4- Matrix Analysis: (10hrs)

Review of matrix theory, linear transformation, Eigen values and Eigen vectors, Laplace transform of matrices, application of electric circuit.

Introduction:

In this chapter, we turn our attention again to matrices, first considered in student have the basic knowledge in Matrix Algebra their applications in engineering.

As the reader will be aware, matrices are arrays of real or complex numbers, and have a special, but not exclusive, relationship with systems of linear equations. An (incorrect) initial impression often formed by users of mathematics is that mathematicians have something of an obsession with these systems and their solution. However, such systems occur quite naturally in the process of numerical solution of ordinary differential equations used to model everyday engineering processes. Systems of linear first-order differential equations with constant coefficients are at the core of the state-space representation of linear system models. Identification, analysis and indeed design of such systems can conveniently be performed in the state-space representation, with this form assuming a particular importance in the case of multivariable systems.

In all these areas, it is convenient to use a matrix representation for the systems under Consideration, since this allows the system model to be manipulated following the rules of matrix algebra. A particularly valuable type of manipulation is simplification in some sense. Such a simplification process is an example of a system transformation, carried out by the process of matrix multiplication. At the heart of many transformations are the eigenvalues and eigenvectors of a square matrix. In addition to providing the means by which simplifying transformations can be deduced, system eigenvalues provide vital information on system stability, fundamental frequencies, speed of decay and long-term system behavior. For this reason, we devote a substantial amount of space to the process of their calculation, both by hand and by numerical means when necessary. Our treatment of numerical methods is intended to be purely indicative rather than complete, because a comprehensive matrix algebra computational tool kit, such as MATLAB, is now part of the essential armory of all serious users of mathematics.

In addition to developing the use of matrix algebra techniques, we also demonstrate the techniques and applications of matrix analysis, focusing on the state-space system model widely used in control and systems engineering. Here we encounter the idea of a function of a matrix, in particular the matrix exponential, and we see again the role of the eigenvalues in its calculation. This chapter also includes a section on singular value decomposition and the pseudo inverse, together with a brief section on Lyapunov stability of linear systems using quadratic forms.

1.2 Review of matrix algebra

This section contains a summary of the definitions and properties associated with matrices and determinants. A full account can be found in previous study. It is assumed that readers, prior to embarking on this chapter, have a fairly thorough understanding of the material summarized in this section.

Definitions

(a) An array of real numbers

$$\boldsymbol{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \dots & a_{mn} \end{bmatrix}$$

is called an $m \times n$ **matrix** with m rows and n columns. The a_{ij} is referred to as the i, jth **element** and denotes the element in the ith row and jth column. If $m = n$ then $\boldsymbol{A}$ is called a **square matrix** of order n. If the matrix has one column or one row then it is called a **column vector** or a **row vector** respectively.

(b) In a square matrix $\boldsymbol{A}$ of order n the diagonal containing the elements $a_{11}, a_{22}, \dots, a_{nn}$ is called the **principal** or **leading** diagonal. The sum of the elements in this diagonal is called the **trace** of $\boldsymbol{A}$, that is

$$\text{trace}\,\boldsymbol{A} = \sum_{i=1}^{n} a_{ii}$$

(c) A **diagonal matrix** is a square matrix that has its only non-zero elements along the leading diagonal. A special case of a diagonal matrix is the **unit** or **identity** matrix $\boldsymbol{I}$ for which $a_{11} = a_{22} = \dots = a_{nn} = 1$.

(d) A **zero** or **null** matrix **0** is a matrix with every element zero.

(e) The **transposed matrix** $\boldsymbol{A}^{\mathrm{T}}$ is the matrix $\boldsymbol{A}$ with rows and columns interchanged, its i, jth element being a_{ji}.

(f) A square matrix $\boldsymbol{A}$ is called a **symmetric matrix** if $\boldsymbol{A}^{\mathrm{T}} = \boldsymbol{A}$. It is called **skew symmetric** if $\boldsymbol{A}^{\mathrm{T}} = -\boldsymbol{A}$.

1.2.2 Basic operations on matrices

In what follows the matrices A, B and C are assumed to have the i, jth elements a_{ij}, b_{ij} and c_{ij} respectively.

Equality

The matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ are **equal**, that is $\boldsymbol{A} = \boldsymbol{B}$, if they are of the same order $m \times n$ and

$$a_{ij} = b_{ij}, \quad 1 \leqslant i \leqslant m, \quad 1 \leqslant j \leqslant n$$

Multiplication by a scalar

If λ is a scalar then the matrix $\lambda\boldsymbol{A}$ has elements λa_{ij}.

