2 = 2A + 3I = 2\begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} + 3\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 5 & 4 \\ 4 & 5 \end{bmatrix}.$$

Multiplying (1.4) by A, we have

$$A^3 = 2A^2 + 3A \tag{1.5}$$

so that

$$A^3 = 2A^2 + 3A = 2\begin{bmatrix} 5 & 4 \\ 4 & 5 \end{bmatrix} + 3\begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} 13 & 14 \\ 14 & 13 \end{bmatrix}.$$

Multiplying (1.5) by A, we have

$$A^4 = 2A^3 + 3A^2 \tag{1.6}$$

so that

$$A^4 = 2A^3 + 3A^2 = 2\begin{bmatrix} 13 & 14 \\ 14 & 13 \end{bmatrix} + 3\begin{bmatrix} 5 & 4 \\ 4 & 5 \end{bmatrix} = \begin{bmatrix} 41 & 40 \\ 40 & 41 \end{bmatrix}.$$

1.2 The Inverse of a Matrix

If A is non-singular, multiplication of (1.1) by A^{-1} will yield

$$A^{-1} = -\frac{1}{a_n}\left(\alpha_{n-1}I_n + \alpha_{n-2}A + \alpha_{n-3}A^2 + ... + \alpha_1 A^{n-2} + A^{n-1}\right). \tag{1.7}$$

Multiplying (1.7) by A^{-1} and substituting for A^{-1}, we have

$$A^{-2} = -\frac{1}{\alpha_n^2}\left[\left(\alpha_{n-1}^2 - \alpha_{n-2}\right)I_n + \left(\alpha_{n-1}\alpha_{n-2} - \alpha_{n-3}\right)A + ... + (\alpha_{n+1}\alpha_2 - 1)A^{n-2} + \alpha_1 A^{n-1}\right]. \tag{1.8}$$

By continuing this process, all negative powers of A may be express as a linear combination $I, A^2, A^3, ..., A^{n-1}$.

Example: Compute A^{-3} for $A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$ using the Cayley-Hamilton theory.

The characteristic equation is $\det(A - \lambda I) = \lambda^2 - 2\lambda - 3 = 0$. Going by the Cayley-Hamilton theory, we have $A^2 - 2A - 3I$ such that $A^2 = 2A + 3I$.

Multiplying (1.4) by A^{-1}, we obtain $A - 2I - 3A^{-1}$ and then

$$A^{-1} = \frac{1}{3}(A - 2I). \tag{1.9}$$

Now,

$$A^{-1} = \frac{1}{3}\left(\begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} - 2\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right) = \frac{1}{3}\begin{bmatrix} -1 & 2 \\ 2 & -1 \end{bmatrix} = \begin{bmatrix} -\frac{1}{3} & \frac{2}{3} \\ \frac{2}{3} & -\frac{1}{3} \end{bmatrix}.$$

Multiplying (1.9) by A^{-1}, we obtain

$$A^{-2} = \frac{1}{3}(I - 2A^{-1}) \tag{1.10}$$

so that

$$A^{-2} = \frac{1}{3}\left(\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - 2\begin{bmatrix} -\frac{1}{3} & \frac{2}{3} \\ \frac{2}{3} & -\frac{1}{3} \end{bmatrix}\right) = \frac{1}{3}\begin{bmatrix} \frac{5}{3} & -\frac{4}{3} \\ -\frac{4}{3} & \frac{5}{3} \end{bmatrix} = \begin{bmatrix} \frac{5}{9} & -\frac{4}{9} \\ -\frac{4}{9} & \frac{5}{9} \end{bmatrix}.$$

When we multiply (1.10) by A^{-1}, we get

$$A^{-3} = \frac{1}{3}\left(A^{-1} - 2A^{-2}\right) \tag{1.11}$$

and

$$A^{-3} = \frac{1}{3}\left(\begin{bmatrix} -\frac{1}{3} & \frac{2}{3} \\ \frac{2}{3} & -\frac{1}{3} \end{bmatrix} - 2\begin{bmatrix} \frac{5}{9} & -\frac{4}{9} \\ -\frac{4}{9} & \frac{5}{9} \end{bmatrix}\right) = \frac{1}{3}\begin{bmatrix} -\frac{13}{9} & \frac{14}{9} \\ \frac{14}{9} & -\frac{13}{9} \end{bmatrix} = \begin{bmatrix} -\frac{13}{27} & \frac{14}{27} \\ \frac{14}{27} & -\frac{13}{27} \end{bmatrix}.$$

2. Diagonalisation of a Real Symmetric Matrix

Recall that

- A symmetric matrix A is such that $A^T = A$.
- An orthogonal matrix A is such that $A^{-1} = A^T$.
- The orthogonality of a matrix implies that it is symmetric but a symmetric matrix is not necessarily orthogonal.

2.1 Diagonalisation Form of a Real Symmetric Matrix

Let $D[\lambda_1, \lambda_2, ..., \lambda_n]$ be a diagonal matrix with diagonal elements $\lambda_1, \lambda_2, ..., \lambda_n$. The matrix $D[\lambda_1, \lambda_2, ..., \lambda_n]$ is called the diagonal form of A and the equation $X_2 = D[\lambda_1, \lambda_2, ..., \lambda_n]X_1$ represents the corresponding operator in the diagonal form.

