LINEAR ALGEBRA II

J. B. Cooper
Johannes Kepler Universität Linz

Contents

1 DETERMINANTS

1.1 Introduction

In this chapter we treat one of the most important themes of linear algebra—that of the determinant. We begin with some remarks which will motivate the formal definition:

I. Recall that the system

$$\begin{aligned} ax + by &= e \\ cx + dy &= f \end{aligned}$$

has the unique solution

$$x = \frac{ed - fb}{ad - bc} \quad y = \frac{af - ce}{ad - bc}$$

provided that the denominator $ad - bc$ is non-zero. If we introduce the notation

$$\det \begin{bmatrix} a & b \\ c & d \end{bmatrix} = ad - bc$$

we can write the solution in the form

$$x = \det \begin{bmatrix} e & b \\ f & d \end{bmatrix} \div \det \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

$$y = \det \begin{bmatrix} a & e \\ c & f \end{bmatrix} \div \det \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

(Note that the numerators are formed by replacing the column of the matrix

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

corresponding to the unknown by the column vector on the right hand side).

Earlier, we displayed a similar formula for the solution of a system of three equations in three unknowns. It is therefore natural to ask whether we can define a function det on the space M_n of $n \times n$ matrices so that the solution of the equation $AX = Y$ is, under suitable conditions, given by the formula

$$x_i = \frac{\det A_i}{\det A}$$

where A_i is the matrix that we obtain by replacing the i-th column of A by Y i.e.

$$A_i = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1,i-1} & y_1 & a_{1,i+1} & \dots & a_{1n} \\ \vdots & & & & & & \vdots & \\ a_{n1} & a_{n2} & \dots & a_{n,i-1} & y_n & a_{n,i+1} & \dots & a_{nn} \end{bmatrix}.$$

II. Recall that

$$ad - bc = \det \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

is the area of the parallelogram spanned by the vectors (a, c) and (b, d). Now if f is the corresponding linear mapping on $\mathbf{R}^2$, this is just the image of the standard unit square (i.e. the square with vertices $(0, 0), (1, 0), (0, 1), (1, 1)$) under f. The natural generalisation would be to define the determinant of an $n \times n$ matrix to be the n-dimensional volume of the image of the standard hypercube in $\mathbf{R}^n$ under the linear mapping induced by the matrix. Although we do not intend to give a rigorous treatment of the volume concept in higher dimensional spaces, it is geometrically clear that it should have the following properties:

a) the volume of the standard hypercube is 1. This means that the determinant of the unit matrix is 1;

b) the volume depends linearly on the length of a fixed side. This means that the function det is linear in each column i.e.

$$\det[A_1 \dots A_i + A_i' \dots A_n] = \det[A_1 \dots A)_i \dots A_n] + \det[A_1 \dots A_i' \dots A_n]$$

and

$$\det[A_1 \dots \lambda A_i \dots A_n] = \lambda \det[A_1 \dots A_n].$$

c) The volume of a degenerate parallelopiped is zero. This means that if two columns of the matrix coincide, then its determinant vanishes.

(Note that the volume referred to here can take on negative values—depending on the orientation of the parallelopiped).

1.2 Existence of the determinant and how to calculate it

We shall now proceed to show that a function with the above properties exists. In fact it will be more convenient to demand the analogous properties for the rows i.e. we shall construct, for each n, a function

$$\det : M_n \to \mathbf{R}$$

with the properties
d1) $\det I_n = 1$;
d2)

$$\det \begin{bmatrix} A_1 \\ \vdots \\ \lambda A_i + \mu A_i' \\ \vdots \\ A_n \end{bmatrix} = \lambda \det \begin{bmatrix} A_1 \\ \vdots \\ A_i \\ \vdots \\ A_n \end{bmatrix} + \mu \det \begin{bmatrix} A_1 \\ \vdots \\ A_i' \\ \vdots \\ A_n \end{bmatrix}.$$

d3) if $A_i = A_j (i \neq j)$, then

$$\det \begin{bmatrix} A_1 \\ \vdots \\ A_i \\ \vdots \\ A_j \\ \vdots \\ A_n \end{bmatrix} = 0.$$

Before we prove the existence of such a function, we shall derive some further properties which are a consequence of d1) - d3):
d4) if we add a multiple of one row to another one, the value of the determinant remains unaltered i.e.

$$\det \begin{bmatrix} A_1 \\ \vdots \\ A_i \\ \vdots \\ A_j \\ \vdots \\ A_n \end{bmatrix} = \det \begin{bmatrix} A_1 \\ \vdots \\ A_i + A_j \\ \vdots \\ A_j \\ \vdots \\ A_n \end{bmatrix};$$

d5) if we interchange two rows of a matrix, then we alter the sign of the determinant i.e.

$$\det \begin{bmatrix} A_1 \\ \vdots \\ A_i \\ \vdots \\ A_j \\ \vdots \\ A_n \end{bmatrix} = -\det \begin{bmatrix} A_1 \\ \vdots \\ A_j \\ \vdots \\ A_i \\ \vdots \\ A_n \end{bmatrix}.$$

d6) if one row of A is a linear combination of the others, then $\det A = 0$. Hence if $r(A) < n$ (i.e. if A is not invertible), then $\det A = 0$.

PROOF. d4)

$$\det \begin{bmatrix} A_1 \\ \vdots \\ A_i + A_j \\ \vdots \\ A_j \\ \vdots \\ A_n \end{bmatrix} = \det \begin{bmatrix} A_1 \\ \vdots \\ A_i \\ \vdots \\ A_j \\ \vdots \\ A_n \end{bmatrix} + \det \begin{bmatrix} A_1 \\ \vdots \\ A_j \\ \vdots \\ A_j \\ \vdots \\ A_n \end{bmatrix} = \det \begin{bmatrix} A_1 \\ \vdots \\ A_i \\ \vdots \\ A_j \\ \vdots \\ A_n \end{bmatrix}$$

by d3).

d5)

$$\det \begin{bmatrix} A_1 \\ \vdots \\ A_i \\ \vdots \\ A_j \\ \vdots \\ A_n \end{bmatrix} = \det \begin{bmatrix} A_1 \\ \vdots \\ A_i + A_j \\ \vdots \\ A_j \\ \vdots \\ A_n \end{bmatrix} \det \begin{bmatrix} A_1 \\ \vdots \\ A_i + A_j \\ \vdots \\ -A_i \\ \vdots \\ A_n \end{bmatrix} = -\det \begin{bmatrix} A_1 \\ \vdots \\ A_j \\ \vdots \\ A_i \\ \vdots \\ A_n \end{bmatrix}.$$

d6) Suppose that $A_i = \lambda_1 A_1 + \cdots + \lambda_{i-1} A_{i-1}$. Then

$$\det \begin{bmatrix} A_1 \\ \vdots \\ A_{i-1} \\ A_i \\ \vdots \\ A_n \end{bmatrix} = \det \begin{bmatrix} A_1 \\ \vdots \\ A_{i-1} \\ (\lambda_1 A_1 + \ldots \lambda_{i-1} A_{i-1}) \\ \vdots \\ A_n \end{bmatrix} = 0$$

since if we expand the expression by using the linearity in the i-th row we obtain a sum of multiples of determinants each of which has two identical rows and these vanish.

Note the fact that with this information we are able to calculate the determinant of a given matrix, despite the fact that it has not yet been defined! We simply reduce the matrix A to Hermitian form $\tilde{A}$ by using elementary transformations. At each step the above rules tell us the effect on the determinant. If there is a zero on the diagonal of $\tilde{A}$ (i.e. if $r(A) < n$), then $\det A = 0$ by d6) above. If not, we can continue to reduce the matrix to the unit matrix by further row operations and so calculate its determinant. In fact, a little reflection shows that most of these calculations are superfluous and that it suffices to reduce the matrix to upper triangular form since the determinant of the latter is the product of its diagonal elements.

We illustrate this by "calculating" the determinant of the 3×3 matrix

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

We have

$$\begin{aligned} \det \begin{bmatrix} 0 & 2 & 3 \\ 1 & 2 & 1 \\ 2 & -3 & 2 \end{bmatrix} &= -\det \begin{bmatrix} 1 & 2 & 1 \\ 0 & 2 & 3 \\ 2 & -3 & 2 \end{bmatrix} \\ &= -\det \begin{bmatrix} 1 & 2 & 1 \\ 0 & 2 & 3 \\ 0 & -7 & 0 \end{bmatrix} \\ &= -2\det \begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & \frac{3}{2} \\ 0 & -7 & 0 \end{bmatrix} \\ &= -2\det \begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & \frac{3}{2} \\ 0 & 0 & \frac{21}{2} \end{bmatrix} \\ &= -21. \end{aligned}$$

In fact, what the above informal argument actually proves is the uniqueness of the determinant function. This fact is often useful and we state it as a Proposition.

Proposition 1 *There exists at most one mapping* $\det : M_n \to \mathbf{R}$ *with the properties d1)-d3) above.*

The main result of this section is the fact that such a function does in fact exist. The proof uses an induction argument on n. We already know that a determinant function exists for $n = 1, 2, 3$. In order to motivate the following proof note the formula

$$a_{11}(a_{22}a_{33} - a_{32}a_{23}) - a_{21}(a_{12}a_{33} - a_{32}a_{13}) + a_{31}(a_{12}a_{23} - a_{22}a_{13})$$

for the determinant of the 3×3 matrix

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}.$$

This is called the development of the determinant along the first column and suggests how to extend the definition to one dimension higher. This will be carried out formally in the proof of the following Proposition:

Proposition 2 *There is a (and hence exactly one) function* $\det : M_n \to \mathbf{R}$ *with the properties d1)-d3) (and so also d4)-d6)).*

Proof. As indicated above, we prove this by induction on n. The case $n = 1$ is clear (take $\det[a] = a$). The step from $n - 1$ to n: we define

$$\det A = \sum_{i=1}^{n} (-1)^{i+1} a_{i1} \det A_{i1}$$

where A_{i1} is the $(n-1) \times (n-1)$ matrix obtained by deleting the first column and the i-throw of A (the induction hypothesis ensures that its determinant is defined) and show that this function satisfies d1), d2) and d3). It is clear that $\det I_n = 1$. We verify the linearity in the k-th row as following. It suffices to show that each term $a_{i1} \det A_{i1}$ is linear in the k-th row. Now if $i \neq k$ a part of the k-th row of A is a row of A_{i1} and so this term is linear by the induction hypothesis. if $i = k$, then $\det A_{i1}$ is independent of the k-throw and a_{i1} depends linearly on it.

It now remains to show that $\det A = 0$ whenever two rows of A are identical, say the k-th and the l-th (with $k < l$). Consider the sum

$$\sum_{i=1}^{n} (-1)^{i+1} a_{i1} \det A_{i1}.$$

then A_{i1} has two identical rows (and so vanishes by the induction hypothesis) except for the cases where $j = k$ or $j = l$. This leaves the two terms

$$(-1)^{k+1} a_{k1} \det A_{k1} \quad \text{and} \quad (-1)^{l+1} a_{l1} \det A_{l1}$$

and they are equal in absolute value, but with opposite signs. (For $a_{k1} = a_{l1}$ and A_{k1} is obtained from A_{l1} by moving one row $(k-l-1)$ places. This can be achieved by the same number of row exchanges and so multiplies the determinant by $(-1)^{k-l-1}$).

The above proof yields the formula

$$\det A = \sum_{i=1}^{n} (-1)^{i+1} a_{i1} \det A_{i1}$$

for the determinant which is called **the development along the first column**. Similarly, one can develop $\det A$ along the j-th column i.e. we have the formula

$$\det A = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det A_{ij}$$

where A_{ij} is the $(n-1) \times (n-1)$ matrix obtained from A by omitting the i-th row and the j-th column. This can be proved by repeating the above proof with this recursion formula in place of the original one.

Example: If we expand the determinant of the triangular matrix

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ 0 & a_{22} & \dots & a_{2n} \\ \vdots & & & \vdots \\ 0 & 0 & \dots & a_{nn} \end{bmatrix}$$

along the first column, we see that its determinant is

$$a_{11} \det \begin{bmatrix} a_{22} & \dots & a_{2n} \\ 0 & \dots & a_{3n} \\ \vdots & & \vdots \\ 0 & \dots & a_{nn} \end{bmatrix}.$$

An obvious induction argument shows that the determinant is $a_{11}a_{22}\dots a_{nn}$, the product of the diagonal elements. In particular, this holds for diagonal matrices.

This provides a justification for the method for calculating the determinant of a matrix by reducing it to triangular form by means of elementary row operations. Note that for small matrices it is usually more convenient to calculate the determinant directly from the formulae given earlier.

1.3 Further properties of the determinant

d7) if $r(A) = n$ i.e. A is invertible, then $\det A \neq 0$.
PROOF. For then the Hermitian form of A has non-zero diagonal elements and so the determinant of A is non-zero.

Combining d5) and d7) we have the following Proposition:

Proposition 3 *An $n \times n$ matrix A is invertible if and only if its determinant is non-zero.*

Shortly we shall see how the determinant can be used to give an explicit formula for the inverse.
d8) The determinant is multiplicative i.e.

$$\det AB = \det A \cdot \det B.$$

PROOF. This is a typical application of the uniqueness of the determinant function. We first dispose of the case where $\det B = 0$. Then $r(B) < n$ and so $r(AB) < n$. In this case the formula holds trivially since both sides are zero.

If $\det B \neq 0$, then the mapping

$$A \mapsto \frac{\det AB}{\det B}$$

is easily seen to satisfy the three characteristic properties d1)-d3) of the determinant function and so *is* the determinant.
d9) Suppose that A is an $n \times n$ matrix whose determinant does not vanish. Then, as we have seen, A is invertible and we now show that the inverse of A can be written down explicitly as follows:

$$A^{-1} = \frac{1}{\det A}(\operatorname{adj} A)$$

where $\operatorname{adj} A$ is the matrix $[(-1)^{i+j} \det A_{ji}]$. (i.e. we form the matrix whose (i, j)-th entry is the determinant of the matrix obtained by removing the i-th row and the j-th column of A, with sign according to the chess-board pattern

$$\begin{matrix} + & - & + & - & \dots \\ - & + & - & + & \dots \\ \vdots & & & & \end{matrix}$$

This matrix is then transposed and the result is divided by $\det A$.

PROOF. We show that $(\operatorname{adj} A)A = (\det A)I$. Suppose that b_{ik} is the (i,k)-th element of the product i.e.

$$b_{ik} = \sum_{j=1}^{n} (-1)^{i+j} a_{jk} \det A_{ji}.$$

If $i = k$ this is just the expansion of $\det A$ along the i-th column i.e. $b_{ii} = \det A$.

If $i \neq k$, it is the expansion of the determinant of the matrix obtained from A by replacing the i-th column with the k-th one and so is 0 (since this is a matrix with two identical columns and so of rank $< n$).

We have discussed the determinant function in terms of its properties with regard to rows. Of course, it would have been just as logical to work with columns and we now show that the result would have been the same. To do this we introduce as a temporary notation the name Det for a function on the $n \times n$ matrices with the following properties:
D1) $\operatorname{Det} I_n = 1$;
D2) Det is linear in the columns;
D3) $\operatorname{Det} A = 0$ whenever two columns of A coincide.

Of course we can prove the existence of such a function exactly as we did for det (exchanging the word "column" for "row" everywhere). Even simpler, we note that if we put

$$\operatorname{Det} A = \det A^t$$

then this will fulfill the required conditions.

All the properties of det carry over in the obvious way. In particular, there is only one function with the above properties and we have the expansions

$$\operatorname{Det} A = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \operatorname{Det} A_{ij}$$

along the i-th row. We shall now prove the following result:

Proposition 4 *For each $n \times n$ matrix A, $\det A = \operatorname{Det} A$.*

In other words, $\det A = \det A^t$ and the notation "Det" is superfluous.

Again the proof is a typical application of the uniqueness. It suffices to show that the function det satisfies conditions d1)-d3). Of course, we have $\det I = 1$. In order to prove the other two assertions, we use induction on the order n and inspect the expansion

$$\operatorname{Det} A = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \operatorname{Det} A_{ij}$$

which is clearly linear in a_{ij} (and so in the i-th row). By the induction hypothesis, it is linear in the other rows (since each of the A_{ij} are). To complete the proof, we need only show that $\operatorname{Det} A$ vanishes if two rows of A coincide. But then $r(A) < n$ and so we have $\operatorname{Det} A = 0$ by the column analogue of property d6).

d11) One can often reduce the computations involved in calculating determinants by using suitable block decompositions. For example, if A has the decomposition

$$\begin{bmatrix} B & C \\ 0 & D \end{bmatrix}$$

where B and D are square matrices, then

$$\det A = \det B \cdot \det D.$$

(Warning: it is not true that if

$$A = \begin{bmatrix} B & C \\ D & E \end{bmatrix}$$

then there is a simple formula such as

$$\det A = \det B \cdot \det E - \det D \cdot \det C$$

which would allow us to calculate the determinant of A from those of B, C, D and E. However, such formulae do exist under suitable conditions, the above being the simplest example).

PROOF. We can assume that A and hence also B and D are invertible (for otherwise both sides vanish). Then if we multiply A on the right by the matrix

$$\begin{bmatrix} I & -B^{-1}C \\ 0 & I \end{bmatrix}$$

which has determinant 1, we get the value

$$\det \begin{bmatrix} B & 0 \\ 0 & D \end{bmatrix}.$$

Now the function

$$B \mapsto \det \begin{bmatrix} B & 0 \\ 0 & D \end{bmatrix} \div \det D$$

fulfills the conditions d1)-d3) and so *is* the determinant function.

Determinants of linear operators Since square matrices are the coordinate versions of linear operators on a vector space V it is tempting to extend the definition of determinants to such operators. The obvious way to do this is to choose some basis $(x_1, \dots, x_n)$ and to define the determinant $\det f$ of f to be the determinant of the matrix of f with respect to this basis. We must then verify that this value is independent of the choice of basis. But if A' is the matrix of f with respect to another basis, we know that

$$A' = S^{-1}AS$$

for some invertible matrix S. Then we have

$$\begin{aligned}\det A' &= \det(S^{-1}AS) \\ &= \det S^{-1} \cdot \det S \cdot \det A \\ &= \det(S^{-1}S) \det A \\ &= \det A.\end{aligned}$$

Of course, it is essential to employ the matrix of f with respect to a *single* basis for calculating the determinant.

Some of the properties of the determinant can now be interpreted as follows:

a) $\det f \neq 0$ if and only if f is an isomorphism;

b) $\det(fg) = \det f \cdot \det g \quad (f, g \in L(V))$;

c) $\det \mathrm{Id} = 1$.

Example: Calculate

$$\det \begin{bmatrix} 6 & 0 & 2 & 0 \\ 4 & 0 & 0 & 2 \\ 0 & 1 & 2 & 0 \\ 2 & 0 & 2 & 2 \end{bmatrix}$$

We have

$$\begin{aligned}\det \begin{bmatrix} 6 & 0 & 2 & 0 \\ 4 & 0 & 0 & 2 \\ 0 & 1 & 2 & 0 \\ 2 & 0 & 2 & 2 \end{bmatrix} &= -\det \begin{bmatrix} 6 & 2 & 0 \\ 4 & 0 & 2 \\ 2 & 2 & 0 \end{bmatrix} \\ &= -8 \det \begin{bmatrix} 3 & 1 & 0 \\ 2 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} \\ &= -8(-3+1) = 16.\end{aligned}$$

Example: Calculate

$$\det \begin{bmatrix} x & 1 & 1 & 1 \\ 1 & x & 1 & 1 \\ 1 & 1 & x & 1 \\ 1 & 1 & 1 & x \end{bmatrix}.$$

Solution: We have

$$\begin{aligned}
\det \begin{bmatrix} x & 1 & 1 & 1 \\ 1 & x & 1 & 1 \\ 1 & 1 & x & 1 \\ 1 & 1 & 1 & x \end{bmatrix} &= \det \begin{bmatrix} 0 & 1-x & 1-x^2 & 1-x^2 \\ 0 & -(1-x) & 0 & 1-x \\ 0 & 0 & -(1-x) & 1-x \\ 1 & 1 & 1 & x \end{bmatrix} \\
&= \det \begin{bmatrix} 1-x & 1-x & 1-x^2 \\ -(1-x) & 0 & 1-x \\ 0 & -(1-x) & 1-x \end{bmatrix} \\
&= (x-1)^3 \det \begin{bmatrix} 1 & 1 & 1+x \\ -1 & 0 & 1 \\ 0 & -1 & 1 \end{bmatrix} \\
&= (x-1)^3(1+x+2) = (x-1)^3(3+x).
\end{aligned}$$

Example: Calculate the determinant d_n of the $n \times n$ matrix

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

We can express the determinant in terms of the following $(n-1) \times (n-1)$ determinants by expanding along the first row:

$$\begin{aligned}
d_n = \det \begin{bmatrix} 2 & -1 & 0 & \dots & 0 \\ -1 & 2 & 0 & \dots & 0 \\ \vdots & & & & \vdots \\ 0 & 0 & 0 & \dots & 2 \end{bmatrix} &+ \det \begin{bmatrix} -1 & 0 & 0 & \dots & 0 \\ -1 & 2 & -1 & \dots & 0 \\ \vdots & & & & \\ \vdots & & & & \\ 0 & 0 & 0 & \dots & 2 \end{bmatrix} \\
= 2d_{n-1} - d_{n-2}
\end{aligned}$$

. It then follows easily by induction that $d_n = n+1$.

Example: Calculate

$$\det \begin{bmatrix} 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & \dots & 1 & 0 \\ \vdots & & & & \vdots \\ 1 & 0 & \dots & 0 & 0 \end{bmatrix}.$$

We have

$$\det \begin{bmatrix} 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & \dots & 1 & 0 \\ \vdots & & & & \vdots \\ 1 & 0 & \dots & 0 & 0 \end{bmatrix} = (-1)^n \det \begin{bmatrix} 0 & \dots & 0 & 1 \\ 0 & \dots & 1 & 0 \\ \vdots & & & \vdots \\ 1 & \dots & 0 & 0 \end{bmatrix}$$

where the left hand matrix is $n \times n$ and the right hand one is $(n-1) \times (n-1)$. From this it follows that the value of the given determinant is

$$(-1)^{n-1}(-1)^{n-2} \dots (-1)^{2-1} = (-1)^{\frac{n(n-1)}{2}}.$$

Example: Calculate

$$\det \begin{bmatrix} x & a & a & \dots & a \\ a & x & a & \dots & a \\ \vdots & & & & \vdots \\ a & a & a & \dots & x \end{bmatrix}.$$

Solution: The required determinant is

$$\begin{aligned} \det \begin{bmatrix} x & a & a & \dots & a \\ a-x & x-a & 0 & \dots & a \\ \vdots & & & & \vdots \\ a-x & 0 & 0 & \dots & x-a \end{bmatrix} &= x(x-a)^{n-1} - (a-x)a(x-a)^{n-2} + \dots + \\ &= x(x-a)^{n-1} + a(x-a)^{n-1} + \dots + \\ &= (x+(n-1)a)(x-a)^{n-1}. \end{aligned}$$

Example: Calculate the determinant of the Vandermonde matrix

$$\begin{bmatrix} 1 & x_1 & x_2^2 & \dots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \dots & ax_2^{n-1} \\ \vdots & & & & \vdots \\ 1 & x_n & x_n^2 & \dots & x_n^{n-1} \end{bmatrix}.$$

Solution: Subtracting from each column x_1 times the one on its left we see that the determinant is equal to

$$\det \begin{bmatrix} 1 & 0 & \dots & 0 \\ 1 & x_2 - x_1 & \dots & x_2^{n-2}(x_2 - x_1) \\ \vdots & & & \vdots \\ 1 & (x_n - x_1) & \dots & x_n^{n-2}(x_n - x_1) \end{bmatrix}$$

which is equal to

$$(x_2 - x_1)(x_3 - x_1)\dots(x_n - x_1) \det \begin{bmatrix} 1 & x_2 & \dots & x_2^{n-2} \\ \vdots & & & \vdots \\ 1 & x_n & \dots & x_n^{n-2} \end{bmatrix}$$

Hence, by induction, the value of the determinant is

$$\prod_{1 \le i < j \le n} (x_j - x_i)$$

(a product of $\frac{n(n-1)}{2}$ terms). (In particular, this determinant is non-zero if the x_i are distinct).

Exercises: 1) Evaluate the determinants of the following matrices:

$$\det \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} \qquad \det \begin{bmatrix} 1 & 1 & 1 \\ a & b & c \\ a^2 & b^2 & c^2 \end{bmatrix}.$$

2) Calculate the determinants of

$$\det \begin{bmatrix} a+b+2c & a & b \\ c & b+c+2a & b \\ c & a & c+a+2b \end{bmatrix}$$

$$\det \begin{bmatrix} x & 1 & 0 & x \\ 0 & x & x & 1 \\ 1 & x & x & 0 \\ x & 0 & 1 & x \end{bmatrix}$$

3) For which values of x does the determinant

$$\det \begin{bmatrix} x & 1 & x \\ 0 & x & 1 \\ 2 & x & 1 \end{bmatrix}$$

vanish?

4) Evaluate the following determinants:

$$\det\begin{bmatrix} 1 & 2 & 3 & \dots & n \\ 2 & 3 & 4 & \dots & 1 \\ \vdots & & & & \vdots \\ n & 1 & 2 & \dots & n-1 \end{bmatrix}$$

$$\det\begin{bmatrix} 0 & 1 & 1 & \dots & 1 \\ -1 & 0 & 1 & \dots & 1 \\ \vdots & & & & \vdots \\ -1 & -1 & -1 & \dots & 0 \end{bmatrix}$$

$$\det\begin{bmatrix} 1 & -a & 0 & \dots & 0 \\ -b & 1 & -a & \dots & 0 \\ 0 & -b & 1 & \dots & 0 \\ \vdots & & & & \vdots \\ 0 & 0 & \dots & -b & 1 \end{bmatrix}$$

$$\det\begin{bmatrix} \lambda_1 & 1 & 0 & \dots & 0 \\ -1 & \lambda_2 & 1 & \dots & 0 \\ \vdots & & & & \vdots \\ 0 & 0 & \dots & -1 & \lambda_n \end{bmatrix}$$

$$\det\begin{bmatrix} \lambda_1 & a & a & \dots & a \\ b & \lambda_2 & a & \dots & a \\ \vdots & & & & \vdots \\ b & b & b & \dots & \lambda_n \end{bmatrix}$$

$$\det\begin{bmatrix} \lambda & a & 0 & \dots & 0 \\ a & \lambda & a & \dots & 0 \\ \vdots & & & & \vdots \\ 0 & 0 & \dots & a & \lambda \end{bmatrix}.$$

5) Show that if P is a projection, then the dimension k of the range of P is determined by the equation

$$2^k = \det(I + P).$$

Use this to show that if $t \mapsto P(t)$ is a continuous mapping from $\mathbf{R}$ into the family of all projections on $\mathbf{R}^n$, then the dimension of the range of $P(t)$ is constant.

6) Show that if A (resp. B) is an $m \times n$ matrix (resp. an $n \times m$ matrix), then $I_m + AB$ is invertible if and only if $I_n + BA$ is.
7) Let A and B be $n \times n$ matrices. Show that

- $\operatorname{adj}(AB) = \operatorname{adj} A \cdot \operatorname{adj} B$;
- $\operatorname{adj}(\operatorname{adj} A) = (\det A)^{n-2} A$;
- $\operatorname{adj}(\lambda A) = \lambda^{n-1} \operatorname{adj} A$.

8) Let A be an invertible matrix whose elements are integers. Show that A^{-1} has the same property if and only if the determinant of A is 1 or -1.
9) Let f be a mapping from M_n into $\mathbf{R}$ which is linear in each row and such that $f(A) = -f(B)$ whenever B is obtained from A by exchanging two rows. Show that there is a $c \in \mathbf{R}$ so that $f(A) = c \cdot \det A$.
10) Show that if A is an $n \times n$ matrix with $A^t = -A$ (such matrices are called **skew symmetric**), then the determinant of A vanishes whenever n is odd.
11) Let A be an $m \times n$ matrix. Show that in the reduction of A to Hermitian form by means of Gaußian elimination, the pivot element never vanishes (i.e. we do not require the use of row exchanges) if and only if $\det A_k \neq 0$ for each $k \leq \min(m, n)$. (A_k is the $k \times k$ matrix $[a_i j]_{1 \leq i \leq k, 1 \leq j \leq k}$.

Deduce that A then has a factorisation of the form LU where L is an invertible $m \times m$ lower triangular matrix and U is an $m \times n$ upper triangular matrix.
12) Show that a matrix A has rank r if and only if there is a $r \times r$ minor of A with non-vanishing determinant and no such $(r+1) \times (r+1)$ minor. (An r-minor of A is a matrix of the form

$$\det \begin{bmatrix} a_{i_1,j_1} & \dots & a_{i_1,j_p} \\ \vdots & & \vdots \\ a_{i_p,j_1} & \dots & a_{i_p,j_p} \end{bmatrix}$$

for increasing sequences

$$1 \leq i_1 < i_1 < \dots < i_r \leq m$$

$$1 \leq j_1 < j_2 < \dots < j_r \leq n).$$

13) Let x and y be continuously differentiable functions on the interval $[a, b]$. Apply Rolle's theorem to the following function

$$t \mapsto \det \begin{bmatrix} 1 & 1 & 1 \\ x(t) & x(a) & x(b) \\ y(t) & y(a) & y(b) \end{bmatrix}$$

Which well-known result of analysis follows?

14) Let

$$\begin{aligned} x_1 &= rc_1c_2\dots c_{n-2}c_{n-1} \\ x_2 &= rc_1\dots c_{n-2}s_{n-1} \\ &\vdots \\ x_j &= rc_1\dots c_{n-j}s_{n-j+1} \\ &\vdots \\ x_n &= rs_1. \end{aligned}$$

where $c_i = \cos\theta_i$, $s_i = \sin\theta_i$ (these are the equations of the transformation to polar coordinates in n dimensions). Calculate the determinant of the Jacobi matrix

$$\frac{\partial(x_1, x_2, \dots, x_n)}{\partial(r, \theta_1, \dots, \theta_{n-1})}.$$

15) Consider the Vandermonde matrix

$$V_n = \begin{bmatrix} 1 & \dots & 1 \\ t_1 & \dots & t_n \\ \vdots & & \vdots \\ t_1^{n-1} & \dots & t_n^{n-1} \end{bmatrix}.$$

Show that $V_nV_n^t$ is the matrix

$$\begin{bmatrix} s_0 & s_1 & \dots & s_{n-1} \\ s_1 & s_2 & \dots & s_n \\ \vdots & & & \vdots \\ s_{n-1} & s_n & \dots & s_{2n-2} \end{bmatrix}$$

where $s_k = \sum_{i=1}^n t_i^k$. Use this to calculate the determinant of this matrix.