Addition

We can only add an $m \times n$ matrix $\boldsymbol{A}$ to another $m \times n$ matrix $\boldsymbol{B}$ and the elements of the sum $\boldsymbol{A} + \boldsymbol{B}$ are

$$a_{ij} + b_{ij}, \quad 1 \leqslant i \leqslant m, \quad 1 \leqslant j \leqslant n$$

Properties of addition

(i) commutative law: $\boldsymbol{A} + \boldsymbol{B} = \boldsymbol{B} + \boldsymbol{A}$

(ii) associative law: $(\boldsymbol{A} + \boldsymbol{B}) + \boldsymbol{C} = \boldsymbol{A} + (\boldsymbol{B} + \boldsymbol{C})$

(iii) distributive law: $\lambda(\boldsymbol{A} + \boldsymbol{B}) = \lambda\boldsymbol{A} + \lambda\boldsymbol{B}$, λ scalar

Matrix multiplication

If $\boldsymbol{A}$ is an $m \times p$ matrix and $\boldsymbol{B}$ a $p \times n$ matrix then we define the product $\boldsymbol{C} = \boldsymbol{AB}$ as the $m \times n$ matrix with elements

$$c_{ij} = \sum_{k=1}^{p} a_{ik} b_{kj}, \quad i = 1, 2, \ldots, m; \quad j = 1, 2, \ldots, n$$

Properties of multiplication

(i) The commutative law is **not satisfied** in general; that is, in general $\boldsymbol{AB} \neq \boldsymbol{BA}$. Order matters and we distinguish between $\boldsymbol{AB}$ and $\boldsymbol{BA}$ by the terminology: **pre**-multiplication of $\boldsymbol{B}$ by $\boldsymbol{A}$ to form $\boldsymbol{AB}$ and **post**-multiplication of $\boldsymbol{B}$ by $\boldsymbol{A}$ to form $\boldsymbol{BA}$.

(ii) Associative law: $\boldsymbol{A}(\boldsymbol{BC}) = (\boldsymbol{AB})\boldsymbol{C}$

(iii) If λ is a scalar then

$(\lambda\boldsymbol{A})\boldsymbol{B} = \boldsymbol{A}(\lambda\boldsymbol{B}) = \lambda\boldsymbol{AB}$

(iv) Distributive law over addition:

$(\boldsymbol{A} + \boldsymbol{B})\boldsymbol{C} = \boldsymbol{AC} + \boldsymbol{BC}$

$\boldsymbol{A}(\boldsymbol{B} + \boldsymbol{C}) = \boldsymbol{AB} + \boldsymbol{AC}$

Note the importance of maintaining order of multiplication.

(v) If $\boldsymbol{A}$ is an $m \times n$ matrix and if $\boldsymbol{I}_m$ and $\boldsymbol{I}_n$ are the unit matrices of order m and n respectively then

$\boldsymbol{I}_m\boldsymbol{A} = \boldsymbol{A}\boldsymbol{I}_n = \boldsymbol{A}$

Properties of the transpose

If $\boldsymbol{A}^{\mathrm{T}}$ is the transposed matrix of $\boldsymbol{A}$ then

(i) $(\boldsymbol{A} + \boldsymbol{B})^{\mathrm{T}} = \boldsymbol{A}^{\mathrm{T}} + \boldsymbol{B}^{\mathrm{T}}$

(ii) $(\boldsymbol{A}^{\mathrm{T}})^{\mathrm{T}} = \boldsymbol{A}$

(iii) $(\boldsymbol{AB})^{\mathrm{T}} = \boldsymbol{B}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}$

1.2.3 Determinants

The determinant of a square $n \times n$ matrix $\boldsymbol{A}$ is denoted by det $\boldsymbol{A}$ or $|\boldsymbol{A}|$.

If we take a determinant and delete row i and column j then the determinant remaining is called the **minor** M_{ij} of the i, jth element. In general we can take any row i (or column) and evaluate an $n \times n$ determinant $|\boldsymbol{A}|$ as

$$|\boldsymbol{A}| = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} M_{ij}$$

A minor multiplied by the appropriate sign is called the **cofactor** A_{ij} of the i, jth element so $A_{ij} = (-1)^{i+j} M_{ij}$ and thus

$$|\boldsymbol{A}| = \sum_{j=1}^{n} a_{ij} A_{ij}$$

Some useful properties

(i) $|\boldsymbol{A}^{\mathrm{T}}| = |\boldsymbol{A}|$

(ii) $|\boldsymbol{AB}| = |\boldsymbol{A}||\boldsymbol{B}|$

(iii) A square matrix $\boldsymbol{A}$ is said to be **non-singular** if $|\boldsymbol{A}| \neq 0$ and **singular** if $|\boldsymbol{A}| = 0$.

1.2.4 Adjoint and inverse matrices

Adjoint matrix

The **adjoint** of a square matrix $\boldsymbol{A}$ is the transpose of the matrix of cofactors, so for a 3×3 matrix $\boldsymbol{A}$

$$\operatorname{adj} \boldsymbol{A} = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix}^{\mathrm{T}}$$

Properties

(i) $\boldsymbol{A}(\operatorname{adj} \boldsymbol{A}) = |\boldsymbol{A}|\boldsymbol{I}$

(ii) $|\operatorname{adj} \boldsymbol{A}| = |\boldsymbol{A}|^{n-1}$, n being the order of $\boldsymbol{A}$

(iii) $\operatorname{adj}(\boldsymbol{AB}) = (\operatorname{adj} \boldsymbol{B})(\operatorname{adj} \boldsymbol{A})$

Inverse matrix

Given a square matrix $\boldsymbol{A}$ if we can construct a square matrix $\boldsymbol{B}$ such that

$$\boldsymbol{BA} = \boldsymbol{AB} = \boldsymbol{I}$$

then we call $\boldsymbol{B}$ the inverse of $\boldsymbol{A}$ and write it as $\boldsymbol{A}^{-1}$.