Theorem 2.1 (Existence)**.** *If A is a real symmetric matrix and $\lambda_1, \lambda_2, ..., \lambda_n$ are its characteristic roots, then there exists an orthogonal matrix U such that $U^T AU = D[\lambda_1, \lambda_2, ..., \lambda_n]$.*

Definition 2.1. A square matrix is called diagonalizable if there is an invertible matrix P such that

$$P^{-1}AP = B \tag{2.1}$$

where B is a diagonal matrix. The matrix P is said to diagonalize A. If A is an $n \times n$, then the following are equivalent

(i) A is diagonalizable, and

(ii) A has linearly independent eigenvectors.

Proof. Suppose that (i) $\Rightarrow$ (ii). Since A is assumed to be diagonalizable, then there exists an invertible matrix

$$P = \begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ p_{n1} & p_{n2} & \cdots & p_{nn} \end{pmatrix}$$

such that $P^{-1}AP = D$ is diagonal, where

$$D = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}$$

and $AP = PD$, i.e.

$$AP = \begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ p_{n1} & p_{n2} & \cdots & p_{nn} \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix} = \begin{pmatrix} \lambda_1 p_{11} & \lambda_2 p_{12} & \cdots & \lambda_n p_{1n} \\ \lambda_1 p_{21} & \lambda_2 p_{22} & \cdots & \lambda_n p_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ \lambda_1 p_{n1} & \lambda_2 p_{n2} & \cdots & \lambda_n p_{nn} \end{pmatrix} \tag{2.2}$$

Let $p_1, p_2, ..., p_n$ denote column vectors of P, then from the result above, the successive columns of AP are $\lambda_1 p_1, \lambda_2 p_2, ..., \lambda_n p_n$. However, the successive columns of AP are $AP_1, AP_2, ..., AP_n$. Thus, we must have

$$AP_1 = \lambda_1 p_1, AP_2 = \lambda_2 p_2, ..., AP_n = \lambda_n p_n. \tag{2.3}$$

The invertibility of P implies that its column vectors are all non-zero and $\lambda_1, \lambda_2, ..., \lambda_n$ are eigenvalues of A while the corresponding eigenvectors are $P_1, P_2, ..., P_n$. These eigenvectors are linearly independent since P is invertible. Thus A has n linearly independent eienvectors.

On the other hand, suppose that (ii) $\Rightarrow$ (i), i.e. assume that A has n linearly independent eigenvectors $P_1, P_2, ..., P_n$ with corresponding eigenvalues $\lambda_1, \lambda_2, ..., \lambda_n$. Let

$$P = \begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ p_{n1} & p_{n2} & \cdots & p_{nn} \end{pmatrix}$$

be the matrix whose column vectors of the product AP are $AP_1, AP_2, ..., AP_n$ but $AP_1 = \lambda_1 p_1, AP_2 = \lambda_2 p_2, ..., AP_n = \lambda_n p_n$ so that

$$AP = \begin{pmatrix} \lambda_1 p_{11} & \lambda_2 p_{12} & \cdots & \lambda_n p_{1n} \\ \lambda_1 p_{21} & \lambda_2 p_{22} & \cdots & \lambda_n p_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ \lambda_1 p_{n1} & \lambda_2 p_{n2} & \cdots & \lambda_n p_{nn} \end{pmatrix} = \begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ p_{n1} & p_{n2} & \cdots & p_{nn} \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix} = PD \tag{2.4}$$

where D is the diagonal matrix with eigenvalues $\lambda_1, \lambda_2, ..., \lambda_n$ on the main diagonal but the column vectors of P are linearly independent. This implies that P is invertible.

Thus, $P^{-1}AP = D$ and A is diagonalizable. $\square$

2.2 Steps for Diagonalizing a Matrix

Step 1 : Find n linearly independent eigenvectors of A, say $P_1, P_2, ..., P_n$.

Step 2 : Form matrix P having $P_1, P_2, ..., P_n$ as its column vectors.

Step 3 : The matrix $P^{-1}AP$ will then be a diagonal matrix D with $P_1, P_2, ..., P_n$ as its successive diagonal entries where λ_i's are the eigenvalues corresponding to the P_i's, $i = 1, 2, ..., n$.

Example: Determine the matrix which diagonalizes $A = \begin{bmatrix} 0 & 0 & -2 \\ 1 & 2 & 1 \\ 1 & 0 & 3 \end{bmatrix}$.

To obtain the eigenvalues, we have

$$\det(A-\lambda I) = \begin{vmatrix} -\lambda & 0 & -2 \\ 1 & 2-\lambda & 1 \\ 1 & 0 & 3-\lambda \end{vmatrix} = -\lambda \begin{vmatrix} 2-\lambda & 1 \\ 0 & 3-\lambda \end{vmatrix} - 0 \begin{vmatrix} 1 & 1 \\ 1 & 3-\lambda \end{vmatrix} - 2 \begin{vmatrix} 1 & 2-\lambda \\ 1 & 0 \end{vmatrix}$$

$$\begin{aligned}
&= -\lambda[(2-\lambda)(3-\lambda) - 0(1)] - 0[1(3-\lambda) - 1(1)] - 2[1(0) - 1(2-\lambda)] \\
&= -\lambda(6 - 5\lambda + \lambda^2) - 0(2-\lambda) - 2(-2+\lambda) \\
&= -6\lambda + 5\lambda^2 - \lambda^3 - 0 + 4 - 2\lambda \\
&= -\lambda^3 + 5\lambda^2 - 8\lambda + 4
\end{aligned}$$

so that the characteristic equation can be given as

$$\phi(\lambda) = \lambda^3 - 5\lambda^2 + 8\lambda - 4 = (x-1)(x-2)^2 = 0$$

and we obtain eigenvalues $\lambda_1 = 1$, $\lambda_2 = 2$ and $\lambda_3 = 2$.