16) Suppose that the square matrix A has a block representation $\begin{bmatrix} B & C \\ D & E \end{bmatrix}$ where B is square and non-singular. Show that

$$\det A = \det B \det(E - DB^{-1}C).$$

Deduce that if D is also square and commutes with B, then $\det A = \det(BE - DC)$.

17) Suppose that $A_0, \dots, A_r$ are complex $n \times n$ matrices and consider the matrix function

$$p(t) = A_0 + A_1t + \dots + A_rt^r.$$

Show that if $\det p(t)$ is constant, then so is $p(t)$ (i.e. A_0 is the only non-vanishing term).

1.4 Applications of the determinant

We conclude this chapter by listing briefly some applications of the determinant:

I. Solving systems of equations—Cramer's rule: returning to one of our original motivations for introducing determinants, we show that if the determinant of the matrix A of the system $AX = Y$ is non-zero, then the unique solution is given by the formulae

$$x_i = \frac{\det A_i}{\det A}$$

where A_i is the matrix obtained by replacing the i-th column of A with the column vector Y. To see this note that we can write the system in the form

$$x_1 \begin{bmatrix} a_{11} \\ \vdots \\ a_{n1} \end{bmatrix} + \dots + \begin{bmatrix} x_1 a_{1i} - y_1 \\ \vdots \\ x_1 a_{ni} - y_n \end{bmatrix} + \dots + x_n \begin{bmatrix} x_1 a_{1n} - y_1 \\ \vdots \\ x_1 a_{nn} - y_n \end{bmatrix} = 0$$

and this just means that the columns of the matrix

$$\begin{bmatrix} a_{11} & \dots & a_{1,i-1} & (x_1 a_{1i} - y_1) & \dots & a_{1n} \\ \vdots & & & & & \vdots \\ a_{n1} & \dots & a_{n,i-1} & (x_i a_{ni} - y_n) & \dots & a_{nn} \end{bmatrix}$$

are linearly independent. Hence its determinant vanishes and, using the linearity in the i-th column, this means that $\det A_i - x_i \det A = 0$.

II. The recognition of bases: If $(x_1, \dots, x_n)$ is a basis for a vector space V (e.g. the canonical basis for $\mathbf{R}^n$), then a set $(x'_1, \dots, x'_n)$ is a basis if and only if the determinant of the transfer matrix $T = [t_{ij}]$ whose columns are the coordinates of the x'_j with respect to the (x_i) is non-zero.

III. Areas, volumes etc.: If ξ, η are points in $\mathbf{R}$, then

$$\det \begin{bmatrix} \eta & 1 \\ \xi & 1 \end{bmatrix} = \eta - \xi$$

is the directed length of the interval from ξ to η.

If $A = (\xi_1, \xi_2)$, $B = (\eta_1, \eta_2)$, $C = (\zeta_1, \zeta_2)$ are points in the plane, then

$$\frac{1}{2} \det \begin{bmatrix} \xi_1 & \xi_2 & 1 \\ \eta_1 & \eta_2 & 1 \\ \zeta_1 & \zeta_2 & 1 \end{bmatrix}$$

is the area of the triangle ABC. The area is positive if the direction $A \rightarrow B \rightarrow C$ is clockwise, otherwise it is negative. (By taking the *signed* area we assure that it is additive i.e. that

$$\triangle ABC = \triangle OAB + \triangle OBC + \triangle OCA$$

regardless of the position of O with respect to the triangle (see figure ??).

If $A = (\xi_1, \xi_2, \xi_3)$, $B = (\eta_1, \eta_2, \eta_3)$, $C = (\zeta_1, \zeta_2, \zeta_3)$, $D = (\upsilon_1, \upsilon_2, \upsilon_3)$ are points in space, then

$$\frac{1}{3!} \det \begin{bmatrix} \xi_1 & \xi_2 & \xi_3 & 1 \\ \eta_1 & \eta_2 & \eta_3 & 1 \\ \zeta_1 & \zeta_2 & \zeta_3 & 1 \\ \upsilon_1 & \upsilon_2 & \upsilon_3 & 1 \end{bmatrix}$$

is the volume of the tetrahedron $ABCD$.
Of course, analogous formulae hold in higher dimensions.

IV. The action of linear mappings on volumes: If f is a linear mapping from $\mathbf{R}^3$ into $\mathbf{R}^3$ with matrix

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

then f multiplies volumes by a factor of $\det f$. For suppose that f maps the tetrahedron $BCDE$ into the tetrahedron $B_1C_1D_1E_1$. The area of $BCDE$ is

$$\frac{1}{3!} \det \begin{bmatrix} \xi_1 & \xi_2 & \xi_3 & 1 \\ \eta_1 & \eta_2 & \eta_3 & 1 \\ \zeta_1 & \zeta_2 & \zeta_3 & 1 \\ \upsilon_1 & \upsilon_2 & \upsilon_3 & 1 \end{bmatrix}$$

where B is the point (ξ_1, ξ_2, ξ_3) etc. and that of the image is

$$\frac{1}{3!} \det \begin{bmatrix} \xi_1^1 & \xi_2^1 & \xi_3^1 & 1 \\ \eta_1^1 & \eta_2^1 & \eta_3^1 & 1 \\ \zeta_1^1 & \zeta_2^1 & \zeta_3^1 & 1 \\ \upsilon_1 & \upsilon_2 & \upsilon_3 & 1 \end{bmatrix}.$$

Now we have

$$\frac{1}{3!} \det \begin{bmatrix} \xi_1^1 & \xi_2^1 & \xi_3^1 & 1 \\ \eta_1^1 & \eta_2^1 & \eta_3^1 & 1 \\ \zeta_1^1 & \zeta_2^1 & \zeta_3^1 & 1 \\ \upsilon_1 & \upsilon_2 & \upsilon_3 & 1 \end{bmatrix} = \frac{1}{3!} \det \begin{bmatrix} \xi_1^1 & \xi_2^1 & \xi_3^1 & 1 \\ \eta_1^1 & \eta_2^1 & \eta_3^1 & 1 \\ \zeta_1^1 & \zeta_2^1 & \zeta_3^1 & 1 \\ \upsilon_1^1 & \upsilon_2^1 & \upsilon_3^1 & 1 \end{bmatrix} \begin{bmatrix} A^t & 0 \\ & \\ & \\ 0 & 1 \end{bmatrix}$$

and so we have that the volume of $B_1C_1D_1E_1$ is $\det A$ times the volume of $BCDE$. It follows from a limiting argument that the same formula holds for arbitrary figures. (This justifies the original geometrical motivation for the existence and properties of the determinant).
Once again, an analogous result holds in higher dimensions.

V. The equations of curves: If $P = (\xi_1^1, \xi_2^1)$ and $Q = (\xi_1^2, \xi_2^2)$ are distinct points in the plane, then the line L through P and Q has equation

$$\det \begin{bmatrix} \xi_1 & \xi_2 & 1 \\ \xi_1^1 & \xi_2^1 & 1 \\ \xi_1^2 & \xi_2^2 & 1 \end{bmatrix} = 0.$$

For if the equation of the line has the form $a\xi_1 + b\xi_2 + c = 0$, then we have

$$\begin{aligned} a\xi_1^1 + b\xi_2^1 + c &= 0 \\ a\xi_1^2 + b\xi_2^2 + c &= 0. \end{aligned}$$

This means that the above three homogeneous equations (in the variables a, b, c) has a non-trivial solution. As we know, this is equivalent to the vanishing of the above determinant.

In exactly the same way one shows that the plane through $(\xi_1^1, \xi_2^1, \xi_3^1)$, $(\xi_1^2, \xi_2^2, \xi_3^2)$ and $(\xi_1^3, \xi_2^3, \xi_3^3)$ has equation

$$\det \begin{bmatrix} \xi_1 & \xi_2 & \xi_3 & 1 \\ \xi_1^1 & \xi_2^1 & \xi_3^1 & 1 \\ \xi_1^2 & \xi_2^2 & \xi_3^2 & 1 \\ \xi_1^3 & \xi_2^3 & \xi_3^3 & 1 \end{bmatrix} = 0.$$

The circle through (ξ_1^1, ξ_2^2), (ξ_1^2, ξ_2^2) and (ξ_1^3, ξ_2^3) has equation:

$$\det \begin{bmatrix} (\xi_1)^2 + (\xi_2)^2 & \xi_1 & \xi_2 & 1 \\ (\xi_1^1)2 + (\xi_2^1)^2 & \xi_1^1 & \xi_2^1 & 1 \\ (\xi_1^2)^2 + (\xi_2^2)^2 & \xi_1^2 & \xi_2^2 & 1 \\ (\xi_1^3)^2 + (\xi_2^3)^2 & \xi_1^3 & \xi_2^3 & 1 \end{bmatrix} = 0.$$

(Note that the coefficient of $\xi_1^2 + \xi_2^2$ is

$$\det \begin{bmatrix} \xi_1^1 & \xi_2^1 & 1 \\ \xi_1^2 & \xi_2^2 & 1 \\ \xi_1^3 & \xi_2^3 & 1 \end{bmatrix}$$

and this fails to vanish precisely when the points are non-collinear).

VI. Orientation: A linear isomorphism f on a vector space V is said to preserve **orientation** if its determinant is positive—otherwise it reverses orientation. This concept is particularly important for isometries and those which preserve orientation are called **proper**. Thus the only proper isometries of the plane are translations and rotations.

Two bases $(x_1, \dots, x_n)$ and $(x'_1, \dots, x'_n)$ have the **same orientation** if the linear mapping which maps x_i onto x'_i for each i preserves orientation. This just means that the transfer matrix from (x_i) to (x'_j) has positive determinant. For instance, in $\mathbf{R}^3$, (e_1, e_2, e_3) and (e_3, e_1, e_2) have the same orientation, whereas that of (e_2, e_1, e_3) is different.

Example Is

$$\begin{bmatrix} \cos\alpha\cos\beta & \sin\alpha\cos\beta & -\sin\beta \\ \cos\alpha\sin\beta & \sin\alpha\sin\beta & \cos\beta \\ -\sin\alpha & \cos\alpha & 0 \end{bmatrix}$$

the matrix of a rotation?

Solution: Firstly the columns are orthonormal and so the matrix induces an isometry. but the determinant is

$$-\sin^2\alpha\cos^2\beta - \cos^2\alpha\sin^2\beta - \sin^2\alpha\sin^2\beta - \cos^2\alpha\cos^2\beta = -1.$$

Hence it is not a rotation.

Example: Solve the system

$$\begin{array}{ccccccc} (m+1)x & + & y & + & z & = & 2-m \\ x & + & (m+1)y & + & z & = & -2 \\ x & + & y & + & (m+1)z & = & m. \end{array}$$

The determinant of the matrix A of the equation is $m^2(m+3)$ which is non-zero, unless $m = 0$ or $m = -3$. Otherwise the solution is, by Cramer's rule,

$$x = \frac{1}{m^2(m+3)} \det \begin{bmatrix} 2-m & 1 & 1 \\ -2 & m+1 & 1 \\ m & 1 & m+1 \end{bmatrix}$$

i.e. $\dfrac{2-m}{m}$. The values of y and z can be calculated similarly.

Exercises: 1) Show that the centre of the circle through the points (ξ_1^1, ξ_2^2), (ξ_1^2, ξ_2^2) and (ξ_1^3, ξ_2^3) has coordinates

$$\frac{\left(\frac{1}{2}\det\begin{bmatrix}(\xi_1^1)^2+(\xi_2^1)^2 & \xi_2^1 & 1\\ (\xi_1^2)^2+(\xi_2^2)^2 & \xi_2^2 & 1\\ (\xi_1^3)^2+(\xi_2^3)^2 & \xi_2^3 & 1\end{bmatrix}, \frac{1}{2}\det\begin{bmatrix}(\xi_1^1)^2+(\xi_2^1)^2 & \xi_1^1 & 1\\ (\xi_1^2)^2+(\xi_2^2)^2 & \xi_1^2 & 1\\ (\xi_1^3)^2+(\xi_2^3)^2 & \xi_1^3 & 1\end{bmatrix}\right)}{\det\begin{bmatrix}\xi_1^1 & \xi_2^2 & 1\\ \xi_1^2 & \xi_2^2 & 1\\ \xi_1^3 & \xi_2^3 & 1\end{bmatrix}}.$$

2) Show that in $\mathbf{R}^n$ the equation of the hyperplane through the affinely independent points $x_1, \dots, x_n$ is

$$\det\begin{bmatrix}\xi_1 & \xi_2 & \dots & \xi_n & 1\\ \xi_1^1 & \xi_2^1 & \dots & \xi_n^1 & 1\\ \vdots & & & & \vdots\\ \xi_1^n & \xi_2^n & \dots & \xi_n^n & 1\end{bmatrix} = 0.$$

3) Let A be an invertible $n \times n$ matrix. Use that fact that if $AX = Y$, then $A\tilde{X} = \tilde{Y}$ where

$$\tilde{X} = \begin{bmatrix}x_1 & 0 & \dots & 0\\ x_2 & 1 & \dots & 0\\ \vdots & & & \vdots\\ x_n & 0 & \dots & 1\end{bmatrix} \qquad \tilde{Y} = \begin{bmatrix}y_1 & a_{12} & \dots & a_{1n}\\ y_2 & a_{22} & \dots & a_{2n}\\ \vdots & & & \vdots\\ y_n & a_{n2} & \dots & a_{nn}\end{bmatrix}$$

to give an alternative proof of Cramer's rule.

4) Let

$$p(t) = a_0 + \dots + a_m t^m$$
$$q(t) = b_0 + \dots + b_n t^n$$

be polynomials whose leading coefficients are non-zero. Show that they have a common root if and only if the determinant of the $(m+n)\times(m+n)$ matrix

$$A = \begin{bmatrix}a_m & a_{m-1} & \dots & a_1 & a_0 & 0 & \dots & 0\\ 0 & a_m & \dots & a_2 & a_1 & a_0 & \dots & 0\\ \vdots & & & & & & & \vdots\\ 0 & 0 & \dots & 0 & a_m & a_{m-1} & \dots & a_0\\ b_n & b_{n-1} & \dots & b_1 & b_0 & 0 & \dots & 0\\ \vdots & & & & & & & \vdots\\ 0 & 0 & \dots & b_n & & \dots & & b_0\end{bmatrix}$$

is non-zero. (This is known as Sylvester's criterium for the existence of a common root). In order to prove it calculate the determinants of the matrices B and BA where B is the $(m+n) \times (m+n)$ matrix

$$\begin{bmatrix} t^{n+m-1} & 0 & 0 & \dots & 0 \\ t^{n+m-2} & 1 & 0 & \dots & 0 \\ t^{n+m-3} & 0 & 1 & \dots & 0 \\ \vdots & & & & \vdots \\ t & 0 & 0 & \dots & 0 \\ 1 & 0 & 0 & \dots & 0 \\ \vdots & & & & \vdots \\ 1 & & \dots & & 1 \end{bmatrix}.$$

2 COMPLEX NUMBERS AND COMPLEX VECTOR SPACES

2.1 The construction of C

When we discuss the eigenvalue problem in the next chapter, it will be convenient to consider complex vector spaces i.e. those for which the complex numbers play the role taken by the reals in the third chapter. We therefore bring a short introduction to the theme of complex numbers.

Complex numbers were stumbled on by the renaissance mathematician Cardano in the famous formulae

$$\lambda_1 = \sqrt[3]{\alpha} + \sqrt[3]{\beta} \quad \lambda_2 = \omega\sqrt[3]{\alpha} + \omega^2\sqrt[3]{\beta} \quad \lambda_3 = \omega^2\sqrt[3]{\alpha} + \omega\sqrt[3]{\beta}$$

where $\alpha = \dfrac{-q + \sqrt{q^2 + 4p^3}}{2}$, $\beta = \dfrac{-q - \sqrt{q^2 + 4p^3}}{2}$, $\omega = \dfrac{-1 + \sqrt{3}i}{2}$ for the roots λ_1,λ_2 and λ_3 of the cubic equation

$$x^3 + 3px = q = 0$$

(which he is claimed to have stolen from a colleague). In the above formulae i denotes the square root of -1 i.e. a number with the property that $i^2 = -1$. This quantity appears already in the solution of the quadratic equation by radicals but only in the case where the quadratic has no real roots. In the cubic equation, its occurrence is unavoidable, even in the case where the discriminant $q^2 + 4p^3$ is positive, in which case the cubic has three real roots. Since the square of a real number is positive, no such number with the defining property of i exists and mathematicians simply calculated formally with expressions of the form $x + iy$ as if the familiar rules of arithmetic still hold for such expressions. Just how uncomfortable they were in doing this is illustrated by the following quotation from Leibniz:

- “The imaginary numbers are a free and marvellous refuge of the divine intellect, almost an amphibian between existence and non-existence.”

The nature of the complex numbers was clarified by Gauß in the nineteenth century with the following geometrical interpretation¿ the real numbers are identified with the points on the x-axis of a coordinate plane. One associates the number i with the point $(0, 1)$ on the plane and, accordingly, the number $x + iy$ with the point (x, y). The arithmetical operations of addition and multiplication are defined geometrically as in the following figure (where the triangles OAB and OCD are similar). A little analytical geometry shows that these operations can be expressed as follows:

addition: $(x, y) + (x_1, y_1) = (x + x_1, y + y_1)$'
multiplication: $(x, y) \cdot (x_1, y_1) = (xx_1 - yy_1, xy_1 + x_1y)$.
Note that these correspond precisely to the expressions obtained by formally adding and multiplying $x + iy$ and $x_1 + iy_1$.

This leads to the following definition: a **complex number** is an ordered pair (x, y) of real numbers. On the set of such numbers we define addition and multiplication by the above formulae. We use the following conventions:

- 1) i denotes the complex number $(0, 1)$ and we identity the real number x with the complex number $(x, 0)$. Then $i^2 = -1$ since

$$(0, 1) \cdot (0, 1) = (-1, 0).$$

  Every complex number (x, y) has a unique representation $x = iy$ where $x, y \in \mathbf{R}$. (It is customary to use letters such as $z, w, \dots$ for complex numbers). If $z = x + iy$ $(x, y \in \mathbf{R})$, then x is called the **real part** of z (written $\Re z$) and y is called the **imaginary part** (written $\Im z$).

- 2) If $z = x + iy$, we denote the complex number $x - iy$ (i.e. the mirror image of z in the x-axis) by $\bar{z}$—the **complex-conjugate** of z. Then the following simple relations holds:

$$\begin{aligned} \overline{z + z_1} &= \bar{z}\overline{z_1}; \\ \overline{zz_1} &= \bar{z} \cdot \overline{z_1}; \\ \Re z &= \frac{1}{2}(z + \bar{z}); \\ \Im z &= \frac{1}{2i}(z - \bar{z}); \\ z \cdot \bar{z} &= |z|^2 \quad \text{where} \quad |z| = \sqrt{x^2 + y^2}. \end{aligned}$$

  $|z|$ is called the **modulus** or **absolute value** of z. It is multiplicative in the sense that $|zz_1| = |z||z_1|$.

- 3) every non-zero complex number z has a unique representation of the form

$$\rho(\cos\theta + i \sin\theta)$$

  where $\rho > 0$ and $\theta \in [0, 2\pi[$. Here $\rho = |z|$ and θ is the unique real number in $[0, 2\pi[$ so that $\cos\theta = \dfrac{x}{\rho}$, $\sin\theta = \dfrac{y}{\rho}$.

We denote the set of complex numbers by $\mathbf{C}$. Of course, *as a set*, it is identical with $\mathbf{R}^2$ and we use the notation $\mathbf{C}$ partly for historical reasons and partly to emphasis the fact that we are considering it not just as a vector space but also with its multiplicative structure.

Proposition 5 *For complex numbers z, z_1, z_2, z_3 we have the relationships*

- $z_1 + z_2 = z_2 + z_1$*;*
- $z_1 + (z_2 + z_3) = (z_1 + z_2) + z_3$*;;*
- $z_1 = 0 = 0 + z_1 = z_1$*'*
- $z_1 + (-z_1) = 0 \quad$ *where* $\quad -z_1 = (-1)z_1$*;*
- $z_1(z_2 + z_3) = z_1 z_2 + z_1 z_3$*;*
- $z_1 z_2 = z_2 z_1$*;*
- $z_1(z_2 z_3) = (z_1 z_2) z_3$*;*
- $z_1 \cdot 1 = 1 \cdot z_1 = z_1$*;*
- *if $z \neq 0$, there is an element z^{-1} so that $z \cdot z^{-1} = 1$ (take $z^{-1} = \dfrac{\bar{z}}{|z|^2}$).*

This result will be of some importance for us since in our treatment of linear equations, determinants, vector spaces and so on, the only properties of the real numbers that we have used are those which correspond to the above list. Hence the bulk of our definitions, results and proofs can be carried over almost *verbatim* to the complex case and, with this justification, we shall use the complex versions of results which we have proved only for the real case without further comment.

It is customary to call a set with multiplication and addition operations with such properties a **field**. A further example of a field is the set $\mathbf{Q}$ of rational numbers.

de Moivre's formula: The formula for multiplication of complex numbers has a particularly transparent form when the complex numbers are represented in polar form: we have

$$\rho_1(\cos\theta_1 + i\sin\theta_1)\rho_2(\cos\theta_2 + i\sin\theta_2) = \rho_1\rho_2(\cos(\theta_1+\theta_2) + i\sin(\theta_1+\theta_2)).$$

This is derived by multiplying out the left hand side and using the addition formulae for the trigonometric functions.

This equation can be interpreted geometrically as follows: multiplication by the complex number $z = \rho(\cos\theta + i\sin\theta)$ has the effect of rotating a second complex number through an angle of θ and multiplying its absolute value by ρ (of course this is one of the similarities considered in the second chapter—in fact, a rotary dilation). As a Corollary of the above formula we have the famous result

$$(\cos\theta + i\sin\theta)^n = \cos n\theta + i\sin n\theta$$

of de Moivre. It is obtained by a simple induction argument on $n \in \mathbf{N}$. Taking complex conjugates gives the result for $-n$ and so it holds for each $n \in \mathbf{Z}$.

From it we can deduce the following fact: if $z = \rho(\cos\theta + i\sin\theta)$ is a non-zero complex number, then there are n solutions of the equation $\zeta^n = z$ ($n \in \mathbf{N}$) given by the complex numbers

$$\rho^{\frac{1}{n}}(\cos\frac{2\pi r + \theta}{n} + \sin\frac{2\pi r + \theta}{n})$$

($r = 0, 1, \dots, n-1$). In particular, there are n roots of unity (i.e. solutions of the equations $\zeta^n = 1$), namely the complex numbers $1, \omega, \omega^2, \dots, \omega^{n-1}$ where $\omega = \cos\frac{2\pi}{n} + i\sin\frac{2\pi}{n}$ is the **primitive n-th root of unity**.

Example: If z is a complex number, not equal to one, show that

$$1 + z + z^2 + \cdots + z^n = \frac{1 - z^{n+1}}{1 - z}.$$

Use this to calculate the sums

$$1 + \cos\theta + \cdots + \cos n\theta$$

and

$$\sin\theta + \cdots + \sin n\theta.$$

Solution: The first part is proved exactly as in the case of the partial sums of a real geometric series. If we set $z = \cos\theta + i\sin\theta$ and take the real part, we get

$$1 + \cos\theta + \cdots + \cos n\theta = \Re\frac{1 - \cos(n+1)\theta - i\sin(n+1)\theta}{1 - \cos\theta - i\sin\theta}$$

which simplifies to the required formula (we leave the details to the reader). The sine part is calculated with the aid of the imaginary part. **Example:** Describe the geometric form of the set

$$C = \{z \in \mathbf{C} : z\bar{z} + \bar{a}z + a\bar{z} + b = 0\}$$

where a is a complex number and b a real number.

Solution Substituting $z = x + iy$ we get

$$C = \{(x, y) : x^2 + y^2 + 2a_1 x + 2a_2 y + b = 0\}$$

(where $a = a_1 + ia_2$) which is a circle, a point or the empty set depending on the values of a_1, a_2 and b.

Exercises on complex numbers: 1) Calculate $\Re z$, $\Im z$, $|z|$, z^{-1} and $\arg z$ where

$$z = 1 - i \quad z = 3 + \sqrt{2}i \quad \frac{1+i}{1-i}.$$

Show that if $z_1, z_2 \in \mathbf{C}$ are such that $z_1 z_2$ and $z_1 + z_2$ are real, then either $z_1 = \overline{z}_2$ or z_1 and z_2 are themselves real.

2) If $x, x_1, y, y_1 \in \mathbf{R}$ and $X = \text{diag}(x, x)$, $Y = \text{diag}(y, y)$ etc. while $\mathbf{J} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$, calculate $(X + \mathbf{J}Y)(X - \mathbf{J}Y)$, $(X + \mathbf{J}Y)(X_1 + \mathbf{J}Y_1)$, $(X + \mathbf{J}Y)^{-1}$ (when it exists) and $\det(X + \mathbf{J}Y)$. (Compare the results with the arithmetic operations in $\mathbf{C}$. This exercise can be used as the basis for an alternative construction of the complex numbers).

3) Use de Moivre's theorem to derive a formula for $\cos^4 \theta$ in terms of $\cos 2\theta$ and $\cos 4\theta$.

4) Show that if n is even, then

$$\cos n\theta = \cos^n \theta - \binom{n}{2} \cos^{n-2} \theta \sin^2 \theta + \dots + (-1)^{\frac{n}{2}} \sin^n \theta;$$

$$\sin n\theta = \binom{n}{1} \cos^{n-1} \theta \sin \theta + \dots + (-1)^{\frac{n}{2}-1} n \cos \theta \sin^{n-1} \theta.$$

What are the corresponding results for n odd?

5) Suppose that $|r| < 1$. Calculate

$$1 + r \cos \theta + r^2 \cos 2\theta + \dots$$

and

$$r \sin \theta + r^2 \sin 2\theta + \dots$$

6) Show that the points z_1, z_2 and z_3 in the complex plane are the vertices of an equilateral triangle if and only if

$$z_1 + \omega z_2 + \omega^2 z_3 = 0$$

or

$$z_1 + \omega^2 z_2 + \omega z_3 = 0$$

where $\omega = e^{\frac{2\pi i}{3}}$.

If z_1, z_2, z_3, z_4 are four complex numbers, what is the geometrical significance of the condition

$$z_1 + iz_2 + (i^2)z_3 + (i^3)z_4 = 0?$$

(Note that i is the primitive fourth root of unity).
7) Show that if z_1, z_2, z_3, z_4 are complex numbers, then

$$(z_1 - z_4)(z_2 - z_3) + (z_2 - z_4)(z_3 - z_1) + (z_3 - z_4)(z_1 - z_2) = 0.$$

Deduce that if A, B, C, D are four points in the plane, then

$$|AD||BC| \leq |BD||CA| + |CD||AB|.$$

8) We defined the complex numbers formally as pairs of real numbers. In this exercise, we investigate what happens if we continue this process i.e. we consider pairs (z, w) of complex numbers. On the set $\mathbf{Q}$ of such pairs we define the natural addition and multiplication as follows:

$$(z_0, w_0)(z_1, w_1) = (z_0 z_1 - w_0 w_1, z_0 w_1 + z_1 w_0).$$

Show that $\mathbf{Q}$ satisfies all of the axioms of a field with the exception of the commutativity of multiplication (such structures are called **skew fields**). Show that if we put $\mathbf{i} = (i, 0)$, $\mathbf{j} = (0, i)$, $\mathbf{k} = (0, 1)$, then $\mathbf{ij} = -\mathbf{ji} = \mathbf{k}$, $\mathbf{jk} = -\mathbf{kj} = \mathbf{i}$ etc. and $\mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 = -1$. Also every element of $\mathbf{Q}$ has a unique representation of the form

$$\xi_0 + (\xi_1 \mathbf{i} + \xi_2 \mathbf{j} + \xi_3 \mathbf{k})$$

with $\xi_0, \xi_1, \xi_2, \xi_3 \in \mathbf{R}$. (The elements of $\mathbf{Q}$ are called **quaternions**).

2.2 Polynomials

The field of complex numbers has one significant advantage over the real field. All polynomials have roots. This result will be very useful in the next chapter—it is known as **the fundamental theorem of algebra** and can be stated in the following form:

Proposition 6 *Let*

$$p(\lambda) = a_0 + \cdots + a_{n-1}\lambda^{n-1} + \lambda^n$$

be a complex polynomial with $n \geq 1$. *Then there are complex numbers* $\lambda_1, \ldots, \lambda_n$ *so that*

$$p(\lambda) = (\lambda - \lambda_1) \ldots (\lambda - \lambda_n).$$

There is no simple algebraic proof of this result which we shall take for granted.

The fundamental theorem has the following Corollary on the factorisation of real polynomials.

Corollar 1 *Let*

$$p(t) = a_0 + \cdots + a_{n-1}t^{n-1} + t^n$$

be a polynomial with real coefficients. Then there are real numbers

$$t_1, \ldots, t_r, \alpha_1, \ldots, \alpha_s, \beta_1, \ldots, \beta_s$$

where $r + 2s = n$ *so that*

$$p(t) = (t - t_1) \ldots (t - t_r)(t^2 - 2\alpha_1 t + \alpha_1^2 + \beta_1^2) \ldots (t^2 - 2\alpha_s t + \alpha_s^2 + \beta_s^2).$$

PROOF. We denote by $\lambda_1, \ldots, \lambda_n$ the complex roots of the polynomial

$$p(\lambda) = a_0 + a_1\lambda + \cdots + \lambda^n.$$

Since $\overline{p(\lambda)} = p(\bar{\lambda})$ (the coefficients being real), we see that a complex number λ is a root if and only if its complex conjugate is also one. Hence we can list the roots of p as follows: firstly the real ones

$$t_1, \ldots, t_r$$

and then the complex ones in conjugate pairs:

$$\alpha_1 + i\beta_1, \alpha_1 - i\beta_1, \ldots, \alpha_s + i\beta_s, \alpha_s - i\beta_s.$$

Then we see that p has the required form by multiplying out the corresponding linear and quadratic terms.

