

MTH 205 - LINEAR ALGEBRA II

Made Available by
SchoolSearch for UJ Connect

Contents

1	Systems of Linear Equations	3
2	Change of Basis and Similarity	14
3	Linear, Bilinear and Quadratic Forms	19
4	Eigenvalues and Eigenvectors	30
5	Characteristic and Minimum Polynomials	35
6	Canonical Forms	42

Chapter 1

Systems of Linear Equations

Definition 1.1 (Linear Equation).

By a linear equation over any arbitrary field K , we refer to an expression of the form

$$a_1x_1 + a_2x_2 + \dots + a_nx_n = b \quad (1.1)$$

where for each i , $a_i, b \in K$ and x_i are variables (or unknowns or indeterminates).

The scalars a_i are called the coefficients of the x_i respectively and b is called the constant term of the linear equation.

Definition 1.2 (Solution of a Linear Equation).

A set of values for the unknowns, say

$$x_1 = c_1, \quad x_2 = c_2, \quad \dots, \quad x_n = c_n$$

where $c_i \in K$, is a solution of (1.1) if the statement obtained by substituting c_i for x_i in (1.1) is true, i.e.

$$a_1c_1 + a_2c_2 + \dots + a_nc_n = b$$

is true.

Note:

- (i) The set of values c_i is then said to satisfy the linear equation (1.1).
- (ii) If we write the set as an ordered set then we can consider it as an n -tuple $c = (c_1, c_2, \dots, c_n)$.
- (iii) Solution of the linear equation (1.1) can be categorized into three. They are:
 - (a) One of the coefficients is not zero, say $a_1 \neq 0$; Then we can rewrite the linear equation as:

$$\begin{aligned} a_1x_1 &= b - a_2x_2 - \dots - a_nx_n \quad \text{i.e.} \\ x_1 &= a_1^{-1}b - a_1^{-1}a_2x_2 - \dots - a_1^{-1}a_nx_n. \end{aligned}$$

By assigning arbitrary values to the unknowns x_2, \dots, x_n , we obtain a value for x_1 ; these values form a solution of the linear equation. Every solution can be found in this way. In this case, we say that the linear equation has many solutions.

- (b) All the coefficients of (1.1) are zero, but the constant term is not. That is, the linear equation is of the form

$$0x_1 + 0x_2 + \dots + 0x_n = b \quad \text{with } b \neq 0.$$

Then the linear equation has no solution and in this case, we say that the linear equation is not consistent.

- (c) All the coefficient of (1.1) are zero and the constant is also zero. That is, the linear equation is of the form:

$$0x_1 + 0x_2 + \dots + 0x_n = 0.$$

In this case, every n -tuple of scalars in K is a solution of the linear equation.

Definition 1.3 (System of Linear Equations) A set of m linear equations in n -unknowns x_1, \dots, x_n given by

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n &= b_2 \\ &\cdot \\ &\cdot \\ &\cdot \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n &= b_m \end{aligned} \tag{1.2}$$

where the a_{ij}, b_i belong to the scalar field K is called a System of Linear Equations.

If each b_i is zero, then the system is said to be homogeneous, otherwise it is non-homogeneous.

Definition 1.4 (Solution of a System of Linear Equations)

An n -tuple $c = (c_1, c_2, \dots, c_n)$ of scalars is a solution or a particular solution of (1.2) if it satisfies each of the linear equations. The set of all such solutions is termed the solution set or the general solution.

Note: The homogeneous system associated with (1.2) always has a solution namely the zero n -tuple $0 = (0, 0, \dots, 0)$ called the zero or trivial solution. Any other solution, if it exists, is called a non zero or non trivial solution.

Theorem 1.5

Let c be a particular solution of the non-homogeneous system (1.2) and suppose that A is the general solution of the associated homogeneous system, then

$$c + A = \{c + a : a \in A\}$$

is the general solution of the non-homogeneous system (1.2).

Proof.

Let B denote the general solution of the non-homogeneous system (1.2). Suppose that $c \in B$ and that $c = (c_1, c_2, \dots, c_n)$. Since c is a solution of (1.2), we have for $i = 1, \dots, m$,

$$a_{i1}c_1 + a_{i2}c_2 + \dots + a_{in}c_n = b_i.$$

Now suppose $a \in A$ and that $a = (a_1, a_2, \dots, a_n)$. Since a is a solution of the associated homogeneous system, we have for $i = 1, \dots, m$,

$$a_{i1}a_1 + a_{i2}a_2 + \dots + a_{in}a_n = 0.$$

Therefore, for $i = 1, \dots, m$,

$$\begin{aligned} & a_{i1}(c_1 + a_1) + a_{i2}(c_2 + a_2) + \dots + a_{in}(c_n + a_n) \\ &= (a_{i1}c_1 + a_{i2}c_2 + \dots + a_{in}c_n) + (a_{i1}a_1 + a_{i2}a_2 + \dots + a_{in}a_n) \\ &= b_i + 0 \\ &= b_i. \end{aligned}$$

That is, $c + a$ is a solution of (1.2). Thus $c + a \in B$ and hence $c + A \subset B$(*)

Now suppose that $p = (p_1, p_2, \dots, p_n)$ is any arbitrary element of B , i.e. a solution of (1.2) then, for $i = 1, \dots, m$,

$$a_{i1}p_1 + a_{i2}p_2 + \dots + a_{in}p_n = b_i.$$

Observe that $p = c + (p - c)$. We claim that $p - c \in A$, for $i = 1, \dots, m$,

$$\begin{aligned} & a_{i1}(p_1 - c_1) + a_{i2}(p_2 - c_2) + \dots + a_{in}(p_n - c_n) \\ &= (a_{i1}p_1 + a_{i2}p_2 + \dots + a_{in}p_n) - (a_{i1}c_1 + a_{i2}c_2 + \dots + a_{in}c_n) \\ &= b_i - b_i \\ &= 0. \end{aligned}$$

Thus, $p - c$ is a solution of the associated (homogeneous) system, i.e., $p - c \in A$.

Thus $p \in c + A$ and hence $B \subset c + A$(**)

Both inclusion (*) and (**) relations give us $B = c + A$.

Theorem 1.6

If in a homogeneous system of linear equations, the coefficients of one of the unknowns are all zero, then the system has a non zero solution.

Proof.

Let x_1, \dots, x_n , be the unknowns of the system and x_i the unknown whose coefficients are all zero. Then each linear equation of the system is of the form

$$a_1x_1 + \dots + a_{i-1}x_{i-1} + 0x_i + a_{i+1}x_{i+1} + \dots + a_nx_n = 0.$$

Then, for example, any n -tuple of the form $(0, \dots, 0, a, 0, \dots, 0)$ with $a \neq 0$ in the i^{th} component is a non zero solution of each linear equation and hence of the system.

Theorem 1.7

On multiplying the i^{th} equation of (1.2) by k_i , $i = 1, \dots, m$ and adding all the equations to obtain the linear combination

$$\begin{aligned} & (k_1a_{11} + k_2a_{21} + \dots + k_ma_{m1})x_1 + \dots + (k_1a_{1n} + k_2a_{2n} + \dots + k_ma_{mn})x_n \\ &= (k_1b_1 + k_2b_2 + \dots + k_mb_m)x_1. \end{aligned} \tag{1.3}$$

We have that any solution of (1.2) is also a solution of (1.3).

Proof.

Let $c = (c_1, c_2, \dots, c_n)$ be a solution of (1.2). Then,

$$a_{i1}c_1 + a_{i2}c_2 + \dots + a_{in}c_n = b_i, \quad i = 1, 2, \dots, m. \tag{1.4}$$

To show that c is a solution of (1.3), we must verify the equation

$$\begin{aligned} & (k_1a_{11} + k_2a_{21} + \dots + k_ma_{m1})c_1 + \dots + (k_1a_{1n} + k_2a_{2n} + \dots + k_ma_{mn})c_n. \\ & = k_1b_1 + k_2b_2 + \dots + k_mb_m. \end{aligned}$$

But this can be rearranged into

$$\begin{aligned} & k_1(a_{11}c_1 + a_{12}c_2 + \dots + a_{1n}c_n) + \dots + k_m(a_{m1}c_1 + \dots + a_{mn}c_n) \\ & = k_1b_1 + \dots + k_mb_m \end{aligned}$$

or by (1.4)

$$k_1b_1 + k_2b_2 + \dots + k_mb_m = k_1b_1 + k_2b_2 + \dots + k_mb_m$$

which is true.

1.8 Gauss Elimination Method

This solution method involves the elimination of an unknown from succeeding equations in the system (1.2). The method consists of a successive application of essentially two steps, viz.:

- (1) Interchange equations so that the first unknowns x_i has a non zero coefficient in the first equation i.e. $a_{11} \neq 0$.
- (2) For each $i > 1$, apply the operation

$$R_i \rightarrow -a_{i1}R_1 + a_{11}R_i.$$

That is, replace the i^{th} linear equation R_i by the operation obtained by multiplying the first equation R_1 by $-a_{i1}$, multiplying the i^{th} equation R_i by a_{11} and then adding.

Applying the two steps to the system (1.2), we find that it is equivalent to the system:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n & = b_1 \\ d_{2j_2}x_{j_2} + \dots + d_{2n}x_n & = f_2 \\ & \dots \\ d_{mj_2}x_{j_2} + d_{mn}x_n & = f_n \end{aligned} \tag{1.5}$$

where $a_{11} \neq 0$, x_{j_2} denotes the first unknown with a non zero coefficient in an equation other than the first.

Note:

- (i) By step 2, $x_{j_2} \neq x_1$.
- (ii) By Theorem 1.7, the system (1.5) has the same solution as (1.2).
- (iii) The equations in (1.5), apart from the first one form a subsystem which has fewer equations and fewer unknowns than the original system (1.2).
- (iv) If an equation of the form

$$0x_1 + \dots + 0x_n = b, \quad b \neq 0$$

occurs, then the system is inconsistent and has no solution.

(v) If an equation of the form

$$0x_1 + \dots + 0x_n = 0$$

occurs, then the equation can be deleted without affecting the solution.

(vi) $a_{11}, d_{2j_2}, \dots, d_{mj_2}$ are called the leading coefficients of the system (1.5)

(vii) If we continue the above process with each new smaller subsystem, we arrive at the conclusion that (1.2) is either inconsistent or it is reducible to an equivalent system in the form:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n &= b_1 \\ a_{2j_2}x_{j_2} + \dots + g_{2n}x_n &= g_2 \\ &\dots \\ a_{rj_r}x_{j_r} + \dots + g_{rn}x_n &= g_r \end{aligned} \tag{1.6}$$

where $1 < j_2 < \dots < j_r$ and the leading coefficients are not zero, that is $a_{11} \neq 0, a_{2j_2} \neq 0, a_{rj_r} \neq 0$.

Definition 1.9 (Echelon form; free variable)

The system (1.6) is said to be the echelon form of (1.2). The unknowns x_i which do not appear at the beginning of any equation ($i \neq 1, j_2, \dots, j_r$) are termed free variable. The number of such free variables is $n - r$.

Theorem 1.10

Consider a system in echelon form (1.6). The solution can be considered in two cases:

- (i) $r = n$. Then the system has a unique solution.
- (ii) $r < n$. Then we can arbitrarily assign values to the $n - r$ free variables and obtain a solution of the system.

Example 1.11

Reduce the following to echelon form and solve

$$\begin{aligned} 2x + 3y - w + 2z &= 6 \\ 3x - y + 2w - z &= 3 \\ 5x + 2y + w + z &= 8 \end{aligned}$$

Solution

We start to reduce the system by applying the operations

$$\begin{aligned} R_2 &\rightarrow -3R_1 + 2R_2 \\ R_3 &\rightarrow -5R_1 + 2R_3 \end{aligned}$$

to obtain

$$\begin{aligned} 2x + 3y - w + 2z &= 6 \\ -11y + 7w - 8z &= -12 \\ -11y + 7w - 8z &= -12 \end{aligned}$$

We next apply the operation

$$R_3 \rightarrow 11R_2 - 11R_3$$

to have

$$\begin{aligned} 2x + 3y - w + 2z &= 6 \\ -11y + 7w - 8z &= -12 \\ 0 &= 22 \end{aligned}$$

The equation $0 = 22$, that is

$$0x + 0y + 0w + 0z = 22,$$

shows that the original system is inconsistent and so has no solution.