To obtain the corresponding eigenvectors, we have

$$(A - \lambda I)X = 0$$

such that

$$\begin{bmatrix} -\lambda & 0 & -2 \\ 1 & 2-\lambda & 1 \\ 1 & 0 & 3-\lambda \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

For $\lambda = 1$, we have

$$\begin{bmatrix} -1 & 0 & -2 \\ 1 & 1 & 1 \\ 1 & 0 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

so that

$$\begin{aligned}
-x_1 - 2x_3 &= 0 \\
x_1 + x_2 + x_3 &= 0 \\
x_1 + 2x_3 &= 0.
\end{aligned}$$

If $x_3 = 1$, then $x_1 = -2$, $x_2 = 1$ and we have $P_1 = \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}$.

For $\lambda = 2$ (twice), we have

$$\begin{bmatrix} -2 & 0 & -2 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

so that

$$-2x_1 - 2x_3 = 0$$
$$x_1 + x_3 = 0$$
$$x_1 + x_3 = 0.$$

If $x_3 = 1$, then $x_1 = -1$. If we take $x_2 = 0$, we have $P_2 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$.

If $x_3 = 0$, then $x_1 = 0$. If we take $x_2 = 1$, we have $P_3 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$.

$\therefore$ the matrix A is diagonalizable by the matrix

$$P = \begin{bmatrix} P_1 & P_2 & P_3 \end{bmatrix} = \begin{bmatrix} -2 & -1 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}.$$

To verify this, we need to show that $P^{-1}AP = D$.
To find P^{-1}, we augment the matrix P and proceed as follows:

$$\begin{bmatrix} -2 & -1 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{2.5}$$

We divide the first row ($R1$) of (2.5) through by -2 so that

$$\begin{bmatrix} 1 & \frac{1}{2} & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} -\frac{1}{2} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

We subtract elements in rows 2 and 3 ($R2$ and $R3$) from corresponding elements in $R1$ and put the results in $R2$ and $R3$ respectively such that

$$\begin{bmatrix} 1 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & -1 \\ 0 & -\frac{1}{2} & 0 \end{bmatrix} \begin{bmatrix} -\frac{1}{2} & 0 & 0 \\ -\frac{1}{2} & -1 & 0 \\ -\frac{1}{2} & 0 & -1 \end{bmatrix}.$$

We add $R2$ and $R3$ to get

$$\begin{bmatrix} 1 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & -1 \\ 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} -\frac{1}{2} & 0 & 0 \\ -\frac{1}{2} & -1 & 0 \\ -1 & -1 & -1 \end{bmatrix}.$$

$R3 - R2$ gives

$$\begin{bmatrix} 1 & \frac{1}{2} & 0 \\ 0 & -\frac{1}{2} & 0 \\ 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} -\frac{1}{2} & 0 & 0 \\ -\frac{1}{2} & 0 & -1 \\ -1 & -1 & -1 \end{bmatrix}.$$

$R1 + R2$ gives

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & -\frac{1}{2} & 0 \\ 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} -1 & 0 & -1 \\ -\frac{1}{2} & 0 & -1 \\ -1 & -1 & -1 \end{bmatrix}.$$

Finally, $-2R_2$ and $-R3$ results to

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} -1 & 0 & -1 \\ 1 & 0 & 2 \\ 1 & 1 & 1 \end{bmatrix}.$$

$$\therefore P^{-1} = \begin{bmatrix} -1 & 0 & -1 \\ 1 & 0 & 2 \\ 1 & 1 & 1 \end{bmatrix} \tag{2.6}$$

Now, we have

$$\begin{aligned} P^{-1}AP &= \begin{bmatrix} -1 & 0 & -1 \\ 1 & 0 & 2 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & -2 \\ 1 & 2 & 1 \\ 1 & 0 & 3 \end{bmatrix} \begin{bmatrix} -2 & -1 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} \\ &= \begin{bmatrix} -1 & 0 & -1 \\ 1 & 0 & 2 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} -2 & -2 & 0 \\ 1 & 0 & 2 \\ 1 & 2 & 0 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix} \\ &= D. \end{aligned}$$

$\therefore P^{-1}AP = D$ is verified.

Remark: There is no preferred order for the columns of P since the ith diagonal entry of $P^{-1}AP$ is an eigenvalue for the ith column vector of P. Changing the order of the columns of P just changes the order of the eigenvalues on the diagonal of $P^{-1}AP$.

Exercise:

1. Determine the matrix which diagonalizes the following matrices:
$A = \begin{pmatrix} 5 & 0 & 0 \\ 1 & 5 & 0 \\ 0 & 1 & 5 \end{pmatrix}$; $B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix}$; $C = \begin{pmatrix} -1 & 0 & 0 \\ -3 & 4 & 0 \\ -3 & 1 & 3 \end{pmatrix}$; $D = \begin{pmatrix} 19 & -9 & -6 \\ 25 & -11 & -9 \\ 17 & -9 & -4 \end{pmatrix}$.