The next result concerns the representation of rational functions. These are functions of the form $\frac{p}{q}$ where p and q are polynomials. By long division every such function can be expressed as the sum of a polynomial and a rational function $\frac{\mathbf{i}ldep}{q}$ where the degree $d(\tilde{p})$ of $\tilde{p}$ (i.e. the index of its highest power) is strictly less than that of q. Hence from now on we shall tacitly assume that this condition is satisfied. Further it is no loss of generality to suppose that the leading coefficient of q is "1".

We consider first the case where q has simple zeros i.e.

$$q(\lambda) = (\lambda - \lambda_1)\dots(\lambda - \lambda_n)$$

where the λ_i are distinct. Then we claim that there are uniquely determined complex numbers $a_1, \dots, a_n$ so that

$$\frac{p(\lambda)}{q(\lambda)} = \frac{a_1}{\lambda - \lambda_1} + \dots \frac{a_n}{\lambda - \lambda_N}$$

for $\lambda \in \mathbf{C} \setminus \{\lambda_1, \dots, \lambda_n\}$.

PROOF. This is equivalent to the equation

$$p(\lambda) = \sum_{i=1}^{n} a_i q_i(\lambda)$$

where $q_i(\lambda) = \frac{q(\lambda)}{\lambda - \lambda_i}$. If this holds for all λ as above then it holds for all λ in $\mathbf{C}$ since both sides are polynomials. Substituting successively λ_1,λ_2, …, λ_n in the equation we see that

$$a_1 = \frac{p(\lambda_1)}{(\lambda_1 - \lambda_2)\dots(\lambda_1 - \lambda_n)}$$

$$\vdots$$

$$a_n = \frac{p(\lambda_n)}{(\lambda_2 - \lambda_n)\dots(\lambda_n - (\lambda_{n-1}}$$

The general result (i.e. where q has multiple zeros) is more complicated to state. We suppose that

$$q(\lambda) = (\lambda - \lambda_1)^{n-1}\dots(\lambda - \lambda_r)^{n_r}$$

where the λ_i are distinct and claim that the rational function can be written as a linear combination of functions of the form $\frac{1}{(\lambda - \lambda_i)^j}$ for $1 \leq i \leq r$ and $1 \leq j \leq n_i$.

PROOF. Write
$$\frac{p(\lambda)}{q(\lambda)} = \frac{p(\lambda)}{(\lambda-\lambda_1)^{n_1}q_1(\lambda)}$$
where $q_1(\lambda) = (\lambda-\lambda_2)^{n_2}\dots(\lambda-\lambda_r)^{n_r}$. We claim that there is a polynomial p_1 with $d(p_1) = d(p) - 1$ and an $a \in \mathbf{C}$ so that
$$\frac{p(\lambda)}{(\lambda-\lambda_1)^{n_1}q_1(\lambda)} = \frac{a}{(\lambda-\lambda_1)^{n_1}} + \frac{p_1(\lambda)}{(\lambda-\lambda_1)^{n_1-1}q_1(\lambda)}$$
from which the proof follows by induction.

For the above equation is equivalent to the following one:
$$\frac{p(\lambda)-aq_1(\lambda)}{q(\lambda)} = \frac{p_1(\lambda)}{(\lambda-\lambda_1)^{n_1-1}q(\lambda)}.$$
Hence it suffices to choose $a \in \mathbf{C}$ so that $p(\lambda) - aq_1(\lambda)$ contains a factor $\lambda - \lambda_1$ and there is precisely one such a namely $a = \dfrac{p(\lambda_1)}{q_1(\lambda)}$.

We remark that the degree function satisfies the following properties:
$$d(p+q) \leq \max(d(p), d(q))$$
with equality if $d(p) \neq d(q))$ and
$$d(pq) = d(p) + d(q)$$
provided that p and q are non-zero.

The standard high school method for the division of polynomials can be used to prove the existence part of the following result:

Proposition 7 *Let p and q be polynomials with $d(p) \geq 1$. Then there are unique polynomials r, s so that*
$$q = ps + r$$
where $r = 0$ or $d(r) < d(p)$.

PROOF. In the light of the above remark, we can confine ourselves to a proof of the uniqueness: suppose that
$$q = ps + r = ps_1 + r_1$$
for suitable s, r, s_1, r_1. Then
$$p(s - s_1) = r - r_1.$$
Now the right hand side is a polynomial of degree strictly less than that of p and hence so is the left hand side. But this can only be the case if $s = s_1$.

■

The above division algorithm can be used to prove an analogue of the Euclidean algorithm for determining the greatest common divisor of two polynomials p, q. We say that for two such polynomials, q is a **divisor** of p (written $q \mid p$) if there is a polynomial s so that $p = qs$. Note that then $d(p) \geq d(q)$ (where $d(p)$ denotes the degree of p). Hence if $p \mid q$ and $q \mid p$, then $d(p) = d(q)$ and it follows that p is a non-zero constant times q (we are tacitly assuming that the polynomials p and q are both non-zero). The **greatest common divisor** of p and q is by definition a common divisor which has as divisor each other divisor of p and q. It is then uniquely determined up to a scalar multiple and we denote it by g.c.d. (p, q). It can be calculated as follows: we suppose that $d(q) \leq d(p)$ and use the division algorithm to write

$$p = qs_1 + r_1$$

with $r_1 = 0$ or $d(r_1) < d(q)$. In the first case, q is the greatest common divisor. Otherwise we write

$$q = s_2 r_1 + r_2$$

then

$$r_1 = s_3 r_2 + r_3$$

and continue until we reach a final equation $r_k = s_{k+2} r_{k+1}$ without remainder. Then r_{k-1} is the greatest common divisor and by substituting backwards along the equations, we can compute a representation of it in the form $mp+nq$ for suitable polynomials m and n.

Lagrange interpolation A further useful property of polynomials is the following interpolation method: suppose that we have $(n+1)$ distinct points $t_0, \dots, t_n$ in $\mathbf{R}$. Then for any complex numbers $a_0, \dots, a_n$ we can find a polynomial p of degree at most n so that $p(t_i) = a_i$ for each i. To do this note that the polynomial

$$p_i(t) = \prod_{i \neq j} \frac{t - t_j}{t_i - t_j}$$

has the property that it takes on the value 1 at t_i and 0 at the other t_j. Then

$$p = \sum_{i=0}^{n} a_i p_i$$

is the required polynomial.

Exercises: 1) Show that a complex number λ_0 is a root of order r of the polynomial p (i.e. $(\lambda - \lambda_0)^r$ divides p) if and only if

$$p(\lambda_0) = p'(\lambda_0) = \cdots = p^{(r-1)}(\lambda_0) = 0.$$

2) Show that, in order to prove the fundamental theorem of algebra, it suffices to prove that every polynomial p of degree ≥ 1 has at least one zero.
3) Prove the statement of 2) (and hence the fundamental theorem of algebra) by verifying the following steps:

- show that if p is a non-constant polynomial, then there is a point λ_0 in $\mathbf{C}$ so that $|p(\lambda_0)| \leq |p(\lambda)| (\lambda \in \mathbf{C})$.
- show that $p(\lambda_0) = 0$.

(If $a_0 = p(\lambda_0)$ is non-zero, consider the Taylor expansion

$$p(\lambda) = a_0 + a_r(\lambda - \lambda_0)^r + \cdots + a_n(\lambda - \lambda_0)^n$$

of p at λ_0 where a_r is the first coefficient after a_0 which does not vanish. Show that there is a point λ_1 so that $|p(\lambda_1)| < |a_0|$ which is a contradiction).
4) Show that the set of rational functions is a field with the natural algebraic operations.

2.3 Complex vector spaces and matrices

We are now in a position to define complex vector spaces.

Definition: A **complex vector space** (or **vector space over C**) is a set V together with an addition and a scalar multiplication i.e. mappings $(x, y) \mapsto x + y$ resp. $(\lambda, x) \mapsto \lambda x$ form $V \times V$ into V resp. from $\mathbf{C} \times C$ into V so that

- $x + y = y + x \quad (x, y \in V)$'
- $x + (y + z) = (x + y) + z \quad (x, y, z \in V)$;
- there is a vector 0 so that $x + 0 = x$ for each $x \in V$;
- for each $x \in V$ there is a vector y so that $x + y = 0$;
- $(\lambda\mu)x = \lambda(\mu x) \quad (\lambda, \mu \in \mathbf{C}, x \in V)$;
- $1 \cdot x = x \quad (x \in V)$;
- $\lambda(x + y) = \lambda x + \lambda y$ and $(\lambda + \mu)x = \lambda x + \mu x \quad (\lambda, \mu \in \mathbf{C}, x, y \in V)$.

The following modifications of our examples of real vector spaces provide us with examples of complex ones:

- $\mathbf{C}^n$—the space of n-tuples $(\lambda_1, \dots, \lambda_n)$ of complex numbers;
- $\mathrm{Pol}_{\mathbf{C}}(n)$—the space of polynomials with complex coefficients of degree at most n;
- $M_{m,n}^{\mathbf{C}}$—the space of $m \times n$ matrices with complex elements.

We can then define linear dependence resp. independence for elements of a complex vector space and hence the concept of basis. Every complex vector space which is spanned by finitely many elements has a basis and so is isomorphic to $\mathbf{C}^n$ where n is the dimension of the space i.e. the cardinality of a basis.

If V and W are complex vector spaces, the notion of a linear mapping from V into W is defined exactly as in the real case (except, of course, for the fact that the homogeneity condition $f(\lambda x) = \lambda f(x)$ must now hold for all *complex* λ). Such a mapping f is determined by a matrix $[a_{ij}]$ with respect to bases $(x_1, \dots, x_n)$ resp. $(y_1, \dots, y_n)$ where the elements of the matrix are now complex numbers and are determined by the equations

$$f(x_j) = \sum_{i=1}^{m} a_{ij} y_i.$$

The theory of chapters I and V for matrices can then be carried over in the obvious way to complex matrices.

Sometimes it is convenient to be able to pass between complex and real vectors and this can be achieved as follows: if V is a complex vector space, then we can regard it as a real vector space simply by ignoring the fact that we can multiply by complex scalars. We denote this space by $V^{\mathbf{R}}$. This notation may seem rather pedantic—but note that if the dimension of V is n then that of $V_{\mathbf{R}}$ is $2n$. This reflects the fact that elements of V can be linearly dependent in V without being so in $V_{\mathbf{R}}$ since there are less possibilities for building linear combinations in the latter. For example, the sequence

$$(1, 0, \dots, 0), (0, 1, \dots, 0), \dots, (0, \dots, 0, 1)$$

is a basis for $\mathbf{C}^n$ whereas the longer sequence

$$(1, 0, \dots, 0), (i, 0, \dots, 0), (0, 1, \dots, 0), \dots, (0, \dots, 0, i)$$

is necessary to attain a basis for the *real* space $\mathbf{C}^n$ which is thus $2n$ dimensional.

On the other hand, it V is a real vector space we can define a corresponding complex vector space $V_{\mathbf{C}}$ as follows: as a set $V_{\mathbf{C}}$ is $V \times V$. It has the natural addition and scalar multiplication is defined by the equation

$$(\lambda + i\mu)(x_1, x_2) = (\lambda x_1 - \mu x_2, \mu x_1 + \lambda x_2).$$

The dimensions of V (as a real space) and $V_{\mathbf{C}}$ (as a complex space) are the same. If $f : V \to W$ is a linear mapping between complex vector space then it is *a fortiori* a linear mapping from $V_{\mathbf{R}}$ into $W_{\mathbf{R}}$. However, a real linear mapping between the latter spaces need not be complex-linear. On the other hand, if $f : V \to W$ is a linear mapping between real vector spaces, we can extend it to a complex linear mapping $f_{\mathbf{C}}$ between $V_{\mathbf{C}}$ and $W_{\mathbf{C}}$ by defining

$$f_{\mathbf{C}}(x_1, x_2) = (f(x_1), f(x_2)).$$

Exercises: 1) Solve the system

$$\begin{array}{ccccl} (1-i)x & - & 9y & & = 0 \\ 2x & + & (1-i)y & & = 1. \end{array}$$

2) Find a Hermitian form for the matrix

$$\begin{bmatrix} i & -(1+i) & 1 \\ 1 & 2 & -1 \\ 2i & 1 & 1 \end{bmatrix}.$$

3) Show that if z_1, z_2, z_3 are complex numbers, then

$$\det \begin{bmatrix} z_1 & \overline{z_1} & 1 \\ z_2 & \overline{z_2} & 1 \\ z_3 & \overline{z_3} & 1 \end{bmatrix}$$

is $4i$ times the area of the triangle with z_1, z_2, z_3 as vertices.
4) Let A and B be real $n \times n$ matrices. Show that if A and B are similar as complex matrices (i.e. if there is an invertible complex matrix P so that $P^{-1}AP = B$) then they are similar as real matrices (i.e. there is a real matrix P with the same property).
5) Let A and B be real $n \times n$ matrices. Show that

$$\overline{\det(A + iB} = \det(A - iB)$$

and that

$$\det \begin{bmatrix} A & -B \\ B & A \end{bmatrix} = |\det(A + iB)|^2.$$

6) (The following exercise shows that complex 2×2 matrices can be used to give a natural approach to the two products in $\mathbf{R}^3$). Consider the space $M_2^{\mathbf{C}}$ of 2×2 complex matrices. If A is such a matrix, say

$$A = \begin{bmatrix} a_{11} & a_{22} \\ alpha_{21} & a_{22} \end{bmatrix}$$

then we write A^* for the matrix

$$\begin{bmatrix} \overline{a_{11}} & \overline{a_{21}} \\ \overline{a_{12}} & \overline{a_{22}} \end{bmatrix}.$$

(The significance of this matrix will be discussed in more detail in Chapter VII). We let E_3 denote the family of those A which satisfy the conditions $A = A^*$ and $\operatorname{tr} A = 0$. The set of such matrices is a *real* vector space (but not a complex one). In fact if $x = (\xi_1, \xi_2, \xi_3)$ is an element of $\mathbf{R}^3$, then the matrix

$$A_x = \begin{bmatrix} \xi_3 & \xi_1 - i\xi_2 \\ \xi_1 + \xi_2 & -\xi_3 \end{bmatrix}$$

is in E_3. Show that this induces an isomorphism between $\mathbf{R}^3$ and E_3. Show that E_3 is not closed under (matrix) products but that if $A, B \in E_3$, then

$$A * B = \frac{1}{2}(AB + BA)$$

is also in E_3 and that

$$A_x * A_y = (x|y)I_2 + A_{x\times y}.$$

7) Complex 2×2 matrices also allow a natural approach to the subject of quaternions (cf. Exercise 8) of section VI.1). Consider the set of matrices of the form

$$\begin{bmatrix} z & w \\ \bar{w} & -\bar{z} \end{bmatrix}$$

where $z, w \in \mathbf{C}$. Show that this is closed under addition and multiplication and that the mapping

$$(z, w) \mapsto \begin{bmatrix} z & w \\ \bar{w} & -\bar{z} \end{bmatrix}$$

is a bijection between $\mathbf{Q}$ and the set of such matrices which preserves the algebraic operations. (Note that under this identification, the special quaternions have the form

$$\mathbf{i}i = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad \mathbf{j} = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \quad \mathbf{k} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

and that the quaternion $\xi_0 + \xi_1\mathbf{i} + \xi_2\mathbf{j} + \xi_3\mathbf{k}$ is represented by the matrix $\xi_0 I_2 + A_x$ where A_x is as in exercise 6)).

8) Use the results of the last two exercises to give new proofs of the following identities involving the vector and scalar products:

$x \times y = -y \times x$;

$\|x \times y\| = \|x\|\|y\| \sin\theta$ where θ is the angle between x and y;

$(x \times y) \times z = (x|z)y - (y|z)x$;

$(x \times y) \times z + (y \times z) \times x + (z \times x) \times y = 0.$

3 EIGENVALUES

3.1 Introduction

In this chapter we discuss the so-called **eigenvalue problem** for operators or matrices. This means that for a given operator $f \in L(V)$ a scalar λ and a non-zero vector x are sought so that $f(x) = \lambda x$ (i.e. the vector x is not rotated by f). Such problems arise in many situations, some of which we shall become acquainted with in the course of this chapter. In fact, if the reader examines the discussion of conic sections in the plane and three dimensional space he will recognise that the main point in the proof was the solution of an eigenvalue problem. The underlying theoretical reason for the importance of eigenvalues is the following: we know that a matrix is the coordinate representation of an operator. Even in the most elementary analytic geometry one soon appreciates the advantage of choosing a basis for which the matrix has a particularly simple form. The simplest possible form is that of a diagonal matrix and the reader will observe that we obtain such a representation precisely when the basis elements are so-called eigenvectors of the operator f i.e. they satisfy the condition $f(x_i) = \lambda_i x_i$ for suitable eigenvalues $\lambda_1, \dots, \lambda_n$ (which then form the diagonal elements of the corresponding matrix). Stated in terms of matrices this comprises what we may call the diagonalisation problem: given an $n \times n$ matrix A can we find an invertible matrix S so that $S^{-1}AS$ is diagonal?

Amongst the advantages that such a diagonalisation brings is the fact that one can then calculate simply and quickly arbitrary powers and thus polynomial functions of a matrix by doing this for the diagonal matrix and then transforming back. We shall discuss some applications of this below. On the other hand, if A is the matrix of a linear mapping in $\mathbf{R}^n$ we can immediately read off the geometrical form of the latter from its diagonalisation.

We begin with the formal definition. If $f \in L(V)$, an **eigenvalue** of f is a scalar λ so that there exists a non-zero x with $f(x) = \lambda x$. The space $\text{Ker}\,(f - \lambda \text{Id})$ is then non-trivial and is called the **eigenspace** of λ and each non-zero element therein is called an **eigenvector**. Our main concern in this chapter will be the following: given an operator f, can we find a basis for V consisting of eigenvectors? In general the answer is no as very simple examples show but we shall obtain a result which, while being much less direct, is still useful in theory and applications.

We can restate the eigenvalue problem in terms of matrices: an eigenvalue resp. eigenvector for an $n \times n$ matrix A is an eigenvalue resp. eigenvector for the operator f_A i.e. λ is an eigenvalue if and only if there exists a non-zero column vector X so that $AX = \lambda X$ and X is then called an eigenvector.

Before beginning a systematic development of the theory, we consider a simple example where an eigenvalue problem arises naturally—in this case in the solution of a linear system of ordinary differential equations:

Example: Consider the coupled system

$$\frac{df}{dt} = 3f + 2g$$
$$\frac{dg}{dt} = f + 2g.$$

If we introduce the (vector-valued) function

$$F(t) = \begin{bmatrix} f(t) \\ g(t) \end{bmatrix}$$

we can write this equation in the form

$$\frac{dF}{dt} = AF$$

where $A = \begin{bmatrix} 3 & 2 \\ 1 & 2 \end{bmatrix}$ and $\dfrac{dF}{dt}$ is, of course, the function $\begin{bmatrix} \frac{df}{dt} \\ \frac{dg}{dt} \end{bmatrix}$.

Formally, this looks very much like one of the simplest of all differential equations—the equation

$$\frac{df}{dt} = af$$

and by analogy, we try the substitution

$$F(t) = e^{\lambda t} X \quad \text{where} \quad X = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

This leads to the solution

$$Ae^{\lambda t} X = AF(t) = F'(t) = \lambda e^{\lambda t} X$$

or $AX = \lambda X$ i.e. an eigenvalue problem for A which we now proceed to solve. The column vector must be a non-trivial solution of the homogeneous system

$$\begin{array}{ccccc} (3-\lambda)x_1 & + & 2x_2 & = & 0 \\ x_1 & + & (2-\lambda)x_2 & = & 0. \end{array}$$

Now we know that such a solution exists if and only if the determinant of the corresponding matrix vanishes. This leads to the equation

$$(3-\lambda)(2-\lambda)-2=0$$

for λ which has solutions $\lambda = 1$ or $\lambda = 4$.

If we solve the corresponding homogeneous systems we get eigenvectors $(-1,1)$ and $(2,1)$ respectively and they form a basis for $\mathbf{R}^2$. hence if

$$S = \begin{bmatrix} -1 & 2 \\ 1 & 1 \end{bmatrix}$$

is the matrix whose columns are the eigenvectors of A we see that

$$S^{-1}AS = \begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix}.$$

Then our differential equation has the solution

$$F(t) = c_1 e^t \begin{bmatrix} -1 \\ 1 \end{bmatrix} + c_2 e^{4t} \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

i.e. $f(t) = -c_1 e^t + 2ce^{4t}$, $g(t) = c_1 e^t + c_2 e^{4t}$ for arbitrary constants c_1, c_2.

3.2 Characteristic polynomials and diagonalisation

The above indicates the following method for characterising eigenvalues:

Proposition 8 *If A is an $n \times n$ matrix, then λ is an eigenvalue of A if and only if λ is a root of the equation*

$$\det(A - \lambda I) = 0.$$

PROOF. λ is an eigenvalue if and only if the homogeneous equation $AX = \lambda X$ has a non-trivial solution and the matrix of this equation is $A - \lambda I$.
We remark that $\det(A - \lambda I) = 0$ is a polynomial equation of degree n.

From this simple result we can immediately draw some useful conclusions:
I. Every matrix over $\mathbf{C}$ has n eigenvalues, the n roots of the above equation (although some of the eigenvalues can occur as repeated roots of the equation).
II. Every real matrix has at least one real eigenvalue if n is odd (in particular if $n = 3$). The example of a rotation in $\mathbf{R}^2$ shows that 2×2 matrices over $\mathbf{R}$ need not have any eigenvalues.
III. If A is a triangular matrix, say

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ 0 & a_{22} & \dots & a_{2n} \\ \vdots & & & \vdots \\ 0 & 0 & \dots & a_{nn} \end{bmatrix},$$

then the eigenvalues of A are just the diagonal elements $a_{11}, \dots, a_{nn}$ (in particular, this holds if A is diagonal). For

$$\det(A - \lambda I) = (a_{11} - \lambda) \dots (a_{nn} - \lambda)$$

with roots $a_{11}, \dots, a_{nn}$.

Owing to the special role played by the polynomial $\det(A - \lambda I)$ in the eigenvalue problem we give it a special name—the **characteristic polynomial** of A—in symbols χ_A. For example, if

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

then $\chi_A(\lambda) = \lambda(1 - \lambda)(\lambda - 3)$.

For an operator $f \in L(V)$, the characteristic polynomial is defined by the equation

$$\chi_f(\lambda) = \det(f - \lambda Id)$$

and the eigenvalues are the roots of this polynomial.

We now turn to the topic of the diagonalisation problem. The connection with the eigenvalue problem is made explicit in the following result:

Proposition 9 *A linear operator $f \in L(V)$ is diagonalisable if and only if V has a basis $(x_1, \dots, x_n)$ consisting of eigenvectors of f.*

Proposition 10 *If an $n \times n$ matrix A has n linearly independent eigenvectors $X_1, \dots, X_n$ and S is the matrix $[X_1 \dots X_n]$, then S diagonalises A i.e.*

$$S^{-1}AS = diag\,(\lambda_1, \dots, \lambda_n)$$

where the λ_i are the corresponding eigenvalues.

PROOF. If the matrix of f with respect to the basis $(x_1, \dots, x_n)$ is the diagonal matrix

$$\begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & & & \vdots \\ 0 & 0 & \dots & \lambda_n \end{bmatrix},$$

then $f(x_i) = \lambda_i x_i$ i.e. each x_i is an eigenvector. Conversely, if (x_i) is a basis so that $f(x_i) = \lambda_i x_i$ for each i, then the matrix of f is as above. The second result is simply the coordinate version.

As already mentioned, the example of a rotation in $\mathbf{R}^2$ shows that the condition of the above theorems need not always hold. The problem is that the matrices of rotations (with the trivial exceptions D_π and D_0) have no (real) eigenvalues. There is no problem if the operator does have n distinct eigenvalues, as the next result shows:

Proposition 11 *Let $f \in L(V)$ be a linear operator in an n dimensional space and suppose that f has r distinct eigenvalues with eigenvectors $x_1, \dots, x_r$. Then $\{x_1, \dots, x_r\}$ is linearly independent. Hence if f has n distinct eigenvalues, it is diagonalisable.*

PROOF. If the x_i are linearly dependent, there is a smallest s so that x_s is linearly dependent on $x_1, \dots, x_{s-1}$, say $x_s = \mu_1 x_1 + \dots \mu_{s-1} x_{s-1}$. If we apply f to both sides and then subtract λ_s times the original equation, we get:

$$0 = \mu_1(\lambda_1 - \lambda_s)x_1 + \dots + \mu_{s-1}(\lambda_{s-1} - \lambda_s)x_{s-1}$$

and this implies that the $x_1, \dots, x_{s-1}$ are linearly dependent which is a contradiction.

Of course, it is not *necessary* for a matrix to have n distinct eigenvalues in order for it to be diagonalisable, the simplest counterexample being the unit matrix.

Estimates for eigenvalues For applications it is often useful to have estimates for the eigenvalues of a given matrix, rather than their precise values. In this section, we bring two such estimates, together with some applications.

Recall that if a matrix A is dominated by the diagonal in the sense that for each i

$$|a_{ii}| - \sum_{j \neq i} |a_{ij}| > 0,$$

then it is invertible (see Chapter IV). This can be used to give the following estimate:

Proposition 12 *Let A be a complex $n \times n$ matrix with eigenvalues $\lambda_1, \dots, \lambda_n$. Put for each i*

$$\alpha_i = \sum_{j \neq i} |a_{ij}|.$$

Then the eigenvalues lie in the region

$$\bigcup \{z \in \mathbf{C} : |z - a_{ii}| \leq \alpha_i\}.$$

PROOF. It is clear that if λ does not lie in one of the above circular regions, then the matrix $(\lambda I - A)$ is dominated by the diagonal in the above sense and so is invertible i.e. λ is not an eigenvalue.

We can use this result to obtain a classical estimate for the zero of polynomials. Consider the polynomial p which maps t onto $a_0 + a_1 t + \cdots + a_{n-1} t^{n-1} + t^n$. The roots of p coincide with the eigenvalues of the companion matrix

$$C = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & & & & \vdots \\ -a_0 & -a_1 & -a_2 & \dots & -a_{n-1} \end{bmatrix}$$

(see Exercise 4) below).

It follows from the above criterium that if λ is a zero of p, then

$$|\lambda| \leq \max(|a_0|, 1 + |a_1|, \dots, 1 + |a_{n-1}|).$$

Our second result shows that the eigenvalues of a small matrix cannot be too large. More precisely, if A is an $n \times n$ matrix and $a = \max_{i,j} |a_{ij}|$, then each eigenvalue λ satisfies the inequality: $|\lambda| \leq na$. For suppose that

$$X = \begin{bmatrix} \xi_1 \\ \vdots \\ \xi_n \end{bmatrix}$$

is a corresponding eigenvector. Then we have

$$(AX|X) = \lambda(X|X)$$

i.e.

$$\lambda \sum_i \xi_i \overline{\xi_i} = \sum_{i,j} a_{ij} \xi_i \overline{\xi_j}.$$

Taking absolute values, we have the inequality

$$\begin{aligned} |\lambda| \sum_i |\xi_i|^2 &\leq \sum_{i,j} |a_{ij}||\xi_i||\xi_j| \\ &\leq a \sum_{i,j} |\xi_i||\xi_j| \\ &= a(\sum_i |\xi_i|)(\sum_j |\xi_j|) \\ &\leq na(\sum_i |\xi_i|^2) \end{aligned}$$

which implies the result. (In the last inequality, we use the Cauchy-Schwarz inequality which implies that $\sum_i |\xi_i| \leq n^{\frac{1}{2}}(\sum_i |\xi_i|^2)^{\frac{1}{2}}$. See the next chapter for details).

We conclude this section with an application of the diagonalisation method:

Difference equations One of many applications of the technique of diagonalising a matrix is the solving of difference equations. Rather than develop a general theory we shall show how to solve a particular equation—the method is, however, quite general.

The example we choose is the difference equation which defines the famous **Fibonacci series**. This is the sequence (f_n) which is defined by the initial conditions $f_1 = f_2 = 1$ and the recursion formula $f_{n+2} = f_n + f_{n+1}$. If we write X_n for the 2×1 matrix $\begin{bmatrix} f_{n+1} \\ f_n \end{bmatrix}$, then we can write the defining conditions in the form

$$X_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \qquad \text{and} \quad X_{n+1} = AX_n$$

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

A simple induction argument shows that the general solution is then $X_n = A^{n-1}X_1$. In order to be able to exploit this representation it is necessary to compute the powers of A. To do this directly would involve astronomical computations. The task is simplified by diagonalising A. A simple calculation shows that A has eigenvalues $\dfrac{1+\sqrt{5}}{2}$ and $\dfrac{1-\sqrt{5}}{2}$, with eigenvectors

$$\begin{bmatrix} \frac{1+\sqrt{5}}{2} \\ 1 \end{bmatrix} \quad \text{and} \begin{bmatrix} \frac{1-\sqrt{5}}{2} \\ 1 \end{bmatrix}.$$

Hence

$$S^{-1}AS = \begin{bmatrix} \frac{1+\sqrt{5}}{2} & 0 \\ 0 & \frac{1-\sqrt{5}}{2} \end{bmatrix}$$

where $S = \begin{bmatrix} \frac{1+\sqrt{5}}{2} & \frac{1-\sqrt{5}}{2} \\ 1 & 1 \end{bmatrix}$.

