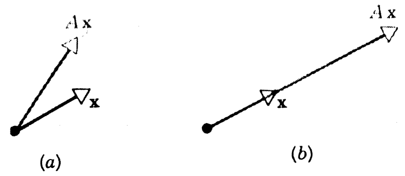
1 Introduction

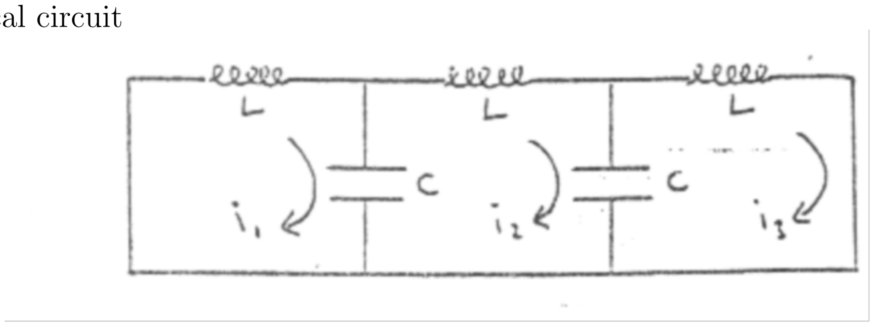
If A is an 3×3 matrix and $\vec{x}$ is a vector in R^3, then there is usually no general geometric relationship between the vector $\vec{x}$ and the vector $A\vec{x}$, see figure (a). However, there are often certain nonzero vectors $\vec{x}$ such that $\vec{x}$ and $A\vec{x}$ are scalar multiples of one another, that is, $A\vec{x} = \lambda\vec{x}$ for some real number λ, see figure (b). Such vectors arise naturally in the study of vibrations, electrical systems, genetics, chemical reactions, quantum mechanics, mechanical stress, economics, and geometry.



Let A be an $n \times n$ square matrix and consider the equation $A\vec{x} = \lambda\vec{x}$. A value of λ for which $A\vec{x} = \lambda\vec{x}$ has a non-trivial solution $\vec{x} \neq \vec{0}$ is called an eigenvalue (or characteristic value) of A and the corresponding solution $\vec{x} \neq \vec{0}$ is an eigenvector (or characteristic vector). The set of all eigenvalues of A is known as the spectrum of A.

Example- Electric Circuit

Consider the electrical circuit



The currents flowing in the loops satisfy

$L\frac{di_1}{dt} + \frac{1}{C}\int (i_1 - i_2)\,dt = 0 \Rightarrow LCi_1'' + i_1 - i_2 = 0$

$L\frac{di_2}{dt} + \frac{1}{C}\int (i_2 - i_3)\,dt + \frac{1}{C}\int (i_2 - i_1)\,dt = 0 \Rightarrow LCi_2'' + 2i_2 - i_3 - i_1 = 0$

$L\frac{di_3}{dt} + \frac{1}{C}\int (i_3 - i_2)\,dt = 0 \Rightarrow LCi_3'' + i_3 - i_2 = 0$

Or

$$\begin{pmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{pmatrix} \begin{pmatrix} i_1 \\ i_2 \\ i_3 \end{pmatrix} = -LC \begin{pmatrix} i_1'' \\ i_2'' \\ i_3'' \end{pmatrix}$$

Look for solutions which are sinusoidal and all have the same frequency, that is,

$i_1 = I_1 \sin(wt + \theta_1), i_2 = I_2 \sin(wt + \theta_2), i_3 = I_3 \sin(wt + \theta_3)$.

Then $i_1'' = -w^2 i_1, i_2'' = -w^2 i_2$ and $i_3'' = -w^2 i_3$ and the problem becomes

$$\begin{pmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{pmatrix} \begin{pmatrix} i_1 \\ i_2 \\ i_3 \end{pmatrix} = w^2 LC \begin{pmatrix} i_1 \\ i_2 \\ i_3 \end{pmatrix},$$ which is an eigenvalue problem with $\lambda = w^2 LC$

$\square$

2 Determination of Eigenvalues and Eigenvectors

$\overrightarrow{Ax} = \lambda \vec{x} \Leftrightarrow \vec{x} = \lambda I \vec{x} \Leftrightarrow \overrightarrow{Ax} - \lambda \vec{x} = \vec{0} \Leftrightarrow (A - \lambda I)\vec{x} = \vec{0}$

For this equation $(A - \lambda I)\vec{x} = \vec{0}$ to have non-trivial solutions $\vec{x} \neq \vec{0}$ we must have $(A - \lambda I)^{-1}$ not existing, for otherwise $(A - \lambda I)\vec{x} = \vec{0} \Rightarrow (A - \lambda I)^{-1}(A - \lambda I)\vec{x} = (A - \lambda I)^{-1}\vec{0} \Rightarrow \overrightarrow{Ix} = \vec{0} \Rightarrow \vec{x} = \vec{0}$ is the unique solution. Hence λ is an eigenvalue if and only if $(A - \lambda I)^{-1}$ does not exist if and only if $\det(A - \lambda I) = 0$

Consider $\det(A - \lambda I)$. After expansion $\det(A - \lambda I)$ is a polynomial of degree n in λ, known as the characteristic polynomial of A. And $\det(A - \lambda I) = 0$ is the characteristic equation of A. Clearly, if multiplicity roots are counted, an $n \times n$ square matrix A has n eigenvalues including real and complex eigenvalues.

Example

Find eigenvalues of $A = \begin{pmatrix} 1 & 4 \\ 2 & 3 \end{pmatrix}$ and the corresponding eigenvectors.

Solution:

$$A - \lambda I = \begin{pmatrix} 1 & 4 \\ 2 & 3 \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1-\lambda & 4 \\ 2 & 3-\lambda \end{pmatrix}$$

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

Hence the eigenvalues are $\lambda = 5$ and $\lambda = -1$

To find the corresponding eigenvectors we solve:

When $\lambda = 5$

$$\begin{pmatrix} -4 & 4 \\ 2 & -2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} = 0 \Leftrightarrow \left(\begin{array}{cc|c} -4 & 4 & 0 \\ 2 & -2 & 0 \end{array}\right) \sim \left(\begin{array}{cc|c} 1 & -1 & 0 \\ 0 & 0 & 0 \end{array}\right) \Rightarrow x_1 = x_2$$

Let $x_2 = k \neq 0$, then the corresponding eigenvectors of A for $\lambda = 5$ are $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = k \begin{pmatrix} 1 \\ 1 \end{pmatrix}, k \neq 0$.

And one of the eigenvectors for $k = 1$ is $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

When $\lambda = -1$

$$\begin{pmatrix} 2 & 4 \\ 2 & 4 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} = 0 \Leftrightarrow \left(\begin{array}{cc|c} 2 & 4 & 0 \\ 2 & 4 & 0 \end{array}\right) \sim \left(\begin{array}{cc|c} 1 & 2 & 0 \\ 0 & 0 & 0 \end{array}\right) \Rightarrow x_1 = -2x_2 \Rightarrow \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} =$$
$$k\begin{pmatrix} -2 \\ 1 \end{pmatrix}, k \neq 0.$$

Let $k = 1$, then $\begin{pmatrix} -2 \\ 1 \end{pmatrix}$ is one of the corresponding eigenvectors of A for $\lambda = -1$.