Properties

(i) If $\boldsymbol{A}$ is non-singular then $|\boldsymbol{A}| \neq 0$ and $\boldsymbol{A}^{-1} = (\operatorname{adj} \boldsymbol{A})/|\boldsymbol{A}|$.

(ii) If $\boldsymbol{A}$ is singular then $|\boldsymbol{A}| = 0$ and $\boldsymbol{A}^{-1}$ does not exist.

(iii) $(\boldsymbol{AB})^{-1} = \boldsymbol{B}^{-1}\boldsymbol{A}^{-1}$.

1.2.5 Linear equations

In this section we reiterate some definitive statements about the solution of the system of simultaneous linear equations

$$\begin{aligned}
a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= b_1 \\
a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n &= b_2 \\
\vdots \qquad\qquad\qquad & \vdots \\
a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n &= b_n
\end{aligned}$$

or, in matrix notation,

$$\begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

that is,

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$$

where $\boldsymbol{A}$ is the matrix of coefficients and $\boldsymbol{x}$ is the vector of unknowns. If $\boldsymbol{b} = 0$ the equations are called **homogeneous**, while if $\boldsymbol{b} \neq 0$ they are called **nonhomogeneous** (or **inhomogeneous**). Considering individual cases:

Case (i)

If $\boldsymbol{b} \neq 0$ and $|\boldsymbol{A}| \neq 0$ then we have a unique solution $\boldsymbol{x} = \boldsymbol{A}^{-1}\boldsymbol{b}$.

Case (ii)

If $\boldsymbol{b} = 0$ and $|\boldsymbol{A}| \neq 0$ we have the trivial solution $\boldsymbol{x} = 0$.

Case (iii)

If $\boldsymbol{b} \neq 0$ and $|\boldsymbol{A}| = 0$ then we have two possibilities: **either** the equations are inconsistent and we have no solution **or** we have infinitely many solutions.

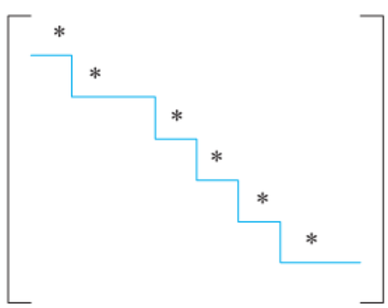
Case (iv)

If $\boldsymbol{b} = 0$ and $|\boldsymbol{A}| = 0$ then we have infinitely many solutions.

Case (iv) is one of the most important, since from it we can deduce the important result that **the homogeneous equation $\boldsymbol{A}\boldsymbol{x} = 0$ has a non-trivial solution if and only if $|\boldsymbol{A}| = 0$.**

1.2.6 Rank of a matrix

The most commonly used definition of the **rank**, rank $\boldsymbol{A}$, of a matrix $\boldsymbol{A}$ is that it is the order of the largest square submatrix of $\boldsymbol{A}$ with a non-zero determinant, a square submatrix being formed by deleting rows and columns to form a square matrix. Unfortunately it is not always easy to compute the rank using this definition and an alternative definition, which provides a constructive approach to calculating the rank, is often adopted. First, using elementary row operations, the matrix $\boldsymbol{A}$ is reduced to **echelon form**



in which all the entries below the line are zero, and the leading element, marked *, in each row above the line is non-zero. The number of non-zero rows in the echelon form is equal to rank $\mathbf{A}$.

When considering the solution of equations (1.1) we saw that provided the determinant of the matrix $\mathbf{A}$ was not zero we could obtain explicit solutions in terms of the inverse matrix. However, when we looked at cases with zero determinant the results were much less clear. The idea of the rank of a matrix helps to make these results more precise. Defining the **augmented matrix** $(\mathbf{A} : \mathbf{b})$ for (1.1) as the matrix $\mathbf{A}$ with the column $\mathbf{b}$ added to it then we can state the results of cases (iii) and (iv) of Section 1.2.5 more clearly as follows:

If $\mathbf{A}$ and $(\mathbf{A} : \mathbf{b})$ have different rank then we have no solution to (1.1). If the two matrices have the same rank then a solution exists, and furthermore the solution will contain a number of free parameters equal to $(n - \text{rank } \mathbf{A})$.