From this it follows that

$$A = \frac{1}{\sqrt{5}} \begin{bmatrix} \frac{1+\sqrt{5}}{2} & \frac{1-\sqrt{5}}{2} \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \frac{1+\sqrt{5}}{2} & 0 \\ 0 & \frac{1-\sqrt{5}}{2} \end{bmatrix} \begin{bmatrix} 1 & -\frac{1-\sqrt{5}}{2} \\ -1 & \frac{1+\sqrt{5}}{2} \end{bmatrix}$$

and

$$A^n = \frac{1}{\sqrt{5}} \begin{bmatrix} \frac{1+\sqrt{5}}{2} & \frac{1-\sqrt{5}}{2} \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \frac{1+\sqrt{5}}{2} & 0 \\ 0 & \frac{1-\sqrt{5}}{2} \end{bmatrix}^n \begin{bmatrix} 1 & -\frac{1-\sqrt{5}}{2} \\ -1 & \frac{1+\sqrt{5}}{2} \end{bmatrix}$$

This leads to the formula

$$f_n = \frac{1}{\sqrt{5}} \left[\left(\frac{1+\sqrt{5}}{2} \right)^n - \left(\frac{1-\sqrt{5}}{2} \right)^n \right].$$

Examples: 1) Calculate $\chi_A(\lambda)$ where

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

Solution:

$$\begin{aligned} \chi_A(\lambda) &= \det \begin{bmatrix} 1-\lambda & 1 & \dots & 1 \\ \vdots & & & \vdots \\ 1 & 1 & \dots & 1-\lambda \end{bmatrix} \\ &= (n-\lambda)\lambda^{n-1}(-1)^{n-1}. \end{aligned}$$

by a result of the previous chapter.

2) Calculate the eigenvalues of the $n \times n$ matrix

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

Solution:

$$\begin{aligned} \chi_A(\lambda) &= \begin{bmatrix} -\lambda & 1 & 0 & \dots & 0 \\ 0 & \lambda & 1 & \dots & 0 \\ \vdots & & & & \vdots \\ 1 & 0 & 0 & \dots & \lambda \end{bmatrix} \\ &= (-1)^{n-1}(\lambda^n - 1). \end{aligned}$$

Hence the eigenvalues are the roots $e^{\frac{2\pi i r}{n}}$ of unity $(r = 0, \dots, n-1)$.
3) Calculate the eigenvalues of the linear mapping

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \mapsto \begin{bmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{bmatrix}$$

on M_2.
Solution: With respect to the basis

$$x_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \quad x_2 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \quad x_3 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \quad x_4 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

the mapping has matrix

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

and this has eigenvalues $1, 1, 1, -1$.
4) Show that $f_{n+1}f_{n-1} - f_n^2 = (-1)^{n-1}$ where f_n is the n-th Fibonacci number.
Solution: Note the

$$\begin{aligned} f_{n+1}f_{n-1} - f_n^2 &= \det \begin{bmatrix} f_{n+1} & f_n \\ f_n & f_{n-1} \end{bmatrix} \\ &= \det \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}^{n-1} \begin{bmatrix} f_2 & f_1 \\ f_1 & f_0 \end{bmatrix} \\ &= (\det \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix})^{n-1} \det \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} \\ &= (-1)^{n-1}. \end{aligned}$$

.

3.3 The Jordan canonical form

As we have seen, not every matrix can be reduced to diagonal form and in this section we shall investigate what can be achieved in the general case. We begin by recalling that failure to be diagonalisable can result from two causes. Firstly, the matrix can fail to have a sufficient number of eigenvalues (i.e. zeroes of χ_A). By the fundamental theorem of algebra, this can only happen in the real case and in this section we shall avoid this difficulty by confining our attention to complex matrices resp. vector spaces. The second difficulty is that the matrix may have n eigenvalues (with repetitions) but may fail to have enough eigenvectors to span the space. A typical example is the shear operator

$$(\xi_1, \xi_2) \mapsto (\xi_1 + \xi_2, \xi_2)$$

with matrix $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$.

This has the double eigenvalue 1 but the only eigenvectors are multiples of the unit vector $(1, 0)$. We will investigate in detail the case of repeated eigenvalues and it will turn out that in a certain sense the shear operator represents the typical situation. The precise result that we shall obtain is rather more delicate to state and prove than the diagonalisable case and we shall proceed by way of a series of partial results. We begin with the following Proposition which allows us to reduce to the case where the operator f has a single eigenvalue.

In order to avoid tedious repetitions we assume from now until the end of this section that f is a fixed operator on a *complex* vector space V of dimension and that f has eigenvalues

$$\lambda_1, \dots, \lambda_1, \lambda_2, \dots, \lambda_r, \dots, \lambda_r$$

where λ_i occurs n_i times. This means that f has characteristic polynomial

$$(\lambda_1 - \lambda)^{n_1} \dots (\lambda_r - \lambda)^{n_r}$$

where $n_1 + \cdots + n_r = n$).

Proposition 13 *There is a direct sum decomposition*

$$V = V_1 \oplus \cdots \oplus V_r$$

where

- *each V_i is f invariant (i.e. $f(V_i) \subset V_i$);*
- *the dimension of V_i is n_i and $(f - \lambda_i Id)^{n_i}|_{V_i} = 0$.*

In particular, the only eigenvalue of $f|_{V_i}$ is λ_i.

PROOF. Fix i. It is clear that

$$\operatorname{Ker}(f-\lambda_i \mathrm{Id}) \subset \operatorname{Ker}(f-\lambda_i \mathrm{Id})^2 \subset \dots$$

Hence there exists a smallest r_i so that

$$\operatorname{Ker}(f-\lambda_i \mathrm{Id})^{r_i} = \operatorname{Ker}(f-\lambda_i \mathrm{Id})^{r_i+1}$$

and so

$$\operatorname{Ker}(f-\lambda_i \mathrm{Id})^{r_i} = \operatorname{Ker}(f-\lambda_i \mathrm{Id})^{r_i+m}$$

for $m \in \mathbf{N}$.

Then we claim that

$$V = \operatorname{Ker}(f-\lambda_i \mathrm{Id})^{r_i} \oplus \operatorname{Im}(f-\lambda_i \mathrm{Id})^{r_i}.$$

Since the sum of the dimensions of these two spaces is that of V, it suffices to show that their intersection is $\{0\}$. But if $y \in \operatorname{Ker}(f-\lambda_i \mathrm{Id})^{r_i}$ and $y = (f-\lambda_i \mathrm{Id})^{r_i}(x)$, then $(f-\lambda_i \mathrm{Id})^{2r_i}(x) = 0$ and so $x \in \operatorname{Ker}(f-\lambda_i \mathrm{Id})^{2r_i} = \operatorname{Ker}(f-\lambda_i \mathrm{Id})^{r_i}$ i.e. $y=0$. It is now clear that if $V_i = \operatorname{Ker}(f-\lambda_i \mathrm{Id})^{r_i}$, then

$$V = V_1 \oplus \dots \oplus V_r$$

is the required decomposition.

If $f \in L(V)$, then the sequences

$$\operatorname{Ker} f \subset \operatorname{Ker} f^2 \subset \dots$$

and

$$f(V) \supset f^2(V) \supset \dots$$

become stationary at points r, s i.e. we have

$$\operatorname{Ker} f \neq \operatorname{Ker} f^2 \neq \dots \neq \operatorname{Ker} f^r = \operatorname{Ker} f^{r+1} = \dots$$

and

$$f(V) \neq f^2(V) \neq \dots \neq f^s(V) = f^{s+1}(V) = \dots$$

Then the above proof actually shows the following:

Proposition 14 *$r = s$ and $V = V_1 \oplus V_2$ where $V_1 = f^r(V)$ and $V_2 = Ker\, f^r$.*

Corollar 2 *If f is such that $Ker\, f = Ker\, f^2$, then $V = f(V) \oplus Ker\, f$.*

Using the above result, we can concentrate on the restrictions of f to the summands. These have the special property that they have only one eigenvalue. Typical examples of matrices with this property are the Jordan matrices which we introduced in the first chapter. Recall the notation

$$J_n(\lambda) = \begin{bmatrix} \lambda & 1 & 0 & \dots & 0 \\ 0 & \lambda & 1 & \dots & 0 \\ \vdots & & & & \vdots \\ 0 & 0 & 0 & \dots & \lambda \end{bmatrix}.$$

In particular, the shear matrix is $J_1(1)$.

The following facts can be computed easily.

1) $J_n(\lambda)$ has one eigenvalue, namely λ, and one eigenvector $(1, 0, \dots, 0)$ (or rather multiples of this vector);

2) If p is a polynomial, then

$$p(J_n(\lambda)) = \begin{bmatrix} p(\lambda) & p'(\lambda) & \dots & \frac{p^{(n-1)}(\lambda)}{(n-1)!} \\ 0 & p(\lambda) & \dots & \\ \vdots & & & \vdots \\ 0 & 0 & \dots & p(\lambda) \end{bmatrix}.$$

3) $J_n(\lambda) - \lambda I$ is the matrix

$$J_n(0) = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & & & & \vdots \\ 0 & 0 & 0 & \dots & 0 \end{bmatrix}.$$

Hence $(J_n(\lambda) - \lambda I)^n = 0$ and $(J_n(\lambda) - \lambda I)^r \neq 0$ if $r < n$.

The next result shows that all operators with only one eigenvalue can be represented by blocks of Jordan matrices:

Proposition 15 *Let g be a linear operator on the n-dimensional space W so that $(g - \lambda I)^n = 0$ for some $\lambda \in \mathbf{C}$. Then there is a decomposition*

$$W = W_1 \oplus \cdots \oplus W_k$$

so that each W_i is g-invariant and has a basis with respect to which the matrix of g is the Jordan matrix $J_{s_i}(\lambda)$ where $s_i = \dim W_i$.

By replacing g by $g - \lambda I$ we can reduce to the following special case which is the one which we shall prove:

Proposition 16 *Let $g \in L(V)$ be nilpotent with $g^r = 0$, $g^{r-1} \neq 0$. Then there is a decomposition*

$$V = V_1 \oplus \cdots \oplus V_k$$

with each V_i g-invariant and a basis for each V_i so that $g|_{V_i}$ has matrix $J_{s_i}(0)$ where $s_i = \dim V_i$.

PROOF. Choose $x_1 \in V$ so that $g^{r-1}(x_1) \neq 0$. Then the vectors

$$x_1, g(x_1), \ldots, g^{r-1}(x_1)$$

are linearly independent. Otherwise there is a greatest k so that $g^k(x_1)$ is linearly dependent on $g^{k+1}(x_1), \ldots, g^{r-1}(x_1)$ say

$$g^k(x_1) = \lambda_{k+1} g^{k+1}(x_1) + \cdots + \lambda_{r-1} g^{r-1}(x_1).$$

But if we apply g^{r-k-1} to both sides we get $g^{r-1}(x_1) = 0$—a contradiction.

Now there are two possibilities: a) $(x_1, g(x_1), \ldots, g^{r-1}(x_1))$ spans V. Then

$$y_1 = g^{r-1}(x_1), y_2 = g^{r-2}(x_1), \ldots, y^r = x_1$$

is a basis for V with respect to which g has matrix $J_r(0)$.
b) $V_1 = [x_1, g(x_1), \ldots, g^{r-1}(x_1)] \neq V$. We then construct a g-invariant subspace V_2 whose intersection with V_1 is the zero-vector. We do this as follows: for each y not in V_1 there is an integer s with $g^{s-1}(y) \notin V_1$, and $g^s(y) \in V_1$ (since $g^i(y)$ is eventually zero). Choose a $y \in V \setminus V_1$ for which this value of s is maximal.

Suppose that $g^s(y) = \sum_{j=0}^{r-1} \lambda_j g^j(x_1)$. Then

$$\begin{aligned} 0 &= g^r(y) \\ &= g^{r-s}(g^s(y)) \\ &= \sum_{j=0}^{r-1} \lambda_j g^{j+r+s}(x_1) \\ &= \sum_{j=0}^{s-1} \lambda_j g^{j+r-s}(x_1). \end{aligned}$$

Hence $\lambda_j = 0$ for $j = 0, \ldots, s-1$ since $x_1, g(x_1), \ldots, g^{r-1}(x_1)$ are linearly independent and so $g^s(y) = \sum_{j=s}^{r-1} \lambda_j g^j(x_1)$. Put $x_2 = y - \sum_{j=s}^{r-1} \lambda_j g^{j-s}(x_1)$. Then $g^s(x_2) = 0$ and by the same argument as above, $\{x_2, g(x_2), \ldots g^{s-1}(x_2)\}$ is linearly independent. Then $V_2 = [x_2, \gamma(x_2), \ldots, g^{s-1}(x_2)]$ is g-invariant and has the desired property.

Now if $V = V_1 \oplus V_2$ we are finished. If not we can proceed in the same manner to obtain a suitable V_3 and so on until we have exhausted V.

We are now in a position to state and prove our general result. Starting with the operator $f \in L(V)$ we first split V up in the form

$$V = V_1 \oplus \cdots \oplus V_r$$

where each V_i is f-invariant and the restriction of $(f - \lambda_i \mathrm{Id})$ to V_i is nilpotent. Applying the second result we get a further splitting

$$V_i = W_1^i \oplus \cdots \oplus W_{k_i}^i$$

and a basis for W_j^i so that the matrix is a Jordan matrix. Combining all of the bases for the various W_i^j we get one for V with respect to which the matrix of f has the form

$$\operatorname{diag}(J(\lambda_1), \dots, J(\lambda_1), J(\lambda_2), \dots, J(\lambda_r), \dots, J(\lambda_r))$$

where we have omitted the subscripts indicating the dimensions of the Jordan matrices.

This result about the existence of the above representation (which is called the **Jordan canonical form** of the operator) is rather powerful and can often be used to prove non-trivial facts about matrices by reducing to the case of Jordan matrices. We use this technique in the following proof of the so-called Cayley-Hamilton theorem which states that a matrix is a "solution" of its own characteristic equation.

Proposition 17 *Let A be an $n \times n$ matrix. Then $\chi_A(A) = 0$.*

PROOF. We begin with the case where A has Jordan form i.e. a block representation $\operatorname{diag}(A_1, \dots, A_r)$ where A_i is the part corresponding to the eigenvalue λ_i. A_i itself can be divided into Jordan blocks i.e.

$$A = \operatorname{diag}(J(\lambda_i), \dots, J(\lambda_i)).$$

Now if p is a polynomial, then $p(A) = \operatorname{diag}(p(A_1), \dots, p(A_r))$ and so it suffices to show that $\chi_A(A_i) = 0$ for each i. But χ_A contains the factor $(\lambda_i - \lambda)^{n_i}$ and so $\chi_A(A_i)$ contains the factor $(\lambda I - A_i)^{n_i}$ and we have seen that this is zero.

We now consider the general case i.e. where A is not necessarily in Jordan form. We can find an invertible matrix S with $\tilde{A} = S^{-1}AS$ has Jordan form. Then $\chi_{\tilde{A}} = \chi_A$ and so

$$\chi_A(A) = \chi_{\tilde{A}}(S) = \chi_{\tilde{A}}(S\tilde{A}S^{-1}) = S\chi_{\tilde{A}}(\tilde{A})S^{-1} = 0.$$

The Cayley-Hamilton theorem can be used to calculate higher powers and inverses of matrices. We illustrate this with a simple example:

Example: If

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

then

$$\chi_A(\lambda) = -\lambda^3 + 3\lambda^2 + 3\lambda - 2$$

and so

$$-A^3 + 3A^2 + 3A - 2I = 0.$$

Hence $A^3 = 3A^2 - 3A - 2I$. From this it follows that

$$\begin{aligned} A^4 &= 3A^3 + 3A^2 - 2A \\ &= 3 \begin{bmatrix} 13 & 18 & 3 \\ 9 & 13 & 21 \\ 9 & 18 & 12 \end{bmatrix} + 3 \begin{bmatrix} 3 & 7 & 7 \\ 2 & 5 & 5 \\ 3 & 3 & 7 \end{bmatrix} - 2 \begin{bmatrix} 2 & -1 & 3 \\ 1 & 0 & 2 \\ 0 & 3 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 44 & 77 & 105 \\ 31 & 57 & 74 \\ 36 & 57 & 85 \end{bmatrix}. \end{aligned}$$

Also $2I = -A^3 + 3A^2 + 3A = A(-A^2 + 3A + 3I)$ i.e.

$$A^{-1} = -\frac{1}{2}(-A^2 + 3A + 3I) = \begin{bmatrix} 3 & -5 & -8 \\ \frac{1}{2} & -1 & -\frac{1}{2} \\ ? & 3 & -\frac{1}{2} \end{bmatrix}.$$

A further interesting fact that can easily be verified with help of the Jordan form is the following:

Proposition 18 *Let A be an $n \times n$ matrix with eigenvalues $\lambda_1, \dots, \lambda_n$ and let p be a polynomial. Then the eigenvalues of $p(A)$ are $p(\lambda_1), \dots, p(\lambda_n)$.*

PROOF. Without loss of generality, we can assume that A has Jordan form and then $p(A)$ is a triangular matrix with diagonal entries $p(\lambda_1), \dots, p(\lambda_n)$.

The above calculations indicate the usefulness of a polynomial p such that $p(A) = 0$. The Cayley-Hamilton theorem provides us with one of degree n. In general, however, there will be suitable polynomials of lower degree. For example, the characteristic polynomial of the identity matrix I_n is $(1 - \lambda)^n$ but $p(I) = 0$ where p is the linear polynomial $p(\lambda) = 1 - \lambda$. Since it is obviously of advantage to take the polynomial of smallest degree with this property, we introduce the following definition:

Definition: Let A be an $n \times n$ matrix with characteristic polynomial‘

$$\chi_A(\lambda) = (\lambda_1 - \lambda)^{n_1} \dots (\lambda_r - \lambda)^{n_r}.$$

Then there exists for each i a smallest m_i ($\leq n_i$) so that $p(A) = 0$ where

$$p(\lambda) = (\lambda_1 - \lambda)^{m_1} \dots (\lambda_r - \lambda)^{m_r}.$$

This polynomial is called the **minimal polynomial** of A and denoted by m_A. In principle it can be calculated by considering the $n_1 \cdot n_2 \dots n_r$ divisors of the characteristic polynomial which contain the factor $(\lambda_1 - \lambda) \dots (\lambda_r - \lambda)$ and determining the one of lowest degree which annihilates A. In terms of the Jordan canonical form of A it is clear that m_i is the order of the largest Jordan matrix in the block corresponding to the eigenvalue λ_i.

We conclude with two simple and typical applications of the Cayley-Hamilton theorem.

I. Suppose that we are given a polynomial p with roots $\lambda_1, \dots, \lambda_n$ and are required to construct a second one whose roots are the square of the λ_i (without calculating these roots explicitly). This can be done as follows: let A be the companion matrix of p so that the eigenvalues of A are the roots of p. Then if $B = A^2$, the eigenvalues of B are the required numbers. Hence $q = \chi_B$ is a suitable polynomial.

II. Suppose that we are given two polynomials p and q whereby the roots of p are $\lambda_1, \dots, \lambda_n$. If A is the companion matrix of p, then the eigenvalues of $q(A)$ are $q(\lambda_1), \dots, q(\lambda_n)$. Hence p and q have a common root if and only if $\det q(A) = 0$. This gives a criterium for the two polynomials to have a common root. For this reason the quantity $\Delta = \det q(A)$ is called the **resultant** of p and q.

The particular case where q is the derivative of p is useful since the existence of a common root for p and p' implies that p has a double root. In this case the expression $\Delta = \det p'(A)$ is called the **discriminant** of p.

Example: For which values of a,b, c is the matrix

$$A = \begin{bmatrix} 0 & 0 & 1 \\ 4 & a & b \\ -2 & 1 & c \end{bmatrix}$$

nilpotent?

Solution: We calculate as follows:

$$A^2 = \begin{bmatrix} -2 & 1 & c \\ 4a - 2b & a^2 + b & 4 + ab + bc \\ 4 - 2c & a + c & -2 + b + c^2 \end{bmatrix}$$

which is never zero.

$$A^3 = \begin{bmatrix} 4-2c & a+c & -2+b+c^2 \\ ? & ? & ? \\ ? & ? & ? \end{bmatrix}$$

(we do not require the entries marked by a question mark). If $A^3 = 0$, we must have $c = 2$, $a = -2$, $b = -2$. Then

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

and one calculates that $A^3 = 0$ i.e. A is nilpotent if and only if a, b and c have the above values.

Example: Show that the matrix

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

is nilpotent and find a basis which reduces it to Jordan form.
Solution: One calculates that

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

and $A^3 = 0$. Now

$$A \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ -3 \\ 1 \end{bmatrix} \quad \text{and} \quad A^2 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} -2 \\ 2 \\ 2 \end{bmatrix}.$$

Then $(-2, 2, 2)$, $(1, -3, 1)$ and $(1, 0, 0)$ are linearly independent and with respect to this basis f_A has matrix

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

Example: Calculate the minimal polynomials of

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

Solution: $\chi_A(\lambda) = (1-\lambda)^3$ and

$$(A - \lambda I) = \begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

It is clear that $(A-I)^2 \neq 0$ and $(A-I)^3 = 0$ i.e. $m_A(\lambda) = (1-\lambda)^3$. $\chi_B(\lambda) = \lambda^2(3-\lambda)$ and one calculates that $B(B-3I) = 0$ i.e. $m_B(\lambda) = \lambda(3-\lambda)$.

Exercises: 1) For which values of a, b and c are the following matrices nilpotent?

$$A = \begin{bmatrix} 2 & -2 & a \\ 2 & -2 & b \\ 1 & -1 & 0 \end{bmatrix} \qquad B = \begin{bmatrix} 0 & a & b \\ 6 & -3 & c \\ -2 & 1 & -3 \end{bmatrix}.$$

2) For each eigenvalue λ of the matrices A and B below calculate the value of r for which $\operatorname{Ker}(A-\lambda I)^r$ becomes stationary:

$$A = \begin{bmatrix} 4 & 6 & 0 \\ -1 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad B = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 3 & 2 & 1 & 0 \\ 4 & 3 & 2 & 1 \end{bmatrix}.$$

3) Solve the differential equations:

$$\begin{aligned} \frac{df}{dt} &= g \\ \frac{dg}{dt} &= -6f + 5g. \end{aligned}$$

4) Diagonalise the matrix

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

and use it to solve the difference equations:

$$\begin{array}{rcrcrcr} a_{n+1} & = & a_n & - & b_n & + & 2c_n \\ b_{n+1} & = & -a_n & + & b_n & + & 2c_n \\ c_{n+1} & = & 2a_n & + & 2b_n & - & 2c_n \end{array}$$

resp. the system of equations:

$$\begin{array}{rcrcrcr} \dfrac{dx}{dt} & = & x & - & y & + & 2z \\ \dfrac{dy}{dt} & = & -x & + & y & + & 2z \\ \dfrac{dz}{dt} & = & 2x & + & 2y & - & 2z. \end{array}$$

5) Let V_1 be a subspace of the vector space V which is invariant under the operator $f \in L(V)$. Show that if f is diagonalisable, then so is the restriction of f to V_1.
6) A **cyclic element** for an operator $f \in L(V)$ is a vector x so that

$$\{x, f(x), f^2(x), \dots, f^{n-1}(x)\}$$

forms a basis for V (where $n = \dim V$). Show that if f is diagonalisable, then it has a cyclic element if and only if its eigenvalues are distinct.
7) Show that if two matrices are diagonalised by the same invertible matrix, then they commute.
8) Let A be a diagonalisable $n \times n$ matrix. Show that a matrix B commutes with every matrix that commutes with A if and only if there is a polynomial p so that $B = p(A)$. Show that if A has distinct eigenvalues, then it suffices for this that B commutes with A.
9) Let $f \in L(V)$ be such that each vector $x \in V$ is an eigenvalue. Show that there is a $\lambda \in \mathbf{R}$ so that $f = \lambda \mathrm{Id}$.
10) Let $f, g \in L(V)$ commute. Show that each eigenspace of f is g-invariant.
11) Suppose that $f \in L(V)$ has n distinct eigenvalues where $n = \dim V$. Show that V contains precisely 2^n f-invariant subspaces.
12) Show that $f \in L(V)$ is nilpotent if and only if its characteristic polynomial has the form $\pm\lambda^n$ where $n = \dim V$. Deduce that if f is nilpotent, then $f^n = 0$. 13) A nilpotent operator f on an n dimensional space V is **nilcyclic** if and only if V has a basis of the form $(x, f(x), f^2(x), \dots, f^{n-1}(x))$ for some $x \in V$ (i.e. x is a cyclic vector as defined in Exercise 6). Show that this is equivalent to each of the following conditions:

- $f^n = 0$ but $f^{n-1} \neq 0$;

- $r(f) = n - 1$;
- $\dim \operatorname{Ker} f = 1$;
- there is no decomposition $V = V_1 \oplus V_2$ where V_1 and V_2 are non-trivial subspaces which are f-invariant.

14) Show that if A is an invertible $n \times n$ matrix, then

$$\chi_{A^{-1}}(\lambda) = (-1)^n \frac{\lambda^n}{\det A} \chi_A(\frac{1}{\lambda}).$$

Show that in general (i.e. without the condition on invertibility of A) we have

$$\operatorname{tr}(\lambda I - A)^{-1} = \frac{\chi'_A(\lambda)}{\chi_A(\lambda)}$$

whenever λ is not an eigenvalue of A.
15) Show that if A is a nilpotent $n \times n$ matrix with $A^k = 0$, then $I - A$ is invertible and

$$(I - A)^{-1} = I + A + \cdots + A^{k-1}.$$

16) Let A be a complex 2×2 matrix which is not a multiple of the unit matrix. Show that any matrix which commutes with A can be written in the form $\lambda I + \mu A$ ($\lambda, \mu \in \mathbf{C}$).
17) Find a Jordan canonical form for the operator

$$D : \operatorname{Pol}(n) \to \operatorname{Pol}(n).$$

18) Show that every $n \times n$ matrix over $\mathbf{C}$ can be represented as a sum $D + N$ where$D$ is a diagonalisable matrix, N is nilpotent and N and D commute. Show that there is only one such representation.
19) A linear mapping $f : V \to V$ has an upper triangular representation (i.e. a basis with respect to which its matrix is upper triangular) if and only if there is a sequence (V_i) of subspaces of V $(i = 0, \dots, n)$ where the dimensional if V_i is i and $V_i \subset V_{i+1}$ for each i so that $f(V_i) \subset V_i$ for each i. Show directly (i.e. without using the Jordan canonical form) that such a sequence exists (for operators on complex vector spaces) and use this to give a proof of the Cayley-Hamilton theorem independent of the existence of a Jordan form. Show that if f and V are real, then such a representation exists if and only if χ_f has n real zeroes.
20) Show that if A is an $n \times n$ matrix with eigenvalues $\lambda_1, \dots, \lambda_n$ then the eigenvalues of $p(A)$ are $p(\lambda_1), \dots, p(\lambda_n)$.
21) Let J be a Jordan block of the form $J_n(\lambda)$. Calculate

- the set of matrices B which commute with A;
- the set of matrices which commute with all matrices which commute with A.

Deduce that a matrix is in the latter set if and only if it has the form $p(A)$ for some polynomial p.
22) Use 21) to show that a matrix B commutes with all matrices which commute with a given matrix A if and only if $B = p(A)$ for some polynomial p (cf. Exercise 7) above).
23) Show that if A_1 and A_2 are commuting $n \times n$ matrices, then their eigenvalues can be ordered as

$$\lambda_1, \dots, \lambda_n \quad \text{resp.} \quad \mu_1, \dots, \mu_n$$

in such a way that for any polynomial p of two variables, the eigenvalues of $p(A_1, A_2)$ are

$$p(\lambda_1, \mu_1), \dots, p(\lambda_n, \mu_n).$$

Generalise to commuting r-tuples $A_1, \dots, A_r$ of matrices.
24) Show that if p and q are polynomials and A is the companion matrix of p, then the nullity of $q(A)$ is the number of common roots of p and q (counted with multiplicities). In the case where $q = p'$, the rank of $p'(A)$ is the number of distinct roots of A.
25) Consider the companion matrix

$$C = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & & & & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ -a_0 & -a_1 & -a_2 & \dots & -a_0 \end{bmatrix}$$

of the polynomial p (cf. a previous exercise) and suppose now that p has repeated roots

$$\lambda_1, \dots, \lambda_1, \lambda_2, \dots, \lambda_i, \dots, \lambda_r, \dots, \lambda_r$$

where λ_i occurs n_i times. Show that C has Jordan form

$$\operatorname{diag}(J_{n_1}(\lambda_1), J_{n_2}(\lambda_2), \dots, J_{n_r}(\lambda_r))$$

and that this is induced by the following generalised Vandermonde matrix:

$$\begin{bmatrix} 1 & 0 & \dots & 1 & \dots & 0 \\ \lambda_1 & 1 & \dots & \lambda_2 & \dots & 0 \\ \lambda_1^2 & 2\lambda_1 & \dots & \lambda_2^2 & \dots & 0 \\ \vdots & & & & & \vdots \\ \lambda_1^{n-1} & (n-1)\lambda_1^{n-2} & \dots & \lambda_2^{n-1} & \dots & \lambda_r^{n-n_r} \end{bmatrix}.$$

(The first n_1 columns are obtained by successive differentiation of the first one and so on).

3.4 Functions of matrices and operators

We have often used the fact that we can substitute square matrices into polynomials. For many applications, it is desirable to be able to do this for more general functions and we discuss briefly some of the possibilities. Suppose firstly that A is a diagonal matrix, say

$$A = \operatorname{diag}(\lambda_1, \dots, \lambda_n).$$

Then if we recall that for a polynomial p,

$$p(A) = \operatorname{diag}(p(\lambda_1), \dots, p(\lambda_n))$$

it is natural to define $x(A)$ to be $\operatorname{diag}(x(\lambda_1), \dots, x(\lambda_n))$ for a suitable function x. In order for this to make sense, it suffices that the domain of definition of x contain the set $\{\lambda_1, \dots, \lambda_n\}$ of eigenvalues of A. In view of the importance of the latter set in what follows, we denote it by $\sigma(A)$. It is called the **spectrum** of A. Of course, substitution satisfies the rules that $(x+y)(A) = x(A)+y(A)$ and $(xy)(A) = x(A)y(A)$.