2. Show that the matrix $P = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 2 & 0 \\ -3 & 5 & 2 \end{pmatrix}$ is not diagonalizable.

2.3 Orthogonal Diagonalization

Definition: If given a matrix A of order n, there exists an orthogonal matrix P such that

$$P^{-1}AP = P^TAP = D \tag{2.7}$$

where D is a diagonal matrix, then matrix A is said to be orthogonally diagonalizable, and P is said to orthogonally diagonalize A.
Remarks: For a matrix, say A, to be orthogonally diagonalizable, the matrix must be symmetric. That means that $A^T = A$. This can be shown thus: $A^T = A \Rightarrow P^TAP = D$, P is orthogonal; D is diagonal $\Rightarrow P^TP = PP^T = I$.

$\therefore A = PDP^T$.

D is diagonal $\Rightarrow D^T = D$.

$\therefore A^T = (PDP^T)^T = PDP^T = A$.

$\therefore A^T = A \Rightarrow A$ is symmetric.

2.4 Diagonalizing a Symmetric Matrix

Theorem 2.2. *If A is a symmetric matrix, then*

(i) the eigenvalues of the matrix are all real.

(ii) the eigenvectors from different eigenspaces are orthogonal.

From the above theorem, a symmetric matrix A can be diagonalized as follows:

1. Find a basis for each eigenspace of A.

2. Apply the Gram-Schmidt process to each of these bases to obtain an orthonormal basis for each eigenspace.

3. Form the matrix P whose columns are the bases vectors constructed in 2. This matrix then orthogonally diagonalizes A.

Example: Find an orthogonal matrix P that diagonalizes $A = \begin{pmatrix} -4 & 2 & -2 \\ 2 & -7 & 4 \\ -2 & 4 & -7 \end{pmatrix}$.

To find the eigenvalues, we have

$$\begin{aligned}
\det(A - \lambda I) &= \begin{vmatrix} -4-\lambda & 2 & -2 \\ 2 & -7-\lambda & 4 \\ -2 & 4 & -7-\lambda \end{vmatrix} \\
&=(-4-\lambda)\begin{vmatrix} -7-\lambda & 4 \\ 4 & -7-\lambda \end{vmatrix} - 2\begin{vmatrix} 2 & 4 \\ -2 & -7-\lambda \end{vmatrix} - 2\begin{vmatrix} 2 & -7-\lambda \\ -2 & 4 \end{vmatrix} \\
&=(-4-\lambda)\left[(-7-\lambda)^2 - 16\right] - 2[2(-7-\lambda)+8] - 2[8+2(-7-\lambda)] \\
&=(-4-\lambda)(\lambda^2+14\lambda+33) - 2(-2\lambda-6) - 2(-2\lambda-6) \\
&=-\lambda^3 - 18\lambda^2 - 81\lambda - 108 \\
&=-(\lambda+12)(\lambda+3)^2.
\end{aligned}$$

The characteristic equation is $\phi(\lambda) = (\lambda+12)(\lambda+3)^2 = 0$ with eigenvalues $\lambda_1 = -12$ and $\lambda_{2,3} = -3$.

For the eigenvalue $\lambda_1 = -12$, we have $(A - \lambda I)X = 0$ such that

$$\begin{pmatrix} 8 & 2 & -2 \\ 2 & 5 & 4 \\ -2 & 4 & 5 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

which is equivalent to

$$\begin{pmatrix} 2 & 0 & -1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

in row reduced echelon form (r.r.e.f.). This gives $P_1' = \begin{pmatrix} 1 \\ -2 \\ 2 \end{pmatrix}$ if $x_3 = 1$. We normalize P_1' to get the required orthonormal eigenvector P_1 as follows:

$$P_1 = \frac{P_1'}{|P_1'|} = \frac{\begin{pmatrix} 1 \\ -2 \\ 2 \end{pmatrix}}{\sqrt{(1)^2 + (-2)^2 + (2)^2}} = \frac{1}{\sqrt{9}}\begin{pmatrix} 1 \\ -2 \\ 2 \end{pmatrix} = \frac{1}{3}\begin{pmatrix} 1 \\ -2 \\ 2 \end{pmatrix} = \begin{pmatrix} \frac{1}{3} \\ -\frac{2}{3} \\ \frac{2}{3} \end{pmatrix}.$$

For the eigenvalue $\lambda_{2,3} = -3$, we have

$$\begin{pmatrix} -1 & 2 & -2 \\ 2 & -4 & 4 \\ -2 & 4 & -4 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

which is equivalent to the r.r.e.f.

$$\begin{pmatrix} 1 & -2 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

The eigenspace spans $P_2' = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$ if $x_3 = 1, x_2 = 1$ and $P_3' = \begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix}$ if $x_3 = 1, x_2 = 2$.