1.3.1 Linear independence

The idea of linear dependence is a general one for any vector space. The vector $\boldsymbol{x}$ is said to be **linearly dependent** on $\boldsymbol{x}_1, \boldsymbol{x}_2, \dots, \boldsymbol{x}_m$ if it can be written as

$$\boldsymbol{x} = \alpha_1\boldsymbol{x}_1 + \alpha_2\boldsymbol{x}_2 + \dots + \alpha_m\boldsymbol{x}_m$$

for some scalars $\alpha_1, \dots, \alpha_m$. The **set of vectors** $\boldsymbol{y}_1, \boldsymbol{y}_2, \dots, \boldsymbol{y}_m$ is said to be **linearly independent** if and only if

$$\beta_1\boldsymbol{y}_1 + \beta_2\boldsymbol{y}_2 + \dots + \beta_m\boldsymbol{y}_m = 0$$

implies that $\beta_1 = \beta_2 = \dots = \beta_m = 0$.

Let us now take a linearly independent set of vectors $\boldsymbol{x}_1, \boldsymbol{x}_2, \dots, \boldsymbol{x}_m$ in V and construct a set consisting of all vectors of the form

$$\boldsymbol{x} = \alpha_1\boldsymbol{x}_1 + \alpha_2\boldsymbol{x}_2 + \dots + \alpha_m\boldsymbol{x}_m$$

We shall call this set $S(\boldsymbol{x}_1, \boldsymbol{x}_2, \dots, \boldsymbol{x}_m)$. It is clearly a vector space, since all the axioms are satisfied.

1.4 The eigenvalue problem

A problem that leads to a concept of crucial importance in many branches of mathematics and its applications is that of seeking non-trivial solutions $x \neq 0$ to the matrix Equation

$$A\,x = \lambda\,x$$

This is referred to as the eigenvalue problem; values of the scalar λ for which non-trivial solutions exist are called eigenvalues and the corresponding solutions $x \neq 0$ are called the eigenvectors. Such problems arise naturally in many branches of engineering. For example, in vibrations the eigenvalues and eigenvectors describe the frequency and mode of vibration respectively, while in mechanics they represent principal stresses and the principal axes of stress in bodies subjected to external forces.

1.4.1 The characteristic equation

The set of simultaneous equations

$$\boldsymbol{A}\boldsymbol{x} = \lambda \boldsymbol{x} \tag{1.4}$$

where $\boldsymbol{A}$ is an $n \times n$ matrix and $\boldsymbol{x} = [x_1 \quad x_2 \quad \dots \quad x_n]^T$ is an $n \times 1$ column vector can be written in the form

$$(\lambda \boldsymbol{I} - \boldsymbol{A})\boldsymbol{x} = 0 \tag{1.5}$$

where $\boldsymbol{I}$ is the identity matrix. The matrix equation (1.5) represents simply a set of homogeneous equations, and we know that a non-trivial solution exists if

$$c(\lambda) = |\lambda \boldsymbol{I} - \boldsymbol{A}| = 0 \tag{1.6}$$

Here $c(\lambda)$ is the expansion of the determinant and is a polynomial of degree n in λ, called the **characteristic polynomial** of $\boldsymbol{A}$. Thus

$$c(\lambda) = \lambda^n + c_{n-1}\lambda^{n-1} + c_{n-2}\lambda^{n-2} + \dots + c_1\lambda + c_0$$

and the equation $c(\lambda) = 0$ is called the **characteristic equation** of $\boldsymbol{A}$. We note that this equation can be obtained just as well by evaluating $|\boldsymbol{A} - \lambda \boldsymbol{I}| = 0$; however, the form (1.6) is preferred for the definition of the characteristic equation, since the coefficient of λ^n is then always +1.

In many areas of engineering, particularly in those involving vibration or the control of processes, the determination of those values of λ for which (1.5) has a non-trivial solution (that is, a solution for which $\boldsymbol{x} \neq 0$) is of vital importance. These values of λ are precisely the values that satisfy the characteristic equation, and are called the **eigenvalues** of $\boldsymbol{A}$.

Example 1.2 Find the characteristic equation for the matrix

$$\boldsymbol{A} = \begin{bmatrix} 1 & 1 & -2 \\ -1 & 2 & 1 \\ 0 & 1 & -1 \end{bmatrix}$$

For matrices of large order, determining the characteristic polynomial by direct expansion of $|\lambda\boldsymbol{I} - \boldsymbol{A}|$ is unsatisfactory in view of the large number of terms involved in the determinant expansion. Alternative procedures are available to reduce the amount of calculation, and that due to Faddeev may be stated as follows.

The method of Faddeev

If the characteristic polynomial of an $n \times n$ matrix $\boldsymbol{A}$ is written as

$$\lambda^n - p_1\lambda^{n-1} - \ldots - p_{n-1}\lambda - p_n$$

then the coefficients $p_1, p_2, \ldots, p_n$ can be computed using

$$p_r = \frac{1}{r}\operatorname{trace} \boldsymbol{A}_r \quad (r = 1, 2, \ldots, n)$$

where

$$\boldsymbol{A}_r = \begin{cases} \boldsymbol{A} & (r = 1) \\ \boldsymbol{A}\boldsymbol{B}_{r-1} & (r = 2, 3, \ldots, n) \end{cases}$$

and

$$\boldsymbol{B}_r = \boldsymbol{A}_r - p_r\boldsymbol{I}, \quad \text{where } \boldsymbol{I} \text{ is the } n \times n \text{ identity matrix}$$

The calculations may be checked using the result that

$$\boldsymbol{B}_n = \boldsymbol{A}_n - p_n\boldsymbol{I} \quad \text{must be the zero matrix}$$

Example 1.3: Using the method of Faddeev, obtain the characteristic equation of the matrix A of Example 1.2.