If A is diagonalisable, say

$$S^{-1}AS = D = \operatorname{diag}(a_{,1}, \dots, \lambda_n),$$

we define $x(A)$ to be

$$S \cdot x(D) \cdot S^{-1} = S \cdot \operatorname{diag}(x(\lambda_1), \dots, x(\lambda_n)) \cdot S^{-1}.$$

The case where A is not diagonalisable turns out to be rather more tricky. Firstly, we note that it suffices to be able to define $x(A)$ for Jordan blocks. For if A has Jordan form

$$S^{-1}AS = \operatorname{diag}(J_1, \dots, J_r)$$

then we can define $x(A)$ to be

$$S \cdot \operatorname{diag}(x(J_1), \dots, x_(J_r)) \cdot S^{-1}$$

once we know how to define the $x(J_i)$. (We are using the notation $\operatorname{diag}(J_1, \dots, J_r)$ for the representation of a Jordan form as a blocked diagonal matrix).

In order to motivate the general definition, consider the case of the square root of the Jordan matrix $J_n(\lambda)$. Firstly, we remark that for $\lambda = 0$, no such square root exists. We show this for the simplest case ($n = 2$) but the same argument works in general.

Example: Show that there is no matrix A so that

$$A^2 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}.$$

Solution: Suppose that such a matrix exists. Then $A^4 = 0$ and so A is nilpotent. But a 2×2 nilpotent matrix must satisfy the equation $A^2 = 0$ (since its characteristic polynomial is λ^2) which is patently absurd. We now turn to the general case and show that if λ is non-zero, then $J_n(\lambda)$ *does* have a square root. In fact, if $\lambda^{1/2}$ denotes one of the square roots of λ, then the matrix

$$A = \lambda^{1/2} \begin{bmatrix} 1 & \binom{\frac{1}{2}}{1}\frac{1}{\lambda} & \binom{\frac{1}{2}}{\frac{1}{\lambda^2}} & \dots & \binom{\frac{1}{2}}{n-1}\frac{1}{\lambda^{n-1}} \\ 0 & 1 & \binom{\frac{1}{2}}{1}\frac{1}{\lambda} & \dots & \binom{\frac{1}{2}}{n-2}\frac{1}{\lambda^{n-2}} \\ \vdots & & & & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{bmatrix}$$

(where for any real number α and $n \in \mathbf{N}$, $\binom{\alpha}{n}$ is the binomial coefficient

$$\frac{\alpha(\alpha-1)\dots(\alpha-n+1)}{n!})$$

satisfies the equation $A^2 = J_n(\lambda)$.

If this matrix seems rather mysterious, notice that the difference between the cases $\lambda = 0$ and $\lambda \neq 0$ lies in the fact that the complex function $z \mapsto z^{1/2}$ is analytic in the neighbourhood of a non-zero λ (i.e. is expressible as a power series in a neighbourhood of λ) whereas this is not the case at 0. In fact, we calculated the above root by writing

$$J_n(\lambda) = \lambda(I + \frac{1}{\lambda}N)$$

where N is the nilpotent matrix $J_n(0)$. We then wrote

$$J_n(\lambda)^{1/2} = \lambda^{1/2}(I + \frac{1}{\lambda}N)^{1/2}$$

and substituted $\frac{1}{\lambda}N$ for z in the Taylor series

$$(1+z)^{1/2} = \sum_{i=0}^{\infty} \binom{\frac{1}{2}}{i} z^i.$$

Of course, the resulting infinite series terminates after n-terms since $N^n = 0$.

If we apply the same method to the Taylor series

$$(1+z)^{-1} = 1 - z + z^2 - z^3 + \dots$$

then we can calculate the inverse of $J_n(\lambda)$ for $\lambda \neq 0$. The reader can check that the result coincides with that given above.

This suggest the following method for defining $x(A)$ where, for the sake of simplicity, we shall assume that x is entire. This will ensure that x has a Taylor expansion around each λ in the spectrum of A. As noted above, we use the Jordan form

$$S^{-1}AS = \operatorname{diag}(J_1, \dots, J_r)$$

where J_i is $J_{n_i}(\lambda_i)$. We define $x(J_i)$ as follows. x has the Taylor expansion

$$x(\lambda) = x(\lambda_i) + x'(\lambda_i)(\lambda - \lambda_i) + \dots$$

around λ_i. If we substitute formally we get the expression

$$x(J_i) = x(\lambda_i)I + x'(\lambda_i)N + \dots + \frac{x^{(n-1)}(\lambda_i)}{(n-1)!}N^{n-1}$$

where $J_i = \lambda_i I + N$ (i.e. $N = J_{n_i}(0)$).

We now give a purely algebraic form of the definition. Suppose that A is an $n \times n$ matrix with distinct eigenvalues $\lambda_1, \dots, \lambda_r$ and minimal polynomial

$$m(\lambda) = \prod (\lambda - \lambda_i)^{m_1}.$$

We shall define $x(A)$ for functions x which are defined on a neighbourhood of the set of eigenvalues of A and have derivatives up to order $m_i - 1$ at λ_i for each i. It is clear that there is a polynomial p so that

$$p^{(k)}(\lambda_i) = x^{(k)}(\lambda_i)$$

for each i and $k \leq m_i - 1$.

Perhaps the easiest way to do this is as follows. Firstly we construct a polynomial L_{ik} (of minimal degree) which vanishes, together with its derivatives up to order $m_j - 1$ at λ_j ($j \neq i$) and is such that the k-th derivative at λ_i is one, while all other ones vanish there (up to order $m_i - 1$). We denote this polynomial by L_{ik} (it can easily be written down explicitly but as we shall not require this directly we leave its computation as an exercise for the reader).

Then the polynomial which we require is

$$p = \sum_{i,k} x_{ik} L_{ik}$$

where $x_{ik} = x^{(k)}(\lambda_i)$. Hence if we write P_{ik} for the operator $L_{ik}(A)$, then

$$x(A) = \sum_{i,k} x_{ik} P_{ik}.$$

The P_{ik} are called the **components of** A.

In the special case where A has n distinct eigenvalues, then χ_A has the form

$$(\lambda - \lambda_1) \dots (\lambda - \lambda_n).$$

The only relevant L's are the L_{i0}'s which we denote simply by L_i. Thus L_i is the Lagrange interpolating polynomial

$$\prod_{j \neq i} \frac{\lambda - \lambda_j}{\lambda_k - \lambda_j}$$

which takes on the value 1 at λ_i and the value 0 at the other λ's. We note also the fact that the sum of the L_i's is the constant function one and that $L_i^2 = L_i$ (both of these when the functions are evaluated at the eigenvalues). In this case, the components $P_i = L_i(A)$ satisfy the equations $P_i^2 = P_i$ (i.e. they are projections) and their sum is the identity operator. The most important example of such a function of a matrix is the exponential function. Since the latter is entire, we can substitute *any* matrix and we denote the result by $\exp(A)$ or e^A. We note some of its simple properties:

- if A is diagonalisable, say $A = SDS^{-1}$ where $D = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$, then
$$\exp A = S \cdot \operatorname{diag}(e^{\lambda_1}, \dots, e^{\lambda_n}) \cdot S^{-1};$$
- if $A = D + N$ where D is diagonalisable and N is nilpotent, both commuting, then $\exp A = \exp D \cdot \exp N$ and
$$\exp N = \sum_{k=0}^{\infty} \frac{N^k}{k!}$$
where the series breaks off after finitely many terms;
- $\exp A \cdot \exp(-A) = I$ (and so $\exp A$ is always invertible);
- if $(\lambda_1, \dots, \lambda_n)$ are the eigenvalues of A, then $(e^{\lambda_1}, \dots, e^{\lambda_n})$ are the eigenvalues of $\exp A$;
- if A and B commute, then so do $\exp A$ and $\exp B$ and we have the formula
$$\exp(A + B) = \exp A \cdot \exp B.$$
In particular, if $s, t \in \mathbf{R}$, then $\exp(s + t)A = \exp sA \cdot \exp tB$;

- the function $t \mapsto \exp tA$ from $\mathbf{R}$ into the set of $n \times n$ matrices is differentiable and
$$\frac{d}{dt}(\exp tA) = A \cdot \exp(tA).$$

We remark that the statement in (6) means that the elements of the matrix $\exp tA$, as functions of t, are differentiable. The derivative on the left hand side is then the matrix obtained by differentiating its elements.

This property is particularly important since it means that the general solution of the system
$$\frac{dX}{dt} = AX$$
of differential equations where X is the column matrix
$$X(t) = \begin{bmatrix} x_1(t) \\ \vdots \\ xi_n(t) \end{bmatrix}$$
whose elements are smooth functions is given by the formula
$$X(t) = \exp tA \cdot X_0$$
where X_0 is the column matrix
$$\begin{bmatrix} x_1(0) \\ \vdots \\ x_n(0) \end{bmatrix}$$
of initial conditions.

Example: Consider the general linear ordinary differential equation of degree n
$$x^{(n)} + a_{n-1}x^{(n-1)} + \cdots + a_0 x = 0$$
with constant coefficients. We can write this equation in the form
$$\frac{dX}{dt} = AX$$
where X is the column vector
$$\begin{bmatrix} x(t) \\ x'(t) \\ \vdots \\ x^{(n-1)}(t) \end{bmatrix}$$

and A is the companion matrix

$$A = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & & & & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ -a_0 & -a_1 & -a_2 & \dots & -a_{n-1} \end{bmatrix}$$

of the polynomial

$$p(t) = t^n + a_{n-1}t^{n-1} + \dots + a_0.$$

As we know, the characteristic polynomial of this matrix is p and so its eigenvalues are the roots $\lambda_1, \dots, \lambda_n$ of p.

We suppose that these λ_i are all distinct. Then A is diagonalisable and the diagonalising matrix is the Vandermonde matrix

$$V(\lambda_1, \dots, \lambda_n) = \begin{bmatrix} 1 & \dots & 1 \\ \lambda_1 & \dots & \lambda_n \\ \vdots & & \vdots \\ \lambda_1^{n-1} & \dots & \lambda_n^{n-1} \end{bmatrix}.$$

Hence if we write S for this matrix, we have

$$\exp At = S \cdot \operatorname{diag}(e^{\lambda_1 t}, \dots, e^{\lambda_n t}) \cdot S^{-1}$$

and the solution of the above equation can be read off from the formula

$$X(t) = S \cdot (e^{\lambda_1 t}, \dots, e^{\lambda_n t}) \cdot S^1 X_0$$

where X_0 is the column matrix of initial values

$$\begin{bmatrix} x(0) \\ x'(0) \\ \vdots \\ x^{(n-1)}(0) \end{bmatrix}.$$

The reader can check that this provides the classical solution.

In principle, the case of repeated roots for p can be treated similarly. Instead of the above diagonalisation, we use the reduction of A to its Jordan form. The details form one of the main topics in most elementary books on differential equations.

Exercises: 1) Calculate $\exp A$ where

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 2 & 0 & 2 \\ 0 & 1 & 0 \end{bmatrix} \quad A = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \quad A = \begin{bmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{bmatrix} \quad A = \begin{bmatrix} 0 & t \\ -t & 0 \end{bmatrix}.$$

2) Solve the following system of differential equations:

$$\begin{aligned} \frac{dx_1}{dt} &= x_1 - 12x_3 + e^{-3t} \\ \frac{dx_2}{dt} &= -x_1 + 7x_2 - 20x_3 \\ \frac{dx_3}{dt} &= x_1 + 5x_3 + \cos t. \end{aligned}$$

3) If

$$C = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & & & & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ 1 & 0 & 0 & \dots & 0 \end{bmatrix}.$$

calculate $\exp tC$.

4) Calculate $\exp A$ where A is the matrix of the differentiation operator D on $\mathrm{Pol}\,(n)$.

5) Show that if A is an $n \times n$ matrix all of whose entries are positive, then the same holds for $\exp A$.

6) Show that $\det(\exp A) = \exp(\operatorname{tr} A)$.

7) Show that the general solution of the equation

$$\frac{dX}{dt} = AX + B$$

with initial condition $X(0) = X_0$, where A is a constant $n \times n$ matrix and B is a continuous mapping from $\mathbf{R}$ into the space of $n \times 1$ column matrices, is given by the equation

$$X(t) = \int_0^t \exp((t-s)A)B(s)\,ds + \exp tA \cdot X_0.$$

8) Show that if A is an $n \times n$ matrix with distinct eigenvalues $\lambda_1, \dots, \lambda_r$ and minimal polynomial

$$m(\lambda) = \prod (\lambda - \lambda_i)^{m_i},$$

then two polynomials p and q agree on A (i.e. are such that $p(A) = q(A)$) if and only if for each i

$$p(\lambda_i) = q(\lambda_i) \quad p'(\lambda_i) = q'(\lambda_i) \quad \ldots \quad p^{m_i-1}(\lambda_i) = q^{m_i-1}(\lambda_i).$$

9) Let A and B be $n \times n$ matrices and define a matrix function $X(t)$ as follows:

$$X(t) = e^{At}CB^{Bt}.$$

Show that X is a solution of the differential equation

$$\frac{dX}{dt} = AX + XB$$

with initial condition $X(0) = C$.

Deduce that if the integral

$$Y = -\int_0^\infty e^{At}Ce^{Bt}\,dt$$

converges for given matrices A, B and C, then it is a solution of the equation

$$AY + YB = C.$$

10) Let A and B be $n \times n$ matrices which commute and suppose that $A^2 = A$. Show that $M(s) = Ae^{Bs}$ is a solution of the functional equation $M(s+t) = M(s)M(t)$. Show that if, on the other hand, M is a smooth function which satisfies this equation, then $M(s)$ has the form Ae^{Bs} where $A = M(0)$ and $B = M'(0)$.

11) For which matrices A are the following expressions defined:

$$\sin A, \cos A, \tan A, \sinh A, \cosh A, \tanh A, \ln(I + A)?$$

Show how to use the methods of this section to solve a differential equation of the form

$$\frac{d^2X}{dt^2} = AX$$

with initial values $X(0) = X_0$, $X'(0) = Y_0$. 12) Show that if a function x is such that $x(A)$ is defined, then so is $x(A^t)$ and we have

$$x(A^t) = x(A)^t.$$

13) Show that if A and B are commuting $n \times n$ matrices, then

$$x(B) = x(A) + x'(A)\cdot(B - A) + \cdots + x^{(m)}(A)\cdot\frac{(B - A)^m}{m!}$$

for suitable functions x and integers m.

3.5 Circulants and geometry

An $n \times n$ matrix A is a **circulant** if it has the form

$$\begin{bmatrix} a_0 & a_1 & \dots & a_{n-1} \\ a_{n-1} & a_0 & \dots & a_{n-2} \\ \vdots & & & \vdots \\ a_1 & a_2 & \dots & a_) \end{bmatrix}.$$

Note that this just means that A is a polynomial function of the special circulant

$$C = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & & & & \vdots \\ 1 & 0 & 0 & \dots & 0 \end{bmatrix}.$$

In fact, A is then $p(C)$ where

$$p(t) = a_0 + a_1 t + \dots + a_{n-1} t^{n-1}.$$

The following result gives an alternative characterisation of circulant matrices. It can be verified by direct calculation.

Proposition 19 *An $n \times n$ matrix is circulant if and only if it commutes with C.*

We have already calculated the eigenvalues of C and found them to be

$$1, \omega, \omega^2, \dots, \omega^{n-1}$$

where ω is the primitive root $\cos \frac{2\pi}{n} + i \sin \frac{2\pi}{n}$. The eigenvector corresponding to ω^k is easily seen to be $u_k = \frac{1}{\sqrt{n}}(\omega^k, \omega^{2k}, \dots, \omega^{nk})$. (The reason for the factor $\frac{1}{\sqrt{n}}$ will become apparent later). These eigenvectors are particularly interesting since they are also eigenvectors for all polynomial functions of C i.e. for the circulant matrices.

Here we shall discuss briefly the circulants of the form

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

where there are m "ones" in each row. In other words,

$$A = p(C) \quad \text{where} \quad p(t) = \frac{1}{m}(1 + t + \dots + t^{m-1}).$$

It follows from the above that the eigenvalues of A are $\lambda_1, \dots, \lambda_n$ where $\lambda_n = 1$ and

$$\lambda_k = \frac{1}{m}\frac{1-\omega^{km}}{1-\omega^k}$$

for $k \neq n$.

Then we see immediately

- that A is invertible if and only if m and n are relatively prime;
- that if d is the greatest common divisor of m and n, then the dimension of the kernel of f_A is d and it has a basis consisting of the vectors

$$\frac{1}{n}(\omega^{jd}, \omega^{2jd}, \dots, 1, \omega^{jd}, \dots)$$

for $1 \leq j \leq d-1$.

These results have the following geometrical interpretation. Suppose that $\mathbf{P}$ is an n-gon in $\mathbf{R}^2$. If we identify the points of $\mathbf{R}^2$ with complex numbers, we can specify $\mathbf{P}$ by an n-tuple $(z_1, \dots, z_n)$ of complex numbers (its vertices). For example, the standard square corresponds to the 4-tuple $(o, 1, 1+i, i)$. An $n \times n$ matrix A can be regarded as acting on such polynomials by left multiplication of the corresponding column matrix i.e. we define the polygon $\mathbf{Q} = A(\mathbf{P})$ to be the one with vertices $\zeta_1, \dots, \zeta_n$ where

$$\begin{bmatrix} \zeta_1 \\ \vdots \\ \zeta_n \end{bmatrix} = A \begin{bmatrix} z_1 \\ \vdots \\ z_n \end{bmatrix}.$$

Consider the transformation with matrix

$$A = \frac{1}{m}\begin{bmatrix} 1 & 1 & \dots & 1 & 0 & \dots & 1 \\ 0 & 1 & \dots & 1 & 1 & \dots & 0 \\ \vdots & & & & & & \vdots \\ 1 & 1 & \dots & 1 & 0 & \dots & 1 \end{bmatrix}$$

discussed above. In this case, $\mathbf{Q}$ is the polygon whose vertices are the centroids of the vertices $P_1, \dots, P_m$ resp. $P_2, \dots, P_{m+1}$ and so on. This polygon is called the m-**descendant** of $\mathbf{P}$.

The results on the matrix A that we obtained above can now be expressed geometrically as follows:

If m and n are relatively prime, then every polygon $\mathbf{Q}$ is the m-descendant of a unique polygon $\mathbf{P}$.

A more delicate investigation of the case where the greatest common factor d of m and n is greater than 1 leads to a characterisation of those polygons $\mathbf{P}$ which are m descendants (see the Exercises below).

Exercises: 1) Show that the determinant of the circulant matrix circ $(a_0, \dots, a_{n-1})$ is

$$\prod_{j=0}^{n-1} \sum_{k=0}^{n-1} \omega^{ik} a_k$$

where ω is the primitive n-th root of unity.
2) A matrix A is called r**-circulant** if $CA = AC^r$ where C is the circulant matrix of the text. Characterise such matrices directly and show that if A is r-circulant and B s-circulant, then AB is rs-circulant.
3) Describe those matrices which are polynomial functions of the matrix $J_n(0)$.
4) Suppose that the greatest common factor d of m and n is greater than 1. Identify the kernel and the range of the matrix

$$A = \frac{1}{m} \begin{bmatrix} 1 & 1 & \dots & 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 1 & 1 & \dots & 0 \\ \vdots & & & & & & \vdots \\ 1 & 1 & \dots & 1 & 0 & \dots & 1 \end{bmatrix}$$

and use this to give a characterisation of those polygons which are m-descendants resp. whose m-descendants are the trivial polygon with all vertices at the origin.
5) Diagonalise the following circulant matrices:

- circ $(a, a+h, \dots, a+(n-1)h)$;
- circ $(a, ah, \dots, ah^{n-1})$;
- circ $(0, 1, 0, 1, \dots, 0, 1)$.

6) Show that if A is a circulant matrix and x is a suitable function, then $x(A)$ is also a circulant.

3.6 The group inverse and the Drazin inverse

As a further application of the Jordan form we shall construct two special types of generalised inverse for linear mappings (resp. matrices). These are of some importance in certain applications. The method will be the same in both cases. Firstly we construct the inverse for matrices in Jordan form and then use this to deal with the general case. We begin with the group inverse. A **group inverse** for an $n \times n$ matrix A is an $n \times n$ matrix S so that

$$ASA = A \qquad SAS = S$$

and A and S commute. As a simple example, suppose that A is the diagonal matrix

$$\operatorname{diag}(\lambda_1, \dots, \lambda_r, 0, \dots, 0).$$

Then $S = \operatorname{diag}(\frac{1}{\lambda_1}, \dots, \frac{1}{\lambda_r}, 0, \dots, 0)$ is obviously a group inverse for A. Hence every diagonalisable matrix has a group inverse. (If $\tilde{A} = P^{-1}AP$ is diagonal and $\tilde{S}$ is a group inverse for $\tilde{A}$ as above, then $S = P\tilde{S}P^{-1}$ is a group inverse for A).

More generally, note that if the vector space has the splitting

$$V = f(V) \oplus \operatorname{Ker} f$$

then we can use this to define a generalised inverse $g = P \circ \tilde{f}^{-1}$ where P is the projection onto $f(V)$ along $\operatorname{Ker} f$ and $\tilde{f}$ is the isomorphism induced by f on $f(V)$. In this case, g is easily seen to be a group inverse for f. Note that from the analysis of section 3 of the present chapter, f has such a splitting in the case where $r(f) = r(f^2)$. In terms of the Jordan form this means that the block corresponding to the zero eigenvalues is the zero block i.e. the Jordan form is

$$\operatorname{diag}(J_1(\lambda_1), \dots, J_r(\lambda_r), 0, \dots, 0)$$

where $\lambda_1, \dots, \lambda_r$ are the non-zero eigenvalues. The group inverse is then

$$\operatorname{diag}(J_1(\lambda_1)^{-1}, \dots, J_r(\lambda_r)^{-1}, 0, \dots, 0).$$

In terms of the minimal polynomial, this means that the matrix has a group inverse provided the latter has the form

$$m_A(\lambda) = \lambda^{\epsilon}(\lambda_1 - \lambda)^{n_1} \dots (\lambda_r - \lambda)^{n_r}$$

where ϵ is either 0 or 1.

The Drazin inverse A Drazin inverse for an $n \times n$ matrix A is an $n \times n$ matrix S so that

- $SAS = S$;
- S and A commute;
- $\lambda \in \sigma(A)$ if and only if $\lambda^\dagger \in \sigma(S)$ where $\lambda^\dagger = \frac{1}{\lambda}$ for $\lambda \neq 0$ and $\lambda^\dagger = 0$ for $\lambda = 0$;
- $A^{k+1}S = A^k$ for some positive integer k.

Suppose first that A is a Jordan block $J_n(\lambda)$. Then we define S to be $J_n(\lambda)^{-1}$ if λ is non-zero and to be 0 if $\lambda = 0$. Then S satisfies the above four condition (with $k = n$ if $\lambda = 0$). We denote this S by A^D. Now if A has Jordan form

$$\mathrm{diag}\,(A_1, A_2, \dots, A_r)$$

where each A_i is a Jordan block, we define A^D to be the matrix

$$\mathrm{diag}\,(A_1^D, \dots, A_r^D).$$

Once again, this satisfies the above four condition (this time with the integer k the maximum of the orders (k_i) of the individual blocks corresponding to the zero eigenvalue).

For a general matrix A we choose an invertible P so that $A = P^{-1}\tilde{A}P$ where $\tilde{A}$ has Jordan form. Then we define A^D to be $P^{-1}\tilde{A}^D P$. Of course, A^D is a Drazin inverse for A.

In terms of operators, the Drazin inverse can be described as follows: suppose that $f : V \to V$ is a linear transformation. Then, as we have seen, there is a smallest integer p at which the sequences

$$V \supset f(V) \supset f^2(V) \supset \dots$$

and

$$\mathrm{Ker}\, f \subset \mathrm{Ker}\,(f^2) \subset \mathrm{Ker}(f^3) \subset \dots$$

become stationary. This integer is called the **index** of f and we have the splitting

$$V = f^p(V) \oplus \mathrm{Ker}\,(f^p).$$

f, restricted to $f^p(V)$, is an isomorphism of this space onto itself. The operator f^D is that one which is obtained by composing the inverse of the latter with the projection onto $f^p(V)$ along $\mathrm{Ker}\,(f^p)$.

Exercises: 1) For $f \in L(V)$ we define $A(f) = f(V) \cap \operatorname{Ker} f$. Show that

- $\dim A(f) = r(f) - r(f^2)$;
- $\operatorname{Ker} f \subset f(V)$ if and only if $r(f) = r(f^2) - n(f)$ where $n(f) = \dim \operatorname{Ker} f$;
- $f(V) = \operatorname{Ker} f$ if and only if $f^2 = 0$ and $n(f) = r(f)$;
- $V = f(V) \oplus \operatorname{Ker} f$ if and only if $lr(f) = r(f^2)$.

2) Show that the Drazin inverse is uniquely determined i.e. that there is at most one matrix S so that $SAS = S$, $AS = SA$ and $A^{k+1}S = A^k$ for some positive integer k.
3) Show that the group inverse is uniquely determined i.e. there is at most one matrix S so that $ASA = A$, $SAS = S$ and $AS = SA$.
4) Show that if $f \in L(V)$ has a Drazin inverse f^D, then $f^D(V) = f(V)$, $\operatorname{Ker} f^D = \operatorname{Ker}(f^k)$ and $ff^D = f^D f$ is the projection onto $f(V)$ along $\operatorname{Ker} f^k$.
5) The following exercise shows how the existence of the Drazin inverse can be deduced directly from the existence of the minimal polynomial, without using the Jordan form. Suppose that m_A has the form

$$t \mapsto a_k t^k + \cdots + t^r$$

where t_k is the lowest power of t which occurs with non-vanishing coefficient (we can suppose that $k \geq 1$). The equation $m_A(A) = 0$ can then be written in the form

$$A^k = A^{k+1}B$$

where B is a suitable polynomial in A. Show that $A^{k+r}B^r = A^k$ for each $r \geq 1$ and deduce that $S = A^k B^{k+1}$ satisfies the defining conditions for the Drazin inverse. 6) Show that if A is an $n \times n$ matrix, then there is a polynomial p so that $A^D = p(A)$.

4 EUCLIDEAN AND HERMITIAN SPACES

4.1 Euclidean space

In chapter II we saw that a number of basic geometrical concepts could be defined in terms of the scalar product. We now discuss such products in higher dimensions where, in the spirit of chapter III, we use the axiomatic approach. We shall prove higher dimensional versions of many of the results of chapter II, culminating in the spectral theorem for self-adjoint operators.

Definition: A **scalar product (or inner product)** on a real vector space V is a mapping from $V \times V \to \mathbf{R}$, denoted as

$$(x, y) \mapsto (x|y)$$

so that

- the mapping is bilinear i.e.

  $$(\lambda_1 x_1 + \lambda_2 x_2 | \mu_1 y_1 + \mu_2 y_2) = \lambda_1\mu_1(x_1|y_1) + \lambda_1\mu_2(x_1|y_2) + \lambda_2\mu_1(x_2|y_1) + \lambda_2\mu_2(x_2|y_2)$$

  for $\lambda_1, \lambda_2, \mu_1, \mu_2 \in \mathbf{R}$, $x_1, x_2, y_1, y_2 \in V$;

- it is **symmetric** i.e. $(x|y) = (y|x)$ $(x, y \in V)$;
- it is **positive definite** i.e. $(x|x) > 0$ if $x \neq 0$.

Property (3) implies the following further property:

(4) if $x \in V$ is such that $(x|y) = 0$ for each $y \in V$, then $x = 0$. (Choose $y = x$ in (3)).

A **euclidean vector space** is a (real) vector space V together with a scalar product.

Of course, a vector space can be provided with many distinct scalar products. This can be visualised in $\mathbf{R}^2$ where a scalar product is given by a bilinear form of the type

$$(x, y) \mapsto \sum_{i,j}^{2} a_{ij}\xi_i\eta_j$$

and the positive definiteness means exactly that the corresponding conic section

$$\sum_{i,j}^{2} a_{ij}\xi_i\xi_j = 1$$

is an ellipse. Hence the choice of a scalar product in $\mathbf{R}^2$ is just the choice of an ellipse with centre 0.

The **standard euclidean space** is $\mathbf{R}^n$ with the scalar product

$$(x|y) = \sum_i \xi_i \eta_i.$$

Another example is the space $\operatorname{Pol}(n)$ with the scalar product

$$(p|q) = \int_0^1 p(t)q(t)\,dt.$$

The latter is a subspace of the infinite dimensional space $C([0,1])$ with scalar product

$$(x|y) = \int_0^1 x(t)y(t)\,dt.$$

Using the scalar product we can define the **length** (or **norm**) of a vector x—written $\|x\|$. It is defined by the formula

$$\|x\| = \sqrt{(x|x)}.$$

We then define the **distance** between two points x, y in V to be $\|x - y\|$, the length of their difference. As in $\mathbf{R}^2$ and $\mathbf{R}^3$ there is a close connection between the concepts of norm and scalar product. Of course, the norm is *defined* in terms of the latter. On the other hand, the scalar product can be expressed in terms of the norm as follows:

$$\|x+y\|^2 = (x+y|x+y) = (x|x) + 2(x|y) + (y|y)$$

and so

$$(x|y) = \frac{1}{2}(\|x+y\|^2 - \|x\|^2 - \|y\|^2).$$

As in $\mathbf{R}^2$ the norm and scalar product satisfy the **Cauchy-Schwartz inequality**:

$$|(x|y)| \leq \|x\|\|y\| \quad (x, y \in V).$$

This is named after the discoverers of the classical case

$$\xi_1\eta_1 + \cdots + \xi_n\eta_n \leq (\xi_1^2 + \cdots + \xi_n)^{\frac{1}{2}}(\eta_1^2 + \cdots + \eta_n^2)^{\frac{1}{2}}.$$

To prove the general inequality, we consider the quadratic function

$$T \mapsto (x+ty|x+ty) = t^2\|y\|^2 + 2t(x|y) + \|x\|^2$$

which is non-negative by the positive-definiteness of the product. Hence its discriminant is less than or equal to zero i.e. $4(x|y)^2 - 4(\|x\|^2\|y\|^2) \leq 0$ which reduces to the required inequality. (Note that the same proof shows that the Cauchy-Schwarz inequality is strict i.e. $|(x|y)| < \|x\|\|y\|$ unless x and y are proportional. For the above quadratic is strictly positive if x and y are not proportional and then the discriminant must be negative).