$\square$

<u>Example</u>

Find eigenvalues of $A = \begin{pmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{pmatrix}$ and the corresponding eigenvectors.

Solution:

$$|A - \lambda I| = \begin{vmatrix} 1-\lambda & -1 & 0 \\ -1 & 2-\lambda & -1 \\ 0 & -1 & 1-\lambda \end{vmatrix} = (1-\lambda)(1-\lambda)(2-\lambda) - (1-\lambda) - (1-\lambda)$$
$$= (1-\lambda)\left(\lambda^2 - 3\lambda + 2 - 2\right) = \lambda(1-\lambda)(\lambda-3)$$

When $\lambda = 0$

$$\begin{pmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = 0 \Leftrightarrow \left(\begin{array}{ccc|c} 1 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 1 & 0 \end{array}\right) \sim \left(\begin{array}{ccc|c} 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & -1 & 1 & 0 \end{array}\right) \sim \left(\begin{array}{ccc|c} 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

$$\Rightarrow x_2 = x_3, x_1 = x_2 \Rightarrow x_1 = x_2 = x_3 \Rightarrow \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = k\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, k \neq 0$$

Let $k = 1$, then one of the eigenvectors for $\lambda = 0$ is $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$.

When $\lambda = 1$

$$\begin{pmatrix} 0 & -1 & 0 \\ -1 & 1 & -1 \\ 0 & -1 & 0 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = 0 \Rightarrow \left(\begin{array}{ccc|c} 0 & -1 & 0 & 0 \\ -1 & 1 & -1 & 0 \\ 0 & -1 & 0 & 0 \end{array}\right) \sim \left(\begin{array}{ccc|c} -1 & 1 & -1 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{array}\right) \sim \left(\begin{array}{ccc|c} 1 & -1 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

$$\Rightarrow x_2 = 0, x_1 = x_2 - x_3 = -x_3 \Rightarrow \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = k\begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}, k \neq 0$$

One of the eigenvectors for $\lambda = 1$ is $\begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$.

When $\lambda = 3$

$$\begin{pmatrix} -2 & -1 & 0 \\ -1 & -1 & -1 \\ 0 & -1 & -2 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = 0 \Rightarrow \left(\begin{array}{ccc|c} -2 & -1 & 0 & 0 \\ -1 & -1 & -1 & 0 \\ 0 & -1 & -2 & 0 \end{array}\right) \sim \left(\begin{array}{ccc|c} -1 & -1 & -1 & 0 \\ -2 & -1 & 0 & 0 \\ 0 & -1 & -2 & 0 \end{array}\right)$$

$$\sim \left(\begin{array}{ccc|c} -1 & -1 & -1 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & -1 & -2 & 0 \end{array}\right) \sim \left(\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right) \Rightarrow x_2 = -2x_3, x_1 = -x_2 - x_3 = x_3 \Rightarrow \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} =$$

$$k\begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}, k \neq 0$$

One of the eigenvectors for $\lambda = 3$ is $\begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}$.

$\square$

<u>Example</u>

Find eigenvalues of $A = \begin{pmatrix} 1 & -1 \\ 4 & 1 \end{pmatrix}$ and the corresponding eigenvectors.

Solution:

$|A - \lambda I| = \begin{vmatrix} 1-\lambda & -1 \\ 4 & 1-\lambda \end{vmatrix} = \lambda^2 - 2\lambda + 5 \Rightarrow \lambda = 1 \pm 2i$ For $\lambda = 1 + 2i$, we have

$$\begin{pmatrix} -2i & -1 \\ 4 & -2i \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Leftrightarrow \left(\begin{array}{cc|c} -2i & -1 & 0 \\ 4 & -2i & 0 \end{array}\right) \underset{\frac{1}{-2i}r_1}{\sim} \left(\begin{array}{cc|c} 1 & \frac{1}{2i} & 0 \\ 4 & -2i & 0 \end{array}\right) = \left(\begin{array}{cc|c} 1 & -\frac{i}{2} & 0 \\ 4 & -2i & 0 \end{array}\right) \underset{r_2-4r_1}{\sim}$$

$$\left(\begin{array}{cc|c} 1 & -\frac{i}{2} & 0 \\ 0 & 0 & 0 \end{array}\right) \Rightarrow x_1 = \frac{i}{2}x_2 \Rightarrow \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \frac{i}{2}k \\ k \end{pmatrix} = k\begin{pmatrix} \frac{i}{2} \\ 1 \end{pmatrix}, k \neq 0$$

Notice $\frac{1}{2i} = \frac{-i}{-2i^2} = -\frac{i}{2}$.

An eigenvector for $\lambda = 1 + 2i$ is $\begin{pmatrix} \frac{i}{2} \\ 1 \end{pmatrix}$.

For $\lambda = 1 - 2i$, we have $\begin{pmatrix} 2i & -1 \\ 4 & 2i \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Leftrightarrow \left(\begin{array}{cc|c} 2i & -1 & 0 \\ 4 & 2i & 0 \end{array}\right) \underset{\frac{1}{2i}r_1}{\sim} \left(\begin{array}{cc|c} 1 & -\frac{1}{2i} & 0 \\ 4 & 2i & 0 \end{array}\right) =$

$$\left(\begin{array}{cc|c} 1 & \frac{i}{2} & 0 \\ 4 & 2i & 0 \end{array}\right) \underset{r_2-4r_1}{\sim} \left(\begin{array}{cc|c} 1 & \frac{i}{2} & 0 \\ 0 & 0 & 0 \end{array}\right) \Rightarrow x_1 = -\frac{i}{2}x_2 \Rightarrow \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} -\frac{i}{2}k \\ k \end{pmatrix} = k\begin{pmatrix} -\frac{i}{2} \\ 1 \end{pmatrix}, k \neq 0$$

Notice $-\frac{1}{2i} = -\frac{i}{2i^2} = -\frac{i}{-2} = \frac{i}{2}$.

An eigenvector for $\lambda = 1 - 2i$ is $\begin{pmatrix} -\frac{i}{2} \\ 1 \end{pmatrix}$.

Example

Find eigenvalues of $A = \begin{pmatrix} 5 & 3 \\ -3 & -1 \end{pmatrix}$ and the corresponding eigenvectors.

Solution:

$|A - \lambda I| = \begin{vmatrix} 5-\lambda & 3 \\ -3 & -1-\lambda \end{vmatrix} = (\lambda-2)^2, \lambda = 2$ is the only eigenvalue.

$\begin{pmatrix} 3 & 3 \\ -3 & -3 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Rightarrow \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = k\begin{pmatrix} -1 \\ 1 \end{pmatrix}, k \neq 0$

□

3 Eigenvalue Results

Theorem

Eigenvectors corresponding to distinct eigenvalues are linearly independent.