To obtain the orthonormal basis for P_2', we have

$$P_2 = \frac{P_2'}{|P_2'|} = \frac{\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}}{\sqrt{(0)^2 + (1)^2 + (1)^2}} = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix} = \begin{pmatrix} 0 \\ \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{pmatrix}.$$

Now, we use the Gram-Schmidt process to obtain the orthonormal basis for P_3' as follows

$$P_3'' = P_3' - \frac{P_2' \cdot P_3'}{P_2' \cdot P_2'} P_2' = \begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix} - \frac{\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix}}{\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}} \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$$

$$= \begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix} - \frac{0(2) + 1(2) + 1(1)}{0(0) + 1(1) + 1(1)} \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$$

$$= \begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix} - \frac{3}{2} \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$$

$$= \begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix} - \begin{pmatrix} 0 \\ \frac{3}{2} \\ \frac{3}{2} \end{pmatrix}$$

$$= \begin{pmatrix} 2 \\ \frac{1}{2} \\ -\frac{1}{2} \end{pmatrix}.$$

Furthermore,

$$P_3 = \frac{P_3''}{|P_3''|} = \frac{\begin{pmatrix} 2 \\ \frac{1}{2} \\ -\frac{1}{2} \end{pmatrix}}{\sqrt{(2)^2 + \left(\frac{1}{2}\right)^2 + \left(-\frac{1}{2}\right)^2}} = \frac{1}{\sqrt{\frac{9}{2}}} \begin{pmatrix} 2 \\ \frac{1}{2} \\ -\frac{1}{2} \end{pmatrix} = \frac{\sqrt{2}}{3} \begin{pmatrix} 2 \\ \frac{1}{2} \\ -\frac{1}{2} \end{pmatrix} = \begin{pmatrix} \frac{2\sqrt{2}}{3} \\ \frac{\sqrt{2}}{6} \\ -\frac{\sqrt{2}}{6} \end{pmatrix}.$$

$\therefore$ The matrix A is orthogonally diagonalizable by the matrix

$$P = \begin{pmatrix} P_1 & P_2 & P_3 \end{pmatrix} = \begin{pmatrix} \frac{1}{3} & 0 & \frac{2\sqrt{2}}{3} \\ -\frac{2}{3} & \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{6} \\ \frac{2}{3} & \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{6} \end{pmatrix}$$

such that

$$P^T A P = \begin{pmatrix} -12 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & -3 \end{pmatrix}.$$

Exercise: Orthogonally diagonalize the following: $A = \begin{pmatrix} 2 & 5 \\ 5 & 2 \end{pmatrix}$; $B = \begin{pmatrix} -3 & 1 \\ 1 & -3 \end{pmatrix}$; $C = \begin{pmatrix} 3 & 0 & 7 \\ 0 & 5 & 0 \\ 7 & 0 & 3 \end{pmatrix}$; $D = \begin{pmatrix} 5 & -2 & 4 \\ -2 & 8 & -2 \\ -4 & -2 & 5 \end{pmatrix}$; $E = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$; $F = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 2 & 2 \\ 0 & 2 & 5 \end{pmatrix}$.

3. Quadratic Forms

Our discussion so far has been on linear equations (or systems of linear equations) of the form

$$a_1x_1 + a_2x_2 + ... + a_nx_n = b \tag{3.1}$$

where $a_1x_1 + a_2x_2 + ... + a_nx_n$ are functions of n-variables $x_1, x_2, ..., x_n$. The left hand side of (3.1) is called the linear form where all variables occur to the first power and there are no products of variables in the expression.

Going forward, we want to look at functions in which the terms are squares of variables or products of two variables. These types of functions arise in a variety of applications such as geometry, vibrations of mechanical systems, statistics and electrical engineering.

Definition 3.1. A quadratic form in two variables, say x and y, is an expression that can be written as

$$ax^2 + 2bxy + cy^2 \equiv \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} a & b \\ b & c \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}. \tag{3.2}$$

Examples

1. $2x^2 + 6xy + 7y^2 \Rightarrow a = 2, b = 3, c = 7$ such that

$$2x^2 + 6xy + 7y^2 \equiv \begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} 2 & 3 \\ 3 & 7 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

2. $x^2 - 3xy - 2y^2 \Rightarrow a = 1, b = -3/2, c = -2$ such that

$$x^2 - 3xy - 2y^2 \equiv \begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} 1 & -\frac{3}{2} \\ -\frac{3}{2} & -2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

Definition 3.2. A quadratic form in n-variables $x_1, x_2, ..., x_n$ is an expression that can be written as

$$X^T AX = \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \tag{3.3}$$

Note: We can show that

$$X^T AX = a_{11}x_1^2 + a_{22}x^2 + ... + a_{nn}x_n^2 + \sum_{i \neq j} a_{ij}x_i x_j \tag{3.4}$$

where $\sum_{i \neq j} a_{ij}x_i x_j$ is the sum of all terms of the form $a_{ij}x_i x_j$ for which x_i and x_j are different variables. The terms $a_{ij}x_i x_j$ are called cross product terms of the quadratic form.