From this we can deduce the **triangle inequality**

$$\|x + y\| \leq \|x\| + \|y\|.$$

For

$$\begin{aligned}\|x + y\|^2 &= (x + y|x + y) \\ &= \|x\|^2 + 2(x|y) + \|y\|^2 \\ &\leq \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 \\ &= (\|x\| + \|y\|)^2.\end{aligned}$$

From the Cauchy-Schwarz inequality we see that if x and y are non-zero then the quotient

$$\frac{(x|y)}{\|x\|\|y\|}$$

lies between -1 and 1. Hence there is a unique $\theta \in [0, \pi]$ so that

$$\cos\theta = \frac{(x|y)}{\|x\|\|y\|}.$$

θ is called the **angle** between x and y.

We can also define the concept of orthogonality or perpendicularity: two vectors are **perpendicular** (written $x \perp y$ if $(x|y) = 0$.

As in $\mathbf{R}^2$ and $\mathbf{R}^3$ we use this notion to define a special class of bases:

Definition: An **orthogonal basis** in V is a sequence $(x_1, \dots, x_n)$ of non-zero vectors so that $(x_i|x_j) = 0$ for $i \neq j$. If, in addition, each vector has unit length, then the system is **orthonormal**.

An orthogonal system is automatically linearly independent. For if

$$\lambda_1 x_1 + \dots \lambda_m x_m = 0,$$

then

$$0 = (\lambda_1 \xi_1 + \dots + \lambda_m x_m | x_i) = \lambda_i (x_i|x_i)$$

and so $\lambda_i = 0$ for each i.

Hence an orthonormal system $(x_1, \dots, x_n)$ with n elements in an n-dimensional space is a basis and such bases are called **orthonormal bases**. The classical example is the canonical basis $(e_1, \dots, e_n)$ for $\mathbf{R}^n$.

One advantage of an orthonormal basis is the fact that the coefficient of a vector with respect to the basis can be calculated simply by taking scalar products. In fact $x = \sum_{k=1}^n (x|x_k)x_k$. (This is sometimes called the Fourier series of x). This is proved by a calculation similar to the one above. Also if $x = \sum_{i+1}^n \lambda_i x_i$ and $y = \sum_{k=1}^n \mu_k x_m$, then

$$(x|y) = (\sum_{i=1}^n \lambda_i x_i | \sum_{k=1}^n \mu_k x_k) = \sum_{i,k=1}^n \lambda_i \mu_k (x_i|x_k) = \sum_{i=1}^n \lambda_i \mu_i$$

and, in particular, $\|x\| = \sqrt{\lambda_1^2 + \cdots + \lambda_n^2}$. Thus the scalar product and norm can be calculated from the coordinates with respect to an orthonormal basis exactly as we calculate them in $\mathbf{R}^n$.

Every euclidean space has an orthonormal basis and to prove this we use a construction which has a natural geometrical background and which we have already used in dimensions 2 and 3.

Proposition 20 *If V is an n-dimensional euclidean space, then V has an orthonormal basis $(x_1, \dots, x_n)$.*

PROOF. We construct the basis recursively. For x_1 we take any unit vector. If we have constructed $x_1, \dots, x_r$ we construct x_{r+1} as follows. We take any z which does not lie in the span of $x_1, \dots, x_r$ (of course, if there is no such element we have already constructed a basis). Then define

$$\tilde{x}_{r+1} = z - \sum_{i=1}^r (z|x_i)x_i.$$

$\tilde{x}_{r+1}$ is non-zero and is perpendicular to each x_i $(1 \leq i \leq r)$. Hence if we put

$$x_{r+1} = \frac{\tilde{x}_{r+1}}{\|\tilde{x}_{r+1}\|}$$

then $(x_r, \dots, x_{r+1})$ is an orthonormal system.

The proof actually yields the following result:

Proposition 21 *Every orthonormal system $(x_1, \dots, x_r)$ in V can be extended to an orthonormal basis i.e. there exist elements $x_{r+1}, \dots, x_n$ so that $(x_1, \dots, x_n)$ is an orthonormal basis for V.*

The standard way to construct an orthonormal basis is called the **Gram-Schmid** process and consists in applying the above method in connection with a given basis $(y_1, \dots, y_n)$, using y_1 for x_1 and, at the r-th step using y_{r+1} for z. This produces an orthonormal basis $x_1, \dots, x_n$ of the form

$$\begin{aligned} x_1 &= b_{11}y_1 \\ x_2 &= b_{21}y_1 + b_{22}y_2 \\ &\vdots \\ x_n &= b_{n1}y_1 + \dots + b_{nn}y_n \end{aligned}$$

where the diagonal elements b_{ii} are non-zero. If we apply this method to the case where the space is $\mathbf{R}^n$ (identified with the space of row vectors) and the basis $(y_1, \dots, y_n)$ consists of the rows of an invertible $n \times n$ matrix A, we obtain a lower triangular matrix

$$B = \begin{bmatrix} b_{11} & 0 & \dots & 0 \\ b_{21} & b_{22} & \dots & 0 \\ \vdots & & & \vdots \\ b_{n1} & b_{n2} & \dots & b_{nn} \end{bmatrix}$$

and a matrix Q whose rows form an orthonormal basis for $\mathbf{R}^n$ (such matrices are called **orthonormal**) so that $Q = BA$. Since B is invertible and its inverse L is also a lower triangular matrix, we obtain the following result on matrices:

Proposition 22 *Any $n \times n$ invertible matrix A has a representation of the form $A = LQ$ where Q is an orthonormal matrix and L is lower triangular.*

We can use this fact to prove a famous inequality for the determinant of the $n \times n$ matrix. We have

$$|\det A| \leq \prod_i (\sum_j |a_{ij}|^2)^{\frac{1}{2}}$$

(i.e. the determinant is bounded by the product of the euclidean norms of the columns. This is known as **Hadamard's inequality**).

Proof. We can write the above equation as $L = AQ^t$. Hence

$$l_{ii} = \sum_j a_{ij}q_{ij} \leq (\sum_j a_{ij}^2)^{\frac{1}{2}} (\sum_j q_{ij}^2)^{\frac{1}{2}}$$

the last step using the Cauchy-Schwarz inequality. Hence (since $\sum_j q_{ij}^2 = 1$),

$$|\det A| = |\det LU| = |\det L| = \prod_i |l_{ii}| \leq \prod_i \left(\sum_j a_{ij}^2\right)^{1/2}$$

(We have used the fact that if the matrix U is orthonormal, then its determinant is ± 1. This follows as in the 2-dimensional case from the equation $U^t U = I$—see below).

In the context of euclidean space, those linear mapping which preserve distance (i.e. are such that $\|f(x)\| = \|x\|$ for $x \in V$) are of particular interest. As in the two and three dimensional cases, we can make the following simple remarks (note that we only consider **linear** isometries):

I. If f is an isometry, then f preserves scalar products i.e. $(f(x)|f(y)) = (x|y)$ $(x, y \in V)$. For

$$\begin{aligned}(f(x)|f(y)) &= \frac{1}{2}(\|f(x+y)\|^2 - \|f(x)\|^2 - \|f(y)\|^2) \\ &= \frac{1}{2}(\|x+y\|^2 - \|x\|^2 - \|y\|^2) \\ &= (x|y).\end{aligned}$$

On the other hand, this property implies that f is an isometry. (Take $x = y$).

II. An isometry from V into V_1 is automatically injective and so is surjective if and only if $\dim V = \dim V_1$. In particular, any isometry of V into itself is a bijection.

III. An isometry maps orthonormal systems onto orthonormal systems. In particular, if $\dim V = \dim V_1$, then f maps orthonormal bases onto orthonormal bases. On the other hand, if f maps *one* orthonormal basis $(x_1, \dots, x_n)$ onto an orthonormal system $(y_1, \dots, y_n)$ in V_1, then f is an isometry. For if $x = \sum_k \lambda_k x_k$, then $f(x) = \sum_k \lambda_k y_k$ and so

$$\|f(x)\|^2 = \sum_k \lambda_k^2 = \|x\|^2.$$

IV. By this criterium, if A is an $n \times n$ matrix, then since f_A maps the canonical basis for $\mathbf{R}^n$ onto the columns of A, we see that f_A is an isometry if and only if the columns of A form an orthonormal basis. This can be conveniently expressed in the equation $A^t \cdot A = I$ (or equivalently, $A^t = A^{-1}$ or $A \cdot A^t = I$). Matrices with this property are called **orthonormal** as we noted above.

Typical isometries in $\mathbf{R}^n$ are **reflections** i.e. operators with matrices of the form

$$\begin{bmatrix} 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 1 & 0 & \dots & 0 & 0 \\ \vdots & & & & & \vdots \\ 0 & 0 & 0 & \dots & 1 & 0 \\ 0 & 0 & 0 & \dots & 0 & -1 \end{bmatrix}$$

with respect to some orthonormal basis and **rotations** i.e. those with matrices of the form

$$\begin{bmatrix} 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 1 & 0 & \dots & 0 & 0 \\ \vdots & & & & & \vdots \\ 0 & 0 & 0 & \dots & \cos\theta & -\sin\theta \\ 0 & 0 & 0 & \dots & \sin\theta & \cos\theta \end{bmatrix}$$

with respect to some orthonormal basis.

Example: Show that the mapping

$$(A|B) = \operatorname{tr}(AB^t)$$

is a scalar product on M_n.
Solution: The bilinearity and symmetry:

$$\begin{aligned} (A_1 + A_2|B) &= \operatorname{tr}((A_1 + A_2)B^t) \\ &= \operatorname{tr}(A_1B^t + A_2B^t) \\ &= \operatorname{tr}(A_1B^t) + \operatorname{tr}(A_2B^t) \\ &= (A_1|B) + (A_2|B). \end{aligned}$$

$$\begin{aligned} (A|B) &= \operatorname{tr}(AB^t) \\ &= \operatorname{tr}(AB^t)^t \\ &= \operatorname{tr}(BA^t) \\ &= (B|A). \end{aligned}$$

Positive definiteness:

$$(A|A) = \operatorname{tr}(AA^t) = \sum_{i,j} a_{ij}^2 > 0$$

for $A \neq 0$.

We remark that the basis $(e_{ij} : i, j = 1, \dots, n)$ is then orthonormal where e_{ij} is the matrix with a 1 in the (i, j)-th position and zeroes elsewhere.

Example: Show that the mapping

$$(x|y) = 4\xi_1\eta_1 - 2\xi_1\eta_2 - 2\xi_2\eta_1 + 3\xi_2\eta_2$$

is a scalar product on $\mathbf{R}^2$.
Solution: The matrix of the quadratic form is

$$\begin{bmatrix} 4 & -2 \\ -2 & 3 \end{bmatrix}$$

with eigenvalues the roots of $\lambda^2 - 7\lambda + 8$ which are positive.

We calculate an orthonormal basis with respect to this scalar product by applying the Gram-Schmidt process to $x_1 = (1, 0)$, $x_2 = (0, 1)$. This gives

$$\begin{aligned} e_1 &= \frac{x_1}{\|x_1\|} = (\frac{1}{2}, 0); \\ y_2 &= (0, 1) - ((0, 1)|(\frac{1}{2}, 0))(\frac{1}{2}, 0) = \frac{1}{2}, 1); \\ e_2 &= \frac{1}{\sqrt{2}}(2, 1). \end{aligned}$$

Example: Construct an orthonormal basis for Pol (2) with scalar product

$$(p|q) = \int_0^1 p(t)q(t)\, dt$$

by applying the Gram-Schmidt method to $(1, t, t^2)$.
Solution:

$$\begin{aligned} \tilde{x}_0 &= 1, x_0 = 1; \\ \tilde{x}_1(t) &= t - (t|1)1 = t = -\frac{1}{2}; \\ x_1(t) &= \sqrt{12}(t - \frac{1}{2}); \\ \tilde{x}_2(t) &= t^2 - (t^2|x_0)x_0(t) - (t^2|x_1)x_1(t) \\ &= t^2 - t + \frac{1}{6}; \\ x_2(t) &= 6\sqrt{5}(t^2 - t + \frac{1}{6}). \end{aligned}$$

We calculate the Fourier series of $1 + t + t^2$ with respect to this basis. We have

$$\begin{aligned} \lambda_1 &= \int_0^1 (t^2 + t + 1)\, dt = \frac{11}{6} \\ \lambda_2 &= \sqrt{12} \int_0^1 (t^2 + t + 1)(t - \frac{1}{2})\, dt = \frac{1}{\sqrt{3}} \\ \lambda_3 &= 6\sqrt{5} \int_0^1 (t^2 + t + 1)(t^2 - t + \frac{1}{6})\, dt = \frac{1}{6\sqrt{5}} \end{aligned}$$

and so

$$t^2+t+1=\frac{11}{6}+2(t-\frac{1}{2})+(t^2-t+\frac{1}{6}).$$

Example If $(e_1,\dots,e_n)$ is an orthonormal basis for V and $x_1,\dots,x_n\in V$, then

$$\det\begin{bmatrix}(x_1|e_1) & \dots & (x_1|e_n)\\ \vdots & & \vdots\\ (x_n|e_1) & \dots & (x_n|e_n)\end{bmatrix}^2=\det\begin{bmatrix}(x_1|x_1) & \dots & (x_1|x_n)\\ \vdots & & \vdots\\ (x_n|x_1) & \dots & (x_n|x_n)\end{bmatrix}.$$

(And hence the right hand side is non-negative and vanishes if and only if the x_i are linearly dependent).
Solution: This follows from the equalities

$$(x_i|x_j)=\sum_{k=1}^{n}(x_i|e_k)(x_j|e_k)$$

which can be written in the matrix form

$$\begin{bmatrix}(x_1|e_1) & \dots & (x_1|e_n)\\ \vdots & & \vdots\\ (x_n|e_1) & \dots & (x_n|e_n)\end{bmatrix}\begin{bmatrix}(x_1|e_1) & \dots & (x_1|e_n)\\ \vdots & & \vdots\\ (x_n|e_1) & \dots & (x_n|e_n)\end{bmatrix}=\begin{bmatrix}(x_1|x_1) & \dots & (x_1|x_n)\\ \vdots & & \vdots\\ (x_n|x_1) & \dots & (x_n|x_n)\end{bmatrix}.$$

Exercises: 1)

- Apply the Gram-Schmidt process to obtain an orthonormal system from the following sets:

$$(1,0,1),(2,1,-1),(-1,1,0)\quad\text{in}\quad\mathbf{R}^3;$$
$$(1,2,1,2),(1,1,1,0),(2,1,0,1)\quad\text{in}\quad\mathbf{R}^4.$$

- Calculate the orthogonal projection of $(1,3,-2)$ on $[(2,1,3),(2,0,5)]$.

2) Show that if the a_i are positive, then

$$(\sum_{i=1}^{n}a_i)(\sum_{i+1}^{n}\frac{1}{a_i})\geq n^2$$
$$(\sum_{i=1}^{n}a_i)^2\leq n(\sum_{i=1}^{n}a_i^2).$$

3) If $x_1, \dots, x_m$ are points in $\mathbf{R}^n$, then the set of points which are equidistant from $x_1, \dots, x_m$ is an affine subspace which is of dimension $n - m$ if the x_i are affinely independent.
4) Let $(x_0, \dots, x_n)$ be affinely independent points in $\mathbf{R}^n$. Then there is exactly one hypersphere through these points and its equation is

$$\det \begin{bmatrix} 1 & 1 & \dots & 1 \\ X_0 & X_1 & \dots & X \\ \vdots & & & \vdots \\ (x_0|x_0) & (x_1|x_1) & \dots & (x|x) \end{bmatrix} = 0$$

where X is the column vector corresponding to x etc.
5) Suppose that $a_1, \dots, a_n, b_1, \dots, b_n$ are positive numbers. Show that either

$$\frac{a_1}{b_1} + \dots + \frac{a_n}{b_n} \geq n$$

or

$$\frac{b_1}{a_1} + \dots + \frac{b_n}{a_n} \geq n.$$

6) Show that a sequence $(x_1, \dots, x_n)$ in a euclidean space V is an orthogonal basis if an only if for each $x \in V$,

$$\|x\|^2 = \sum_{i=1}^{n} |(x|x_i)|^2.$$

7) Calculate an orthonormal basis of polynomials for $\mathrm{Pol}\,(3)$ by applying the Gram-Schmidt process to the basis $(1, t, t^2, t^3)$ with respect to the scalar product

$$(p|q) = \int_0^\infty p(t)q(t)e^{-t}\,dt.$$

8) Let p_n be the polynomial

$$\frac{d^n}{dt^n}(t^2 - 1)^n.$$

Show that

$$\int_{-1}^{1} t^k p_n(t)\,dt = 0 \quad (k < n)$$

and hence that the system (p_n) is orthogonal for the corresponding scalar product. (This shows that the system (p_n) is, up to norming factors, the sequence obtained by applying the Gram-Schmidt process to the sequence (t^n). These functions are called the **Legendre polynomials).**

9) Approximate $\sin x$ by a polynomial of degree 3 using the orthonormal basis of the last exercise. Use this to check the accuracy by calculating an approximate value of $\sin 1$. 10) Show that for an $n \times n$ matrix A the following inequality holds:

$$\sum_{i,j=1}^{n} a_{ij}^2 \geq \frac{1}{n}(\operatorname{tr} A)^2.$$

For which matrices do we have equality?
11) Show that if (ξ_i) and (η_i) are vectors in $\mathbf{R}^n$, then

$$(\sum_i \xi_i^2)(\sum_i \eta_i^2) = (\sum_i \xi_i \eta_i)^2 - \sum_{1 \leq i < j \leq n} (\xi_i \eta_j - \xi_j \eta_i)^2.$$

(This is a quantitative version of the Cauchy-Schwarz equations).
12) Consider the euclidean space M_n with the scalar product discussed in the example abovje. If A is an $n \times n$ matrix one can apply the Gram-Schmidt process to the sequence $I, A, A_2, A^3, \dots$. This will stop at the smallest k so that A^k is a linear combination of the earlier. This provides a specific description of the minimal polynomial of A. The reader is invited to fill in the details.
13) Let $(x_1, \dots, x_k)$ be a linearly independent set in the euclidean space V. Show that if $x \in V$, then its orthogonal projection x_0 onto the linear span of the above vectors is given by the formula

$$x_0 = \sum_{r+1}^{k} \lambda_r x_r$$

where $\lambda_j = \dfrac{\det G_j}{\det G}$. Here G is the matrix $[g_{ij}]$ where $g_{ij} = (x_i | x_j)$ and G_j is the matrix obtained from G by replacing the j-th column by ??
14) Show that if $x_1, \dots, x_n$ is a basis for the euclidean space V, then the Gram-Schmidt process, applied to this basis, leads to the system (y_k) where

$$y_k = (d_{k-1} d_k)^{\frac{1}{2}} D$$

where $d_n = \det G$ (G as in 13)) and

$$D_n = \det \begin{bmatrix} (x_1|x_1) & \dots & (x_1|x_n) \\ \vdots & & \vdots \\ (x_{n-1}|x_1) & \dots & (x_{n-1}|x_n) \\ x_1 & \dots & x_n \end{bmatrix}.$$

(The last expression is to be understood as the linear combination of the x's obtained by formally expanding the "determinant" along the last row).
15) Use Hadamard's inequality to show that

$$|\det A|^2 \leq K^n \cdot n^{\frac{n}{2}}$$

where A is an $n \times n$ matrix and $K = \max_{i,j} |a_{ij}|$.
16) Show that

$$\min \int_0^\infty e^{-t}(1 + a_1 t + \dots a_n t^n)^2 \, dt = \frac{1}{n+1};$$
$$\min \int_0^1 (1 + a_1 t + \dots + a_n t^n)^2 \, dt = \frac{1}{(n+1)^2}$$

where the minima are each taken over the possible choices of the coefficients $a_1, \dots, a_n$.

4.2 Orthogonal decompositions

In our discussion of vector spaces we saw that each subspace has a complementary subspace which determines a splitting of the original space. In general this complementary subspace is not unique but, as we shall now see, the structure of a euclidean space allows us to choose a unique one in a natural way.

Definition: If V_1 is a subspace of the euclidean space V, then the set

$$V_1^\perp = \{x \in V : (x|y) = 0 \quad \text{for each} \quad y \in V_1\}$$

of vectors which are orthogonal to each vector in V_1 forms a subspace of V called the **orthogonal complement** of V_1.

Proposition 23 *V_1 and $V_1^\perp$ are complementary subspaces. More precisely, if $(x_1, \dots, x_n)$ is an orthonormal basis for V so that $V_1 = [x_1, \dots, x_r]$, then $V_1^\perp = [x_{r+1}, \dots, x_n]$. Hence $V = V_1 \oplus V_1^\perp$.*

PROOF. It suffices to check that $y = \sum_{i+1}^n \lambda_i x_i$ is orthogonal to each x_i $(i = 1, \dots, r)$ and so to V_1 if and only if $\lambda - i = 0$ for $i = 1, \dots, r$ i.e. $y \in [x_{r+1}, \dots, x_n]$.

We shall write $V = V \perp V_1$ to denote the fact that $V = V_1 \oplus V_2$ and $V_1 \perp V_2$.

If $V = V_1 \perp V_2$ and $(x_1, \dots, x_n)$ is an orthonormal basis with $V_1 = [x_1, \dots, x_r]$, then

$$P_{V_1}(x) = \sum_{r=1}^{r} (x|x_i)x_i$$

is the projection of x onto V_1 along $V_1^\perp$. It is called the **orthogonal projection** of x onto V_1 and has the following important geometric property: $P_{V_1}(x)$ is the nearest point to x in V_1 i.e.

$$\|x - P_{V_1}(x)\| < \|x - z\| \quad (z \in V_1, z \neq P_{V_1}(x)).$$

For if $z \in V_1$, we have

$$\|x - z\|^2 = \|x - P_{V_1}(x)\|^2 + \|P_{V_1}(x) - z\|^2$$

since $x - P_{V_1}(x) \perp P_{V_1}(z)$.

The point $x - 2P_{V_1}(x)$ is then the **mirror image** (or **reflection**) of x in $V_1^\perp$. The linear mapping $\mathrm{Id} - 2P_{V_1}$ is called a **reflection**.

In the same way we define an orthogonal decomposition

$$V = V_1 \perp \cdots \perp V_r.$$

This means that $V = V_1 \oplus \cdots \oplus V_r$ and $V_i \perp V_j$ (i.e. $x \perp y$ if $x \in V_i$, $y \in V_j$) when $i \neq j$.

Then if P_i is the orthogonal projection onto V_i,

- $P_1 + \cdots + P_r = \mathrm{Id}$;
- $P_i P_j = 0 \quad (i \neq j)$.

Such a sequence of projections is called a **partition of unity**.

Orthogonal splittings of the space are obtained by partitioning an orthonormal basis $(x_1, \dots, x_n)$ i.e. we define V_i to be $[x_{k_{i-1}}, \dots, x_{k_i - 1}]$ where $1 = k_0 < k_1 < \cdots < k_r = n + 1$.

On the other hand, if we have such a splitting $V_1 \perp \cdots \perp V_r$, then we can construct an orthonormal basis for V by combining orthonormal bases for the components $V_1, \dots, V_n$.

Exercise: 1) Consider the space M_n with the scalar product

$$(A|B) = \mathrm{tr}\,(B^t A).$$

Show that M_n is the orthogonal direct sum of the symmetric and the anti-symmetric matrices.

4.3 Self-ajdoint mappings— the spectral theorem

One of the most important consequences of the existence of a scalar product is that the fact that it induces a certain symmetry on the linear operators on the space. If $f : V \to V_1$ we say that $g : V_1 \to V$ is **ajdoint** to f if $(f(x)|y) = (x|g(y))$ for $x \in V$ and $y \in V_1$. We shall presently see that such a g always exists. Furthermore it is unique. The general construction is illustrated by the following two examples:
I. Consider the mapping $f : \mathbf{R}^2 \to \mathbf{R}^2$ defined by the matrix

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{22} & a_{22} \end{bmatrix}.$$

A simple calculation shows that

$$(f(x)|y) = a_{11}\xi_1\eta_1 + a_{12}\xi_2\eta_1 + a_{21}\xi_1\eta_2 + a_{22}\xi_2\eta_2.$$

If g is the mapping with matrix A^t, then analogously

$$(x|g(y)) = a_{11}\eta_1\xi_1 + a_{21}\eta_2\xi_1 + a_{12}\eta_1\xi_2 + a_{22}\eta_2\xi_2$$

and both are equal.
II. Now consider the mapping $f : \mathbf{R}^n \to \mathbf{R}^m$ defined by the matrix $A = [a_{ij}]$. We know that a_{ij} is the i-th coordinate of $f(e_j)$ i.e. that $a_{ij} = (f(e_j)|e_i)$ (remember the method of calculating coordinates with respect to an orthonormal basis). Hence if a mapping $g : \mathbf{R}^m \to \mathbf{R}^n$ exists which is adjoint to f, then

$$a_{ij} = (f(e_j)|e_i) = (e_j|g(e_i))$$

and by the same reasoning the latter is the (j,i)-th entry of g. In other words the only possible choice for g is the mapping with matrix $B = A^t$. A simple calculation shows that this mapping does in fact satisfy the required conditions. Since these considerations are perfectly general, we state and prove them in the form of the following Proposition:

Proposition 24 *If $f : V \to V_1$ is a linear mapping there is exactly one linear mapping $g : V_1 \to V$ which is adjoint to f. If f has matrix A with respect to orthonormal bases (x_i) resp. (y_j) then g is the mapping with matrix A^t with respect to (y_j) and (x_i).*

PROOF. Suppose that $x \in V$ and $y \in V_1$. Then

$$\begin{aligned}
(f(x)|y)) &= (f(\sum_{j=1}^{n}(x|x_j)x_j)|\sum_{k=1}^{m}(y|y_k)y_k) \\
&= (\sum_{i-1}^{m}(\sum_{j=1}^{n} a_{ij}(x|x_j)y_i|\sum(y|y_k)y_k) \\
&= \sum_{i=1}^{m}\sum_{j=1}^{n}\sum_{k=1}^{m} a_{ij}(x|x_j)(y|y_k)(y_i|y_k) \\
&= \sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}(x|x_j)(y|y_i).
\end{aligned}$$

Similarly, we have

$$(x|g(y)) = \sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}(x|x_j)(y|y_i)$$

if g is the operator with matrix as in the formulation and so

$$(f(x)|y) = (x|g(y))$$

i.e. g is adjoint to f.
Naturally, we shall denote g by f^t and note the following simple properties:

- $(f+g)^t = f^t + g^t$;
- $(\lambda f)^t = \lambda f^t$;
- $(fg)^t = g^t f^t$;
- $(f^t)^t = f$;
- an operator f is an isometry if and only if $f^t f = \mathrm{Id}$. In particular, if $V = V_1$, this means that $f^t = f^{-1}$.

We prove the last statement. We have seen that f is an isometry if and only if $(f(x)|f(y)) = (x|y)$ for each x, y and this can be restated in the form

$$(f^t f(x)|y) = (x|y)$$

i.e. $f^t f = \mathrm{Id}$.

A mapping $f : V \to V$ is said to be **self-adjoint** if $f^t = f$ i.e.

$$(f(x)|y) = (x|f(y))$$

for $x, y \in V$. This is equivalent to f being represented by a **symmetric** A with respect to an orthonormal basis. In particular, this will be the case if f is represented by a diagonal matrix. The most important result of this chapter is the following which states that the converse is true:

Proposition 25 *Let $f : V \to V$ be self-adjoint. Then V possesses an orthonormal basis consisting of eigenvectors of f. With respect to this basis, f is represented by the diagonal matrix $diag(\lambda_1, \dots, \lambda_n)$ where the λ_i are the eigenvalues of f.*

In terms of matrices this can be stated in the following form:

Proposition 26 *Let A be a symmetric $n \times n$ matrix. Then there exists an orthonormal matrix U and real numbers $\lambda_1, \dots, \lambda_n$ so that*

$$U^t A U = U^{-1} A U = diag(\lambda_1, \dots, \lambda_n).$$

In particular, a symmetric matrix is diagonalisable and its characteristic polynomial has n real roots.

This follows immediately from the above Proposition, since the transfer matrix between orthonormal bases is orthonormal.

Before proceeding with the proof we recall that we have essentially proved this result in the two dimensional case in our treatment of conic sections in Chapter II.

One of the consequences of the result is the fact that every symmetric matrix A has at least *one* eigenvalue. Indeed this is the essential part of the proof as we shall see and before proving this we reconsider the two dimensional case. The symmetric matrix $\begin{bmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{bmatrix}$ induces the quadratic form

$$\phi_1(x) = (f_A(x)|x)$$

and so define the conic section

$$\mathbf{Q}_1 = \{x : \phi_1(x) = 1\}.$$

If we compare this to the form $\phi_2(x) = (x|x)$ which defines the unit circles $\mathbf{Q}_2 = \{x : \phi_2(x) = 1\}$ and consider the case where $\mathbf{Q}_1$ is an ellipse, then we see that the eigenvectors of A are just the major and minor axes of the ellipse and these can be characterised as those directions x for which the ratio $\dfrac{\phi_1(x)}{\phi_2(x)}$ is extremal. Hence we can reduce the search for eigenvalues to one for the extremal value of a suitable functional, a problem which we can solve with the help of elementary calculus. This simple idea will allow us to prove the following result which is the core of the proof. We use the Proposition that a continuous function on a closed, bounded subset of $\mathbf{R}^n$ is bounded and attains its supremum.