Proof:

The proof is by induction on the number of distinct eigenvalues.

Any single eigenvector by itself, being nonzero, is linearly independent.

Assume any $k-1$ eigenvectors associated with $k-1$ distinct eigenvalues are linearly independent.

Suppose $\vec{v}_1, \cdots, \vec{v}_k$ are eigenvectors associated, respectively, with k distinct eigenvalues $\lambda_1, \cdots, \lambda_k$.

Consider $c_1\overrightarrow{v_1} + \cdots + c_k\overrightarrow{v_k} = \overrightarrow{0}$, claim $c_1 = \cdots = c_k = 0$.

Suppose not, then without loss of generality suppose, say $c_1 \neq 0$. Then

$(A - \lambda_1 I)(c_1\vec{v}_1 + \cdots + c_k\vec{v}_k) = (A - \lambda_1 I) c_1\vec{v}_1 + \cdots + (A - \lambda_1 I) c_k\vec{v}_k = (A - \lambda_1 I)\overrightarrow{0}$

$\Rightarrow (Ac_1\vec{v}_1 - \lambda_1 I c_1\vec{v}_1) + (Ac_2\overrightarrow{v_2} - \lambda_1 I c_2\vec{v}_2) + \cdots + (Ac_k\overrightarrow{v_k} - \lambda_1 I c_k\vec{v}_k) = \overrightarrow{0}$

$\Rightarrow (c_1A\overrightarrow{v_1} - c_1\lambda_1\overrightarrow{v_1}) + (c_2A\overrightarrow{v_2} - \lambda_1 c_2\overrightarrow{v_2}) + \cdots + (c_kA\overrightarrow{v_k} - c_k\lambda_1 I\overrightarrow{v_k}) = \overrightarrow{0}$

$\Rightarrow (c_1\lambda_1\vec{v}_1 - c_1\lambda_1\vec{v}_1) + (c_2\lambda_2\vec{v}_1 - c_2\lambda_1\vec{v}_1)\cdots + (c_k\lambda_k\vec{v}_1 - c_k\lambda_1\vec{v}_1) = \overrightarrow{0}$

$\Rightarrow c_2(\lambda_2 - \lambda_1)\overrightarrow{v_2}\cdots + c_k(\lambda_k - \lambda_1)\overrightarrow{v_k} = \overrightarrow{0}$

By induction hypothesis $\overrightarrow{v_2}, \cdots, \overrightarrow{v_k}$ are linearly independent, so $c_2(\lambda_2 - \lambda_1) = \cdots = c_k(\lambda_k - \lambda_1) = 0$.

Then $\lambda_2 - \lambda_1 \neq 0, \cdots, \lambda_k - \lambda_1 \neq 0 \Rightarrow c_2 = \cdots = c_k = 0$. It follows that $c_1\overrightarrow{v_1} = \overrightarrow{0}$.

$c_1\overrightarrow{v_1} = \overrightarrow{0} \Rightarrow c_1 = 0$. It contradicts to $c_1 \neq 0$, implying $\overrightarrow{v_1}, \cdots, \overrightarrow{v_k}$ are linearly independent.

□

It is also possible to have a repeated eigenvalue which does lead to more than one linearly independent eigenvector.

Example

Let $J_n, n = 2, 3, \cdots$ be a square matrix with all its entries being 1 , that is,

$$J_n = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & 1 \end{pmatrix}$$

(a) Find all eigenvalues of J_n.

(b) Find n linearly independent eigenvectors.

Solution:

Observe that $\begin{pmatrix} 1 & 1 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & 1 \end{pmatrix} \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} = \begin{pmatrix} 1\times 1 + \cdots + 1 \times 1 \\ \vdots \\ 1\times 1 + \cdots + 1 \times 1 \end{pmatrix} = \begin{pmatrix} n \\ \vdots \\ n \end{pmatrix} = n \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$ so n is an eigenvalue of $J_n$and one of its eigenvector is $\begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$.

In addition, see that $\begin{pmatrix} 1 & 1 & \ldots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & 1 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} 1\times 1 - 1 \times 1 \\ \vdots \\ 1\times 1 - 1 \times 1 \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} = 0 \begin{pmatrix} 1 \\ -1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$

So 0 is also an eigenvalue of $J_n, n = 2, 3, \cdots$ and one of its eigenvector is $\begin{pmatrix} 1 \\ -1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$.

For $\lambda = 0, J_n - 0I_n = J_n$.

Solve $\left(\begin{array}{cccc|c} 1 & 1 & \cdots & 1 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & 1 & \cdots & 1 & 0 \end{array}\right) \sim \left(\begin{array}{cccc|c} 1 & 1 & \cdots & 1 & 0 \\ 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 \end{array}\right)$.

So $x_1 + x_2 + \cdots + x_n = 0$. Then $x_n = -x_1 - \cdots - x_{n-1}$

Let $x_1 = a_1, x_2 = a_2, \cdots, x_{n-1} = a_{n-1}$ where $a_1, a_2, \cdots, a_{n-1}$ are arbitrary real numbers.

Then

$$\begin{pmatrix} x_1 \\ \vdots \\ x_{n-1} \\ x_n \end{pmatrix} = \begin{pmatrix} x_1 \\ \vdots \\ x_{n-1} \\ -x_1 - \cdots - x_{n-1} \end{pmatrix} = \begin{pmatrix} a_1 \\ \vdots \\ a_{n-1} \\ -a_1 - \cdots - a_{n-1} \end{pmatrix}$$

$$= \begin{pmatrix} a_1 \\ 0 \\ \vdots \\ 0 \\ -a_1 \end{pmatrix} + \begin{pmatrix} 0 \\ a_2 \\ 0 \\ \vdots \\ 0 \\ -a_2 \end{pmatrix} + \cdots + \begin{pmatrix} 0 \\ \vdots \\ 0 \\ a_{n-1} \\ -a_{n-1} \end{pmatrix}$$

$$= a_1 \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ -1 \end{pmatrix} + a_2 \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \\ -1 \end{pmatrix} + \cdots + a_{n-1} \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ -1 \end{pmatrix}.$$

Observe that

$$a_1 \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ -1 \end{pmatrix} + a_2 \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \\ -1 \end{pmatrix} + \cdots + a_{n-1} \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} \Rightarrow \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_{n-1} \\ -a_1 - a_2 - \cdots - a_{n-1} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}$$

$\Rightarrow a_1 = 0, a_2 = 0, \cdots, a_{n-1} = 0$

It follows that $\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \\ -1 \end{pmatrix}, \cdots, \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ -1 \end{pmatrix}$ are linearly independent.

Claim that $\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}, \cdots, \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$ are linearly independent.