Example If $X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$; $\begin{pmatrix} 1 & 2 & -1 \\ 2 & 7 & 3 \\ -1 & 3 & -3 \end{pmatrix}$, show that $X^T AX$ is a quadratic form.

$$X^TAX = \begin{pmatrix} x_1 & x_2 & x_3 \end{pmatrix} \begin{pmatrix} 1 & 2 & -1 \\ 2 & 7 & 3 \\ -1 & 3 & -3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

$$\begin{aligned}
(X^TA)X &= \begin{pmatrix} x_1 + 2x_2 - x_3 & 2x_1 + 7x_2 + 3x_3 & -x_1 + 3x_2 - 3x_3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \\
&= (x_1 + 2x_2 - x_3)x_1 + (2x_1 + 7x_2 + 3x_3)x_2 + (-x_1 + 3x_2 - 3x_3)x_3 \\
&= (x_1^2 + 2x_1x_2 - x_1x_3) + (2x_1x_2 + 7x_2^2 + 3x_2x_3) + (-x_1x_3 + 3x_2x_3 - 3x_3^2) \\
&= x_1^2 + 7x_2^2 - 3x_3^2 + 4x_1x_2 - 2x_1x_3 + 6x_2x_3 \\
&= a_{11}x_1^2 + a_{22}x^2 + \ldots + a_{nn}x_n^2 + \sum_{i \neq j} a_{ij}x_ix_j \Rightarrow \text{quadratic form.}
\end{aligned}$$

Note

(1) The coefficients of the squared terms (x_1^2, x_2^2, x_3^2) appear on the main diagonal of $A_{3\times 3}$

(2) The coefficients of the cross-product terms are each split in half and appear in the off-diagonal positions in matrix A as follows:
$x_1x_2 \quad a_{12};\ a_{21}$
$x_1x_3 \quad a_{13};\ a_{31}$
$x_2x_3 \quad a_{23};\ a_{32}$

(3) $x_ix_j = x_jx_i$

3.1 Maximum and Minimum Values of a Quadratic Form Subject to Constraint

Theorem 3.1. *Suppose A is a symmetric matrix of order n and its eigenvalues in ascending size order are $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$. If X is constrained such that $\| X \| = (x_1^2 + x_2^2 + \cdots + x_n^2)^{\frac{1}{2}} = 1$, then*

i) $\lambda_1 \leq X^TAX \leq \lambda_n$

ii) $X^TAX = \lambda_n$ *if X is an eigenvector of A corresponding to λ_n and* $X^TAX = \lambda_1$ *if X is an eigenvector of A corresponding to λ_1.*

From the theorem above, it follows that subject to the constraint $\| X \| = 1$, the quadratic form X^TAX has a maximum value of λ_n (the largest eigenvalue) and a minimum value λ_1 (the lowest eigenvalue), i.e. X^TAX is maximum for λ_n and minimum if otherwise.

Example
Find the maximum and minimum values of the quadratic form $x_1^2 + x_2^2 + 4x_1x_2$ subject to the constraint $x_1^2 + x_2^2 = 1$.

The quadratic form can be written as

$$x_1^2 + x_2^2 + 4x_1x_2 = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \tag{3.5}$$

For $A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$, $|A - \lambda I| = (1-\lambda)^2 - 4 = \lambda^2 - 2\lambda - 3 = 0 \Rightarrow \lambda = 3, \lambda = -1$. $\therefore X^TAX$ is maximum when $\lambda = 3$ and minimum when $\lambda = -1$ subject to the constraint $x_1^2 + x_2^2 = 1$. To find values for x_1 and x_2 at the maximum and minimum, we have to find the eigenvectors corresponding to these eigenvalues and normalize these eigenvectors to satisfy the constraint $x_1^2 + x_2^2 = 1$.

We have the corresponding eigenvector for $\lambda = 3$ as $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and for $\lambda = -1$ as $\begin{bmatrix} -1 \\ 1 \end{bmatrix}$.

$\therefore$ The bases for the eigenspaces are $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} -1 \\ 1 \end{bmatrix}$ for $\lambda = 3$ and $\lambda = -1$ respectively.

Normalizing the eigenvectors, we have $\begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}$ and $\begin{bmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}$. $\left(\text{Normalizing } x_i \Rightarrow \dfrac{x_i}{\sqrt{\sum\limits_{i=1} x_i^2}} \right)$.

$\therefore$ Subject to the constraint $x_1^2 + x_2^2 = 1$, the maximum values for x_1 and x_2 are $\frac{1}{\sqrt{2}}$ and $\frac{1}{\sqrt{2}}$ respectively when $\lambda = 3$; and the minimum values for x_1 and x_2 are $\frac{1}{\sqrt{2}}$ and $-\frac{1}{\sqrt{2}}$ respectively when $\lambda = -1$.

Now, the maximum value of the quadratic form $x_1^2 + x_2^2 + 4x_1x_2$ is

$$\left(\frac{1}{\sqrt{2}}\right)^2 + \left(\frac{1}{\sqrt{2}}\right)^2 + 4\left(\frac{1}{\sqrt{2}}\right)\left(\frac{1}{\sqrt{2}}\right) = \frac{1}{2} + \frac{1}{2} + 4\left(\frac{1}{2}\right) = 3.$$

The minimum value of the quadratic form is

$$\left(-\frac{1}{\sqrt{2}}\right)^2 + \left(\frac{1}{\sqrt{2}}\right)^2 + 4\left(-\frac{1}{\sqrt{2}}\right)\left(\frac{1}{\sqrt{2}}\right) = \frac{1}{2} + \frac{1}{2} - 4\left(\frac{1}{2}\right) = -1.$$

Note: Alternate bases for the eigenspaces can be obtained by multiplying the basis vectors by -1. Thus, the maximum value when $\lambda = 3$ also occurs if $x_{1,2} = -\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}$, and the minimum value when $\lambda = -1$ also occurs if $x_{1,2} = -\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}$.