Proposition 27 *Lemma Let $f : V \to V$ be self-adjoint. Then there exists an $x \in V$ with $\|x\| = 1$ and $\lambda \in \mathbf{R}$ so that $f(x) = \lambda x$ (i.e. f has an eigenvalue).*

PROOF. We first consider the case where $V = \mathbf{R}^n$ with the natural scalar product. Then f is defined by an $n \times n$ symmetric matrix A. Consider the function

$$\phi : x \mapsto \frac{(f(x)|x)}{(x|x)}$$

on $V \setminus \{0\}$. Then $\phi(\lambda x) = \phi(x)$ for $\lambda \neq 0$. There exists an $x_1 \in V$ so that $\|x_1\| = 1$ and $\phi(x_1) \geq \phi(x)$ for $x \in V$ with $\|x\| = 1$. Hence, by the homogeneity,

$$\phi(x_1) \geq \phi(x) \quad (x \in V \setminus \{0\}).$$

We show that x_1 is an eigenvector with $f(x_1) = \lambda_1 x_1$ where $\lambda_1 = \phi(x_1) = (f(x_1)|x_1)$. To do this choose $y \neq 0$ in V. The function

$$\psi : t \mapsto \phi(x_1 + ty)$$

has a minimum for $t = 0$. We show that $\psi'(x)$ exists and its value is $2(z|f(x_1)) - 2(y|x_1)\lambda_1$. Hence this must vanish for each y i.e.

$$(y|f(x_1) - \lambda_1 x_1) = 0.$$

Since this holds for each y, $f(x_1) = \lambda_1 x_1$. To calculate the derivative of ψ we compute the limit as t tends to zero of the difference quotient

$$\frac{1}{t}(\psi(t) - \psi(0)) = \frac{1}{t}(\phi(x_1 + ty) - \phi(x_1)).$$

But this is the limit of the expression

$$\frac{1}{t}[\frac{(x_1|x_1)[(f(x_1)|x_1) + 2t(f(x_1)|y)) + t^2(f(y)|y)] - (f(x_1)|x_1)[(x_1|x_1) + 2t(x_1|y) + t^2(y|y)]}{(x_1 + ty|x_1 + ty)(x_1|x_1)}]$$

which is easily seen to be

$$2(y|f(x_1)) - 2(y|x_1)\lambda_1.$$

In order to prove the general case (i.e. where f is an operator on an abstract euclidean space), we consider an orthonormal basis $(x_1, \dots, x_n)$ for V and let A be the matrix of f with respect to this basis. Then A has an eigenvalue and of course this means that f also has one.

We can now continue to the proof of the main result:

Proof. The proof is an induction argument on the dimension of V. For $\dim V = 1$, the result is trivial (all 1×1 matrices are diagonal!) The step $n-1 \to n$: By the Lemma, there exists an eigenvector x_1 with $\|x_1\| = 1$ and eigenvalue λ_1. Put $V_1 = \{x_1\}^\perp = \{x \in V : (x|x_1) = 0\}$. Then V_1 is $(n-1)$-dimensional and $f(V_1) \subset V_1$ since if $x \in V_1$, then $(f(x)|x_1) = (x|f(x_1)) = (x|\lambda_1 x_1) = 0$ and so $f(x) \in V_1$. Hence by the induction hypothesis there exists an orthonormal basis $(x_2, \dots, x_n)$ consisting of eigenvectors for f. Then $(x_1, \dots, x_n)$ is the required orthonormal basis for V.

The above proof implies the following useful characterisation of the largest resp. smallest eigenvalue of f;

Corollar 3 *Let $f : V \to V$ be a self-adjoint linear mapping with eigenvalues $\lambda_1, \dots, \lambda_n$ numbered so that $\lambda_1 \leq \dots \leq \lambda_n$. Then*

$$\lambda_1 = \min\{\phi(x) : x \in V \setminus \{0\}\}$$
$$\lambda_n = \max\{\phi(x) : x \in V \setminus \{0\}\}.$$

This can be generalised to the following so-called minimax characterisation of the k-th eigenvalue:

$$\lambda_k = \min \max\{(f(x)|x) : x \in V \setminus \{0\}, (x|y_1) = \dots = (x|y_r) = 0\}$$

the minimum being taken over all finite sequences $y_1, \dots, y_r$ of unit vectors where $r = n - k$.

We remark here that it follows from the proofs of the above results that if f is such that $(f(x)|x) \geq 0$ for each x in V, then its eigenvalues are all non-negative (and they are positive if $(f(x)|x) > 0$ for non-zero x). Such f are called **positive semi-definite** resp. **positive definite** and will be examined in some detail below. It follows form the above minimax description of the eigenvalues that

$$\lambda_k(f + g) \geq \lambda_k(f)$$

whenever f is self-adjoint and g is positive semi-definite (with strict inequality when g is positive definite). As we have seen, a symmetric operator is always diagonalisable. Using this fact, we can prove the following weaker result for arbitrary linear operators between euclidean spaces.

Proposition 28 *Let $f : V \to V_1$ be a linear mapping. Then there exist orthonormal bases $(x_1, \dots, x_n)$ for V and $(y_1, \dots, y_m)$ for V_1 so that the matrix of f with respect to these bases has the form*

$$\begin{bmatrix} A_1 & 0 \\ 0 & 0 \end{bmatrix}$$

where $A_1 = diag(\mu_1, \dots, \mu_r)$, r is the rank of f and $\mu_1, \dots, \mu_r$ are positive scalars.

The corresponding result for matrices is the following:

Proposition 29 *If A is an $m \times n$ matrix then there exist orthonormal matrices U_1 and U_2 so that $U_1 A U_2$ has the above form.*

Proof. We prove the operator form of the result. Note that the mapping $f^t f$ on V is self-adjoint since $(f^t f)^t = f^t f^{tt} = f^t f$. We can thus choose an orthonormal basis (x_i) consisting of eigenvectors for $f^t f$. If λ_i is the corresponding eigenvalue we can number the x_i so that the first r eigenvalues are non-zero but the following ones all vanish. Then each λ_i is positive $(i = 1, \dots, r)$ and $\|f(x_i)\| = \sqrt{\lambda_i}$. For

$$\lambda_i = \lambda_i(x_i|x_i) = (\lambda_i x_i|x_i) = (f^t f(x_i)|x_i) = (f(x_i)|f(x_i)) > 0.$$

Also if $i \neq j$ then $f(x_i)$ and $f(x_j)$ are perpendicular since

$$(f(x_i)|f(x_j)) = (f^t f(x_i)|x_j) = \lambda_i(x_i|x_j) = 0.$$

Hence if we put

$$y_1 = \frac{f(x_1)}{\sqrt{\lambda_1}}, \dots, y_r = \frac{f(x_r)}{\sqrt{\lambda_r}}$$

then $(y_1, \dots, y_r)$ is an orthonormal system in V_1. We extend it to an orthonormal basis $(y_1, \dots, y_m)$ for V_1. The matrix of f with respect to these bases clearly has the desired form.

The proof shows that the μ_i are the square roots of the eigenvalues of $f^t f$ and so are uniquely determined by f. They are call the **singular values** of f.

Example: Let V be a euclidean space with a basis (x_i) which is not assumed to be orthonormal. Show that if $f \in L(V)$ has matrix A with respect to this basis, then f is self-adjoint if and only if $A^t G = GA$ where $G = [(x_i|x_j)]_{i,j}$.
Solution: f is self-adjoint if and only if

$$(f(\sum_{i=1}^n \lambda_i x_i | \sum_{j=1}^n \mu_j x_j))) = (\sum_{i=1}^n \lambda_i x_i | f(\sum_{j=1}^n \mu_j x_j))))$$

for all choices of scalars $\lambda_1, \dots, \lambda_n, \mu_1, \dots, \mu_n$ and this holds if and only if

$$(f(x_i)|x_j) = (x_i|f(x_j))$$

for each i and j. Substituting the values of $f(x_i)$ and $f(x_j)$ one gets the required equation $A^t G = GA$.

Exercises: 1) Calculate the adjoints of the following operators:

- $(\xi_1, \xi_2, \xi_3) \mapsto (\xi_1 + \xi_2 + \xi_3, \xi_1 + 2\xi_2 + 2\xi_3, \xi_1 - 2\xi_2 - 2\xi_3)$ (on $\mathbf{R}^3$ with the usual scalar product);
- the differentiation operator D on $\mathrm{Pol}\,(n)$;
- $p \mapsto (t \mapsto tp(t))$ from $\mathrm{Pol}\,(n)$ into $\mathrm{Pol}\,(n+1)$.

2) Show that if $f \in L(V)$, then

$$\mathrm{Ker}\, f^t = f(V)^\perp \quad f^t(V) = (\mathrm{Ker}\, f)^\perp.$$

3) Show that for every self-adjoint operator f there are orthogonal projections $P_1, \dots, P_r$ and real scalars $\lambda_1, \dots, \lambda_r$ so that

- $P_i P_j = 0$ if $i \neq j$;
- $P_1 + \cdots + P_r = \mathrm{Id}$;
- $f = \lambda_1 P_1 + \cdots + \lambda_r P_r$.

Calculate the P_i explicitly for the operator on $\mathbf{R}^3$ with matrix

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

4) Show that an operator f is an orthogonal projection if and only if $f^2 = f$ and f is self-adjoint. Let f be a self-adjoint operator so that $f^m = f^n$ for distinct positive integers m and n. Show that f is an orthogonal projection.
5) Show that if f is an invertible, self-adjoint operator, then λ is an eigenvalue of f if and only if $\frac{1}{\lambda}$ is an eigenvalue of f^{-1}. Show that f and f^{-1} have the same eigenvectors.

4.4 Conic sections

As mentioned above, the theory of conic sections in the plane was the main source of the ideas which lie behind the spectral theorem. We now indicate briefly how the latter can be used to give a complete classification of higher dimensional conic sections. The latter are defined as follows:

Definition: A **conic section** in a euclidean space V is a set of the form

$$\mathbf{Q} = \{x \in V : (f(x)|x) + 2(b|x) + c = 0\}$$

where $b \in V$, $c \in \mathbf{R}$ and f is a self-adjoint operator on V. In order to simplify the analysis we assume that the conic section is **central** i.e. has a point of central symmetry (for example, in $\mathbf{R}^2$ an ellipse is central whereas a parabola is not). If we then take this point to be the origin, it follows that $x \in \mathbf{Q}$ if and only if $-x \in \mathbf{Q}$. Analytically this means that $b = 0$ i.e. A has the form

$$\{x \in V; (f(x)|x) + c = 0\}.$$

Proposition 30 *If $\mathbf{Q}$ is a conic section of the above form, then there is an orthonormal basis $(x_1, \dots, x_n)$ for V and real numbers $\lambda_1, \dots, \lambda_n$ so that*

$$\mathbf{Q} = \{x = \xi_1 x_1 + \dots \xi_n x_n : \lambda_1 \xi_1^2 + \dots \lambda_n \xi_n^2 + c = 0\}.$$

This result can be restated as follows: let f be the isometry from V into $\mathbf{R}^n$ which maps $(x\cdot 1, \dots, x_n)$ onto the canonical basis $(e_1, \dots, e_n)$. Then

$$f(\mathbf{Q}) = \{(\xi_1, \dots, \xi_n) \in \mathbf{R}^n : \lambda_1 \xi_1^2 + \dots + \lambda_n \xi_n^2 + c = 0\}.$$

PROOF. Choose for $(x_1, \dots, x_n)$ an orthonormal basis for V so that $f(x_i) = \lambda_i x_i$. Then

$$(f(x)|x) = \lambda_1 \xi_1^2 + \dots + \lambda_n \xi_n^2$$

if $x = \xi_1 x_1 + \dots \xi_n x_n$.

As an application, we classify the central conics in $\mathbf{R}^3$. Up to an isometry they have the form

$$\{(\xi_1, \xi_2, \xi_3) : \lambda_1 \xi_1^2 + \lambda_2 \xi_2^2 + \lambda_3 \xi_3^2 + c = 0\}.$$

By distinguishing the various possibilities for the signs of the λ's, we obtain the following types: $\lambda_1, \lambda_2, \lambda_3$ all positive. Then we can reduce to the form

$$\frac{\xi_1^2}{a^2} + \frac{\xi_2^2}{b^2} + \frac{\xi_3^2}{c^2} + d = 0$$

and this is an ellipsoid for $d < 0$, a point for $d = 0$ and the empty set for $d > 0$. $\lambda_1, \lambda_2, \lambda_3$ all negative. This can be reduced to the first case by multiplying by -1. λ_1, λ_2 positive, λ_3 negative. Then we can write the equation in the form

$$\frac{\xi_1^2}{a^2} + \frac{\xi_2^2}{b^2} - \frac{\xi_3^2}{c^2} + d = 0.$$

(the cases $\lambda_1, \lambda_3 > 0, \lambda_2 < 0$ resp. $\lambda_2, \lambda_3 > 0, \lambda_1 < 0$ can be reduced to the above by permuting the unknowns). The above equation represents a circular cone ($d = 0$) or a one-sheeted or two sheeted hyperboloid (depending on the sign of d). The cases where at least one of the λ's vanishes can be reduced to the two-dimensional case.

Exercises: 1) Show that each non-central conic

$$\{x \in \mathbf{R}^3 : (f(x)|x) + 2(g|x) + c = 0\}$$

where $b \neq 0$ is isometric to one of the following:

$$\{x : \frac{\xi_1^2}{a^2} + \frac{\xi_2^2}{b^2} - 2\xi_3 + d = 0\}$$
$$\{x : \frac{\xi_1^2}{a^2} - \frac{\xi_2^2}{b^2} - 2\xi_3 + d = 0\}$$
$$\{x : \xi_1^2 - 2x_2 + d = 0\}$$

and discuss their geometrical forms.
2) If

$$\mathbf{Q} = \{x \in \mathbf{R}^3 : \frac{\xi_1^2}{a^2} + \frac{\xi_2^2}{b^2} + \frac{\xi_3^2}{c^2} = 1\}$$

is an ellipsoid with $a \leq b \leq c$ show that the number of planes through 0 which cut $\mathbf{Q}$ in a circle is 1 (if $a = b$ or $b = c$ but $a \neq c$), 2 (if a, b, c are distinct) or infinite (if $a = b = c$). 3) Diagonalise the quadratic form

$$\phi(x) = \sum_{i=1}^{n} (\xi_i - \xi_j)^2$$

on $\mathbf{R}^n$.

4.5 Hermitian spaces

For several reasons, it is useful to reconsider the theory developed in this chapter in the context of complex vector space. Amongst other advantages this will allow us to give a purely algebraic proof of the central result—the diagonalisation of symmetric matrices. This is because the existence of an eigenvalue in the complex case follows automatically from the fundamental theorem of algebra.

We begin by introducing the concept of a **hermitian vector space** i.e. a vector space V over $\mathbf{C}$ with a mapping

$$(\ |\ ) : V \times V \to \mathbf{C}$$

so that

- $(\lambda x + y|z) = \lambda(x|z) + (y|z)$ (linearity in the first variable);
- $(x|x) \geq 0$ and $(x|x) = 0$ if and only if $x = 0$;
- $(x|y) = \overline{(y|x)}$ $(x, y \in V)$.

Examples: All the examples of euclidean spaces can be "complexified" in the natural and obvious ways. Thus we have: a) the standard scalar product

$$((\lambda_1, \dots, \lambda_n))|(\mu_1, \dots, \mu_n)) = \lambda_1\overline{\mu_1} + \dots + \lambda_n\overline{\mu_n}$$

on $\mathbf{C}^n$;
b) the scalar product

$$(p|q) = \int_0^1 p(t)\overline{q(t)}\, dt$$

on the space $\mathrm{Pol}_{\mathbf{C}}(n)$ of polynomials with complex coefficients. Note the rather unexpected appearance of the complex conjugation in condition 2) above. This means that the scalar product is no longer bilinear and is used in order to ensure that condition 1) can hold. However, the product *is* **sesqui-linear** (from the classical Greek for one and a half) i.e. satisfies the condition

$$(\sum_{i=1}^n \lambda_i x_i | \sum_{j=1}^n \mu_j y_j) = \sum_{i,j=1}^n \lambda_i \overline{\mu_j}(x_i|y_j).$$

All of the concepts which we have introduced for euclidean spaces can be employed, with suitable changes usually necessitated by the sesquilinearity of the scalar product, for hermitian spaces. We shall review them briefly: The length of a vector x is $\|x\| = \sqrt{(x|x)}$. Once again we have the inequality

$$|(x|y)| \leq \|x\|\|y\| \quad (x, y \in V).$$

(Since there is a slight twist in the argument we give the proof. Firstly we have (for each $t \in \mathbf{R}$),

$$0 \leq (x + ty | x + ty) = \|x\|^2 + 2t\Re(x|y) + t^2\|y\|^2$$

and, as in the real case, this gives the inequality

$$\Re(x|y) \leq \|x\|\|y\|.$$

Now there is a complex number λ with $|\lambda| = 1$ and $\lambda(x|y) > 0$. If we apply the above inequality with x replaced by λx we get

$$|(x|y)| = |(\lambda x|y)| = \lambda(x|y) = \Re\lambda(x|y) = \Re(\lambda x|y) \leq \|\lambda x\|\|y\| = \|x\|\|y\|.$$

Using this inequality, one proves just as before that the distance function satisfies the triangle inequality i.e.

$$\|x + y\| \leq \|x\| + \|y\|.$$

For reasons that we hope will be obvious we do not attempt to define the angle between two vectors in hermitian space but the concept of orthogonality continues to play a central role. Thus we say that x and y are **perpendicular** if $(x|y) = 0$ (written $x \perp y$). Then we can define orthonormal systems and bases as before and the Gram-Schmidt method can be used to show that every hermitian space V has an orthonormal basis $(x_1, \dots, x_n)$ and the mapping

$$(\lambda_1, \lambda_2, \dots, \lambda_n) \mapsto \lambda_1 x_1 + \dots + \lambda_n x_n$$

is an isometry from $\mathbf{C}^n$ onto V i.e. we have the formulae

$$\begin{aligned}(x|y) &= \lambda_1\overline{\mu_1} + \dots + \lambda_n\overline{\mu_n} \\ \|x\|^2 &= |\lambda_1|^2 + \dots + |\lambda_n|^2\end{aligned}$$

if $x = \lambda_1 x_1 + \dots + \lambda_n x_n$, $y = \mu_1 x_1 + \dots \mu_n x_n$.

Exercises: 1) Let f be the linear mapping on the two dimensional space $\mathbf{C}^2$ with matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. Then the scalar product on $\mathbf{C}^2$ defines a norm there and so a norm for operators. Show that the norm of the above operator is given by the formula

$$\|f\|^2 = \frac{1}{2}(h^2 + \sqrt{[(h^2 - 4|\det A|^2]}).$$

(Calculate the singular values of A).
2) Show that if H is a hermitian matrix of the form $A+iB$ where A and B are real and A is non-singular, then we have the formula

$$(\det H)^2 = (\det A)^2 \det(I + A^{-1}BA^{-1}B).$$

3) Show that if U is a complex $n \times n$ matrix of the form $P+iQ$ where P and Q are real, then U is unitary (i.e. such that $U^*U = I$) if and only if P^tQ is symmetric and $P^tP + Q^tQ = I$.

4.6 The spectral theorem—complex version

If $f \in L(V, V_1)$ there is exactly one mapping $g : V_1 \to V$ so that

$$(f(x)|y) = (x|g(y)) \quad (x \in V, y \in V_1).$$

We denote this mapping by f^*. The proof is exactly the same as for the real case, except that we use the formula

$$a_{ij} = (f(x_j)|y_i) = (x_j|g(y_j)) = \overline{(g(y_i)|x_j)}$$

for the elements of the matrix A of f with respect to the orthonormal bases $(x_1, \dots, x_n)$ resp. $(y_1, \dots, y_m)$ to show that the matrix of f^* is A^*, the $n \times m$ matrix obtained from A by taking the complex conjugates of elements and then transposing.

The linear mapping f on V is **hermitian** if $f^* = f$ i.e. if $(f(x)|y) = (x|f(y))$ $(x, y \in V)$. This means that the matrix A of f with respect to an orthonormal basis satisfies the condition $A = A^*$ (i.e. $a_{ij} = \overline{a_{ji}}$ for each i, j). Such matrices are also called **hermitian**. $f : V \to V_1$ is **unitary** if $(f(x)|f(y)) = (x|y)$ $(x, y \in V)$. This is equivalent to the condition that $f^* f = \mathrm{Id}$. Hence the matrix U of f (with respect to orthonormal bases) must satisfy the condition $U^* U = I$ (i.e. the columns of U are an orthonormal system in $\mathbf{C}^n$). If $\dim V = \dim V_1$ $(= n$ say), then U is an $n \times n$ matrix and the above condition is equivalent to the equation $U^* = U^{-1}$. Such matrices are called **unitary**.

We now proceed to give a purely algebraic proof of the so-called spectral theorem for hermitian operators. We begin with some preliminary results on eigenvalues and eigenvectors:

Lemma 1 *If $f \in L(V)$ with eigenvalue λ, then*

- *λ is real if f is hermitian;*
- *$|\lambda| = 1$ if f is unitary.*

Proof. 1) if the non-zero element x is a corresponding eigenvector, then we have

$$(f(x)|x) = (\lambda x|x) = \lambda(x|x)$$

and

$$((f(x)|x) = (x|f(x)) = (x|\lambda x) = \bar{\lambda}(x|x)$$

and so $\lambda = \bar{\lambda}$. 2) Here we have

$$(x|x) = (f(x)|f(x)) = (\lambda x|\lambda x) = |\lambda|^2(x|x)$$

and so $\|\lambda\|^2 = 1$.

Proposition 31 *Lemma If λ_1, λ_2 are distinct eigenvalues of the hermitian mapping f with corresponding eigenvectors x_1, x_2, then $x_1 \perp x_2$.*

PROOF. We have $(f(x_1)|x_2) = (x_1|f(x_2))$ and so

$$(\lambda_1 x_1|x_2) = (x_1|\lambda_2 x_2) \quad \text{i.e.} \quad (\lambda_1 - \lambda_2)(x_1|x_2) = 0.$$

(Recall that $\lambda_2 = \bar{\lambda}_2$ by the Lemma above). Hence $(x_1|x_2) = 0$.

Proposition 32 *Lemma If $f \in L(V)$, then $\mathrm{Ker}\, f = \mathrm{Ker}\,(f^* f)$.*

PROOF. Clearly $\mathrm{Ker}\, f \subset \mathrm{Ker}\, f^* f$. Now if $x \in \mathrm{Ker}\, f^* f$, then $f^* f(x) = 0$ and so $(f^* f(x)|x) = 0$ i.e. $(f(x)|f(x)) = 0$. Thus $f(x) = 0$ i.e. $x \in \mathrm{Ker}\, f$.

In passing we note that this means that if A is an $m \times n$ matrix, then $r(A) = r(A^* A) = r(A^*)$.

Corollar 4 *If $f \in L(V)$ is hermitian and $r \in \mathbf{N}$, then $\mathrm{Ker}\, f = \mathrm{Ker}\, f^r$.*

PROOF. Applying the last result repeatedly, we get the chain of equalities

$$\mathrm{Ker}\, f = \mathrm{Ker}\, f^2 = \cdots = \mathrm{Ker}\, f^{2^k} = \ldots$$

Each $\mathrm{Ker}\, f^r$ lies between two terms of this series and so coincides with $\mathrm{Ker}\, f$.

We now come to our main result:

Proposition 33 *If $f \in L(V)$ is a hermitian mapping, then there exists an orthonormal basis (x_i) for V so that each x_i is an eigenvector for f. With respect to this basis, f has the matrix $\mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ where the λ_i are the (real) eigenvalues of f.*

PROOF. Let $V_i = \mathrm{Ker}\,(f - \lambda_i \mathrm{Id})$. It follows from the above corollary (applied to $f - \lambda_1 \mathrm{Id}$) that V is the direct sum

$$V = V_1 \oplus \mathrm{Im}(f - \lambda_1 \mathrm{Id}).$$

(Recall from Chapter VII that we have such a splitting exactly when the kernel of a mapping coincides with the kernel of its square). A simple induction argument shows that V is the direct sum

$$V = V_1 \oplus V_2 \oplus \cdots \oplus V_r.$$

However, we know, by the Lemma above that $V_i \perp V_j$ if $i \neq j$ and so in fact we have

$$V = V_1 \perp \cdots \perp V_r.$$

Hence we can construct the required basis by piecing together orthonormal bases for the various V_i.

Corollar 5 *If A is a hermitian $n \times n$ matrix, then there exists a unitary $n \times n$ matrix U and real numbers $\lambda_1, \ldots, \lambda_n$ so that*

$$U^{-1} A U = \mathit{diag}\,(\lambda_1, \ldots, \lambda_n).$$

A useful sharpening of the spectral theorem is as follows:

Proposition 34 *Let f and g be commuting hermitian operators on V. Then V has an orthonormal basis $(x_1, \ldots, x_n)$ whereby the x_i are simultaneously eigenvectors for f and g.*

PROOF. Let $\lambda_1, \ldots, \lambda_r$ be the distinct eigenvalues of f and put $V_i = \mathrm{Ker}\,(f - \lambda_i \mathrm{Id})$. Then $g(V_i) \subset V_i$. For if $x \in V_i$, then $f(x) = \lambda_i x$ and $f(g(x)) = g(f(x)) = g(\lambda_i x) = \lambda_i g(x)$ i.e. $g(x) \in V_i$. Hence if we apply the above result to g on V_i, we can find an orthonormal basis for the latter consisting of eigenvectors for g (regarded as an operator on V_i). Of course they are also eigenvectors for f and so the basis that we get for V by piecing together these sub-bases has the required properties.

For matrices, this means that if A and B are hermition $n \times n$ matrices, then there is a unitary matrix U so that both $U^* A U$ and $U * B U$ are diagonal (real) matrices.

Example: We diagonalise the matrix $A = \begin{bmatrix} 1+i & 1 \\ 1 & 1+i \end{bmatrix}$.
Solution: The characteristic polynomial is

$$(1+i)^2 - 2(1+i)\lambda + \lambda^2 - 1$$

with roots $2+i$ and i. The corresponding eigenvectors are $\frac{1}{\sqrt{2}}(1,-1)$ and $\frac{1}{\sqrt{2}}(1,1)$. hence $U^* A U = \mathrm{diag}\,(i, 2+i)$ where $U = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$.

Example: Show that if A is hermitian, then $I + iA$ is invertible and $U = (I - iA)(I + iA)^{-1}$ is unitary.
Solution: We can diagonalise A as

$$V^* A V = \mathrm{diag}\,(\lambda_1, \ldots, \lambda_n)$$

where the λ_i are real. Then

$$V^* A V = \mathrm{diag}\,(1 + i\lambda_1, \ldots, 1 + i\lambda_n)$$

and the right hand side is invertible. Hence so is the left hand side. This implies that $(I + iA)^{-1}$ exists. Then

$$V^*(I - iA)(I + iA)V = \operatorname{diag}\left(\frac{1 - i\lambda_1}{1 + i\lambda_1}, \dots, \frac{1 - i\lambda_n}{1 + i\lambda_n}\right).$$

The right hand side is clearly unitary. **Exercises:** 1) Show that an operator f on a hermitian space is an orthogonal projection if and only if $f = f^* f$. 2) Let A be a complex $n \times n$ matrix, p a polynomial. Show that if $p(A^*A) = 0$, then $p(AA^*) = 0$.
3) Let p be a complex polynomial in two variables. Show that if A is an $n \times n$ complex matrix so that $p(A, A^*) = 0$, then $p(\lambda, \bar{\lambda}) = 0$ for any eigenvalue λ of A. What can you deduce about the eigenvalues of a matrix A which satisfies one of the conditions:

$$A^* = cA \quad (c \in \mathbf{R}) \quad A^*A = A^* + A \quad A^*A = -I?$$

4) Let A be an $n \times n$ complex matrix with eigenvector X and eigenvalue λ_1. Show that there is a unitary matrix U so that $U^{-1}AU$ has the block form

$$\begin{bmatrix} \lambda_1 & \\ & B \\ 0 & \end{bmatrix}$$

for some $n \times (n-1)$ matrix B. Deduce that there is a unitary matrix $\tilde{U}$ so that $\tilde{U}^{-1}A\tilde{U}$ is upper triangular and that if A is hermitian, then the latter matrix is diagonal. (This exercise provides an alternative proof of the spectral theorem).

4.7 Normal operators

Normal operators are a generalisation of hermitian ones. Consider first the diagonal matrix

$$A = \operatorname{diag}(\lambda_1, \dots, \lambda_n).$$

A need not necessarily be hermitian (indeed this is the case precisely when the λ_i are real). However, it *does* satisfy the weaker condition that $AA^* = A^*A$ i.e. that A and A^* commute. We say that such A are **normal**. Similarly, an operator f on V is **normal** if $f^*f = ff^*$. Note that unitary mappings are examples of normal mappings—they are not usually hermitian. We shall now show that normal mappings have diagonal representations. In order to do this, we note that any $f \in L(V)$ has a unique representation in the form $f = g + ih$ where g and h are hermitian (compare the representation of a complex number z in the form $x + iy$ with x and y real). Indeed if $f = g+ih$, then $f^* = g-ih$ and so $f+f^* = 2g$ i.e. $g = \frac{1}{2}(f+f^*)$. Similarly, $h = \dfrac{1}{2i}(f - f^*)$. This proves the uniqueness. On the other hand, it is clear that if g and h are as in the above formula, then they are hermitian and $f = g + ih$.

The fact that normal operators are diagonalisable will follow easily from the following simple characterisation: f is normal if and only if g and h commute.

PROOF. Clearly, if g and h commute, then so do $f = g+ih$ and $f^* = g-ih$. On the other hand, if f and f^* commute then so do g and h since both are linear combinations of f and f^*.

Proposition 35 *If $f \in L(V)$ is normal, then V has an orthonormal basis $(x_1, \dots, x_n)$ consisting of eigenvectors of f.*

PROOF. We write $f = g + ih$ as above and find an orthonormal basis $(x_1, \dots, x_n)$ consisting of vectors which are simultaneously eigenvectors for g and h, say $g(x_i) = \lambda_i x_i$, $h(x_i) = \mu_i x_i$. Then

$$f(x_i) = g(x_i) + ih(x_i) = (\lambda_i + \mu_i)(x_i)$$

as claimed.

■

Corollar 6 *If $f \in L(V)$ is unitary, then there exists an orthonormal basis $(x_1, \dots, x_n)$ so that the matrix of f has the form diag$(e^{i\theta_1}, \dots, e^{i\theta_n})$ for suitable $\theta_1, \dots, \theta_n \in \mathbf{R}$.*

We close this section with the classification of the isometries of $\mathbf{R}^n$. This generalises the results of Chapter II on those of $\mathbf{R}^2$ and $\mathbf{R}^3$. The method we use is a standard one for deducing results about the real case from the complex one and can also be employed to deduce the spectral theorem for self-adjoint operators on euclidean spaces from the corresponding one for hermitian operators.