Consider $x_1 \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ -1 \end{pmatrix} + x_2 \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \\ -1 \end{pmatrix} + \cdots + x_{n-1} \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ -1 \end{pmatrix} + x_n \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} \cdots\cdots$(A)

Both sides of (A) time J_n, then

$$x_1 J_n \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ -1 \end{pmatrix} + x_2 J_n \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \\ -1 \end{pmatrix} + \cdots + x_{n-1} J_n \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ -1 \end{pmatrix} + x_n J_n \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} = J_n \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Since $J_n \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ -1 \end{pmatrix} = \overrightarrow{0}, J_n \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \\ -1 \end{pmatrix} = \overrightarrow{0}, \cdots, J_n \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ -1 \end{pmatrix} = \overrightarrow{0}$, it follows that $x_n n \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$

then $x_n = 0$.

Thus, $x_1 \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ -1 \end{pmatrix} + x_2 \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \\ -1 \end{pmatrix} + \cdots + x_{n-1} \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$.

As $\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \\ -1 \end{pmatrix}, \cdots, \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ -1 \end{pmatrix}$ are linearly independent, $x_1 = x_2 = \cdots = x_{n-1} = 0$

It follows that $x_1 = x_2 = \cdots = x_{n-1} = x_n = 0$ and $\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \\ -1 \end{pmatrix}, \cdots, \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ are

linearly independent.

$\square$

<u>Theorem</u>

If A is a real symmetric matrix, then

(a) the eigenvalues are all real (but not necessarily distinct).

(b) eigenvectors corresponding to distinct eigenvalues are orthogonal.

Proof:

(a)

Suppose $A\vec{v} = \lambda\vec{v}$ where $\vec{v} \neq \overrightarrow{0}$

Consider $\bar{\vec{v}}^t A \overrightarrow{v} = \bar{\vec{v}}^t \overrightarrow{\lambda v} = \lambda \bar{\vec{v}}^t \vec{v} = \lambda \left(\overline{x_1} x_1 + \cdots + \overline{x_n} x_n \right)$ where $\bar{\vec{v}}^t = (\overline{x_1}, \cdots, \overline{x_n})$ if $\vec{v} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$.

Then $\overline{\bar{\vec{v}}^t A \vec{v}} = \overline{\bar{\vec{v}}^t}\, \overline{A\vec{v}} = \overline{\overline{\vec{v}^t}} \bar{A} \bar{\vec{v}} = \vec{v}^t A \bar{\vec{v}}$

Observe that $\vec{v}^t A \bar{\vec{v}}$ is a 1×1 matrix, so $\vec{v}^t A \bar{\vec{v}} = \left(\vec{v}^t A \bar{\vec{v}} \right)^t = \overline{\vec{v}}^t A^t (\vec{v}^t)^t = \bar{\vec{v}}^t A \overrightarrow{v}$

As a result, we have $\overline{\bar{\vec{v}}^t A \vec{v}} = \bar{\vec{v}}^t A \overrightarrow{v}$. It follows that $\bar{\vec{v}}^t A \overrightarrow{v}$ is real.

We conclude that $\lambda = \frac{\bar{\vec{v}}^t A \overrightarrow{v}}{\overline{x_1} x_1 + \cdots + \overline{x_n} x_n}$ is real (note that $\overline{x_1} x_1 + \cdots + \overline{x_n} x_n$ is real).

(b) Suppose $\overrightarrow{Av_1} = \lambda_1 \overrightarrow{v_1}, \overrightarrow{Av_2} = \lambda_2 \overrightarrow{v_2}$ where $\lambda_1 \neq \lambda_2, \overrightarrow{v_1}, \overrightarrow{v_2} \neq \overrightarrow{0}$

$\lambda_1 \overrightarrow{v_1}^{t-} \overrightarrow{v_2} = \left(\lambda_1 \overrightarrow{v_1} \right)^t \overrightarrow{v_2} = \left(A \overrightarrow{v_1} \right)^t \overrightarrow{v_2} = \left(\overrightarrow{v_1} A^t \right) \overrightarrow{v_2} = \left(\overrightarrow{v_1}^t A \right) \overrightarrow{v_2} = \overrightarrow{v_1}^t \left(A \overrightarrow{v_2} \right) = \vec{v}_1^t \left(\lambda_2 \overrightarrow{v_2} \right) = \lambda_2 \overrightarrow{v_1}^t \overrightarrow{v_2}$

$\Rightarrow (\lambda_1 - \lambda_2) \overrightarrow{v_1}^t \overrightarrow{v_2} = 0 \underset{\lambda_1 - \lambda_2 \neq 0}{\Rightarrow} \overrightarrow{v_1}^t \overrightarrow{v_2} = 0$

$\square$

<u>Theorem</u>

If A is an $n \times n$ real symmetric matrix, then it is possible to find n linearly independent eigenvectors (even though the eigenvalues may not be distinct).

Proof: Omitted.

Example

Consider $J_3 = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$, which is real symmetric.

$J_3 = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$ has eigenvalue 3 of multiplicity 1 and eigenvalue 0 of multiplicity 2, all of them are real.

For eigenvalue 3, one of its corresponding eigenvectors is $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$.

For eigenvalue 0, its independent eigenvectors are $\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}$.

Observe that $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} = 0, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}\begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} = 0$

$\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}$ are linearly independent.

□

An $n \times n$ matrix is said to be diagonalizable if there is a non-singular $n \times n$ matrix P such that $D = P^{-1}AP$,where D is a diagonal matrix.

We observe that:

$\det(D - \lambda I) = \det\left(P^{-1}AP - \lambda I\right) = \det\left(P^{-1}AP - \lambda P^{-1}IP\right) = \det\left[P^{-1}(A - \lambda I)P\right]$

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

Thus, A and $P^{-1}AP = D$ have the same characteristic polynomial thus the same eigenvalues. We also notice that the eigenvalus of D are just the main diagonal entries.

Theorem

If $n \times n$ matrix A has n linearly independent eigenvectors then it is diagonalizable.

Remark:

Let A be a $p \times m$ matrix and B an $m \times n$ matrix. Recall that the product of AB can be evaluated in two different ways:

$$A=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \ddots & \ddots & \vdots \\ a_{p1} & a_{p2} & \cdots & a_{pm} \end{pmatrix} = A = (\vec{a}_1\,|\overrightarrow{a_2}|\cdots\,|\,\overrightarrow{a_m}), \text{ where } \overrightarrow{a_1} = \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{p1} \end{pmatrix}, \cdots \overrightarrow{a_m} = \begin{pmatrix} a_{1m} \\ a_{2m} \\ \vdots \\ a_{pm} \end{pmatrix}$$

$$B=\begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \ddots & \ddots & \vdots \\ b_{m1} & b_{m2} & \cdots & b_{mn} \end{pmatrix} = B = \left(\vec{b}_1\left|\vec{b}_2\right|\cdots\,|\,\vec{b}_n\right), \text{ where } \vec{b}_1 = \begin{pmatrix} b_{11} \\ b_{21} \\ \vdots \\ b_{m1} \end{pmatrix}, \cdots, \overrightarrow{b_n} = \begin{pmatrix} b_{1n} \\ b_{2n} \\ \vdots \\ b_{mn} \end{pmatrix}$$