1.4.2 Eigenvalues and eigenvectors

The roots of the characteristic equation ( ) are called the eigenvalues of the matrix A (the terms latent roots, proper roots and characteristic roots are also sometimes used). By the Fundamental Theorem of Algebra, a polynomial equation of degree n has exactly n roots, so that the matrix A has exactly n eigenvalues λ_i, i = 1, 2, . . . , n. These eigenvalues may be real or complex, and not necessarily distinct. Corresponding to each eigenvalue λ_i, there is a non-zero solution $x = e_i$ of ( ); e_i is called the eigenvector of A corresponding to the eigenvalue λ_i. (Again the terms latent vector, proper vector and characteristic vector are sometimes seen, but are generally obsolete.) We note that if $x = e_i$ satisfies ( ) then any scalar multiple β_i

e_i of e_i also satisfies ( ), so that the eigenvector e_i may only be determined to within a scalar multiple.

Example 1.4: Determine the eigenvalues and eigenvectors for the matrix A.

$$\boldsymbol{A} = \begin{bmatrix} 1 & 1 & -2 \\ -1 & 2 & 1 \\ 0 & 1 & -1 \end{bmatrix}$$

Example 1.5 Find the eigenvalues and eigenvectors of

$$\boldsymbol{A} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

1.4.4 Repeated eigenvalues

In the examples considered so far the eigenvalues λ_i $(i = 1, 2, \ldots)$ of the matrix $\boldsymbol{A}$ have been distinct, and in such cases the corresponding eigenvectors can be found and are linearly independent. The matrix $\boldsymbol{A}$ is then said to have a full set of linearly independent eigenvectors. It is clear that the roots of the characteristic polynomial $c(\lambda)$ may not all be distinct; and when $c(\lambda)$ has $p \leqslant n$ distinct roots, $c(\lambda)$ may be factorized as

$$c(\lambda) = (\lambda - \lambda_1)^{m_1}(\lambda - \lambda_2)^{m_2} \ldots (\lambda - \lambda_p)^{m_p}$$

indicating that the root $\lambda = \lambda_i$, $i = 1, 2, \ldots, p$, is a root of order m_i, where the integer m_i is called the **algebraic multiplicity** of the eigenvalue λ_i. Clearly $m_1 + m_2 + \ldots + m_p = n$. When a matrix $\boldsymbol{A}$ has repeated eigenvalues, the question arises as to whether it is possible to obtain a full set of linearly independent eigenvectors for $\boldsymbol{A}$. We first consider two examples to illustrate the situation.

Example 1.6 Determine the eigenvalues and corresponding eigenvectors of the matrix

$$\boldsymbol{A} = \begin{bmatrix} 3 & -3 & 2 \\ -1 & 5 & -2 \\ -1 & 3 & 0 \end{bmatrix}$$

Example 1.7 Determine the eigenvalues and corresponding eigenvectors for the matrix

$$\boldsymbol{A} = \begin{bmatrix} 1 & 2 & 2 \\ 0 & 2 & 1 \\ -1 & 2 & 2 \end{bmatrix}$$

1.4.6 Some useful properties of eigenvalues

The following basic properties of the eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_n$ of an $n \times n$ matrix $\boldsymbol{A}$ are sometimes useful. The results are readily proved either from the definition of eigenvalues as the values of λ satisfying (1.4), or by comparison of corresponding characteristic polynomials (1.6). Consequently, the proofs are left to Exercise 10.

Property 1.1

The sum of the eigenvalues of $\boldsymbol{A}$ is

$$\sum_{i=1}^{n} \lambda_i = \text{trace}\,\boldsymbol{A} = \sum_{i=1}^{n} a_{ii}$$

Property 1.2

The product of the eigenvalues of $\boldsymbol{A}$ is

$$\prod_{i=1}^{n} \lambda_i = \det \boldsymbol{A}$$

where $\det\boldsymbol{A}$ denotes the determinant of the matrix $\boldsymbol{A}$.

Property 1.3

The eigenvalues of the inverse matrix $\boldsymbol{A}^{-1}$, provided it exists, are

$$\frac{1}{\lambda_1},\quad \frac{1}{\lambda_2},\quad \dots,\quad \frac{1}{\lambda_n}$$

Property 1.4

The eigenvalues of the transposed matrix $\boldsymbol{A}^{\mathrm{T}}$ are

$$\lambda_1,\quad \lambda_2,\quad \dots,\quad \lambda_n$$

as for the matrix $\boldsymbol{A}$.