Example/Exercise

Find the maximum and minimum values of the following quadratic forms subject to the constraint $x_1^2 + x_2^2 + x_3^2 = 1$ and determine the values of x_1, x_2, x_3 at which the maximum and minimum occur.

i) $x_1^2 + x_2^2 + 2x_3^2 - 2x_1x_2 + 4x_1x_3 + 4x_2x_3$

ii) $2x_1^2 + x_2^2 + x_3^2 + 2x_2x_3 + 2x_1x_2$

iii) $3x_1^2 + 2x_2^2 + 3x_3^2 + 2x_1x_2$

Solution

Given $x_1^2 + x_2^2 + 2x_3^2 - 2x_1x_2 + 4x_1x_3 + 4x_2x_3$.

First, we determine a symmetric matrix A such that X^TAX is a quadratic form, i.e. $A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$.

The matrix is given as $A = \begin{pmatrix} 1 & -1 & 2 \\ -1 & 1 & 2 \\ 2 & 2 & 2 \end{pmatrix}$ where $a_{11} = a_{22} = 1$; $a_{33} = 2$; $a_{12} + a_{21} = -2 \Rightarrow a_{12} = a_{21} = -1$; $a_{13} + a_{31} = 4 \Rightarrow a_{13} = a_{31} = 2$; $a_{23} + a_{32} = 4 \Rightarrow a_{23} = a_{32} = 2$. A is symmetric since $A^T = A$.

Second, we determine the maximum and minimum eigenvalues of A.

$|A - \lambda I| = \begin{vmatrix} 1-\lambda & -1 & 2 \\ -1 & 1-\lambda & 2 \\ 2 & 2 & 2-\lambda \end{vmatrix} = -\lambda^3 + 4\lambda^2 + 4\lambda - 16 = -(\lambda - 4)(\lambda - 2)(\lambda + 2)$. Solving the characteristic equation $(\lambda - 4)(\lambda - 2)(\lambda + 2) = 0$, we have $\lambda_1 = 4$, $\lambda_2 = 2$ and $\lambda_3 = -2$. So, the quadratic form is maximum for $\lambda_1 = 4$ and minimum for $\lambda_3 = -2$.

Third, we obtain the eigenvectors corresponding to the maximum and minimum eigenvalues.

The eigenvector corresponding to $\lambda = 4$ is $\begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}$ while the eigenvector corresponding to $\lambda = -2$ is $\begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}$.

Fourth, we normalize the eigenvectors to satisfy $x_1^2 + x_2^2 + x_3^2 = 1$.

For $\begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}$, we have $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \frac{1}{\sqrt{6}} \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \\ \frac{2}{\sqrt{6}} \end{pmatrix} = \begin{pmatrix} \frac{\sqrt{6}}{6} \\ \frac{\sqrt{6}}{6} \\ \frac{\sqrt{6}}{3} \end{pmatrix}$.

For $\begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}$, we have $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{3}} \end{pmatrix} = \begin{pmatrix} \frac{\sqrt{3}}{3} \\ \frac{\sqrt{3}}{3} \\ -\frac{\sqrt{3}}{3} \end{pmatrix}$.

Finally, we find the maximum and minimum values of the quadratic form.

The maximum value of the quadratic form $x_1^2 + x_2^2 + 2x_3^2 - 2x_1x_2 + 4x_1x_3 + 4x_2x_3$ is

$$\left(\frac{1}{\sqrt{6}}\right)^2 + \left(\frac{1}{\sqrt{6}}\right)^2 + 2\left(\frac{2}{\sqrt{6}}\right)^2 - 2\left(\frac{1}{\sqrt{6}}\right)\left(\frac{1}{\sqrt{6}}\right) + 4\left(\frac{1}{\sqrt{6}}\right)\left(\frac{2}{\sqrt{6}}\right) + 4\left(\frac{1}{\sqrt{6}}\right)\left(\frac{2}{\sqrt{6}}\right)$$
$$= \frac{1}{6} + \frac{1}{6} + \frac{8}{6} - \frac{2}{6} + \frac{8}{6} + \frac{8}{6}$$
$$= 4.$$

The minimum value of the quadratic form is

$$\left(\frac{1}{\sqrt{3}}\right)^2 + \left(\frac{1}{\sqrt{3}}\right)^2 + 2\left(-\frac{1}{\sqrt{3}}\right)^2 - 2\left(\frac{1}{\sqrt{3}}\right)\left(\frac{1}{\sqrt{3}}\right) + 4\left(\frac{1}{\sqrt{3}}\right)\left(-\frac{1}{\sqrt{3}}\right) + 4\left(\frac{1}{\sqrt{3}}\right)\left(-\frac{1}{\sqrt{3}}\right)$$
$$= \frac{1}{3} + \frac{1}{3} + \frac{2}{3} - \frac{2}{3} - \frac{4}{3} - \frac{4}{3}$$
$$= -2.$$

3.2 Positive Definite Matrices and Quadratic Forms

Definition 3.3.

1. A quadratic form X^TAX is called positive definite if $X^TAX > 0$ for all $X \neq 0$.
2. A symmetric matrix A is called a positive definite matrix if the quadratic form X^TAX is positive definite, i.e. $X^TAX > 0$.