We observe that

$$AB=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \ddots & \ddots & \vdots \\ a_{p1} & a_{p2} & \cdots & a_{pm} \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \ddots & \ddots & \vdots \\ b_{m1} & b_{m2} & \cdots & b_{mn} \end{pmatrix} = \left(\ A\vec{b}_1\ |\ \ A\overrightarrow{b_2}\ |\ \ \cdots\ \ |\,\overrightarrow{A}\vec{b}_n\ \right)$$

$$= (b_{11}\overrightarrow{a_1} + b_{21}\overrightarrow{a_2} + \cdots + b_{m1}\overrightarrow{a_m}\,|b_{12}\overrightarrow{a_1} + b_{22}\overrightarrow{a_2} + \cdots + b_{m2}\overrightarrow{a_m}|\cdots\,|\,b_{1n}\overrightarrow{a_1} + b_{2n}\overrightarrow{a_2} + \cdots + b_{mn}\overrightarrow{a_m})$$

Proof:

Suppose $\vec{x}_1,\cdots,\vec{x}_n$ are linearly independent eigenvectors with corresponding eigenvalues $\lambda_1,\cdots,\lambda_n$ such that $A\overrightarrow{x_1}=\lambda_1\overrightarrow{x_1},\cdots,A\overrightarrow{x_n}=\lambda_n\overrightarrow{x_n}$ Construct $n\times n$ matrix $P=(\vec{x}_1\,|\overrightarrow{x_2}|\cdots\,|\,\overrightarrow{x_n})$ with $\overrightarrow{x_1},\cdots,\overrightarrow{x_n}$ as its columns. Then eigenvectors of A. Then $AP=(A\vec{x}_1\,|A\vec{x}_2|\cdots\,|\,A\vec{x}_n)=(\lambda_1\vec{x}_{11}\,|\lambda_2\vec{x}_2|\cdots\,|\,\lambda_n\vec{x}_n)$

$$= (\lambda_1\vec{x}_1 + 0\overrightarrow{x_2} + \cdots + 0\overrightarrow{x_n}|\,0\overrightarrow{x_1} + \lambda_2\overrightarrow{x_2} + \cdots + 0\overrightarrow{x_n}|\ldots|0\overrightarrow{x_1} + \cdots + 0\overrightarrow{x_{n-1}} + \lambda_n\vec{x}_n)$$

$$= (\vec{x}_1\,|\overrightarrow{x_2}|\ldots\,|\,\overrightarrow{x_n}) \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & \cdots & \lambda_n \end{pmatrix} = PD, \text{ where } D = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & \cdots & \lambda_n \end{pmatrix} = \operatorname{diag}(\lambda_1,\cdots,\lambda_n).$$

As $\vec{x}_1,\cdots,\vec{x}_n$ are linearly independent, P^{-1} exist. Thus we have $P^{-1}AP=P^{-1}PD=D$

Example

Show that $A=\begin{pmatrix} 1 & 4 \\ 2 & 3 \end{pmatrix}$ is diagonalizable.

Proof:

For $A=\begin{pmatrix} 1 & 4 \\ 2 & 3 \end{pmatrix}$, we have eigenvalues $\lambda_1=5, \lambda_2=-1$(they are distinct).

For $\lambda_1=5$, one of the corresponding eigenvector is $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

For $\lambda_2=-1$, one of the corresponding eigenvector is $\begin{pmatrix} 2 \\ -1 \end{pmatrix}$.

Let $P=\begin{pmatrix} 1 & 2 \\ 1 & -1 \end{pmatrix}$, as the eigenvectors of A are linearly independent, $P^{-1}=\begin{pmatrix} 1 & 2 \\ 1 & -1 \end{pmatrix}^{-1}$ exists and

$P^{-1} = \begin{pmatrix} \frac{1}{3} & \frac{2}{3} \\ \frac{1}{3} & -\frac{1}{3} \end{pmatrix}$. Also $P^{-1}AP = \begin{pmatrix} \frac{1}{3} & \frac{2}{3} \\ \frac{1}{3} & -\frac{1}{3} \end{pmatrix} \begin{pmatrix} 1 & 4 \\ 2 & 3 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} 5 & 0 \\ 0 & -1 \end{pmatrix}$.

It concludes that $A = \begin{pmatrix} 1 & 4 \\ 2 & 3 \end{pmatrix}$ is diagonalizable.

□

<u>Example</u>

Is $A = \begin{pmatrix} 5 & 3 \\ -3 & -1 \end{pmatrix}$ diagonalizable?

Solution: $|A - \lambda I| = \begin{vmatrix} 5-\lambda & 3 \\ -3 & -1-\lambda \end{vmatrix} = (\lambda - 2)^2$ and $\lambda = 2$ is the only eigenvalue.

$$\begin{pmatrix} 3 & 3 \\ -3 & -3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Rightarrow \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = k \begin{pmatrix} -1 \\ 1 \end{pmatrix}, k \neq 0$$

As there is only one linearly independent eigenvector, $A = \begin{pmatrix} 5 & 3 \\ -3 & -1 \end{pmatrix}$ is not diagonalizable.

□

<u>Example</u>

Diagonalize $A = \begin{pmatrix} 1 & 3 & 3 \\ -3 & -5 & -3 \\ 3 & 3 & 1 \end{pmatrix}$, if possible.

Proof:

$|A - \lambda I| = \begin{vmatrix} 1-\lambda & 3 & 3 \\ -3 & -5-\lambda & -3 \\ 3 & 3 & 1-\lambda \end{vmatrix} = -\lambda^3 - 3\lambda^2 + 4 = -(\lambda - 1)(\lambda + 2)^2$ When $\lambda = 1$

$$\begin{pmatrix} 0 & 3 & 3 \\ -3 & -6 & -3 \\ 3 & 3 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Leftrightarrow \left(\begin{array}{ccc|c} 3 & 3 & 0 & 0 \\ -3 & -6 & -3 & 0 \\ 0 & 3 & 3 & 0 \end{array}\right) \sim \left(\begin{array}{ccc|c} 1 & 1 & 0 & 0 \\ 0 & -3 & -3 & 0 \\ 0 & 3 & 3 & 0 \end{array}\right) \sim$$
$$\left(\begin{array}{ccc|c} 1 & 1 & 0 & 0 \\ 0 & -1 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

$$\Rightarrow x_2 = -x_3, x_1 = -x_2 = x_3 \Rightarrow \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = k \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}, k \neq 0$$

When $\lambda = -2$

$$\begin{pmatrix} 3 & 3 & 3 \\ -3 & -3 & -3 \\ 3 & 3 & 3 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Rightarrow \left(\begin{array}{ccc|c} 3 & 3 & 3 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right) \sim \left(\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right) \Rightarrow x_3 = t, x_2 = s, x_1 = -s-t$$

$$\Rightarrow \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -s-t \\ s \\ t \end{pmatrix} = t\begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} + s\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, t^2+s^2 \neq 0$$

$A = \begin{pmatrix} 1 & 3 & 3 \\ -3 & -5 & -3 \\ 3 & 3 & 1 \end{pmatrix}$ has three linearly independent eigenvectors, therefore $A = \begin{pmatrix} 1 & 3 & 3 \\ -3 & -5 & -3 \\ 3 & 3 & 1 \end{pmatrix}$ is diagonalizable and $\begin{pmatrix} -1 & -1 & 1 \\ 0 & 1 & -1 \\ 1 & 0 & 1 \end{pmatrix}^{-1}\begin{pmatrix} 1 & 3 & 3 \\ -3 & -5 & -3 \\ 3 & 3 & 1 \end{pmatrix}\begin{pmatrix} -1 & -1 & 1 \\ 0 & 1 & -1 \\ 1 & 0 & 1 \end{pmatrix} = \begin{pmatrix} -2 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 1 \end{pmatrix}$.