Property 1.5

If k is a scalar then the eigenvalues of $k\boldsymbol{A}$ are

$$k\lambda_1, \quad k\lambda_2, \quad \ldots, \quad k\lambda_n$$

Property 1.6

If k is a scalar and $\boldsymbol{I}$ the $n \times n$ identity (unit) matrix then the eigenvalues of $\boldsymbol{A} \pm k\boldsymbol{I}$ are respectively

$$\lambda_1 \pm k, \quad \lambda_2 \pm k, \quad \ldots, \quad \lambda_n \pm k$$

Property 1.7

If k is a positive integer then the eigenvalues of $\boldsymbol{A}^k$ are

$$\lambda_1^k, \quad \lambda_2^k, \quad \ldots, \quad \lambda_n^k$$

Example 1.9 Obtain the eigenvalues and corresponding orthogonal eigenvectors of the symmetric matrix

$$\boldsymbol{A} = \begin{bmatrix} 2 & 2 & 0 \\ 2 & 5 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

and show that the normalized eigenvectors form an orthonormal set.

1.4.7 Symmetric matrices

A square matrix $\boldsymbol{A}$ is said to be **symmetric** if $\boldsymbol{A}^{\mathrm{T}} = \boldsymbol{A}$. Such matrices form an important class and arise in a variety of practical situations. Two important results concerning the eigenvalues and eigenvectors of such matrices are

(a) the eigenvalues of a real symmetric matrix are real;

(b) for an $n \times n$ real symmetric matrix it is always possible to find n linearly independent eigenvectors $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n$ that are mutually orthogonal so that $\boldsymbol{e}_i^{\mathrm{T}}\boldsymbol{e}_i = 0$ for $i \neq j$.

If the orthogonal eigenvectors of a symmetric matrix are normalized as

$$\hat{\boldsymbol{e}}_1, \hat{\boldsymbol{e}}_2, \ldots, \hat{\boldsymbol{e}}_n$$

then the **inner (scalar) product** is

$$\hat{\boldsymbol{e}}_i^{\mathrm{T}}\hat{\boldsymbol{e}}_j = \delta_{ij} \quad (i, j = 1, 2, \ldots, n)$$

where δ_{ij} is the Kronecker delta defined in Section 1.3.2.

The set of normalized eigenvectors of a symmetric matrix therefore forms an orthonormal set (that is, it forms a mutually orthogonal normalized set of vectors).

1.6 Reduction to canonical form

In this section we examine the process of reduction of a matrix to canonical form.

Specifically, we examine methods by which certain square matrices can be reduced or transformed into diagonal form. The process of transformation can be thought of as a change of system coordinates, with the new coordinate axes chosen in such a way that the system can be expressed in a simple form. The simplification may, for example, be a transformation to principal axes or a decoupling of system equations. We will see that not all matrices can be reduced to diagonal form. In some cases we can only achieve the so-called Jordan canonical form, but many of the advantages of the diagonal form can be extended to this case as well.

The transformation to diagonal form is just one example of a similarity transform. Other such transforms exist, but, in common with the transformation to diagonal form, their purpose is usually that of simplifying the system model in some way.

1.6.1 Reduction to diagonal form

For an $n \times n$ matrix $\boldsymbol{A}$ possessing a full set of n linearly independent eigenvectors $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n$ we can write down a **modal matrix** $\boldsymbol{M}$ having the n eigenvectors as its columns:

$$\boldsymbol{M} = [\boldsymbol{e}_1 \quad \boldsymbol{e}_2 \quad \boldsymbol{e}_3 \quad \ldots \quad \boldsymbol{e}_n]$$

The diagonal matrix having the eigenvalues of $\boldsymbol{A}$ as its diagonal elements is called the **spectral matrix** corresponding to the modal matrix $\boldsymbol{M}$ of $\boldsymbol{A}$, often denoted by $\boldsymbol{\Lambda}$. That is,

$$\boldsymbol{\Lambda} = \begin{bmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{bmatrix}$$

with the ijth element being given by $\lambda_i \delta_{ij}$, where δ_{ij} is the Kronecker delta and $i, j = 1, 2, \ldots, n$. It is important in the work that follows that the pair of matrices $\boldsymbol{M}$ and $\boldsymbol{\Lambda}$ are written down correctly. If the ith column of $\boldsymbol{M}$ is the eigenvector $\boldsymbol{e}_i$ then the element in the (i, i) position in $\boldsymbol{\Lambda}$ must be λ_i, the eigenvalue corresponding to the eigenvector $\boldsymbol{e}_i$.