Example 1.12

Reduce the following to echelon form and solve.

$$\begin{aligned} 2x + 3y - w + 2z &= 6 \\ 3x - y + 2w - z &= 3 \\ x + 2y - 2w + 3z &= 4 \\ 4x + 6y - 2w + 4z &= 12 \\ 2x - y + w + z &= 3 \end{aligned}$$

Solution:

We start to reduce the system by applying the operations:

$$\begin{aligned} R_2 &\rightarrow -3R_1 + 2R_2 \\ R_3 &\rightarrow -R_1 + 2R_3 \\ R_4 &\rightarrow -4R_1 + 2R_4 \\ R_5 &\rightarrow -2R_1 + 2R_5 \end{aligned}$$

to obtain

$$\begin{aligned} 2x + 3y - w + 2z &= 6 \\ -11y + 7w - 8z &= -12 \\ y - 3w + 4z &= 2 \\ 0 &= 0 \\ -8y + 4w - 2z &= -6 \end{aligned}$$

We now interchange R_4 and R_5 to get

$$\begin{aligned} 2x + 3y - w + 2z &= 6 \\ -11y + 7w - 8z &= -12 \\ y - 3w + 4z &= 2 \\ -8y + 4w - 2z &= -6 \\ 0 &= 0 \end{aligned}$$

Next, apply the operations

$$\begin{aligned} R_3 &\rightarrow -R_2 - 11R_3 \\ R_4 &\rightarrow 8R_2 - 11R_4 \end{aligned}$$

to obtain

$$\begin{aligned}2x + 3y - w + 2z &= 6 \\-11y + 7w - 8z &= -12 \\26w - 36z &= -10 \\-660z &= -660 \\0 &= 0\end{aligned}$$

This is in the echelon form and the system is a consistent one, since there is no equation of the form $0 = b$ with $b \neq 0$. Furthermore, in its echelon form there are four equations in four unknowns and so, the system has a unique solution. By the fourth equation, $z = 1$. Substituting $z = 1$ into the third equation, we obtain $w = 1$. Substituting $z = 1$, $w = 1$ into the second equation, we obtain $y = 1$. Substituting $z = 1$, $w = 1$, $y = 1$ into the first equation we obtain $x = 1$. Thus the 4-tuple $(1, 1, 1, 1)$ is the unique solution of the system.

Example 1.13

Reduce the following to echelon form and solve

$$\begin{aligned}2x + 3y - w + 2z &= 6 \\3x - y + 2w - z &= 3 \\5x + 2y + w + z &= 9.\end{aligned}$$

Solution

We start to reduce the system by applying the operations

$$\begin{aligned}R_2 &\rightarrow -3R_1 + 2R_2 \\R_3 &\rightarrow -5R_1 + 2R_3\end{aligned}$$

to obtain

$$\begin{aligned}2x + 3y - w + 2z &= 6 \\-11y + 7w - 8z &= -12 \\-11y + 7w - 8z &= -12.\end{aligned}$$

We next apply the operation

$$R_3 \rightarrow 11R_2 - 11R_3$$

to have

$$\begin{aligned}2x + 3y - w + 2z &= 6 \\-11y + 7w - 8z &= -12 \\0 &= 0.\end{aligned}$$

The system is consistent and since there are more unknowns than equations in echelon form, the system has an infinite number of solutions, There are two free variables w and z . A particular solution may be obtained by assigning w and z the values $(1, 1)$. Substituting, we have that $y = 1$ and $x = 1$. The 4-tuple $(1, 1, 1, 1)$ is a particular solution of the solution.

Note 1.14

Any homogeneous system of linear equation is consistent, since it has the zero solution $0 = (0, 0, \dots, 0)$. Thus it can always be reduced to an equivalent homogeneous system in echelon form (1.6) where $b_i = 0$ for all $i = 1, \dots, r$.

The solution is of two possibilities viz.:

- (i) $r = n$. Then the system has only the zero solution.
- (ii) $r < n$. Then the system has a non-zero solution.

Thus we observe that a homogeneous system of linear equations with more unknowns (n) than equations (r) has a non-zero solution. This is because in echelon form, $r < n$ and hence the system has a non-zero solution.

Example 1.15

Consider the homogeneous system

$$\begin{aligned} x + y - z &= 0 \\ 3x + y - 5z &= 0 \\ 2x - y + z &= 0 \end{aligned}$$

The echelon form is

$$\begin{aligned} x + y - z &= 0 \\ -2y - 2z &= 0 \\ -12z &= 0 \end{aligned}$$

Thus the system has only the trivial solution since there are three equations in three unknowns.

Example 1.16

By reducing the system

$$\begin{aligned} x + 2y - z &= 6 \\ 2x + 5y + 2z &= 0 \\ x + 3y + 3z &= 0 \\ x + 4y + 7z &= 0 \end{aligned}$$

to echelon form, we obtain

$$\begin{aligned} x + 2y - z &= 6 \\ y + 4z &= 0 \end{aligned}$$

The system has a non zero solution, since we have two equations in three unknowns in the echelon form.

Note 1.17

The system of linear equations (1.2) is equivalent to the matrix equation

$$\begin{pmatrix} a_{11} & a_{12} & \cdot & \cdot & \cdot & a_{1n} \\ a_{21} & a_{22} & \cdot & \cdot & \cdot & a_{2n} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{m1} & a_{m2} & \cdot & \cdot & \cdot & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \dots \\ b_m \end{pmatrix} \quad (1.7)$$

or simply

$$AX = B$$

where $X = (x_i)$, $B = (b_i)$ and $A = (a_{ij})$.

This means that every solution of (1.2) is a solution of the matrix equation (1.7). We further note that the associated homogeneous system of (1.2) is equivalent to the matrix equation $AX = 0$.

Definition 1.18 (Coefficient and Augmented Matrices)

The matrix $A = (a_{ij})$ in (1.7) is the coefficient matrix of the system (1.2) and the matrix:

$$\begin{pmatrix} a_{11} & a_{12} & \cdot & \cdot & \cdot & a_{1n} \\ a_{21} & a_{22} & \cdot & \cdot & \cdot & a_{2n} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{m1} & a_{m2} & \cdot & \cdot & \cdot & a_{mn} \end{pmatrix} \quad (1.8)$$

is called the augmented matrix of (1.2). It is obtained by adjoining the column of constants to A . Denote the augmented matrix by (A, B) .

We observe that the system (1.2) is completely determined by its augmented matrix. For example, the system given in Example 1.12 has the matrix form:

$$\begin{pmatrix} 2 & 3 & -1 & 2 \\ 3 & -1 & 2 & -1 \\ 1 & 2 & -2 & 3 \\ 4 & 6 & -2 & 4 \\ 2 & -1 & 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ w \\ z \end{pmatrix} = \begin{pmatrix} 6 \\ 3 \\ 4 \\ 12 \\ 3 \end{pmatrix}.$$

Its coefficient matrix and augmented matrix are respectively the matrices:

$$\begin{pmatrix} 2 & 3 & -1 & 2 \\ 3 & -1 & 2 & -1 \\ 1 & 2 & -2 & 3 \\ 4 & 6 & -2 & 4 \\ 2 & -1 & 1 & 1 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 2 & 3 & -1 & 2 & 6 \\ 3 & -1 & 2 & -1 & 3 \\ 1 & 2 & -2 & 3 & 4 \\ 4 & 6 & -2 & 4 & 12 \\ 2 & -1 & 1 & 1 & 3 \end{pmatrix}.$$

1.19 Solution of the Matrix Form

We apply elementary row operations to the augmented matrix with a view to reducing it to its echelon form or to its row canonical form.

For example, we can use the augmented matrix in Definition 1.18 to find the solution of the system in Example 1.12. To do this, just perform the following operations on the augmented matrix following the numbering.

- (1) $R_2 \rightarrow -3R_1 + 2R_2$
- (2) $R_3 \rightarrow -R_1 + 2R_3$
- (3) $R_4 \rightarrow -4R_1 + 2R_4$
- (4) $R_5 \rightarrow -2R_1 + 2R_5$

$$(5) R_4 \iff R_2$$

$$(6) R_3 \rightarrow -R_2 - 11R_3$$

$$(7) R_4 \rightarrow 8R_2 - 11R_4$$

$$(8) R_4 \rightarrow -12R_3 + 26R_4$$

We arrive at the echelon form:

$$\begin{pmatrix} 2 & 3 & -1 & 2 & 6 \\ & -11 & 7 & -8 & -12 \\ & & 26 & -36 & -10 \\ & & & -660 & -660 \\ & & & & 0 \end{pmatrix}$$

Starting from row 4, we have

$$-660z = -660 \implies z = 1.$$

Substituting in row 3, we find that $w = 1$. Row 2 give $y = 1$ and row 1 give $x = 1$.

Theorem 1.20

Suppose that u_1, u_2, \dots, u_n are solutions of a homogeneous system of linear equations $AX = 0$. Then every linear combination of the u_i is also a solution of $AX = 0$.

Proof.

Let

$$a_1u_1 + a_2u_2 + \dots + a_nu_n$$

(where the a_i are scalars) be a linear combination of the u_i .

We are given that $Au_1 = 0, Au_2 = 0, \dots, Au_n = 0$. Hence,

$$\begin{aligned} A(a_1u_1 + a_2u_2 + \dots + a_nu_n) &= a_1Au_1 + a_2Au_2 + \dots + a_nAu_n \\ &= a_1 \cdot 0 + a_2 \cdot 0 + \dots + a_n \cdot 0 \\ &= 0, \end{aligned}$$

which implies that $a_1u_1 + a_2u_2 + \dots + a_nu_n$ is a solution of the homogeneous system $AX = 0$.

1.21 Tutorial Questions

(1) Solve each of the following system.

(a)

$$2x + 3y = 3$$

$$x - 2y = 5$$

$$3x + 2y = 7$$

(b)

$$x + 2y - 3z + 2w = 2$$

$$2x + 5y - 8z + 6w = 5$$

$$3x + 4y - 5z + 2w = 4$$

- (2) What condition(s) must be placed on e , f and g so that the following system in unknowns x , y and z has: (i) no solution (ii) more than one solution (iii) a unique solution.

$$\begin{aligned}x + 2y - 3z &= e \\2x - y + 4z &= f \\4x + 3y - 2z &= g\end{aligned}$$

- (3) Determine the values of b so that the following system in unknowns x , y and z has: (i) no solution (ii) more than one solution (iii) a unique solution.

$$\begin{aligned}3x - 2y + 2z &= 5 \\2x + y - bz &= 5 \\5x - by - z &= 16\end{aligned}$$

- (4) Determine whether each of the following system has a non zero solution.

(i)

$$\begin{aligned}x + 2y - 2z &= 0 \\x - 8y + 8z &= 0 \\3x - 2y + 4z &= 0\end{aligned}$$

(ii)

$$\begin{aligned}x + 3y - 2z &= 0 \\2x - 3y + z &= 0 \\3x - 2y + 2z &= 0\end{aligned}$$

- (5) What condition(s) can you place on b so that the following system has a non-zero solution?

$$\begin{aligned}bx + 12y &= 0 \\3x + by &= 0.\end{aligned}$$

Solution to Tutorial Questions

- (1) (a) $x = 3$, $y = -1$. (b) $x = -z + 2w$, $y = 1 + 2z - 2w$
(2) $g - 2e - f = 0$.
(3) (i) None (ii) None (iii) any b
(4) (i) No (ii) No
(5) $b = \pm 6$.

Chapter 2

Change of Basis and Similarity

Definition 2.1(Transition Matrix)

Let $\{e_1, e_2, \dots, e_n\}$ be a basis of a vector space V and let $\{f_1, f_2, \dots, f_n\}$ be another basis. Suppose that

$$\begin{aligned} f_1 &= a_{11}e_1 + a_{12}e_2 + \dots + a_{1n}e_n \\ f_2 &= a_{21}e_1 + a_{22}e_2 + \dots + a_{2n}e_n \\ &\cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\ f_n &= a_{n1}e_1 + a_{n2}e_2 + \dots + a_{nn}e_n \end{aligned} \tag{2.1}$$

This means that

$$\begin{pmatrix} f_1 \\ f_2 \\ \dots \\ f_n \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} \begin{pmatrix} e_1 \\ e_2 \\ \dots \\ e_n \end{pmatrix}$$

or $F = AE$.

Then the transpose P of the matrix of coefficients A is termed the transition matrix from the former basis $\{e_i\}$ to the present basis $\{f_i\}$. While the inverse of P , that is P^{-1} is the transition matrix from the basis $\{f_i\}$ back to the basis $\{e_i\}$.

Example 2.2

Let $\{e_i\} = \{e_1 = (0, 1), e_2 = (0, 1)\}$ and $\{f_i\} = \{f_1 = (1, 1), f_2 = (2, 0)\}$.

Find

- (i) the transition matrix P_1 from $\{e_i\}$ to $\{f_i\}$.
- (ii) the transition matrix P_2 from $\{f_i\}$ to $\{e_i\}$.
- (iii) and verify that $P_2 = P_1^{-1}$.