□

Example

Diagonalisation is useful if it is required to raise a matrix A to a high power, that is, A^m. If $P^{-1}AP = D$, then $A = PDP^{-1}$ and

$A^m = (PDP^{-1})\cdots(PDP^{-1}) = PDP^{-1}PDP^{-1}\cdots PDP^{-1}PDP^{-1} = PDIDI\cdots IDIDP^{-1} = PD^mP^{-1}$

and

$$D = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & \cdots & \lambda_n \end{pmatrix} = \operatorname{diag}(\lambda_1, \cdots, \lambda_n) \Rightarrow D^m = \begin{pmatrix} \lambda_1^m & 0 & \cdots & 0 \\ 0 & \lambda_2^m & \cdots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & \cdots & \lambda_n^m \end{pmatrix} = \operatorname{diag}(\lambda_1^m, \cdots, \lambda_n^m)$$

□

An important application of diagonalisation occurs in the solution of systems of differential equations.

Example

Consider a system of n first order linear differential equations for the dependent variables $x_1, x_2, \cdots, x_n$ as follows.

$$\begin{cases} \frac{dx_1}{dt} = a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ \frac{dx_2}{dt} = a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \ldots\ldots(E), \\ \vdots \\ \frac{dx_n}{dt} = a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{mn}x_n \end{cases}$$ where the a 's are given constants. The system (E) can be wri

as the following matrix form $\frac{d\vec{u}}{dt} = \overrightarrow{Au} \cdots$ (F), where $\vec{u} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ and $A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ . & & & . \\ . & & & . \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}$.

Suppose A is diagonalizable, then there exists nonsingular matrix P such that $D = P^{-1}AP$, where D is a

Let $\begin{pmatrix} y_1(t) \\ \vdots \\ y_n(t) \end{pmatrix} = P^{-1} \begin{pmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{pmatrix}$, then

$$\frac{d}{dt}\begin{pmatrix} y_1(t) \\ \vdots \\ y_n(t) \end{pmatrix} = \frac{d}{dt}P^{-1}\begin{pmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{pmatrix} = P^{-1}\frac{d}{dt}\begin{pmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{pmatrix} = P^{-1}A\begin{pmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{pmatrix} = DP^{-1}\begin{pmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{pmatrix} = D\begin{pmatrix} y_1(t) \\ \vdots \\ y_n(t) \end{pmatrix}$$

Observe that D is diagonal so each row of the system is of the form $\frac{d}{dt}y_i(t) = d_j y_i(t)$.

The solution of $\frac{d}{dt}\begin{pmatrix} y_1(t) \\ \vdots \\ y_n(t) \end{pmatrix} = D\begin{pmatrix} y_1(t) \\ \vdots \\ y_n(t) \end{pmatrix}$ is easy to obtain. Once $\begin{pmatrix} y_1(t) \\ \vdots \\ y_n(t) \end{pmatrix}$ is known, then

$$\begin{pmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{pmatrix} = P\begin{pmatrix} y_1(t) \\ \vdots \\ y_n(t) \end{pmatrix}$$

□

<u>Example</u>

(a) Use Gauss-Jordan elimination to compute the inverse of $\begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ 0 & 0 & 1 \end{pmatrix}$.

(b) Let A be a 3×3 matrix which has eigenvalues $-1, 1, 0$ with corresponding eigenvectors $\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$. With the help of (a), find A^{10}.

Solution:

(a)

$$\left(\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 2 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{array}\right) \underset{r_2 - r_1}{\sim} \left(\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & -1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{array}\right) \underset{r_1 - r_3}{\sim} \left(\begin{array}{ccc|ccc} 1 & 1 & 0 & 1 & 0 & -1 \\ 0 & 1 & 0 & -1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{array}\right)$$

$$\underset{r_1-r_2}{\sim}\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 2 & -1 & -1 \\ 0 & 1 & 0 & -1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{array}\right)$$

So $\begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ 0 & 0 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 2 & -1 & -1 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

(b)

We have $A\begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$.

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ 0 & 0 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} 2 & -1 & -1 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$\Rightarrow A^{10} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}^{10}\begin{pmatrix} 2 & -1 & -1 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} =$$

$$\begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} (-1)^{10} & 0 & 0 \\ 0 & 1^{10} & 0 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} 2 & -1 & -1 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

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

□

Example

(a) Let B be a 3×3 matrix and $\lambda_1, \lambda_2, \lambda_3$ are the eigenvalues of B. Write down the characteristic equation of B and show that $\lambda_1\lambda_2\lambda_3 = \det B$.

(b) Let $A_{3\times3}$ be a real symmetric matrix. Suppose

$det A = 1, A\begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ -2 \\ -1 \end{pmatrix}, A\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} = \begin{pmatrix} -2 \\ 0 \\ 2 \end{pmatrix},$

(i) use part (a), or otherwise, to find the eigenvalues of A;

(ii) find a matrix P having its columns to be the unit eigenvectors of A and show that $P^TP = I$;

(iii) show that A is diagonalizable and find the matrix A.

Solution:

(a)

$$\det(\lambda I - B) = \det\begin{pmatrix} \lambda - b_{11} & -b_{12} & -b_{13} \\ -b_{21} & \lambda - b_{22} & -b_{23} \\ -b_{31} & -b_{32} & \lambda - b_{33} \end{pmatrix} = (\lambda - \lambda_1)(\lambda - \lambda_2)(\lambda - \lambda_3)$$

$= \lambda^3 - (\lambda_1 + \lambda_2 + \lambda_3)\lambda^2 + (\lambda_1\lambda_2 + \lambda_1\lambda_3 + \lambda_2\lambda_3)\lambda - \lambda_1\lambda_2\lambda_3$

Let $\lambda = 0$, then $\det(-B) = -\lambda_1\lambda_2\lambda_3 \Rightarrow (-1)^3 \det B = -\lambda_1\lambda_2\lambda_3 \Rightarrow \det B = \lambda_1\lambda_2\lambda_3$.

(b)

(i) $A\begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ -2 \\ -1 \end{pmatrix}, A\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} = \begin{pmatrix} -2 \\ 0 \\ 2 \end{pmatrix} \Rightarrow A\begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ -2 \\ -1 \end{pmatrix} = (-1)\begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}, A\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} = \begin{pmatrix} -2 \\ 0 \\ 2 \end{pmatrix} = (-2)\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$.