Example 1.14 Obtain a modal matrix and the corresponding spectral matrix for the matrix $\boldsymbol{A}$ of

$$\boldsymbol{A} = \begin{bmatrix} 1 & 1 & -2 \\ -1 & 2 & 1 \\ 0 & 1 & -1 \end{bmatrix}$$

Returning to the general case, if we premultiply the matrix $\boldsymbol{M}$ by $\boldsymbol{A}$, we obtain

$$\begin{aligned}\boldsymbol{AM} &= \boldsymbol{A}[\boldsymbol{e}_1 \quad \boldsymbol{e}_2 \quad \dots \quad \boldsymbol{e}_n] = [\boldsymbol{Ae}_1 \quad \boldsymbol{Ae}_2 \quad \dots \quad \boldsymbol{Ae}_n] \\ &= [\lambda_1\boldsymbol{e}_1 \quad \lambda_2\boldsymbol{e}_2 \quad \dots \quad \lambda_n\boldsymbol{e}_n]\end{aligned}$$

so that

$$\boldsymbol{AM} = \boldsymbol{M\Lambda} \tag{1.18}$$

Since the n eigenvectors $\boldsymbol{e}_1, \boldsymbol{e}_2, \dots, \boldsymbol{e}_n$ are linearly independent, the matrix $\boldsymbol{M}$ is non-singular, so that $\boldsymbol{M}^{-1}$ exists. Thus premultiplying by $\boldsymbol{M}^{-1}$ gives

$$\boldsymbol{M}^{-1}\boldsymbol{AM} = \boldsymbol{M}^{-1}\boldsymbol{M\Lambda} = \boldsymbol{\Lambda} \tag{1.19}$$

indicating that the similarity transformation $\boldsymbol{M}^{-1}\boldsymbol{AM}$ reduces the matrix $\boldsymbol{A}$ to the **diagonal** or **canonical form** $\boldsymbol{\Lambda}$. Thus a matrix $\boldsymbol{A}$ possessing a full set of linearly independent eigenvectors is reducible to diagonal form, and the reduction process is often referred to as the **diagonalization** of the matrix $\boldsymbol{A}$. Since

$$\boldsymbol{A} = \boldsymbol{M\Lambda M}^{-1} \tag{1.20}$$

it follows that $\boldsymbol{A}$ is uniquely determined once the eigenvalues and corresponding eigenvectors are known. Note that knowledge of the eigenvalues and eigenvectors alone is not sufficient: in order to structure $\boldsymbol{M}$ and $\boldsymbol{\Lambda}$ correctly, the association of eigenvalues and the *corresponding* eigenvectors must also be known.

Example : Verify results (1.19) and (1.20) for the matrix A

$$\boldsymbol{A} = \begin{bmatrix} 1 & 1 & -2 \\ -1 & 2 & 1 \\ 0 & 1 & -1 \end{bmatrix}$$

1.7 Functions of a matrix

Let $\boldsymbol{A}$ be an $n \times n$ constant square matrix, so that

$$\boldsymbol{A}^2 = \boldsymbol{AA},\ \boldsymbol{A}^3 = \boldsymbol{AA}^2 = \boldsymbol{A}^2\boldsymbol{A}, \qquad \text{and so on}$$

are all defined. We can then define a function $f(\boldsymbol{A})$ of the matrix $\boldsymbol{A}$ using a power series representation. For example,

$$f(\boldsymbol{A}) = \sum_{r=0}^{p} \beta_r \boldsymbol{A}^r = \beta_0 \boldsymbol{I} + \beta_1 \boldsymbol{A} + \ldots + \beta_p \boldsymbol{A}^p \tag{1.26}$$

where we have interpreted $\boldsymbol{A}^0$ as the $n \times n$ identity matrix $\boldsymbol{I}$.

Example 1.25 Given the 2×2 square matrix

$$\boldsymbol{A} = \begin{bmatrix} 1 & -1 \\ 2 & 3 \end{bmatrix}$$

determine $f(\boldsymbol{A}) = \sum_{r=0}^{2} \beta_r \boldsymbol{A}^r$ when $\beta_0 = 1$, $\beta_1 = -1$ and $\beta_2 = 3$.

Theorem 1.3 **Cayley–Hamilton theorem**

A square matrix $\boldsymbol{A}$ satisfies its own characteristic equation; that is, if

$$\lambda^n + c_{n-1}\lambda^{n-1} + c_{n-2}\lambda^{n-2} + \ldots + c_1\lambda + c_0 = 0$$

is the characteristic equation of an $n \times n$ matrix $\boldsymbol{A}$ then

$$\boldsymbol{A}^n + c_{n-1}\boldsymbol{A}^{n-1} + c_{n-2}\boldsymbol{A}^{n-2} + \ldots + c_1\boldsymbol{A} + c_0\boldsymbol{I} = \boldsymbol{0} \qquad \textbf{(1.28)}$$

where $\boldsymbol{I}$ is the $n \times n$ identity matrix.

Example 1.27 Verify the Cayley–Hamilton theorem for the matrix

$$\boldsymbol{A} = \begin{bmatrix} 3 & 4 \\ 1 & 2 \end{bmatrix}$$

Example 1.28 Given that the matrix

$$\boldsymbol{A} = \begin{bmatrix} 0 & 1 \\ -2 & -3 \end{bmatrix}$$

has eigenvalues $\lambda_1 = -1$ and $\lambda_2 = -2$ calculate $\boldsymbol{A}^5$ and $\boldsymbol{A}^r$, where r is an integer greater than 2.

Example 1.29 Given that the matrix

$$\boldsymbol{A} = \begin{bmatrix} 0 & 1 \\ -1 & -2 \end{bmatrix}$$

has eigenvalues $\lambda_1 = \lambda_2 = -1$, determine $\boldsymbol{A}^r$, where r is an integer greater than 2.