Solution

(i)

$$\begin{aligned} f_1 &= (1, 1) = e_1 + e_2 \\ f_2 &= (2, 0) = 2e_1 + 0e_2 \end{aligned}$$

Then

$$\begin{pmatrix} f_1 \\ f_2 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 2 & 0 \end{pmatrix} \begin{pmatrix} e_1 \\ e_2 \end{pmatrix}$$

Hence the transition matrix P_1 from the basis $\{e_i\}$ to $\{f_i\}$ is

$$P_1 = \begin{pmatrix} 1 & 1 \\ 2 & 0 \end{pmatrix}^t = \begin{pmatrix} 1 & 2 \\ 1 & 0 \end{pmatrix}$$

(ii)

$$\begin{aligned} e_1 &= (1, 0) = 0f_1 + \frac{1}{2}f_2 \\ e_2 &= (0, 1) = f_1 - \frac{1}{2}f_2 \end{aligned}$$

Then

$$\begin{pmatrix} e_1 \\ e_2 \end{pmatrix} = \begin{pmatrix} 0 & \frac{1}{2} \\ 1 & -\frac{1}{2} \end{pmatrix} \begin{pmatrix} f_1 \\ f_2 \end{pmatrix}$$

Hence the transition matrix P_2 from the basis $\{f_i\}$ to $\{e_i\}$ is

$$P_2 = \begin{pmatrix} 0 & \frac{1}{2} \\ 1 & -\frac{1}{2} \end{pmatrix}^t = \begin{pmatrix} 0 & 1 \\ \frac{1}{2} & -\frac{1}{2} \end{pmatrix}$$

(iii)

$$P_1P_2 = \begin{pmatrix} 1 & 2 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ \frac{1}{2} & -\frac{1}{2} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I$$

Theorem 2.3

Let P be the transition matrix from a basis $\{e_i\}$ to a basis $\{f_i\}$ in a vector space V . Then, for any vector $u \in V$,

$$P[u]_f = [u]_e \quad \underline{\text{OR}} \quad [u]_f = P^{-1}[u]_e.$$

Example 2.4

Let $\{e_i\}$ and $\{f_i\}$ be as in Example 2.2 and let $u = (2, 3)$. Show that

(i) $P[u] = [u]_e$ and

(ii) $[u]_f = P^{-1}[u]_e$.

Solution

$$u = (2, 3) = 2e_1 + 3e_2.$$

Then $[u]_e = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$.

Also, $u = (2, 3) = f_1 - \frac{1}{2}f_2$ and

$$[u]_f = \begin{pmatrix} 3 \\ -\frac{1}{2} \end{pmatrix}.$$

$$(i) P[u]_f = \begin{pmatrix} 1 & 2 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 3 \\ -\frac{1}{2} \end{pmatrix} = \begin{pmatrix} 2 \\ 3 \end{pmatrix} = [u]_e \text{ where } P = P_1 \text{ in Example 2.2.}$$

$$(ii) P^{-1}[u]_e = \begin{pmatrix} 0 & 1 \\ \frac{1}{2} & -\frac{1}{2} \end{pmatrix} \begin{pmatrix} 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 3 \\ -\frac{1}{2} \end{pmatrix} = [u]_f \text{ where } P^{-1} = P_2 \text{ in Example 2.2.}$$

Theorem 2.5

Let P be the transition matrix from a basis $\{e_i\}$ to a basis $\{f_i\}$ in a vector space V . Then for any linear operator T on V ,

$$[T]_f = P^{-1}[T]_e P.$$

Proof.

For any vector $u \in V$,

$$P^{-1}[T]_e P[u]_f = P^{-1}[T]_e [u]_e = P^{-1}[T(u)]_e = [T(u)]_f.$$

But,

$$[T]_f [u]_f = [T(u)]_f; \text{ hence}$$

$$P^{-1}[T]_e P[u]_f = [T]_f [u]_f.$$

Since the mapping $u \rightarrow [u]_f$ is onto K^n , $P^{-1}[T]_e P X = [T]_f X$ for every $X \in K^n$. Hence $P^{-1}[T]_e P = [T]_f$.

Example 2.6

Let $\{e_i\}$ and $\{f_i\}$ be as in Example 2.2 and T be the linear operator on \mathbb{R}^2 defined by

$$T(x, y) = (x + y, x - y),$$

where \mathbb{R} is the set of real numbers. Show that $[T]_f = P^{-1}[T]_e P$.

Solution

$$T(e_1) = T(1, 0) = (1, 1) = e_1 + e_2$$

$$T(e_2) = T(0, 1) = (1, -1) = e_1 - e_2$$

$$\therefore [T]_e = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

$$T(f_1) = T(1, 1) = (2, 0) = 0f_1 + f_2$$

$$T(f_2) = T(2, 0) = (2, 2) = 2f_1 + 0f_2$$

$$\therefore [T]_f = \begin{pmatrix} 0 & 2 \\ 1 & 0 \end{pmatrix}.$$

$$\begin{aligned} \text{Now, } P^{-1}[T]_e P &= \begin{pmatrix} 0 & 1 \\ \frac{1}{2} & -\frac{1}{2} \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 1 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \\ &= [T]_f. \end{aligned}$$

Definition 2.7 (Similar Matrices)

Let A and B be square matrices for which there exists an invertible matrix P such that $B = P^{-1}AP$, then B is said to be similar to A or we say that B is obtained from A by a similarity transformation.

Theorem 2.8

Similarity of matrices is an equivalence relation.

Proof.

(i) We first show that A is similar to A . The identity matrix I is invertible and $I = I^{-1}$. Since $A = I^{-1}AI$, A is similar to A .

(ii) We next show that if A is similar to B , then B is similar to A . Since A is similar to B , there exists an invertible matrix P such that $A = P^{-1}BP$. Hence

$$B = PAP^{-1} = (P^{-1})^{-1}AP^{-1}$$

and P^{-1} is invertible. Thus B is similar to A .

(iii) We finally show that if A is similar to B and B is similar to C , then A is similar to C . Since A is similar to B there exists an invertible matrix P such that $A = P^{-1}BP$ and since B is similar to C there exists an invertible matrix Q such that $B = Q^{-1}CQ$.

Hence

$$A = P^{-1}BP = P^{-1}(Q^{-1}CQ)P = (QP)^{-1}C(QP)$$

and QP is invertible. Thus A is similar to C .

From (i) to (iii), similarity of matrices from an equivalence relation.

Theorem 2.9

Two matrices A and B represent the same linear operator T if and only if they are similar to each other.

Definition 2.10 (Diagonalization)

A linear operator T is said to be diagonalizable if for some basis $\{e_i\}$ it is represented by a diagonal matrix. The basis $\{e_i\}$ is then said to diagonalize T .

Theorem 2.11

Let A be a matrix representation of a linear operator T . Then T is diagonalizable if and only if there exists an invertible matrix P such that $P^{-1}AP$ is a diagonal matrix.

Definition 2.12

Two $m \times n$ matrices A and B over K is said to be equivalent if there exists an m -square invertible matrix Q and an n -square invertible matrix P such that $B = QAP$.

Tutorial Questions 2.13

(1) Let

$$\begin{aligned} \{f_i\} &= \{f_1 = (1, 2), f_2 = (2, 3)\} \text{ and} \\ \{g_i\} &= \{g_1 = (1, 3), g_2 = (1, 4)\} \end{aligned}$$

be basis of \mathbb{R}^2 .

- (i) Find the transition matrix P and Q from $\{f_i\}$ to $\{g_i\}$ and from $\{g_i\}$ to $\{f_i\}$ respectively.
 - (ii) Verify that $Q = P^{-1}$.
 - (iii) Show that $[v]_f = P[v]_g$ for any vector $v \in \mathbb{R}^2$
 - (iv) Show that $[T]_g = P^{-1}[T]_fP$ for each of the operators:
 - (a) $T(x, y) = (2x - 3y, x + y)$
 - (b) $T(x, y) = (5x + y, 3x - 2y)$.
- (2) Suppose that $\{e_1, e_2\}$ is a basis of V and T is a linear operator for which $T(e_1) = 3e_1 - 2e_2$ and $T(e_2) = e_1 + 4e_2$. Let $\{f_1, f_2\}$ be the basis of V for which $f_1 = e_1 + e_2$ and $f_2 = 2e_1 + 3e_2$. Find the matrix of T in the basis $\{f_1, f_2\}$.
- (3) Consider the bases $\{f_i\} = \{f_1 = 1, f_2 = i\}$ and $\{g_i\} = \{g_1 = 1 + i, g_2 = 1 + 2i\}$ if the complex field \mathbb{C} over the real field \mathbb{R} .
- (i) Find the transition matrices P and Q from $\{f_i\}$ to $\{g_i\}$ and from $\{g_i\}$ to $\{f_i\}$ respectively.
 - (ii) Verify that $Q = P^{-1}$.
 - (iii) Show that $[T]_g = P^{-1}[T]_fP$ for the operator $T_z = \bar{z}$, where \bar{z} is the complex conjugate of z .
- (4) Suppose that $\{e_i\}$, $\{f_i\}$ and $\{g_i\}$ are basis of V and that P and Q are the transition matrices from $\{e_i\}$ to $\{f_i\}$ and from $\{f_i\}$ to $\{g_i\}$ respectively. Show that PQ is the transition matrix from $\{e_i\}$ to $\{g_i\}$.
- (5) Show that equivalence of matrices is an equivalence relation.

Solution to Tutorial Questions 2.14

(1) (i) $P = \begin{pmatrix} 3 & 5 \\ -1 & -2 \end{pmatrix}; \quad Q = \begin{pmatrix} 2 & 5 \\ -1 & -3 \end{pmatrix}.$

(2) $\begin{pmatrix} 8 & 11 \\ -2 & -1 \end{pmatrix}.$

(3) (i) $P = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}; \quad \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix}.$

Chapter 3

Linear, Bilinear and Quadratic Forms

Definition 3.1 (Linear Form)

Let V be a vector space over a field K . A mapping $\phi : V \rightarrow K$ is termed a linear form (or linear functional) if for every $u, v \in V$ and every $a, b \in K$,

$$\phi(au + bv) = a\phi(u) + b\phi(v).$$

Example 3.2

- (1) Let $\pi_i : K^n \rightarrow K$ be the i^{th} projection mapping, i.e. $\pi_i(a_1, a_2, \dots, a_n) = a_i$. Then π_i is linear and so it is a linear form on K^n .
- (2) Let V be a vector space of polynomials in t over \mathbb{R} . Let $\phi : V \rightarrow \mathbb{R}$ be the integral operator defined by

$$\phi(p(t)) = \int_0^1 p(t) dt.$$

Then ϕ is linear and hence it is a linear form on V .

- (3) Let V be the vector space of n -square matrices over K . Let $\phi : V \rightarrow K$ be the trace mapping.

$$\phi(A) = a_{11} + a_{22} + \dots + a_{nn},$$

where $A = (a_{ij})$. This map is linear and so it is a linear form on V .

Theorem 3.3

The set of all linear forms on a vector space V over a field K is also a vector space over K with addition and scalar multiplication defined by

$$\begin{aligned}(\phi + \rho)(v) &= \phi(v) + \rho(v) \text{ and} \\(k\phi)(v) &= k\phi(v)\end{aligned}$$

where ϕ and ρ are linear forms on V and $k \in K$.

Definition 3.4 (Dual Space)

The vector space formed from the set of all linear forms on a vector space V is called the dual space of V . We shall denote it by V^1 .

Note that if $V = K^n$ (the vector space of n -tuples which we write as column vectors), then the

dual space V^1 can be identified with the space of row vectors. In particular, any linear form $\phi = (a_1, \dots, a_n)$ in V^1 has the representation:

$$\phi(x_1, \dots, x_n) = (a_1, a_2, \dots, a_n) \begin{pmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ x_n \end{pmatrix}$$

or simply $\phi(x_1, \dots, x_n) = a_1x_1 + a_2x_2 + \dots + a_nx_n$. This expression is termed a linear form.

Theorem 3.5

Suppose that V is a vector space of dimension n over K , then the dimension of the dual space V^1 is also n .

Theorem 3.6

Suppose that $\{v_1, \dots, v_n\}$ is a basis of V over K . Let $\phi_1, \dots, \phi_n \in V^1$ be the linear forms defined by:

$$\phi_i(v_j) = \delta_{(ij)} = \begin{cases} 1, & \text{if } i = j \\ 0 & i \neq j. \end{cases}$$

Then $\{\phi_1, \dots, \phi_n\}$ is a basis of V^1 .