So we have, say $\lambda_1 = -1, \lambda_2 = -2$.

$\det A = 1 = \lambda_1\lambda_2\lambda_3, \lambda_1 = -1, \lambda_2 = -2 \Rightarrow 1 = 2\lambda_3 \Rightarrow \lambda_3 = \frac{1}{2}$. As a result, $\lambda_1 = -1, \lambda_2 = -2, \lambda_3 = \frac{1}{2}$.

(Observe that the real symmetric A has distinct simple (of multiplicity 1) eigenvalues, therefore, A has the orthonormal (orthogonal unit) eigenvectors which are linearly independent.)

(ii)

For $\lambda_1 = -1, A\begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} = (-1)\begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} \Rightarrow A\begin{pmatrix} \frac{1}{\sqrt{6}} \\ \frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \end{pmatrix} = (-1)\begin{pmatrix} \frac{1}{\sqrt{6}} \\ \frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \end{pmatrix}$.

For $\lambda_1 = -2, A\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} = (-2)\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \Rightarrow A\begin{pmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ -\frac{1}{\sqrt{2}} \end{pmatrix} = (-2)\begin{pmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ -\frac{1}{\sqrt{2}} \end{pmatrix}$.

Suppose $A\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \frac{1}{2}\begin{pmatrix} x \\ y \\ z \end{pmatrix}$. Since A is real symmetric, we have

$\begin{cases} \begin{pmatrix} x \\ y \\ z \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} = 0 \\ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} = 0 \end{cases} \Leftrightarrow \begin{cases} x + 2y + z = 0 \\ x - z = 0 \end{cases} \Leftrightarrow \begin{cases} y = -z \\ x = z \end{cases}$. Let $z = 1$, then $\begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$ is an eigenvector of A for $\lambda_3 = \frac{1}{2}$

and the corresponding unit eigenvector is $\begin{pmatrix} \frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{pmatrix}$.

Let $P = \begin{pmatrix} \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \\ \frac{2}{\sqrt{6}} & 0 & -\frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \end{pmatrix}$. Observe that $P^T = \begin{pmatrix} \frac{1}{\sqrt{6}} & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \end{pmatrix}$.

And $P^T P = \begin{pmatrix} \frac{1}{\sqrt{6}} & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \\ \frac{2}{\sqrt{6}} & 0 & -\frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$.

(Observe matrix P which is formed by the orthonormal eigenvectors of A has its inverse just its transpose) (iii)

$$A\begin{pmatrix} \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \\ \frac{2}{\sqrt{6}} & 0 & -\frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \\ \frac{2}{\sqrt{6}} & 0 & -\frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \end{pmatrix} \begin{pmatrix} -1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & \frac{1}{2} \end{pmatrix}$$

$$\Rightarrow A = \begin{pmatrix} \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \\ \frac{2}{\sqrt{6}} & 0 & -\frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \end{pmatrix} \begin{pmatrix} -1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & \frac{1}{2} \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \\ \frac{2}{\sqrt{6}} & 0 & -\frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \end{pmatrix}^{-1}$$

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

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

$$= \begin{pmatrix} -\frac{1}{6} - 1 + \frac{1}{6} & -\frac{2}{6} - \frac{1}{6} & -\frac{1}{6} + 1 + \frac{1}{6} \\ -\frac{2}{6} - \frac{1}{6} & -\frac{4}{6} + \frac{1}{6} & -\frac{2}{6} - \frac{1}{6} \\ -\frac{1}{6} + 1 + \frac{1}{6} & -\frac{2}{6} - \frac{1}{6} & -\frac{1}{6} - 1 + \frac{1}{6} \end{pmatrix} = \begin{pmatrix} -1 & -\frac{1}{2} & 1 \\ -\frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \\ 1 & -\frac{1}{2} & -1 \end{pmatrix}$$

$\square$

-End-

1 Quadratic Forms

1.1 Introduction

A quadratic form is an expression of the form

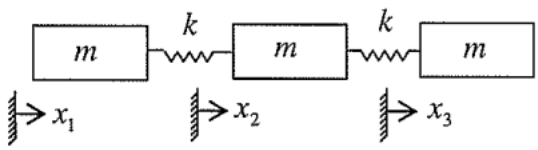
$$q = \sum_{\mathrm{i}=1}^{n} \sum_{\mathrm{j}=1}^{n} a_{\mathrm{i}j} x_{\mathrm{i}} x_{\mathrm{j}}$$

Example

$q = 3x_1^2 + 4x_1x_2 - x_2^2$

Example

The 3-mass, 2-spring system considered previously



has kinetic energy

$$T = m\left(\frac{x_1^2}{2} + \frac{x_2^2}{2} + \frac{x_3^2}{2}\right)$$

and potential energy

$$V = m\left[k\left(x_2 - x_1\right)^2/2 + k\left(x_3 - x_2\right)^2/2\right]$$

$$V = m\left[kx_1^2/2 + kx_2^2 + kx_3^2/2 - kx_1x_2 - kx_2x_3\right]$$

T is a quadratic form in x_1, x_2 and x_3 while V is a quadratic form in x_1, x_2 and x_3.

Example

The general equation of a conic is

$ax^2 + 2by + cy^2 + dx + ey + f = 0$

Associated with every quadratic form is a symmetric matrix, since

$$q = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j$$

May be written as

$$q = \underline{x}^T A \underline{x} \quad A = [a_{ij}]$$

Where $A = A^T$.

Example

$$\begin{aligned} q &= 3x_1^2 + 4x_1x_2 - x_2^2 \\ &= [x_1 x_2] \begin{bmatrix} 3 & 2 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \end{aligned}$$

Note however that it is also possible to write q in terms of non-symmetric matrices, eg.

$$q = [x_1 x_2] \begin{bmatrix} 3 & 4 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

or

$$q = [x_1 x_2] \begin{bmatrix} 3 & 1 \\ 3 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

etc.

Example

$$\begin{aligned} q &= x_1^2 + 2x_2^2 + 7x_3^2 - 2x_1x_2 + 4x_1x_3 - 2x_2x_3 \\ &= [x_1 x_2 x_3] \begin{bmatrix} 1 & -1 & 2 \\ -1 & 2 & -1 \\ 2 & -1 & 7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \end{aligned}$$

If in the general case $a_{\mathrm{ij}} \neq a_{\mathrm{ji}}$, we define $c_{\mathrm{ij}} = (a_{\mathrm{ij}} + a_{\mathrm{ji}})/2$, then $c_{\mathrm{ij}} = c_{\mathrm{ji}}$ and the coefficient of $(x_i x_j + x_j x_i)$ is $c_{ij} + c_{ji} = a_{ij} + a_{ji}$ as required. Then

$$q = \underline{x}^T C \underline{x} \text{ with } C = C^T$$

1.2 Definitieness

It is of interest in many applications to know if the sign of q remains the same whatever the values of the elements of $\underline{x}$.