Theorem 3.2. *A symmetric matrix A is positive definite if and only if all the eigenvalues of A are positive.*

For instance, the matrix $\begin{bmatrix} 1 & 0 & 0 \\ 3 & 4 & 0 \\ 3 & 1 & 2 \end{bmatrix}$ is positive definite since the eigenvalues are $\lambda_1 = 4 > 0$; $\lambda_2 = 2 > 0$ and $\lambda_3 = 1 > 0$.

3.3 Criterion for Determination of Positive Definiteness of a Symmetric Matrix

Definition 3.4 (Principal submatrices)**.** Given a square matrix $A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$.

The principal submatrices of A are the submatrices formed from the first i-row and i-column. That is

$$A_1 = \begin{bmatrix} a_{11} \end{bmatrix};\ A_2 = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix};\ A_3 = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix};\ \ldots\ ;\ A_n = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}.$$

Theorem 3.3. *A symmetric matrix A is positive definite if and only if the determinant of every principal submatrix is positive.*

For example, the matrix $A = \begin{bmatrix} 2 & -1 & -3 \\ -1 & 2 & 4 \\ -3 & 4 & 9 \end{bmatrix}$ is positive definite since

$$A_1 = \begin{vmatrix} 2 \end{vmatrix} = 2 > 0; \qquad A_2 = \begin{vmatrix} 2 & -1 \\ -1 & 2 \end{vmatrix} = 3 > 0; \qquad A_3 = \begin{vmatrix} 2 & -1 & -3 \\ -1 & 2 & 4 \\ -3 & 4 & 9 \end{vmatrix} = 1 > 0.$$

Exercise

Examine the following matrices for positive definiteness: 1. $\begin{pmatrix} 1 & -1 & -1 \\ -1 & 2 & 4 \\ -1 & 4 & 6 \end{pmatrix}$ 2. $\begin{pmatrix} 4 & 2 & -2 \\ 2 & 4 & 2 \\ -2 & 2 & 4 \end{pmatrix}$

3. $\begin{pmatrix} 1 & 2 & 0 \\ 2 & 1 & 3 \\ 0 & 3 & -3 \end{pmatrix}$ 4. $\begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 3 & 3 & 2 \end{pmatrix}$.

3.4 Diagonalization of Quadratic Forms

Theorem 3.4 (The Principal Axes Theorem)**.** *Every quadratic form is diagonalizable. In essence, if X^TAX is a quadratic form in $X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$, then there exists an orthogonal matrix Q such that $X^TAX = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \ldots + \lambda_n y_n^2$, where $Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = Q^T \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ and $\lambda_1, \lambda_2, \ldots, \lambda_n$.*

Recall: Q is the matrix whose columns are the unit eigenvectors of the matrix A.

Example: Diagonalize the quadratic form $h(x_1, x_2, x_3) = 2x_1x_2 + 2x_1x_3 + 2x_2x_3$.

The matrix of the quadratic form is

$$A = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$$

with eigenvalues $\lambda_1 = 2$, $\lambda_{2,3} = -1$ and corresponding unit eigenvectors

$$\begin{bmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{bmatrix}, \begin{bmatrix} -\frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \end{bmatrix}, \begin{bmatrix} 0 \\ \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{bmatrix}$$

such that

$$Q = \begin{bmatrix} \frac{1}{\sqrt{3}} & -\frac{2}{\sqrt{6}} & 0 \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} \end{bmatrix}.$$

If we set $Y = Q^T X$, we have

$$y_1 = \frac{1}{\sqrt{3}}(x_1 + x_2 + x_3), y_2 = \frac{1}{\sqrt{6}}(-2x_1 + x_2 + x_3), y_3 = \frac{1}{\sqrt{2}}(x_2 - x_3).$$

In terms of the variables, y_1, y_2, y_3, the quadratic form is given as

$$2y_1^2 - y_2^2 - y_3^2.$$

Exercise: Diagonalize each of the following quadratic forms:

1) $f(x) = x_1^2 - 6x_1x_2 + x_2^2$

2) $g(x) = 2x_1x_2 + 2x_2x_3$

4. Bilinear Form

5. Canonical Forms

References

[1] H. Anton, C. Rorres (1994). *Elementary Linear Algebra, 7th Edition*. John Wiley and Sons Inc., New York.

[2] S.D. Conte, C. de Boor (1980). *Elementary Numerical Analysis- An Algorithmic Approach, 3rd Edition*. The McGraw-Hill Book Company, New York.

[3] F.E. Holm (1972). *Introduction to Linear Algebra: Models, Methods and Theory*. The Macmillan Publishing Company, New York.

[4] S.R.K. Iyenger, R.K. Jain (2010). *Mathematical Methods, 2nd Edition*. Narosa Publishing House, New Delhi.

[5] S. Lipschutz, M.L. Lipson (2004). *Schaum's Outline of Theory and Problems of Linear Algebra, 3rd Edition*. The McGraw-Hill Companies.

[6] M.R. Spiegel (1977). *Schaum's Series Outline of Theory and Problems of Advanced Mathematics for Engineers and Scientists*. McGraw-Hill Inc., New York.

[7] A. Tucker (1988). *A Unified Introduction to Linear Algebra: Models, Methods and Theory*. The Macmillan Publishing Company, New York.