Proof.

We first show that $\{\phi_1, \dots, \phi_n\}$ spans V^1 . Let ϕ be an arbitrary element of V^1 , and suppose that

$$\phi(v_1) = k_1, \quad \phi(v_2) = k_2, \quad \dots, \quad \phi(v_n) = k_n.$$

Set

$$\rho = k_1\phi_1 + \dots + k_n\phi_n.$$

Then, for $i = 1, 2, \dots, n$,

$$\begin{aligned} \rho(v_i) &= (k_1\phi_1 + \dots + k_n\phi_n)(v_i) \\ &= k_1\phi_1(v_i) + \dots + k_i\phi_i(v_i) + \dots + k_n\phi_n(v_i) \\ &= k_i. \end{aligned}$$

Thus $\phi(v_i) = \rho(v_i)$ for $i = 1, \dots, n$. Since ϕ and ρ agree on the basis vectors,

$$\phi = \rho = k_1\phi_1 + \dots + k_n\phi_n.$$

Hence, $\{\phi_1, \dots, \phi_n\}$ spans V^1 .

It remains to be shown that $\{\phi_1, \dots, \phi_n\}$ is linearly independent. Suppose that

$$a_1\phi_1 + a_2\phi_2 + \dots + a_n\phi_n = 0.$$

Applying each side to v_i for $i = 1, 2, \dots, n$.

$$\begin{aligned} 0 = 0(v_i) &= (a_1\phi_1 + \dots + a_n\phi_n)(v_i) \\ &= a_1\phi_1(v_i) + \dots + a_i\phi_i(v_i) + \dots + a_n\phi_n(v_i) \\ &= a_i \end{aligned}$$

That is, $a_1 = 0, \dots, a_n = 0$. Hence $\{\phi_1, \dots, \phi_n\}$ is linearly independent and so it is a basis of V^1 .

Definition 3.7 (Dual Basis)

The above basis $\{\phi_i\}$ in Theorem 3.6 is termed the basis dual to $\{v_i\}$ or the dual basis.

Example 3.8

Consider the following basis of \mathbb{R}^2 ;

$$\{v_i\} = \{v_1 = (1, 1), v_2 = (2, 1)\}.$$

Find the dual basis $\{\phi_1, \phi_2\} = \{\phi_i\}$.

Solution:

We seek linear forms $\phi_1(x, y) = a_1x + b_1y$ and $\phi_2(x, y) = a_2x + b_2y$ such that

$$\phi_1(v_1) = \delta_{11} = 1, \quad \phi_1(v_2) = \delta_{12} = 0, \quad \phi_2(v_1) = \delta_{21} = 0, \quad \phi_2(v_2) = \delta_{22} = 1.$$

Thus

$$\begin{aligned} \phi_1(v_1) &= \phi(1, 1) = a_1 + b_1 = 1 \\ \phi_1(v_2) &= \phi(2, 1) = 2a_1 + b_1 = 0 \end{aligned}$$

Solving, we have $a_1 = -1, b_1 = 2$.

$$\begin{aligned} \phi_2(v_1) &= \phi_2(1, 1) = a_2 + b_2 = 0 \\ \phi_2(v_2) &= \phi_2(2, 1) = 2a_2 + b_2 = 1. \end{aligned}$$

Solving, we have $a_2 = 1, b_2 = -1$.

Hence the dual basis is $\{\phi_1(x, y) = 2y - x; \phi_2(x, y) = x - y\}$.

Theorem 3.9

Let $\{v_1, \dots, v_n\}$ be a basis of V and let $\{\phi_1, \dots, \phi_n\}$ be the dual basis of V^1 . Then for any vector $u \in V$,

$$u = \phi_1(u)v_1 + \phi_2(u)v_2 + \dots + \phi_n(u)v_n \dots \dots \dots (a)$$

and for any linear form $\rho \in V^1$,

$$\rho = \rho(v_1)\phi_1 + \rho(v_2)\phi_2 + \dots + \rho(v_n)\phi_n \dots \dots \dots (b)$$

Proof.

Suppose that

$$u = a_1v_1 + a_2v_2 + \dots + a_nv_n \dots \dots \dots (c)$$

Then, for $i = 1, 2, \dots, n$,

$$\begin{aligned} \phi_i(u) &= a_1\phi_i(v_1) + \dots + a_i\phi_i(v_i) + \dots + a_n\phi_i(v_n) \\ &= a_i. \end{aligned}$$

That is,

$$\phi_1(u) = a_2, \quad \phi_2(u) = a_2, \dots, \phi_n(u) = a_n.$$

Substituting these results into (c), we obtain (a).

Next we prove (b). Applying the linear form ρ to both sides of (a), we have,

$$\begin{aligned} \rho(u) &= \phi_1(u)\rho(v_1) + \phi_2(u)\rho(v_2) + \dots + \phi_n(u)\rho(v_n) \\ &= \rho(v_1)\phi_1(u) + \rho(v_2)\phi_2(u) + \dots + \rho(v_n)\phi_n(u) \\ &= (\rho(v_1)\phi_1 + \rho(v_2)\phi_2 + \dots + \rho(v_n)\phi_n)(u). \end{aligned}$$

Since this holds for every $u \in V$,

$$\rho = \rho(v_1)\phi_1 + \rho(v_2)\phi_2 + \dots + \rho(v_n)\phi_n$$

as required.

Theorem 3.10

Let $\{v_1, \dots, v_n\}$ and $\{w_1, \dots, w_n\}$ be bases of V and let $\{\phi_1, \dots, \phi_n\}$ and $\{\rho_1, \dots, \rho_n\}$ be the basis of V^1 dual to $\{v_i\}$ and $\{w_i\}$ respectively. Suppose that P is the transition matrix from $\{v_i\}$ to $\{w_i\}$. Then $(P^{-1})^t$ is the transition matrix from $\{\phi_i\}$ to $\{\rho_i\}$.

Definition 3.11 (Bilinear Forms)

Let V be a vector space of finite dimension over a field K . A bilinear form on V is a mapping $f : V \times V \rightarrow K$ which satisfies for all $a, b \in K$ and all $u_i, v_i \in V$,

$$(i) \quad f(au_1 + bu_2, v) = af(u_1, v) + bf(u_2, v).$$

$$(ii) \quad f(u, av_1 + bv_2) = af(u, v_1) + bf(u, v_2).$$

Note. Condition (i) express the fact that f is linear in the first variable and condition (ii) express that f is linear in the second variable.

Example 3.12

(a) Let ϕ and ρ be arbitrary linear functionals on V . Let $f : V \times V \rightarrow K$ be defined by $f(u, v) = \phi(u)\rho(v)$. Then f is bilinear because ϕ and ρ are each linear.

(b) Let f be the dot product on \mathbb{R}^n ; that is,

$$f(u, v) = u \cdot v = a_1b_1 + a_2b_2 + \dots + a_nb_n$$

where $u = (a_i)$ and $v = (b_i)$. Then f is bilinear form on \mathbb{R}^n .

(c) Let $A = (a_{ij})$ be any $n \times n$ matrix over K . Then A may be viewed as a bilinear form f on K^n by defining:

$$f(X, Y) = X^t A Y = (x_1, x_2, \dots, x_n) \begin{pmatrix} a_{11} & a_{12} & \cdot & \cdot & \cdot & a_{1n} \\ a_{21} & a_{22} & \cdot & \cdot & \cdot & a_{2n} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \cdot & \cdot & \cdot & a_{nn} \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \cdot \\ \cdot \\ y_n \end{pmatrix}$$

$$\begin{aligned}
&= \sum_{i,j=1}^n a_{ij}x_iy_j \\
&= a_{11}x_1y_1 + a_{12}x_2y_2 + \dots + a_{1n}x_ny_n.
\end{aligned}$$

Note. The last expression in Example 3.12(c) which is a formal expression in variables x_i, y_i is termed the bilinear polynomial corresponding to the matrix A .

Definition 3.13 (Vector Spaces $B(V)$)

Let $B(V)$ denote the set of all bilinear forms on V . A vector space structure is placed on $B(V)$ by defining $f + g$ and kf by:

$$\begin{aligned}
(f + g)(u, v) &= f(u, v) + g(u, v) \\
(kf)(u, v) &= kf(u, v)
\end{aligned}$$

for any $f, g \in B(V)$ and any $k \in K$.

Theorem 3.14

Let V be a vector space of dimension n over K . Let $\{\phi_1, \dots, \phi_n\}$ be a basis of the dual space V^\vee . Then $\{f_{ij} : i, j = 1, \dots, n\}$ is a basis of $B(V)$ where f_{ij} is defined by

$$f_{ij}(u, v) = \phi_i(u)\phi_j(v).$$

Proof.

Let $\{e_1, \dots, e_n\}$ be the basis of V dual to $\{\phi_i\}$. We first show that $\{f_{ij}\}$ spans $B(V)$. Let $f \in B(V)$ and suppose that $f(e_i, e_j) = a_{ij}$. We claim that

$$f = \sum_{ij} a_{ij}f_{ij}.$$

It suffices to show that $f(e_s, e_t) = (\sum a_{ij}f_{ij})(e_s, e_t)$, $s, t = 1, \dots, n$.

We have

$$\begin{aligned}
\left(\sum a_{ij}f_{ij}\right)(e_s, e_t) &= \sum a_{ij}f_{ij}(e_s, e_t) \\
&= \sum a_{ij}\phi_i(e_s)\phi_j(e_t) \\
&= \sum a_{ij}\delta_{is}\delta_{jt} \\
&= a_{st} \\
&= f(e_s, e_t)
\end{aligned}$$

as required. Hence $\{f_{ij}\}$ spans $B(V)$.

It remains to be shown that $\{f_{ij}\}$ is linearly independent. Suppose that:

$$\sum a_{ij}f_{ij} = 0.$$

Then for $s, t = 1, \dots, n$.

$$0 = 0(e_s, e_t) = \left(\sum a_{ij}f_{ij}\right)(e_s, e_t) = a_{st}.$$

The last step follows as above. Thus $\{f_{ij}\}$ is independent and hence is a basis of $B(V)$.

Note 3.15

Let f be a bilinear form on V and let $\{e_1, \dots, e_n\}$ be a basis of V . Suppose that $u, v \in V$ and that

$$\begin{aligned} u &= a_1e_1 + \dots + a_n e_n \\ v &= b_1e_1 + \dots + b_n e_n. \end{aligned}$$

Then,

$$\begin{aligned} f(u, v) &= f(a_1e_1 + \dots + a_n e_n, b_1e_1 + \dots + b_n e_n) \\ &= a_1b_1f(e_1, e_1) + a_1b_2f(e_1, e_2) + \dots + a_nb_n f(e_n, e_n) \\ &= \sum_{i,j=1}^n a_i b_j f(e_i, e_j). \end{aligned}$$

Thus f is completely determined by the n^2 values $f(e_i, e_j)$.

Definition 3.16 (Matrix Representation)

The matrix $A = (a_{ij})$ where $a_{ij} = f(e_i, e_j)$ is called the matrix representation of f relative to the basis $\{e_i\}$ or simply, the matrix of f in $\{e_i\}$.

Observe that it represents f in the sense that

$$\begin{aligned} f(u, v) &= \sum a_i b_j f(e_i, e_j) \\ &= (a_1, \dots, a_n) A \begin{pmatrix} b_1 \\ b_2 \\ \cdot \\ \cdot \\ b_n \end{pmatrix} \\ &= [u]_e^t A [v]_e, \text{ for all } u, v \in V. \end{aligned}$$

Example 3.17

Let $u = (x_1, x_2, x_3)$ and $v = (y_1, y_2, y_3)$ and let

$$f(u, v) = 3x_1y_1 - 2x_1y_2 + 5x_2y_1 + 7x_2y_2 - 8x_2y_3 + 4x_3y_2 - x_3y_3.$$

Express f in matrix notation.

Solution

Let A be the 3×3 matrix whose ij -entry is the coefficient of $x_i y_j$. Then

$$f(u, v) = (x_1, x_2, x_3) \begin{pmatrix} 3 & -2 & 0 \\ 5 & 7 & -8 \\ 0 & 4 & -1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$$

Theorem 3.18

Let P be the transition matrix from one basis to another. If A is the matrix of f in the original basis, then

$$B = P^t A P$$

is the matrix of f in the new basis.