Example 1.30 Calculate e^{At} and $\sin \boldsymbol{A}t$ when

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

From this we can deduce that for an $n \times n$ matrix $\boldsymbol{A}$ we may write

$$f(\boldsymbol{A}) = \sum_{r=0}^{\infty} \beta_r \boldsymbol{A}^r$$

as

$$f(\boldsymbol{A}) = \sum_{r=0}^{n-1} \alpha_r \boldsymbol{A}^r \qquad \textbf{(1.34a)}$$

which generalizes the result (1.33). Again the coefficients $\alpha_0, \alpha_1, \dots, \alpha_{n-1}$ are obtained by solving the n equations

$$f(\lambda_i) = \sum_{r=0}^{n-1} \alpha_r \lambda_i^r \quad (i = 1, 2, \dots, n) \tag{1.34b}$$

where $\lambda_1, \lambda_2, \dots, \lambda_n$ are the eigenvalues of $\boldsymbol{A}$. If $\boldsymbol{A}$ has repeated eigenvalues, we differentiate as before, noting that if λ_i is an eigenvalue of multiplicity m then the first $m - 1$ derivatives

$$\frac{\mathrm{d}^k}{\mathrm{d}\lambda_i^k} f(\lambda_i) = \frac{\mathrm{d}^k}{\mathrm{d}\lambda_i^k} \sum_{r=0}^{n-1} \alpha_r \lambda_i^r \quad (k = 1, 2, \dots, m-1)$$

are also satisfied by λ_i.

Example 1.31 Using the result (1.35), calculate $\boldsymbol{A}^k$ for the matrix

$$\boldsymbol{A} = \begin{bmatrix} 0 & 1 \\ -2 & -3 \end{bmatrix}$$

$$\frac{d}{dt}(e^{\boldsymbol{A}t}) = \boldsymbol{A}e^{\boldsymbol{A}t} = e^{\boldsymbol{A}t}\boldsymbol{A}$$

$$\int_0^t e^{\boldsymbol{A}\tau}d\tau = \boldsymbol{A}^{-1}[e^{\boldsymbol{A}t} - \boldsymbol{I}] = [e^{\boldsymbol{A}t} - \boldsymbol{I}]\boldsymbol{A}^{-1}$$

$$e^{\boldsymbol{A}(t_1+t_2)} = e^{\boldsymbol{A}t_1}e^{\boldsymbol{A}t_2}$$

It follows from the power series definition that

$$e^{\boldsymbol{A}t}e^{\boldsymbol{B}t} = e^{(\boldsymbol{A}+\boldsymbol{B})t}$$

if and only if the matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ commute; that is, if $\boldsymbol{AB} = \boldsymbol{BA}$.

$$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{A}(t) = \left[\frac{\mathrm{d}}{\mathrm{d}t}a_{ij}(t)\right]$$

and

$$\int \boldsymbol{A}(t)\,\mathrm{d}t = \left[\int a_{ij}(t)\,\mathrm{d}t\right]$$

Example 1.32 Evaluate $\mathrm{d}\boldsymbol{A}/\mathrm{d}t$ and $\int \boldsymbol{A}\,\mathrm{d}t$ for the matrix

$$\begin{bmatrix} t^2+1 & t-3 \\ 2 & t^2+2t-1 \end{bmatrix}$$

Note that in general

$$\frac{d}{dt}[\mathbf{A}(t)]^n \neq n\mathbf{A}^{n-1}\frac{d\mathbf{A}}{dt}$$

35 Show that the matrix

$$\mathbf{A} = \begin{bmatrix} 5 & 6 \\ 2 & 3 \end{bmatrix}$$

satisfies its own characteristic equation.

36 Given

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

use the Cayley–Hamilton theorem to evaluate

(a) $\mathbf{A}^2$ (b) $\mathbf{A}^3$ (c) $\mathbf{A}^4$

38 Given

$$\mathbf{A} = \begin{bmatrix} 2 & 3 & 1 \\ 3 & 1 & 2 \\ 1 & 2 & 3 \end{bmatrix}$$

compute $\mathbf{A}^2$ and, using the Cayley–Hamilton theorem, compute

$$\mathbf{A}^7 - 3\mathbf{A}^6 + \mathbf{A}^4 + 3\mathbf{A}^3 - 2\mathbf{A}^2 + 3\mathbf{I}$$

39 Evaluate $e^{\mathbf{A}t}$ for

(a) $\mathbf{A} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$ (b) $\mathbf{A} = \begin{bmatrix} 1 & 0 \\ 1 & 2 \end{bmatrix}$

42 Given

$$\mathbf{A} = \begin{bmatrix} t^2 + 1 & t - 1 \\ 5 & 0 \end{bmatrix}$$

evaluate $\mathbf{A}^2$ and show that

$$\frac{d}{dt}(\mathbf{A}^2) \neq 2\mathbf{A}\frac{d\mathbf{A}}{dt}$$