$q = \underline{x}^T A \underline{x}$ is positive definite if $q > 0$ for all $\underline{x} \neq \underline{0}$

$q = \underline{x}^T A \underline{x}$ is <u>negative definite</u> if $q < 0$ for all $\underline{x} \neq \underline{0}$

[q is <u>positive semi-definite</u> if $q \geq 0$

q is <u>negative semi-definite</u> if $q \leq 0$

q is <u>indefinite</u> if it can take both positive and negative values]

The associated real symmetric matrix A is also said to be positive or negative definite if $q > 0$ or $q < 0$ respectively.

We may determine whether q has constant sign or not by "completing the square".

<u>Example</u>

$$
\begin{aligned}
q &= x_1^2 + 2x_2^2 + 7x_3^2 - 2x_1x_2 + 4x_1x_3 - 2x_2x_3 \\
&= (x_1 - x_2 + 2x_3)^2 + 2x_2^2 + 7x_3^2 - 2x_2x_3 - x_2^2 - 4x_3^2 + 4x_2x_3 \\
&= (x_1 - x_2 + 2x_3)^2 + x_2^2 + 3x_3^2 + 2x_2x_3 \\
&= (x_1 - x_2 + 2x_3)^2 + (x_2 + x_3)^2 + 2x_3^2
\end{aligned}
$$

Thus $q > 0$ for any choice of x_1, x_2, x_3 and both q

And the matrix $A = \begin{bmatrix} 1 & -1 & 2 \\ -1 & 2 & -1 \\ 2 & -1 & 7 \end{bmatrix}$ are positive definite.

<u>Example</u>

$q = 3x_1^2 + 4x_1x_2 - x_2^2$

$= 3\left(x_1^2 + 4x_1x_2/3 - x_2^2/3\right)$

$= 3\left[(x_1 + 2x_2/3)^2 - x_2^2/3 - 4x_2^2/9\right]$

$q = 3\left(x_1 + 2x_2/3\right)^2 - 7x_2^2/3$

Thus q and its associated symmetric matrix are indefinite.

We shall now reduce q to a sum of square using matrix and eigenvalue ideas.

Recall that for a real symmetric matrix it is always possible to find n linearly independent eigevectors. If the eigenvalues are distinct these eigenvectors will also be orthogonal. If they are not distinct it is in fact still possible to obtain n orthogonal eigenvectors (proof not required).

If these eigenvectors $\underline{x}_{\mathrm{i}}$ are normalized such that $\underline{x}_{\mathrm{i}}^T \underline{x}_{\mathrm{i}} = 1$ (orthonormal eigenvectors) then the matrix P whose columns are the orthonormalised eigenvectors will have the property that

$$P^T P = I$$

Since the ij-th element of $P^T P$ is

$$\underline{x}_{\mathrm{i}}^T \underline{x}_{\mathrm{j}} = \delta_{\mathrm{ij}} = \begin{cases} 1 & \mathrm{i} = \mathrm{j} \\ 0 & \mathrm{i} \neq \mathrm{j} \end{cases}$$

Hence $\quad P^T = P^{-1}$

Any matrix with this property is said to be orthogonal.

Thus it is possible to diagonalise a symmetric matrix using an orthogonal matrix P i.e.

$$P^T A P = D$$

Given a quadratic form

$$q = \underline{x}^T A \underline{x}$$

We may make the change of variable

$$\underline{x} = P\underline{y}$$

where P is the orthogonal matrix whose columns are the orthogonal eigenvectors of A to give,

$$q = (P\underline{y})^T A (P\underline{y})$$

$$= \underline{y}^T P^T A P \underline{y}$$

$$= \underline{y}^T D \underline{y}$$

$$q = \sum_{i=1}^{n} \lambda_{\mathrm{i}} {y_i}^2$$

This shows that q may be written as a sum of squares with the coefficients of the squared terms being the eigenvalues of A.

This gives the result:

Theorem

A quadratic form and its associated matrix are positive definite iff all the eigenvalues of the associated symmetric matrix are positive.

Example

Consider the quadratic form

$$
\begin{aligned}
q &= 2x_1^2 - 2x_1x_2 + 2x_2^2 \\
&= [x_1 x_2] \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \underline{x}^T A \underline{x}
\end{aligned}
$$

The eigenvalues and eigenvectors of A are

$\lambda_1 = 1 \quad \underline{x}_1 = [1,1]^T; \lambda_2 = 3 \quad \underline{x}_2 = [-1,1]^T$

The normalized eigenvectors are

$$
\begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix} \text{ and } \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}
$$

Giving

$$
P = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} \text{ and } P^T = P^{-1} = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}
$$

Then $P^TAP = \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix} = D$ and if we make

the change of variable $\underline{x} = P\underline{y}$ or $\underline{y} = P^{-1}\underline{x} = P^T\underline{x}$

$$
\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} (y_1 - y_2)/\sqrt{2} \\ (y_1 + y_2)/\sqrt{2} \end{bmatrix}
$$

we have

$$
q = y_1^2 + 3{y_2}^2 \quad \left(= \underline{y}^T D \underline{y}\right)
$$

showing that q and A are positive definite.

It may be verified directly that

$$q = 2x_1^2 - 2x_1x_2 + 2x_2^2$$

$$= 2\left(y_1 - y_2\right)^2/2 - 2\left(y_1 - y_2\right)/\sqrt{2}\left(y_1 + y_2\right)/\sqrt{2} + 2\left(y_1 + y_2\right)^2/2$$

$$q = y_1^2 + 3y_2^2$$

Or

$$q = \left(x_1 + x_2\right)^2/2 + 3\left(x_2 - x_1\right)^2/2$$

Note that "completing the square" does not necessarily give the same factorization. In this case we might have:

$$q = 2x_1^2 - 2x_1x_2 + 2x_2^2$$

$$= 2\left(x_1^2 - x_1x_2 + x_2^2\right)$$

$$= 2\left[\left(x_1 - x_2/2\right)^2 + x_2^2 - x_2^2/4\right]$$

$$q = 2\left(x_1 - x_2/2\right)^2 + 3x_2^2/2$$

<u>Example</u>

For any constant $c > 0$ the points (x_1, x_2) satisfying

$$2x_1^2 - 2x_1x_2 + 2x_2^2 = c$$

Form a conic in the plane.

With the change of variable given in the previous example, we see that the equation becomes

$$y_1^2 + 3y_2^2 = c$$

Which we recognize as the equation of an ellipse. The transformation

$$\underline{x} = P\underline{y}$$

where

$$P = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$$

represents a rotation of the axes through $45°$ (section 2.1)

Equivalent definitions of positive definiteness of a matrix A are:

A) All leading principle minors are positive; that is, let

$$A_1 = a_{11}, A_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}, A_3 = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix},$$

..., $A_n = \det(A)$ be the leading principle minors, then A is positive definite iff $A_1, A_2, \ldots, A_n$ are all positive.

B) A is positive definite iff when A has been reduced to echelon form (only by replacement row reduction), all the pivots (diagonal elements) are positive.

We shall not prove that these definitions are equivalent but only note that they can be useful tests in practice.