Proof

Let $\{e_i\}$ be the original basis, $\{e_i^1\}$ the new basis and $u, v \in V$. Since P is the transition matrix from $\{e_i\}$ to $\{e_i^1\}$, we have

$$P[u]_{e^1} = [u]_e \text{ and}$$

$$P[v]_{e^1} = [v]_e \text{ hence}$$

$$[u]_e^t = [u]_{e^1}^t P^t.$$

Thus

$$\begin{aligned} f(u, v) &= [u]_e^t A [v]_e \\ &= [u]_{e^1}^t P^t A P [v]_{e^1}. \end{aligned}$$

Since u and v are arbitrary elements of V , $P^t A P$ is the matrix of f in the new basis $\{e_i^1\}$.

Definition 3.19 (Congruent Matrices)

A matrix B is said to be congruent to a matrix A if there exists an invertible matrix P such that $B = P^t A P$.

Definition 3.20 (Alternating Bilinear Forms)

A bilinear form f on V is said to be alternating if

$$f(v, v) = 0 \tag{3.1}$$

for every $v \in V$.

Note: If f is alternating, then

$$\begin{aligned} 0 &= f(u + v, u + v) \\ &= f(u, u) + f(u, v) + f(v, u) + f(v, v) \end{aligned}$$

and so

$$f(u, v) = -f(v, u) \text{ for every } u, v \in V. \tag{3.2}$$

Definition 3.21 (Skew Symmetric Forms)

A bilinear form which satisfies equation (3.2) is said to be skew symmetric or anti-symmetric.

Definition 3.22 (Symmetric Bilinear Forms)

A bilinear form f on V is said to be symmetric if $f(u, v) = f(v, u)$ for every $u, v \in V$.

Theorem 3.23

If f is a bilinear form, then it is symmetric if and only if A is symmetric, where A is the matrix representation of f .

Proof

If f is symmetric and A is its matrix representation, then

$$f(X, Y) = X^t A Y = (X^t A Y)^t = Y^t A^t X$$

(where we employ the fact that $X^t A Y$ is a scalar and therefore equals its transpose), and

$$Y^t A^t X = f(X, Y) = f(Y, X) = Y^t A X.$$

Since this is true for all vectors X, Y it follows that $A = A^t$, *i.e.* A is symmetric.

Theorem 3.24 (Converse of Theorem 3.23)

Let f be a symmetric bilinear form on V over K in which $1 + 1 \neq 0$. Then V has a basis $\{v_1, \dots, v_n\}$ in which f is represented by a diagonal matrix, *i.e.*

$$f(v_i, v_j) = 0 \text{ for } i \neq j.$$

Definition 3.25 (Quadratic Forms)

A mapping $q : V \rightarrow K$ is called a quadratic form if

$$q(v) = f(v, v)$$

for some symmetric bilinear form f on V .

Theorem 3.26

Let q be the quadratic form associated with the symmetric bilinear form f . Then if $1 + 1 \neq 0$, we have

$$f(u, v) = \frac{1}{2} \{q(u + v) - q(u) - q(v)\} \quad (3.3)$$

Proof.

$$\begin{aligned} q(u + v) - q(u) - q(v) &= f(u + v, u + v) - f(u, u) - f(v, v) \\ &= f(u, u) + f(u, v) + f(v, u) + f(v, v) - f(u, u) - f(v, v) \\ &= 2f(u, v) \end{aligned}$$

If $1 + 1 \neq 0$, we can divide by 2 to obtain the required identity.

Definition 3.27 (Polar Form)

The formular (3.3) is called the polar form of f .

Note 3.28(i) Quadratic Polynomial

If f is represented by a symmetric matrix $A = (a_{ij})$, then q is represented in the form

$$\begin{aligned} q(x) &= f(X, X) \\ &= X^t A X \\ &= (x_1, \dots, x_n) \begin{pmatrix} a_{11} & a_{12} & \cdot & \cdot & \cdot & a_{1n} \\ a_{21} & a_{22} & \cdot & \cdot & \cdot & a_{2n} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \cdot & \cdot & \cdot & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ x_n \end{pmatrix} \\ &= \sum_{i,j} a_{ij} x_i x_j \\ &= a_{11} x_1^2 + a_{22} x_2^2 + \dots + a_{nn} x_n^2 + 2 \sum_{i \neq j} a_{ij} x_i x_j. \end{aligned}$$

This formal expression in variables x_i is termed the quadratic polynomial corresponding to the symmetric matrix A .

(ii) Diagonal Representation

If the matrix A is diagonal, then q has the diagonal representation

$$\begin{aligned} q(x) &= X^t A X \\ &= a_{11} x_1^2 + a_{22} x_2^2 + \dots + a_{nn} x_n^2 \end{aligned}$$

that is, the quadratic polynomial representing q will contain no terms of the form $x_i x_j$, $i \neq j$. By Theorem 3.24, every quadratic form has such a representation when $1 + 1 \neq 0$.

Tutorial Questions 3.29

- (1) Let $\phi : \mathbb{R}^3 \rightarrow \mathbb{R}$ and $\rho : \mathbb{R}^3 \rightarrow \mathbb{R}$ be the linear forms defined by

$$\phi(x, y, z) = 2x - 3y + z \quad \text{and} \quad \rho(x, y, z) = 4x - 2y + 3z.$$

Find: (i) $\phi + \rho$ (ii) 3ϕ (iii) $2\phi - 5\rho$.

- (2) Let ϕ be the linear form on \mathbb{R}^2 defined by $\phi(2, 1) = 15$ and $\phi(1, -2) = -10$. Find $\phi(x, y)$ and in particular, find the value of $\phi(-2, 7)$.

- (3) Find the dual basis of each of the following bases of \mathbb{R}^3 .

(i) $\{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$.

(ii) $\{(1, -2, 3), (1, -1, 1), (2, -4, 7)\}$.

- (4) Let V be the vector space of polynomials over \mathbb{R} of degree ≤ 2 . Let ϕ_1, ϕ_2 and ϕ_3 be the linear forms on V defined by

$$\phi_1(f(t)) = \int_0^1 f(t) dt, \quad \phi_2(f(t)) = f'(1), \quad \phi_3(f(t)) = f(0).$$

Here $f(t) = a + bt + ct^2 \in V$ and $f'(t)$ denotes the derivative of $f(t)$. Find the basis $\{f_1(t), f_2(t), f_3(t)\}$ of V which is dual to $\{\phi_1, \phi_2, \phi_3\}$.

(5) Show that each of the following mappings from V to K are linear forms

(i) $\pi_i : K^n \rightarrow K$ defined by $\pi_i(a_1, \dots, a_n) = a_i$.

(ii) $\phi : V \rightarrow K$ defined by $\phi(p(t)) = \int_0^1 p(t) dt$.

(iii) $\phi : V \rightarrow K$ defined by $\phi(A) = a_{11} + \dots + a_{nn}$, where $A = (a_{ij})$.

(6) Let $u = (x_1, x_2)$ and $v = (y_1, y_2)$. Determine which of the following are bilinear forms on \mathbb{R}^2 .

(i) $f(u, v) = 2x_1y_2 - 3x_2y_1$.

(ii) $f(u, v) = 3x_2y_2$.

(iii) $f(u, v) = 1$.

(iv) $f(u, v) = x_1 + y_2$.

(v) $f(u, v) = x_1x_2 + y_1y_2$.

(vi) $f(u, v) = 0$.

(7) Let f be the bilinear form on \mathbb{R}^2 defined by

$$f((x_1, x_2), (y_1, y_2)) = 3x_1y_1 - 2x_1y_2 + 4x_2y_1 - x_2y_2.$$

(i) Find the matrix A of f in the basis $\{u_1 = (1, 1), u_2 = (1, 2)\}$.

(ii) Find the matrix B of f in the basis $\{u_1 = (1, -1), u_2 = (3, 1)\}$.

(iii) Find the transition matrix P from $\{u_i\}$ to $\{v_i\}$ and verify that $B = P^tAP$.

(8) Let V be the vector space of 2×2 matrices over \mathbb{R} . Let

$$M = \begin{pmatrix} 1 & 2 \\ 3 & 5 \end{pmatrix}$$

and

$$f(A, B) = \text{tr}(A^tMB)$$

where $A, B \in V$ and tr denotes 'trace.'

(i) Show that f is a bilinear form on V .

(ii) Find the matrix of f in the basis $\left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\}$.

(9) Find the symmetric matrix belonging to each of the following quadratic polynomials:

(i) $q(x, y, z) = 2x^2 - 8xy + y^2 - 16xz + 14yz + 5z^2$

(ii) $q(x, y, z) = x^2 - xz + y^2$

(iii) $q(x, y, z) = xy + y^2 + 4xz + z^2$

(iv) $q(x, y, z) = xy + yz$

Solutions to Tutorial Questions 3.30

(1) (i) $6x - 5y + 4z$ (ii) $6x - 9y + 3z$ (iii) $-16x + 4y - 13z$

(2) $\phi(x, y) = 4x + 7y$; $\phi(-2, 7) = 41$.

(3) (i) $\{\phi_1(x, y, z) = x; \phi_2(x, y, z) = y; \phi_3(x, y, z) = z\}$

(ii) $\{\phi_1(x, y, z) = -3x - 5y - 2z; \phi_2(x, y, z) = 2x + y; \phi_3(x, y, z) = x + 2y + z\}$

(4) $\{f_1(t) = \frac{3}{2}(2t - t^3); f_2(t) = \frac{1}{4}(-2y + 3t^2); f_3(t) = \frac{1}{2}(2 - 6t + 3t^2)\}$.

(6) (i) yes (ii) no (iii)yes (iv)no (v)no (vi)yes.

(7) (i) $A = \begin{pmatrix} 4 & 1 \\ 7 & 3 \end{pmatrix}$; $B = \begin{pmatrix} 0 & -4 \\ 20 & 32 \end{pmatrix}$; $P = \begin{pmatrix} 3 & 5 \\ -2 & -2 \end{pmatrix}$.

(8) (ii) $\begin{pmatrix} 1 & 0 & 2 & 0 \\ 0 & 1 & 0 & 2 \\ 3 & 0 & 4 & 0 \\ 0 & 3 & 0 & 4 \end{pmatrix}$.

(9) (i) $\begin{pmatrix} 2 & -4 & -8 \\ -4 & 1 & 7 \\ -8 & 7 & 5 \end{pmatrix}$; (ii) $\begin{pmatrix} 1 & 0 & -\frac{1}{2} \\ 0 & 1 & 0 \\ -\frac{1}{2} & 0 & 0 \end{pmatrix}$;

(iii) $\begin{pmatrix} 0 & \frac{1}{2} & 2 \\ \frac{1}{2} & 1 & 0 \\ 2 & 0 & 1 \end{pmatrix}$; (iv) $\begin{pmatrix} 0 & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & \frac{1}{2} & 0 \end{pmatrix}$.

Chapter 4

Eigenvalues and Eigenvectors

Definition 4.1 (Eigenvalues, Eigenvectors, Eigenspace)

- (i) Let $T : V \rightarrow V$ be a linear operator on a vector space V over a field K . A scalar $\lambda \in K$ is called an eigenvalue of T if there exists a non-zero vector $v \in V$ for which

$$T(v) = \lambda v. \quad (4.1)$$

- (ii) Every vector v satisfying (4.1) is called an eigenvector of T belonging to the eigenvalue λ .
- (iii) The set of all such vectors which is a subspace of V is called the eigenspace of λ .

Note

- (i) If v is an eigenvector, then each scalar multiple kv is also an eigenvector since:

$$T(kv) = kT(v) = k(\lambda v) = \lambda(kv).$$

- (ii) The terms: characteristic value and characteristic vector or proper value and proper vector are frequently used instead of eigenvalue and eigenvector.

Examples 4.2

- (i) Let $I : V \rightarrow V$ be the identity mapping. Then, for every $v \in V$, $I(v) = v = 1v$. Hence 1 is an eigenvalue of I . Every vector in V is an eigenvector belonging to I .
- (ii) Let D be the differential operator on the vector space V of differentiable functions. We have $D(e^{5t}) = 5e^{5t}$. Hence 5 is an eigenvalue of D with eigenvector e^{5t} .
- (iii) If A is an n -square matrix over K , then an eigenvalue of A means an eigenvalue of A viewed as an operator on K^n . That is, $\lambda \in K$ is an eigenvalue of A if, for some non zero (column) vector $v \in K^n$, $Av = \lambda v$; where v is an eigenvector of A belonging to λ .

For example, if $A = \begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix}$, then we seek a scalar λ and a non-zero vector $X = \begin{pmatrix} x \\ y \end{pmatrix}$ such that $Ax = \lambda x$, i.e. $\begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \lambda \begin{pmatrix} x \\ y \end{pmatrix}$. This matrix equation is equivalent to the homogeneous system

$$\begin{aligned} x + 2y &= \lambda x \\ 3x + 2y &= -\lambda y \end{aligned}$$

or

$$\begin{aligned}(\lambda - 1)x - 2y &= 0 \\ -3x + (\lambda - 2)y &= 0\end{aligned}\tag{4.2}$$

This homogeneous system has a non zero solution if and only if the determinant of the matrix coefficients is zero, i.e.

$$\begin{vmatrix} \lambda - 1 & -2 \\ -3 & \lambda - 2 \end{vmatrix} = \lambda^2 - 3\lambda - 4 = (\lambda - 4)(\lambda + 1) = 0.$$

Thus λ is an eigenvalue of A if and only if $\lambda = 4$ or $\lambda = -1$. Setting $\lambda = 4$ in (4.2),

$$\begin{aligned}3x - 2y &= 0 \\ -3x + 2y &= 0\end{aligned}$$

or simply $3x - 2y = 0$. Thus $v = \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$ is a non zero eigenvector belonging to the eigenvalue $\lambda = 4$ and every other eigenvector belonging to $\lambda = 4$ is a multiple of v .

Setting $\lambda = -1$ in (4.2),

$$\begin{aligned}-2x - 2y &= 0 \\ -3x - 3y &= 0\end{aligned}$$

or simply $x + y = 0$. Thus $w = \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ is a non zero eigenvector belonging to the eigenvalue $\lambda = -1$ and every other eigenvector belonging to $\lambda = -1$ is a multiple of w .

Theorem 4.3

Let $T : V \rightarrow V$ be a linear operator on a vector space V over K . Then $\lambda \in K$ is an eigenvalue of T if and only if the operator $\lambda I - T$ is singular. The eigenspace of λ is then the kernel of $\lambda I - T$.

Proof.

λ is an eigenvalue of T if and only if there exists a non zero v such that

$$T(v) = \lambda v \text{ or } (\lambda I)v - T(v) = 0 \text{ or } (\lambda I - T)(v) = 0.$$

i.e. $\lambda I - T$ is singular. We also have that v is in the eigenspace of λ if and only if the above relation hold; hence v is in the kernel of $\lambda I - T$.

Theorem 4.4

Non-zero eigenvectors belonging to distinct eigenvalues are linearly independent.

Proof.

Let v_1, \dots, v_n be non zero eigenvectors of an operator $T : V \rightarrow V$ belonging to distinct eigenvalues $\lambda_1, \dots, \lambda_n$. We prove the theorem by induction on n .

If $n = 1$, then v_1 is linearly independent since $v_1 \neq 0$.

Assume $n > 1$ and suppose that

$$a_1v_1 + a_2v_2 + \dots + a_nv_n = 0\tag{4.3}$$

where the a_i are scalars. Applying T to the above relation, we obtain by linearity

$$a_1T(v_1) + a_2T(v_2) + \dots + a_nT(v_n) = T(0) = 0.$$

But by hypothesis $T(v_i) = \lambda_i v_i$, hence

$$a_1\lambda_1v_1 + a_2\lambda_2v_2 + \dots + a_n\lambda_nv_n = 0. \quad (4.4)$$

On the other hand, multiplying (4.3) by λ_n ,

$$a_1\lambda_nv_1 + a_1\lambda_nv_1 + \dots + a_n\lambda_nv_n = 0. \quad (4.5)$$

Now, subtracting (4.5) from (4.4),

$$a_1(\lambda_1 - \lambda_n)v_1 + a_2(\lambda_2 - \lambda_n)v_2 + \dots + a_{n-1}(\lambda_{n-1} - \lambda_n)v_{n-1} = 0.$$

By induction, each of the above coefficients is 0. Since the λ_i are distinct, $\lambda_i - \lambda_n \neq 0$ for $i \neq n$. Hence $a_1 = \dots = a_{n-1} = 0$.

Substituting this into (4.3) we get $a_nv_n = 0$ and hence $a_n = 0$. Thus the v_i are linearly independent.

Example

Consider the functions $e^{a_1t}, e^{a_2t}, \dots, e^{a_nt}$ where a_1, a_2, \dots, a_n are distinct real numbers. If D is the differential operator, then

$$D(e^{a_kt}) = a_k e^{a_kt}.$$

Accordingly, $e^{a_1t}, \dots, e^{a_nt}$ are eigenvectors of D belonging to the distinct eigenvalues a_1, \dots, a_n and so by (4.4) e^{a_it} are linearly independent.

Theorem 4.5

A linear operator $T : V \rightarrow V$ can be represented by a diagonal matrix B if and only if V has a basis consisting of eigenvectors of T . In this case the diagonal elements of B are the corresponding eigenvalues.

Proof.

Let the dimension of V be n . Then T can be represented by a diagonal matrix

$$\begin{pmatrix} k_1 & 0 & \cdot & \cdot & \cdot & 0 \\ 0 & k_2 & \cdot & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdot & \cdot & \cdot & k_n \end{pmatrix}$$

if and only if there exists a basis $\{v_1, \dots, v_n\}$ of V for which

$$T(v_1) = k_1v_1, \quad T(v_2) = k_2v_2, \quad T(v_n) = k_nv_n$$

that is, such that the vectors v_1, \dots, v_n are eigenvectors of T belonging respectively to eigenvalues k_1, \dots, k_n .

Note 4.6

- (i) An n -square matrix A is similar to a diagonal matrix B if and only if A has n linearly independent eigenvectors. In this case the diagonal elements of B are the corresponding eigenvalues.

- (ii) If we let P be the matrix A whose columns are the n independent eigenvectors of A , then $B = P^{-1}AP$. For example, if we consider the matrix A as in Example 4.2(iii), A has two independent eigenvectors $\begin{pmatrix} 2 \\ 3 \end{pmatrix}$ and $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$. Set

$$P = \begin{pmatrix} 2 & 1 \\ 3 & -1 \end{pmatrix} \text{ and so } P^{-1} = \begin{pmatrix} \frac{1}{5} & \frac{1}{5} \\ \frac{3}{5} & -\frac{2}{5} \end{pmatrix}.$$

Then A is similar to the diagonal matrix

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

As expected, the diagonal elements 4 and -1 of the diagonal matrix B are the eigenvalues corresponding to the given eigenvectors.

Tutorial Questions 4.7

- (1) For each matrix below, find all eigenvalues and linearly independent eigenvectors.

$$(i) A = \begin{pmatrix} 2 & 2 \\ 1 & 3 \end{pmatrix}, \quad (ii) B = \begin{pmatrix} 4 & 2 \\ 3 & 3 \end{pmatrix}; \quad (iii) C = \begin{pmatrix} 5 & -1 \\ 1 & 3 \end{pmatrix}.$$

Find invertible matrices P_1, P_2 and P_3 such that $P_1^{-1}AP_1, P_2^{-1}BP_2$ and $P_3^{-1}CP_3$ are diagonal.

- (2) For each matrix below, find all eigenvalues and a basis for each eigenspace.

$$(i) A = \begin{pmatrix} 3 & 1 & 1 \\ 2 & 4 & 2 \\ 1 & 1 & 3 \end{pmatrix}; \quad (ii) B = \begin{pmatrix} 1 & 2 & 2 \\ 1 & 2 & -1 \\ -1 & 1 & 4 \end{pmatrix}.$$

- (3) Let $A = \begin{pmatrix} 2 & -1 \\ 1 & 4 \end{pmatrix}$ and $B = \begin{pmatrix} 3 & -1 \\ 13 & -3 \end{pmatrix}$; find all eigenvalues and linearly independent eigenvectors of each matrix if:

- (i) they are matrices over the real field \mathbb{R} ;
(ii) they are matrices over the complex field \mathbb{C} .

- (4) Show that similar matrices A and A^t have the same eigenvalues.
(5) Show that matrices A and A^t have the same eigenvalues.
(6) Suppose that v is also an eigenvector of operator S and T . Show that v is also an eigenvector for the operator $aS + bT$ where a and b are scalars.
(7) For each of the following operators $T : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ find all eigenvalues and a basis for each eigenspace.

- (i) $T(x, y, z) = (x + y + z, 2y + z, 2y + 3z)$
(ii) $T(x, y, z) = (x + y, y + z, -2y - z)$
(iii) $T(x, y, z) = (x - y, 2x + 3y + 2z, x + y + 2z)$

Solutions to Tutorial Questions 4.8

- (1) (i) $t_1 = 1, u = (2, -1); t_2 = 4, v = (1, 1)$
(ii) $t_1 = 1, u = (2, -3); t_2 = 6, v = (1, 1)$
(iii) $t_1 = 4, u = (1, 1). P_1 = \begin{pmatrix} 2 & 1 \\ -1 & 1 \end{pmatrix}; P_2 = \begin{pmatrix} 2 & 1 \\ -3 & 1 \end{pmatrix}.$

P_3 does not exist since C has only one independent eigenvector and so cannot be diagonalized.

- (2) (i) $t_1 = 2, \quad u = (1, -1, 0), \quad v = (1, 0, -1)$
 $t_2 = 6, \quad w = (1, 2, 1).$
(ii) $t_1 = 3, \quad u = (1, 1, 0), \quad v = (1, 0, 1)$
 $t_2 = 1, \quad w = (2, -1, 1).$
- (3) (i) For $A, \quad t = 3, \quad u = (1, -1)$ B has no eigenvalues in \mathbb{R} .
(ii) For A , it is the same as in (i).
For $B, \quad t_1 = 2i, \quad u_1 = (1, 3, -2i), \quad u_2 = (3 + 2i, 13)$
 $t_2 = -2i, \quad u_1 = (1, 3 + 2i), \quad u_2 = (3 - 2i, 13)$
- (7) (i) $t_1 = 1, \quad u = (1, 0, 0); \quad t_2 = 4, \quad v = (1, 1, -2).$
(ii) $t_1 = 1, \quad u = (1, 0, 0).$ No other eigenvalues in \mathbb{R} .
(iii) $t_1 = 1, \quad u = (1, 0, -1), \quad t_2 = 2, \quad v = (2, -2, -1), \quad t_3 = 3, \quad w = (1, -2, -1).$

Chapter 5

Characteristic and Minimum Polynomials

Definition 5.1 (Characteristic Matrix, Characteristic Polynomial, Characteristic Equation and Monic Polynomial)

- (i) Let A be an n -square matrix over a field K . The matrix $tI_n - A$, where I_n is the n -square identity matrix and t is an indeterminate, is called the **characteristic matrix of A** .
- (ii) The determinant of $tI_n - A$, denoted by $\Delta_A(t)$ or $\det(tI_n - A)$, which is a polynomial in t is called the **characteristic polynomial of A** .
- (iii) The equation

$$\Delta_A(t) = \det(tI_n - A) = 0.$$

is called the **characteristic equation of A** .

- (iv) A polynomial is **monic** if its leading coefficient is 1.

Example 5.2

If $A = \begin{pmatrix} 1 & 3 \\ 2 & 1 \end{pmatrix}$ then its characteristic matrix is

$$tI_2 - A = \begin{pmatrix} t-1 & -3 \\ -2 & t-1 \end{pmatrix}$$

while its characteristic polynomial is

$$\det(tI_2 - A) = \begin{vmatrix} t-1 & -3 \\ -2 & t-1 \end{vmatrix} = t^2 - 2t - 5$$

and its characteristic equation is $t^2 - 2t - 5 = 0$.

Theorem 5.3

The characteristic polynomial of a matrix A is a monic polynomial of degree n and the coefficient of t^{n-1} is the negative of the trace of A . Moreover the constant term of the polynomial is $(-1)^n |A|$ where n is the order of A .

Example 5.4

(1) From the matrix in Example 5.2,

$$\Delta_A(t) = t^2 - 2t - 5$$

which is a monic polynomial of degree 2. Here $n = 2$. The coefficient of t^{n-1} which is the coefficient of t is -2 . The trace of A is $1 + 1 = 2$. Therefore, the negative of the trace of A is the coefficient of t . The constant term is

$$\Delta_A(0) = (-1)^n |A| = (-1)^2 (-5) = -5.$$

(2) Find the characteristic polynomial of $A = \begin{pmatrix} 2 & 3 \\ 4 & 5 \end{pmatrix}$.

Solution

Here $n = 2$.

Coefficient of $t^2 = 1$.

Coefficient of $t = -(2 + 5) = -7$.

Constant term $= (-1)^2 |A| = (-1)^2 (-2) = -2$.

Therefore the characteristic polynomial is $t^2 - 7t - 2$.

Theorem 5.5 (Cayley-Hamilton Theorem)

Every matrix is a zero of its characteristic polynomial.

Example 5.6

The characteristic polynomial of the matrix A in Example 5.2 is $\Delta(t) = t^2 - 2t - 5$.

As expected from the Cayley-Hamilton Theorem, A is a zero of $\Delta(t)$ since:

$$\Delta(A) = \begin{pmatrix} 1 & 3 \\ 2 & 1 \end{pmatrix}^2 - 2 \begin{pmatrix} 1 & 3 \\ 2 & 1 \end{pmatrix} - 5 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Theorem 5.7

Let A be an n -square matrix over a field K . A scalar $\lambda \in K$ is an eigenvalue of A if and only if λ is a root of the characteristic polynomial of A .

Proof.

We recall that λ is an eigenvalue of A if and only if $\lambda I - A$ is singular. Furthermore, $\lambda I - A$ is singular if and only if $|\lambda I - A| = 0$, i.e. λ is a root of $\Delta(t)$. Hence the theorem.

Corollary 5.8

If the characteristic polynomial $\Delta(t)$ of an n -square matrix A is a product of distinct linear factors i.e.

$$\Delta(t) = (t - a_1)(t - a_2)\dots(t - a_n)$$

where a_1, \dots, a_n are distinct roots of $\Delta(t)$, then A is similar to a diagonal matrix whose diagonal elements are the a_i .

Example 5.9

Let $A = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 2 & -5 \\ 0 & 1 & -2 \end{pmatrix}$.

Is it diagonalizable over: (i) the real field \mathbb{R} ; (ii) the complex field \mathbb{C} ?

Solution

The characteristic polynomial of A is $\Delta(t) = (t - 3)(t^2 + 1)$.

- (i) If A is a matrix over the real field \mathbb{R} , then A has only one eigenvalue 3. Since 3 has only one independent eigenvector, A is not diagonalizable.
- (ii) If A is a matrix over the complex field \mathbb{C} , then A has three distinct eigenvalues 3, i and $-i$. Thus, there exists an invertible matrix P over the complex field \mathbb{C} for which

$$P^{-1}AP = \begin{pmatrix} 3 & 0 & 0 \\ 0 & i & 0 \\ 0 & 0 & -i \end{pmatrix} \text{ and } A \text{ is diagonalizable.}$$

Theorem 5.10

Similar matrices have the same characteristic polynomial.

Proof.

Suppose that A and B are similar matrices, say $B = P^{-1}AP$ where P is invertible. We show that A and B have the same characteristic polynomial. Using $tI = P^{-1}IP$, we have

$$\begin{aligned} |tI - B| &= |tI - P^{-1}AP| \\ &= |P^{-1}tIP - P^{-1}AP| \\ &= |P^{-1}(tI - A)P| \\ &= |P^{-1}||tI - A||P| \end{aligned}$$

Since determinants are scalars, commute and since $|P^{-1}||P| = 1$, we obtain

$$|tI - B| = |tI - A|.$$

Theorem 5.11 (Minimum Polynomial)

A non zero polynomial $m(t)$ is called a minimum polynomial of A if

- (i) $m(t)$ is a monic polynomial for which $m(A) = 0$ and
- (ii) of all such polynomial $m(t)$ is of the lowest degree.

Theorem 5.12

The minimum polynomial of a matrix A exists and is unique.

Proof

By Theorem 5.5, A is a zero of some non zero polynomial; the characteristic polynomial. Let n be the lowest degree for which a polynomial $f(t)$ exists such that $f(A) = 0$. Dividing $f(t)$ by its

leading coefficient, we obtain a monic polynomial $m(t)$ of degree n which has A as zero. Suppose that $m^1(t)$ is another monic polynomial of degree n which has A as a zero, i.e. $m^1(A) = 0$. Then the difference $m(t) - m^1(t)$ is a non zero polynomial of degree less than n which has A as a zero.

This contradicts the original assumption on n , hence $m(t)$ is a unique minimum polynomial.

Theorem 5.13

The minimum polynomial $m(t)$ of a matrix A divides every polynomial which has A as a zero. In particular, $m(t)$ divides the characteristic polynomial of A .

Proof

Suppose that $f(t)$ is a polynomial for which $f(A) = 0$. By the division algorithm, there exist polynomials $q(t)$ and $r(t)$ for which

$$f(t) = m(t)q(t) + r(t); \quad r(t) = 0 \quad \text{or} \quad \deg r(t) < \deg m(t).$$

Substituting $t = A$ in this equation and using that $f(A) = 0$ and $m(A) = 0$, we obtain $r(A) = 0$. If $r(t) \neq 0$, then $r(t)$ is a polynomial of degree less than $m(t)$ which has A as a zero; this contradicts the definition of the minimum polynomial. Thus $r(t) = 0$ and so

$$f(t) = m(t)q(t),$$

i.e. $m(t)$ divides $f(t)$.

Theorem 5.14

Let $m(t)$ be the minimum polynomial of an n -square matrix A , then the characteristic polynomial of A divides $(m(t))^n$.

Theorem 5.15

The characteristic polynomial $\Delta(t)$ and the minimum polynomial $m(t)$ of a matrix A have the same irreducible factors.

Proof.

Suppose that $f(t)$ is an irreducible polynomial. If $f(t)$ divides $m(t)$, since $m(t)$ divides $\Delta(t)$, then $f(t)$ divides $\Delta(t)$.

On the other hand, if $f(t)$ divides $\Delta(t)$ then, by Theorem 5.14, $f(t)$ divides $(m(t))^n$. But $f(t)$ is irreducible; hence $f(t)$ also divides $m(t)$. Thus $m(t)$ and $\Delta(t)$ have the same irreducible factors.

Example 5.16

Find the minimum polynomial $m(t)$ of $m = \begin{pmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & -2 & 4 \end{pmatrix}$.

Solution

The characteristic polynomial of A is $\Delta(t) = (t - 3)(t - 2)^3$.

The minimum polynomial $m(t)$ must divide $\Delta(t)$. Also, each irreducible factor of $\Delta(t)$, i.e.

$t - 2$ and $t - 3$ must be a factor of $m(t)$. Thus $m(t)$ is exactly one of the following:

$$\begin{aligned} m_1(t) &= (t - 3)(t - 2) \\ m_2(t) &= (t - 3)(t - 2)^2 \\ m_3(t) &= (t - 3)(t - 2)^3. \end{aligned}$$

Clearly, $m_3(A) = \Delta(A) = 0$ by Theorem 5.5.

It is easy to verify that $m_1(A) \neq 0$ and $m_2(A) = 0$. Thus $m_2(t) = (t - 3)(t - 2)^2$ is the minimum polynomial of A .

Theorem 5.17

Let $M = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$

where A and B are square matrices, then the minimum polynomial $m(t)$ of M is the least common multiple of the minimum polynomial $m_1(t)$ and $m_2(t)$ of A and B respectively.

Proof.

Since $m(t)$ is the minimum polynomial of M , then

$$m(M) = \begin{pmatrix} m(A) & 0 \\ 0 & m(B) \end{pmatrix} = 0$$

and hence $m(A) = m(B) = 0$. Since $m_1(t)$ is the minimum polynomial of A , $m_1(t)$ divides $m(t)$. Similarly, $m_2(t)$ divides $m(t)$. Thus $m(t)$ is a multiple of $m_1(t)$ and $m_2(t)$.

Now let $f(t)$ be another multiple of $m_1(t)$ and $m_2(t)$; then

$$f(M) = \begin{pmatrix} f(A) & 0 \\ 0 & f(B) \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} = 0.$$

But $m(t)$ is the minimum polynomial of M ; hence $m(t)$ divides $f(t)$. Thus $m(t)$ is the least common multiple of $g(t)$ and $h(t)$.

Note

Theorem 5.17 can be generalized by induction. For let

$$M = \begin{pmatrix} A_1 & 0 & \dots & 0 \\ 0 & A_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & A_r \end{pmatrix}$$

where the A_i are square matrices. The minimum polynomial is the least common multiple of the minimum polynomials of the A_i .

Example 5.18

Use Theorem 5.17 to solve Example 5.16.

Solution

$$M = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} \text{ where } A = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 1 \\ -2 & 4 \end{pmatrix}.$$

The minimum polynomials of A and B are $(t - 2)^2$ and $(t - 2)(t - 3)$ respectively. The least

common multiple is $(t - 3)(t - 2)^2$ which is the minimum polynomial of M .

Theorem 5.19 (Characteristic Polynomial of T).

Let $T : V \rightarrow V$ be a linear operator on a vector space V with finite dimension. We define the characteristic polynomial $\Delta(t)$ of T to be the characteristic polynomial of any matrix representation of T .

Note 5.20

- (i) By Theorem 5.10, $\Delta(t)$ is independent of the particular basis in which the matrix representation is computed.
- (ii) The degree of $\Delta(t)$ is equal to the dimension of V .
- (iii) We have theorems for T which are similar to the ones we had for matrices:
 - (a) For Theorem 5.5, we have: T is a zero of its characteristic polynomial.
 - (b) For Theorem 5.7, we have: A scalar $\lambda \in K$ is an eigenvalue of T if and only if λ is a root of the characteristic polynomial of T .

Example 5.21

Let V be the vector space of functions which has $\{\sin \theta, \cos \theta\}$ as a basis and let D be the differential operator on V . Find the characteristic polynomial of D .

Solution

$$\begin{aligned} D(\sin \theta) &= \cos \theta = 0(\sin \theta) + 1(\cos \theta) \\ D(\cos \theta) &= -\sin \theta = -1(\sin \theta) + 0(\cos \theta). \end{aligned}$$

The matrix A of D in the above basis is $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$.

Thus

$$\det(tI - A) = \begin{vmatrix} t & -1 \\ 1 & t \end{vmatrix} = t^2 + 1.$$

Definition 5.22 (Minimum Polynomial of T)

The minimum polynomial $m(t)$ of a linear operator T is the polynomial of lowest degree with 1 as the leading coefficient which has T as a zero.

Note

- (i) For any polynomial $f(t)$, $f(T) = 0$ if and only if $f(A) = 0$ where A is any matrix representation of T . Accordingly, T and A have the same minimum polynomial.
- (ii) All the theorems we had on the minimum polynomial of a matrix also hold for the minimum polynomial of a matrix also hold for the minimum polynomial of the operator T .

Tutorial Questions 5.23

(1) Show that a matrix A and its transpose A^t have the same characteristic polynomial. Show the same for minimum polynomial.

(2) For each matrix below, find a polynomial for which the matrix is a root:

$$(i) A = \begin{pmatrix} 3 & -7 \\ 4 & 5 \end{pmatrix} \quad (ii) B = \begin{pmatrix} 5 & -1 \\ 8 & 3 \end{pmatrix} \quad (iii) C = \begin{pmatrix} 2 & 3 & -2 \\ 0 & 5 & 4 \\ 1 & 0 & -1 \end{pmatrix}.$$

(3) Find the characteristic and minimum polynomial of each matrix below:

$$(i) A = \begin{pmatrix} 2 & 5 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 4 & 2 & 0 \\ 0 & 0 & 3 & 5 & 0 \\ 0 & 0 & 0 & 0 & 7 \end{pmatrix}; \quad (ii) B = \begin{pmatrix} 3 & 1 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 & 0 \\ 0 & 0 & 3 & 1 & 0 \\ 0 & 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 & 3 \end{pmatrix}$$

$$(iii) C = \begin{pmatrix} \lambda & 0 & 0 & 0 & 0 \\ 0 & \lambda & 0 & 0 & 0 \\ 0 & 0 & \lambda & 0 & 0 \\ 0 & 0 & 0 & \lambda & 0 \\ 0 & 0 & 0 & 0 & \lambda \end{pmatrix}; \quad (iv) D = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & 2 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

Solutions to Tutorial Questions 5.24

- (2) (i) $f(t) = t^2 - 8t + 43$
(ii) $g(t) = t^2 - 8t + 23$
(iii) $h(t) = t^3 - 6t^2 + 5t - 12$
- (3) (i) $\Delta(t) = (t - 2)^3(t - 7)^2$; $m(t) = (t - 2)^2(t - 7)$
(ii) $\Delta(t) = (t - 3)^5$; $m(t) = (t - 3)^3$
(iii) $\Delta(t) = (t - \lambda)^5$; $m(t) = (t - \lambda)$
(iv) $\Delta(t) = (t - 1)^3(t - 2)^3$; $m(t) = (t - 1)(t - 2)$.

Chapter 6

Canonical Forms

Theorem 6.1 (Jordan Canonical Form)

Let $T : V \rightarrow V$ be a linear operator whose characteristic and minimum polynomials are respectively:

$$\Delta(t) = (t - \lambda_1)^{n_1} \dots (t - \lambda_r)^{n_r}$$

and

$$m(t) = (t - \lambda_1)^{m_1} \dots (t - \lambda_r)^{m_r}$$

where the λ_i are distinct scalars. Then T has a block diagonal matrix representation of J whose diagonal entries are of the form

$$J_{ij} = \begin{pmatrix} \lambda_i & 1 & 0 & \dots & 0 & 0 \\ 0 & \lambda_i & 1 & \dots & 0 & 0 \\ 0 & 0 & 0 & \dots & \lambda_i & 1 \\ 0 & 0 & 0 & \dots & 0 & \lambda_i \end{pmatrix}.$$

For each λ_i the corresponding blocks J_{ij} have the following properties:

- (i) There is at least one J_{ij} of order m_i , all other J_{ij} are of the order $\leq m_i$.
- (ii) The sum of the order of the J_{ij} is n_i .
- (iii) The number of J_{ij} equals the geometric multiplicity of λ_i .
- (iv) The number of J_{ij} of each possible order is the uniquely determined by T .

Definition 6.2 (Jordan Canonical Form; Jordan Block)

A diagonal block matrix J whose diagonal blocks are the J_{ij} is called the Jordan canonical form of the operator T . A diagonal block J_{ij} is called a Jordan block belonging to the eigenvalue λ_i .

Example 6.3

- (1) Suppose that the characteristic and minimum polynomials of an operator T are respectively

$$\Delta(t) = (t - 2)^4(t - 3)^3$$

and

$$m(t) = (t - 2)^2(t - 3)^2.$$

$$\begin{pmatrix} 2 & 1 & 0 & . & . \\ 0 & 2 & 1 & . & . \\ 0 & 0 & 2 & . & . \\ . & . & . & . & . \\ . & . & . & . & 5 \\ . & . & . & . & . \\ . & . & . & . & . \\ . & . & . & . & . \\ . & . & . & . & . \\ . & . & . & . & . \end{pmatrix}, \quad \begin{pmatrix} 2 & 1 & . & . & . \\ 0 & 2 & . & . & . \\ . & . & . & . & . \\ . & . & . & . & 2 \\ . & . & . & . & . \\ . & . & . & . & . \\ . & . & . & . & . \\ . & . & . & . & 5 \\ . & . & . & . & . \\ . & . & . & . & . \\ . & . & . & . & 5 \end{pmatrix}, \quad \begin{pmatrix} 2 & . & . & . & . \\ . & . & . & . & . \\ . & . & . & . & . \\ . & . & . & 2 & . \\ . & . & . & . & . \\ . & . & . & . & . \\ . & . & . & . & 2 \\ . & . & . & . & . \\ . & . & . & . & . \\ . & . & . & . & 5 \\ . & . & . & . & . \\ . & . & . & . & . \\ . & . & . & . & . \end{pmatrix} \quad \begin{matrix} | \\ \dots \\ | \end{matrix}$$

$(iv) \qquad (v) \qquad (vi)$

(3) Determine all possible Jordan canonical forms J for matrix of order 5 whose minimal polynomial is $m(t) = (t - 2)^2$.

Solution

J must have one Jordan block of order 2 and the others must be of order 2 or 1. Thus there are only two possibilities.

$$J = \begin{pmatrix} 2 & 1 & . & . & . \\ 0 & 2 & . & . & . \\ . & . & . & . & . \\ . & . & . & 2 & 1 \\ . & . & . & 0 & 2 \\ . & . & . & . & . \\ . & . & . & . & 2 \\ . & . & . & . & . \end{pmatrix}, \quad J = \begin{pmatrix} 2 & 1 & . & . & . \\ 0 & 2 & . & . & . \\ . & . & . & . & . \\ . & . & . & . & 2 \\ . & . & . & . & . \\ . & . & . & . & . \\ . & . & . & . & 2 \\ . & . & . & . & . \\ . & . & . & . & 2 \\ . & . & . & . & . \end{pmatrix}.$$

Definition 6.4 (*T*-Cyclic Subspace)

Let T be a linear operator on a vector space V of finite dimension over K . Suppose $v \in V$ and $v \neq 0$. The set of all vectors of the form $f(T)(v)$, where $f(t)$ ranges over all polynomials over K , is a T -invariant subspace of V called the T -cyclic subspace of V generated by v . We denote it by $Z(v, T)$ and denote the restriction of T to $Z(v, T)$ by T_v . We could equivalently define $Z(v, T)$ as the intersection of all T -invariant subspaces of V containing v .

Definition 6.5 (*T*-annihilator)

Consider the sequence

$$v, T(v), T^2(v), T^3(v), \dots$$

of powers of T acting on v . Let k be the lowest integer such that $T^k(v)$ is a linear combination of those vectors which precede it in the sequence, say

$$T^k(v) = -a_{k-1}T^{k-1}(v) - \dots - a_1T(v) - a_0v.$$

Then

$$m_v(t) = t^k + a_{k-1}t^{k-1} + \dots + a_1t + a_0$$

is the unique monic polynomial of lowest degree for which $m_v(T)(v) = 0$. We call $m_v(t)$ the T -annihilator of v and $Z(v, T)$.

Theorem 6.6

Let $Z(v, T)$, T_v and $m_v(t)$ be defined as above. Then

- (i) The set $\{v, T(v), \dots, T^{k-1}(v)\}$ is a basis of $Z(v, T)$; hence $\dim Z(v, T) = k$.
- (ii) The minimal polynomial of T_v is $m_v(t)$.
- (iii) The matrix representation of T_v in the above basis is

$$C = \begin{pmatrix} 0 & 0 & 0 & \dots & \dots & \dots & 0 & -a_0 \\ 1 & 0 & 0 & \dots & \dots & \dots & 0 & -a_1 \\ 0 & 1 & 0 & \dots & \dots & \dots & 0 & -a_2 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & \dots & \dots & 0 & -a_{k-2} \\ 0 & 0 & 0 & \dots & \dots & \dots & 1 & -a_{k-1} \end{pmatrix}$$

Definition 6.7 (Companion Matrix)

The above matrix C is called the companion matrix of the polynomial $m_v(t)$.

Theorem 6.8

Let $T : V \rightarrow V$ be a linear operator whose minimal polynomial is $(f(t))^n$ where $f(t)$ is a monic irreducible polynomial. Then V is the direct sum

$$V = Z(v_1, T) \oplus \dots \oplus Z(v_r, T)$$

of T -cyclic subspaces $Z(v_i, T)$ with corresponding T -annihilators:

$$(f(t))^{n_1}, (f(t))^{n_2}, \dots, (f(t))^{n_r}, \quad n = n_1 \geq n_2 \geq \dots \geq n_r.$$

Any other decomposition of V into T -cyclic subspace has the same number of components and the same set of T -annihilators.

Theorem 6.9

Let $T : V \rightarrow V$ be a linear operator with minimal polynomial

$$m(t) = (f_1(t))^{m_1}(f_2(t))^{m_2} \dots (f_s(t))^{m_s}$$

where the $f_i(t)$ are distinct monic irreducible polynomials. Then T has a unique block diagonal matrix representation.

$$\begin{pmatrix} c_{11} & & & & & \\ & \dots & & & & \\ & & c_1 r_1 & & & \\ & & & \dots & & \\ & & & & c_{s1} & \\ & & & & & \dots \\ & & & & & & c_s r_s \end{pmatrix}$$

where the c_{ij} are companion matrices. In particular, the c_{ij} are the companion matrices of the polynomials $(f_i(t))^{n_{ij}}$, where

$$m_1 = n_{11} \geq n_{12} \geq \dots \geq n_{1r_1}, \dots, m_s = n_{s1} \geq n_{s2} \geq \dots \geq n_{sr_s}.$$

Definition 6.10 (Rational Canonical Form)

The above matrix representation of T is called its rational canonical form.

Examples 6.11

- (1) Let V be a vector space of dimension b over \mathbb{R} and let T be a linear operator whose minimal polynomial is

$$m(t) = (t^2 - t + 3)(t - 2)^2.$$

Find the possible rational canonical forms of T .

Solution

The rational canonical form of T is one of the following direct sum of companion matrices:

- (i) $C(t^2 - t + 3) \oplus C(t^2 - t + 3) \oplus C((t - 2)^2)$.
- (ii) $C(t^2 - t + 3) \oplus C((t - 2)^2) \oplus C((t - 2)^2)$.
- (iii) $C(t^2 - t + 3) \oplus C((t - 2)^2) \oplus C(t - 2) \oplus (t - 2)$

where $C(f(t))$ is the companion matrix of $f(t)$, that is:

$$\begin{pmatrix} 0 & -3 & \cdot & \cdot & \cdot & \cdot & \cdot \\ 1 & 1 & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & 0 & -3 & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & 0 & 1 & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 0 & -4 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 1 & 4 \end{pmatrix}, \quad \begin{pmatrix} 0 & -3 & \cdot & \cdot & \cdot & \cdot & \cdot \\ 1 & 1 & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & 0 & -4 & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & 1 & 4 & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 0 & -4 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 1 & 4 \end{pmatrix},$$

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

- (2) Let V be a vector space of dimension 7 over \mathbb{R} and let $T : V \rightarrow V$ be a linear operator with minimal polynomial

$$m(t) = (t^2 + 2)(t + 3)^3.$$

Find all the possible rational canonical forms for T .

Solution

The sum of the degrees of the companion matrices must add up to 7. Also, one companion matrix must be $t^2 + 2$ and one must be $(t + 3)^3$. Thus the rational canonical form of T is exactly one of the following direct sum of companion matrices.

- (i) $C(t^2 + 2) \oplus C(t^2 + 2) \oplus C((t + 3)^3)$
 (ii) $C(t^2 + 2) \oplus C((t + 3)^3) \oplus C((t + 3)^2)$
 (iii) $C(t^2 + 2) \oplus C((t + 3)^3) \oplus C(t + 3) \oplus C(t + 3)$

that is:

$$\begin{pmatrix} 0 & -2 & & & & & & & \\ 1 & 0 & & & & & & & \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & & & \\ & & & 0 & -2 & & & & \\ & & & \cdot & 1 & 0 & & & \\ & & \cdot & \cdot & \cdot & \cdot & \cdot & & \\ & & & & & & 0 & 0 & -27 \\ & & & & & & \cdot & 1 & 0 & -27 \\ & & & & & & \cdot & 0 & 1 & -9 \end{pmatrix}, \quad \begin{pmatrix} 0 & -2 & & & & & & & \\ 1 & 0 & & & & & & & \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & & & \\ & & & 0 & 0 & -27 & & & \\ & & & \cdot & 1 & 0 & -27 & & \\ & & & \cdot & 0 & 1 & -9 & & \\ & & & & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ & & & & & & & & 0 & -9 \\ & & & & & & & & \cdot & 1 & -6 \end{pmatrix},$$

(i) (ii)

$$\begin{pmatrix} 0 & -2 & & & & & & & \\ 1 & 0 & & & & & & & \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & & & \\ & & & 0 & 0 & -27 & & & \\ & & & \cdot & 1 & 0 & -27 & & \\ & & & \cdot & 0 & 1 & -9 & & \\ & & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \\ & & & & & & & & -3 & \\ & & & & & & & \cdot & \cdot & \cdot & \\ & & & & & & & & & & -3 \end{pmatrix}.$$

(iii)

Tutorial Questions 6.12

- (1) Find all possible Jordan canonical forms for those matrices whose characteristic polynomial $\Delta(t)$ and minimal polynomial $m(t)$ are as follows:
- (i) $\Delta(t) = (t - 2)^4(t - 3)^2$, $m(t) = (t - 2)^2(t - 3)^2$
 (ii) $\Delta(t) = (t - 7)^5$, $m(t) = (t - 7)^2$

(iii) $\Delta(t) = (t - 2)^7, m(t) = (t - 2)^3$

(iv) $\Delta(t) = (t - 3)^4(t - 5)^4, m(t) = (t - 3)^2(t - 5)^2$.

Find all possible rational canonical forms for:

(i) 6×6 matrices with minimum polynomial $m(t) = (t^2 + 3)(t + 1)^2$

(ii) 6×6 matrices with minimum polynomial $m(t) = (t + 1)^3$

(iii) 8×8 matrices with minimum polynomial $m(t) = (t^2 + 2)^2(t + 3)^2$.

(3) Let A be a 4×4 matrix with minimum polynomial $m(t) = (t^2 + 1)(t^2 - 3)$. Find the rational canonical form for A if A is a matrix over:

(i) the rational field \mathbb{Q}

(ii) the real field \mathbb{R}

(iii) the complex field \mathbb{C} .

(4) Find the rational canonical form for the Jordan block $\begin{pmatrix} \lambda & 1 & 0 & 0 \\ 0 & \lambda & 1 & 0 \\ 0 & 0 & \lambda & 1 \\ 0 & 0 & 0 & \lambda \end{pmatrix}$.