We require the following simple Lemma:

Lemma 2 *Let V be a subspace of $\mathbf{C}^n$ with the property that if $z \in V$, then $\Re z$ and $\Im z$ also belong to V (where if $z = (z_1, \dots, z_n)$, then*

$$\Re z = (\Re z_1, \dots, \Re z_n) \quad \Im z = (\Im z_1, \dots, \Im z_n)).$$

Then V has an orthonormal basis consisting of real vectors (i.e. vectors z with $\Im z = 0$.]

PROOF. We consider the set of all real vectors in V. Then it follows from the assumptions that this is a euclidean space whose real dimension coincides with the complex one of V. If $(x_1, \dots, x_r)$ is an orthonormal basis for this space, it is also one for V.

Proposition 36 *Let f be an isometry of the euclidean space V. Then there is an orthonormal basis $(x_1, \dots, x_n)$ for V with respect to which f has block matrix of the form*

$$\begin{bmatrix} I_r & 0 & 0 & \dots & 0 \\ 0 & -I_s & 0 & \dots & 0 \\ 0 & 0 & D_{\theta_1} & \dots & 0 \\ 0 & 0 & \vdots & & \vdots \\ 0 & 0 & 0 & \dots & D_{\theta_t} \end{bmatrix}$$

for suitable integers r, s, t with $r + s + 2t = n$ where D_θ denotes the matrix

$$\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}.$$

PROOF. Let A be the matrix of f with respect to an orthonormal basis. We regard A as a unitary operator on $\mathbf{C}^n$. Since its non-real eigenvalues occur in complex conjugate pairs there is a (complex) unitary matrix U so that

$$U^{-1}AU = \operatorname{diag}(1, 1, \dots, 1, -1, \dots, -1, e^{i\theta_1}, e^{-i\theta_1}, \dots, e^{i\theta_t}, e^{-i\theta_t}).$$

From this it follows that $\mathbf{C}^n$ splits into the direct sum

$$V_1 \oplus V_{-1} \oplus (W_1 \oplus W_1') \oplus \dots \oplus (W_t \oplus W_t')$$

where

$$\begin{aligned} V_1 &= \{x : f_A(x) = x\} \\ V_{-1} &= \{x : f_A(x) = -x\} \\ W_i &= \{x : f_A(x) = e^{i\theta_i} x\} \\ W_i' &= \{x : f_A(x) = e^{-i\theta_i} x\}. \end{aligned}$$

Now consider V_1. Then since A is real, $z \in V_1$ if and only if $\Re z \in V_1$ and $\Im z \in V_1$. Hence by the Lemma above we can choose an orthonormal basis $(x_1, \dots, x_r)$ for V_1 consisting of real vectors. Similarly, we can construct an orthonormal basis for V_{-1} consisting of real vectors $(y_1, \dots, y_s)$.

For the same reason, the mapping

$$z \mapsto \bar{z} = \Re z - \Im z$$

maps W_i onto W_i'. Hence if $(z_1, \dots, z_t)$ is an orthonormal basis for W_i, $(\bar{z}_1, \dots, \bar{z}_t)$ is an orthonormal basis for W_i'. Now f_A maps the two dimensional space spanned by z_i and $\bar{z}_i$ into itself and has matrix

$$\begin{bmatrix} e^{i\theta_i} & 0 \\ 0 & e^{-i\theta_i} \end{bmatrix}$$

with respect to this basis.

We introduce the real bases w_i and w_i' where

$$w_i = \frac{z_i + \bar{z}_i}{\sqrt{2}} \quad w_i' = \frac{z_i - \bar{z}_i}{i\sqrt{2}}$$

and consider the mapping f_A on the two dimensional *real* space spanned by these vectors. Then a simple calculation shows that w_i and w_i' are perpendicular (since $z_i \perp \bar{z}_i$) and that the matrix of f_A with respect to this basis is

$$\begin{bmatrix} \cos\theta_i & -\sin\theta_i \\ \sin\theta_i & \cos\theta_i \end{bmatrix}.$$

If we combine all the basis elements obtained in this way, we obtain an orthonormal basis of the required sort.

Exercises: 1) Show that if f is a linear operator on a hermitian space V, then either of the following conditions is equivalent to the fact that f is normal:

- $(f(x)|f(y)) = (f^*(x)|f^*(y)) \quad (x, y \in V)$;

- $\|f(x)\| = \|f^*(x)\| \quad (x \in V)$.

2) Show that an $m \times n$ complex matrix has factorisations

$$A = UB = CV$$

where U (resp. V) is unitary $m \times m$ (resp. $n \times n$) and B and C are positive semi-definite. Show that if A is normal (and so square), then U and B can be chosen in such a way that they commute.
3) Show that a complex $n \times n$ matrix A is normal if and only if there is a complex polynomial p so that $A^* = p(A)$.
4) Let A be a normal matrix and suppose that all of the row sums have a common value c. Show that the same is true of the column sums. What is the common value in this case?
5) Let A be a normal $n \times n$ matrix. Show that the set of eigenvalues of A is symmetric about zero (i.e. is such that if λ is an eigenvalue then so is $-\lambda$) if and only if $\operatorname{tr} A^{2k+1} = 0$ for $k = 0, 1, 2, \dots$.
6) Let U be an isometry of the hermitian space V and put

$$V_1 = \{x : Ux = x\}.$$

Show that if U_n is the mapping

$$U_n = \frac{1}{n}(I + U + \dots + U^{n-1})$$

then for each $x \in V$, $U_n x$ converges to the orthogonal projection of x onto V_1.
7) If $f \in L(V)$ is such that $f^* = -f$, then $\mathrm{Id} + f$ is invertible and $(\mathrm{Id} - f)(\mathrm{Id} + f)^{-1}$ is unitary. Which unitary operators can be expressed in this form?
8) Show that if an $n \times n$ complex matrix is simultaneously unitary and triangular, then it is diagonal.
9) Let A be a normal matrix with the property that distinct eigenvalues have distinct absolute values. Show that if B is a normal matrix, then AB is normal if and only if $AB = BA$. (Note that it is *not* true in general that AB is normal if and only if A and B commute (where A and B are normal matrices) although the corresponding result holds for self-adjoint mappings. The above result implies that this is true in the special case where either A or B is positive definite).
10) Let $f \in L(V)$ be such that if $x \perp y$, then $f(x) \perp f(y)$. Show that there exists an isometry f and a $\lambda \in \mathbf{R}$ so that $f = \lambda g$.
11) (This exercise provides the basis for a direct proof of the spectral theorem for isometries on $\mathbf{R}^n$ without having recourse to the complex result). Suppose

that f is an orthonormal operator on $\mathbf{R}^n$. Show that there exists a one- or two-dimensional subspace of $\mathbf{R}^n$ which is f-invariant. (One can suppose that f has no eigenvalues. Choose a unit vector x so that the angle between x and $f(x)$ is minimum. Show that if y is the bisector of the angle between x and $f(x)$ (i.e. $y = \dfrac{x + f(x)}{2}$), then $f(y)$ lies on the plane through x and $f(x)$—hence the latter is f-invariant).

4.8 The Moore-Penrose inverse

We now return once again to the topic of generalised inverses. Recall that if $f : V \to W$ is a linear mapping, we construct a generalised inverse for f by considering splittings

$$V = V_1 \oplus V_2 \qquad W = W_1 \oplus W_2$$

where V_2 is the kernel of f, W_1 the image and V_1 and W_2 are complementary subspaces. In the absence of any further structure, there is no natural way to choose W_2 and V_1. However, when V and W are euclidean space, the most obvious choices are the orthogonal complements $V_1 = V_2^\perp$ and $W_2 = W_1^\perp$. Then generalised inverse that we obtain in this way is uniquely specified and denoted by $f^\dagger$. It is called the **Moore-Penrose inverse** of f and has the following properties:

- $f^\dagger f f^\dagger = f^\dagger$;
- $f f^\dagger f = f^\dagger$;
- $f^\dagger f$ is the orthogonal projection onto V_1 and so is self-adjoint;
- $f f^\dagger$ is the orthogonal projection onto W_1 and so is self-adjoint.

In fact, these properties characterise $f^\dagger$—it is the only linear mapping from W into V which satisfies them as can easily be seen.

It follows that if $y \in W$, then $x = f^\dagger(y)$ is the "best" solution of the equation $f(x) = y$ in the sense that

$$\|f(x) - y\| \leq \|f(z) - y\|$$

for each $z \in V$ i.e. $f(x)$ is the nearest point to y in $f(V)$. In addition x is the element of smallest norm which is mapped onto this nearest point.

In terms of matrices, these results can be restated as follows: let V and W have orthonormal bases $(x_1, \dots, x_n)$ resp. $(y_1, \dots, y_m)$ and let f have matrix A with respect to them. Then the matrix $A^\dagger$ of $f^\dagger$ satisfies the conditions:

$$AA^\dagger A = A \quad A^\dagger A A^\dagger = A^\dagger \quad AA^\dagger \quad \text{and } A^\dagger A \text{ are self-adjoint.}$$

Of course, $A^\dagger$ is then called the **Moore-Penrose inverse** of A and is uniquely determined by the above equations. The existence of $f^\dagger$ can also be proved elegantly by using the result on singular values from the third paragraph. Recall that we can choose orthonormal bases $(x_1, \dots, x_n)$ resp. $(y_1, \dots, y_m)$ so that the matrix of f has the block form

$$\begin{bmatrix} A_1 & 0 \\ 0 & 0 \end{bmatrix}$$

with $A_1 = \operatorname{diag}(\mu_1, \dots, \mu_r)$. Then $f^\dagger$ is the operator with matrix

$$\begin{bmatrix} A_1^{-1} & 0 \\ 0 & 0 \end{bmatrix}$$

with respect to $(y_1, \dots, y_m)$ resp. $(x_1, \dots, x_n)$. Note that f is injective if and only if $r = n$. In this case, f has matrix

$$\begin{bmatrix} A_1 \\ 0 \end{bmatrix}$$

and $f^\dagger$ is the mapping $(f^t f)^{-1} f^t$ as one can verify by computing the matrix of the latter product.

Of course, the abstract geometric description of the Moore-Penrose inverse is of little help in calculating concrete examples and we mention some explicit formulae which are often useful.

Firstly, suppose that A has block representation $[B\ C]$ where B is an invertible (and hence square) matrix. Then it follows from the results on positive definite matrices that $BB^t + CC^t$ is invertible. The Moore-Penrose inverse of A is then given by the formula

$$A^\dagger = \begin{bmatrix} B^t(BB^t + CC^t)^{-1} \\ C^t(BB^t + CC^t)^{-1} \end{bmatrix}$$

as can be checked by multiplying out.

Example: Consider the matrix

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

Then A is of the above form with $B = I_n$ and one can easily calculate that $A^\dagger$ is the matrix

$$\frac{1}{n+1} \begin{bmatrix} n & -1 & -1 & \dots & -1 \\ -1 & n & -1 & \dots & -1 \\ \vdots & & & & \vdots \\ -1 & -1 & -1 & \dots & n \\ 1 & 1 & 1 & \dots & 1 \end{bmatrix}$$

The Moore-Penrose inverse can also be calculated by means of the following recursion method. We put

$$B = A^* A$$

and define

$$\begin{aligned} C_0 &= I \\ C_1 &= \operatorname{tr}(C_0 B) I - C_0 B \\ C_2 &= \frac{1}{2} \operatorname{tr}(C_1 B) I - C_1 B \end{aligned}$$

and so on. It turns out that $C_r B$ vanishes where r is the rank of A and the earlier values have non-vanishing trace. Then we have the formula

$$A^\dagger = \frac{(r-1) C_{r-1} A^*}{\operatorname{tr}(C_{r-1} B)}.$$

This can be checked by noting that the various steps are independent of the choice of basis. Hence we can choose bases so that the matrix of the operator defined by A has the form

$$\begin{bmatrix} A_1 & 0 \\ 0 & 0 \end{bmatrix}$$

where $A_1 = \operatorname{diag}(\lambda_1, \dots, \lambda_r)$. This is a simple calculation.

As an application, we consider the problem of the least square fitting of data. Let $(t_1, y_1), \dots, (t_n, y_n)$ be points in $\mathbf{R}^2$. We determine real numbers c, d so that the line $y = ct + d$ provides an optimal fit. This means that c and d should be a "solution" of the equation

$$\begin{bmatrix} t_1 & 1 \\ \vdots & \vdots \\ t_n & 1 \end{bmatrix} \begin{bmatrix} c \\ d \end{bmatrix} = \begin{matrix} y_1 \\ \vdots \\ y_n \end{matrix}.$$

If we interpret this in the sense that c and d are to be chosen so that the error

$$(y_1 - ct_1 - d_1)^2 + \cdots + (y_n - ct_n - d_n)^2$$

be as small as possible, then this reduces to calculating the Moore Penrose inverse of

$$A = \begin{bmatrix} t_1 & 1 \\ \vdots & \vdots \\ t_n & 1 \end{bmatrix}$$

since the solution is

$$\begin{bmatrix} c \\ d \end{bmatrix} = A^\dagger \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}.$$

If the t_i are distinct (which we tacitly assume), then $A^\dagger$ is given by the formula

$$A^\dagger = (A^t A)^{-1} A^t.$$

In this case

$$A^t A = \begin{bmatrix} \sum t_i^2 & \bar{t} \\ \bar{t} & n \end{bmatrix}$$

where $\bar{t} = t_1 + \cdots + t_n$.

Example: We consider the concrete case of the data

$$0\ 1\ 2\ 3\ 5$$

for the value of t and

$$0\ 2\ 3\ 4\ 6$$

for the corresponding values of the y's, then we have

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

and so

$$A^t A = \begin{bmatrix} 39 & 11 \\ 11 & 5 \end{bmatrix}$$

and

$$(A^t A)^{-1} = \frac{1}{74} \begin{bmatrix} 5 & -11 \\ -11 & 39 \end{bmatrix}.$$

This leads to the solution

$$\begin{bmatrix} c \\ d \end{bmatrix} = \frac{1}{74} \begin{bmatrix} 5 & -11 \\ -11 & 39 \end{bmatrix} \begin{bmatrix} 0 & 1 & 2 & 3 & 5 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 2 \\ 3 \\ 4 \\ 5 \end{bmatrix} = \begin{bmatrix} \frac{71}{74} \\ \frac{51}{74} \end{bmatrix}.$$

Similar applications of the Moore-Penrose inverse arise in the problem of curve fitting. Here one is interested in fitting lower order curves to given data. In chapter V we saw how the methods of linear algebra could be applied. In practical applications, however, the data will be overdetermined and will not fit the required type of curve exactlyl. In this case, the Moore-Penrose inverse can be used to find a curve whcih provides what is, in a certain sense, a best fit. We illustrate this with an example.

Example: Suppose that we are given a set of points $P_1, \dots, P_n$ in the plane and are looking for an ellipse which passes through them. In order to simplify the arithmetic, we shall assume that the ellipse has equation of the form

$$\alpha\xi_1^2 + \beta\xi_2^2 = 1$$

(i.e. that the principal axes are on the coordinate axes). Then we are requred to find (positive) α and β so that the equations

$$\alpha(\xi_1^i)^2 + \beta(\xi_2^i)^2 = 1$$

are satisfied (where P_i has coordinates (ξ_1^i, ξ_2^i)). This is a linear equation with matrix

$$A = \begin{bmatrix} (\xi_1^1)^2 & (\xi_2^1)^2 \\ \vdots & \vdots \\ (\xi_1^n)^2 & (\xi_2^n)^2 \end{bmatrix}.$$

Our theory would lead us to expect that the vector

$$\begin{bmatrix} a & b \end{bmatrix} = A^\dagger \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}$$

will be a good solution.

Exercises: 1) Show that the least square solution of the system

$$x = b_i \quad (i = 1, \dots, n)$$

is the mean value of $b_1, \dots, b_n$.
2) Calculate explicitly the least square solution of the system

$$y_i = c_i t + d$$

considered above.
3) Suppose that f is an operator on the hermitian space V. Show that if f is surjective, then ff^t is invertible and the Moore-Penrose inverse of f is given by the formula

$$f^\dagger = f^t(ff^t)^{-1}.$$

Interpret this in terms of matrices.
4) Show that $f \in L(V)$ commutes with its Moore-Penrose inverse if and only if the ranges of f and f^* coincide and this is equivalent to the fact that we have splitting

$$V = f(V) \perp \operatorname{Ker}(f)$$

resp. that there is an orthonormal basis with respect to which f has matrix

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

where A is invertible.
5) Show that if A is an $m \times n$ matrix, then there are polynomials p and q so that

$$A^\dagger = A^* p(AA^*) \quad \text{and} \quad A^\dagger = q(A^*A)A^*.$$

6) Show that the Moore-Penrose inverse of A can be written down explicitly with the help of the following integrals:

$$A^\dagger = \int_{-\infty}^{\infty} e^{-(A^*A)^t} A^* \, dt$$
$$A^\dagger = \frac{1}{2\pi i} \int_c \frac{1}{z}(zI - A^*A)^{-1} A^* \, dz$$

(the latter being integrated around a simple closed curve which encloses the non-zero eigenvalues of A^*A. These integrals of matrix-valued functions are to be interpreted in the natural way i.e. they are integrated elementwise).
7) Show that if A is normal with diagonalisation

$$U^*AU = \begin{bmatrix} A_1 & 0 \\ 0 & 0 \end{bmatrix}$$

where $A = \operatorname{diag}(\lambda_1, \dots, \lambda_r)$ with $\lambda_1, \dots, \lambda_r$ the non-zero eigenvalues, then

$$A^\dagger = (A + P)^{-1} - P$$

where

$$P = U^* \begin{bmatrix} 0 & 0 \\ 0 & I_{n-r} \end{bmatrix} U.$$

8) Use 7) to show that if C is a circulant with rank $r = n - p$ and $F^*CF = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$ is its diagonalisation as above, then

$$C^\dagger = C(I + K)^{-1} - K$$

where K is defined as follows. Suppose that the eigenvalues $\lambda_{i_1}, \dots, \lambda_{i_p}$ vanish. Then

$$K = \sum_s K_{i_s} \quad \text{with} \quad K_{i_r} = \frac{1}{n} \operatorname{circ}(1, \o^{-i_r+1}, \dots, \o^{(n-1)(-i_r+1)}).$$

9) Show how to use the Moore-Penrose inverse in obtaining a polynomial of degree at most $n-1$ to fit data

$$(t_1, x_1), \dots, (t_m, x_m)$$

where $m > n$. (The t_i are assumed to be distinct.)
10) Show that an operator f commutes with its Moore-Penrose inverse if and only if its range is the orthogonal complement of its kernel (or, alternatively, if the kernels of f and f^* coincide). Show that in this case, $f^\dagger$ is a polynomial in f.

4.9 Positive definite matrices

We conclude this chapter with a discussion of the important topic of positive definite matrices. Recall the following characterisation:

Proposition 37 *Let f be a self-adjoint operator on V. Then the following are equivalent:*

- *f is positive definite;*
- *all of the eigenvalues of f are positive;*
- *there is an invertible operator g on V so that $f = g^t g$.*

Proof. (1) implies (2): If λ is an eigenvalue,with unit eigenvector x, then

$$0 < (f(x)|x) = (\lambda x|x) = \lambda(x|x) = \lambda.$$

(2) implies (3): Choose an orthonormal basis (x_i) of eigenvectors for f. Then the matrix of f is $\operatorname{diag}(\lambda_1, \dots, \lambda_n)$ where, by assumption, each $\lambda_i > 0$. Let g be the operator with matrix $\operatorname{diag}(\sqrt{\lambda_1}, \dots, \sqrt{\lambda_n})$. Then $f = g^t g$. (3) implies (1): If $f = g^t g$, then

$$(f(x)|x) = (g^t g(x)|x) = (g(x)|g(x)) = \|g(x)\|^2 > 0$$

if $x \neq 0$.

There are corresponding characterisations of positive-semidefinite operators, resp. positive definite operators on hermitian spaces.

Suppose that the $n \times n$ matrix A is positive definite. By the above, A has a factorisation $B^t B$ for some invertible $n \times n$ matrix. We shall now show that B can be chosen to be upper triangular (in which case it is unique). For if $A = [a_{ij}]$, then $a_{11} > 0$ (put $X = (1, 0, \dots, 0)$ in the condition $X^t A X > 0$). Hence there is a matrix L_1 of the form

$$\left.\begin{array}{cccc} \frac{1}{a_{11}} & 0 & \dots & 0 \\ -b_{21} & 1 & \dots & 0 \\ \vdots & & & \vdots \\ -b_{n1} & 0 & & \dots \\ 1 & & & \end{array}\right]$$

so that the first row of $L_1 A$ is

$$\begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

(we are applying the Gaußian elimination method to reduce the first column). Since A is symmetric, we have the inequality

$$L_1 A L_1^t = \begin{bmatrix} 1 & 0 \\ 0 & A_2 \end{bmatrix}$$

where A_2 is also positive definite. Note that the matrix L_1 is lower triangular. Proceeding inductively, we obtain a sequence $L_1, \dots, L_{n-1}$ of such matrices so that if $L = L_{n-1} \dots L_1$, then $LAL^t = I$. Hence A has the factorisation $B^t B$ where $B = (L^{-1})^t$ and so is upper triangular. This is called the **Cholelsky factorisation** of A.

An almost immediate Corollary of the above is the following characterisation of positive definite matrices: A symmetric $n \times n$ matrix A is positive definite if and only if $\det A_k > 0$ for $k = 1, \dots, n$ where A is the $k \times k$ matrix $[a_{ij}]_{i,j=1}^k$.

Proof. Necessity: Note that if A is positive definite then $\det A > 0$ since the determinant is the product of the eigenvalues of A. Clearly each A_k is positive definite if A is (apply the defining condition on A to the vectors of the form $(\xi_1, \dots, \xi_k, 0, \dots, 0)$). Sufficiency: Let A satisfy the above condition. In particular, $a_{11} > 0$. As above we find a lower triangular matrix L_1 with

$$\tilde{A} = L_1 A L_1^t = \begin{bmatrix} 1 & 0 \\ 0 & C \end{bmatrix}$$

for a suitable $(n-1) \times (n-1)$ matrix C. The submatrices $\tilde{A}_k$ of this new matrix are obtained from those of A by multiplying the first row resp. column by $\dfrac{1}{a_{11}}$ (which is positive) or by subtracting the first row (resp. column) from the later ones. This implies that the new matrix $\tilde{A}$ and hence C satisfy the same conditions and can proceed as in the construction of the Cholelsky factorisation to get $A = B^t B$ i.e. A is positive definite.

There are a whole range of inequalities involving positive definite matrices. Many of their proofs are based on the following formula: Let A be a positive definite $n \times n$ matrix. Then

$$\int_{\mathbf{R}^n} e^{-\sum a_{ij}\xi_i\xi_j}\, d\xi_1 \dots d\xi_n = \frac{\pi^{\frac{n}{2}}}{\sqrt{\det A}}.$$

Proof. If we choose new coordinates $(\eta_1, \dots, \eta_n)$ with respect to which A is diagonal, we get the integral

$$\int_{\mathbf{R}^n} e^{-(\lambda_1\eta_1^2 + \dots + \lambda_n\eta_n^2)}\, d\eta_1 \dots d\eta_n$$

whose value is $\prod_{i=1}^n (\int_{\mathbf{R}} e^{-\lambda_i \eta_i^2}\, d\eta_i)$ ($\lambda_1, \dots, \lambda_n$ are the eigenvalues of A). The result now follows from the classical formula

$$\int_{\mathbf{R}} e^{-y^2}\, dy = \sqrt{\pi}$$

by a change of variables.

Exercises: 1) Show that if $f \in L(V)$ is positive definite, then

$$(f(x)|x)(f^{-1}(y)|y) \geq (x|y)^2$$

for $x, y \in V$.
2) Let f and g be self-adjoint operators where g is positive definite. The generalised eigenvalue problem for f and g is the equation $f(x) = \lambda g(x)$ (where, as usual, only non-zero x are of interest). Show that the space has a basis of eigenvectors for this problem (put $g = h^t h$ where h is invertible and note that the problem is equivalent to the usual eigenvalue problem for $(h^{-1})^t f h^{-1}$).
3) Show that every operator f on a euclidean space has uniquely determined representations

$$f = hu = u_1 h_1$$

where u and U_1 are isometries and h, h_1 are positive semi-definite. Show that f is then normal if and only if h and u commute, in which case $u_1 = u$ and $h_1 = h$.
4) Show that if A is a real, positive definite matrix, then

$$(\det A)^{\frac{1}{n}} = \min_{\det B = 1, B \geq 0} \operatorname{tr}(AB).$$

Hence deduce that

$$(\det(A+B))^{\frac{1}{n}} \geq (\det A)^{\frac{1}{n}} + (\det B)^{\frac{1}{n}}.$$

5) Let A be a positive definite $n \times n$ matrix and denote by A_i the matrix obtained by deleting the i-th row and the i-th column. Show that

$$\frac{\det(A+B)}{\det(A_i + B_i)} \geq \frac{\det A}{\det A_i} + \frac{\det B}{\det B_i}$$

(where B is also positive definite and B_i is defined as for A_i). 6) Show that if f is a positive definite operator with eigenvalues

$$\lambda_1 > \cdots > \lambda_n,$$

then

$$\prod \lambda_i \leq \prod (f(x_i)|x_i)$$
$$\sum \lambda_i \leq \sum (f(x_i)|x_i)$$

for any orthonormal basis (x_i).
7) Let f and g be operators on the euclidean space V. Show that they satisfy the condition $\|f(x)\| \leq \|g(x)\|$ for each $x \in V$ if and only if $g^t g - f^t f$ is positive semi-definite. Show that if they are normal and commute, this is equivalent to the fact that they have simultaneous matrix representation $\operatorname{diag}(\lambda_1, \dots \lambda_n)$ and $\operatorname{diag}(\mu_1, \dots, \mu_n)$ (with respect to an orthonormal basis) where $|\lambda_i| \geq |\mu_i|$ for each i.
8) Show that if A is a symmetric matrix so that there points $a < b$ such that

$$(A - aI)(A - bI)$$

is positive definite, then A has no eigenvalues between a and b.

5 MULTILINEAR ALGEBRA

In this chapter we bring a brief introduction to the topic of multilinear algebra. This includes such important subjects as tensors and multilinear forms. As usual, we employ a coordinate-free approach but show how to manipulate with coordinates *via* suitable bases in the spaces considered. We begin with the concept of the dual space.

5.1 Dual spaces

If V is a vector space, then the **dual space** of V is the space $L(V;\mathbf{R})$ of linear mappings from V into $\mathbf{R}$. As we have seen, this is a linear space with the natural operations. We denote it by V^*. For example, if V is the space $\mathbf{R}^n$, then, as we know, the space of linear mappings from $\mathbf{R}^n$ into $\mathbf{R}$ can be identified with the space $M_{1,n}$ of $1 \times n$ matrices, where $y = [\eta_1, \dots, \eta_n]$ induces the linear mapping

$$f_y : (\xi_1, \dots, \xi_n) \mapsto \xi_1\eta_1 + \cdots + \xi_n\eta_n.$$

Since $M_{1,n}$ is naturally isomorphic to $\mathbf{R}^n$ we can express this in the equation $(\mathbf{R}^n)^* = \mathbf{R}^n$ which we regard as a shorthand for the fact that the mapping $y \mapsto f_y$ is an isomorphism from the second space onto the first one.

We note the following simple properties of the dual space:

Proposition 38 *Let x, y be elements of a vector space V. Then*

- *$x \neq 0$ if and only if there is an $f \in V^*$ with $f(x) = 1$;*
- *x and y are linearly independent if and only if there is an $f \in V^*$ so that $f(x) = 1$, $f(y) = 0$ or $f(x) = 0$, $f(y) = 1$.*

PROOF. It suffices to prove these for the special case $V = \mathbf{R}^n$ where they are trivial

If f is a non-zero element in the dual V^* of V, then the subset

$$H^f_\alpha = \{x \in V : f(x) = \alpha\}$$

is called a **hyperplane** in V. It is an affine subspace of dimension one less than that of V, in fact, a translate of the kernel of f. Every hyperplane is of the above form and H^f_α and H^f_β are parallel (i.e. they can be mapped onto each other by a translation). Then the above result can be expressed as follows:

- a point x in V is non-zero if and only if it lies on some hyperplane which does not pass through zero;
- two points x and y in V are linearly independent if and only if there are parallel, but distinct, hyperplanes of the form H^f_α and H^f_0 so that $x \in H^f_\alpha$ and $y \in H^f_0$ or *vice versa.*

The dual basis: Suppose now that V has a basis $(x_1, \dots, x_n)$. For each i there is precisely one $f_i \in V^*$ so that

$$f_i(x_i) = 1 \quad \text{and} \quad f_i(x_j) = 0 \quad (i \neq j).$$

In other words, f_i is that element of the dual space which associates to each $x \in V$ its i-th coefficient with respect to (x_j) for if $x = \sum_{j=1}^n \lambda_j x_j$, then

$$f_i(x) = \sum_{j=1}^n \lambda_j f_i(x_j) = \lambda_i.$$

We claim that this sequence (f_i) is a basis for V^*—called the **dual basis** to (x_i). For if $f \in V^*$, it has the representation $f = \sum_{i=1}^n f(x_i) f_i$ which should be compared to the representation $x = \sum_{i=1}^n f_i(x) x_i$ in V. In order to prove this it suffices to show that

$$f(x) = \sum_{i=1}^n f(x_i) f_i(x)$$

for each $x \in V$ and this follows from an application of f to both sides of the equation $x = \sum_{i=1}^n f_i(x) x_i$.

In order to see that the f_i are linearly independent suppose that the linear combination $\sum_{i=1}^n \lambda_i f_i$ is zero. Then applying this form to x_j and using the defining condition on the f_i we see that $\lambda_j = 0$.

(Of course, the last step is, strictly speaking unnecessary since we already know that V and V^* have the same dimension).

The principle used in this argument will be applied again and so, in order to avoid tedious repetitions, we state an abstract form of it as a Lemma:

Lemma 3 *Let V be a vector space whose elements are functions defined on a set S with values in $\mathbf{R}$ so that the arithmetic operations on V coincide with the natural ones for functions (i.e. $(x+y)(t) = x(t)+y(t)$, $(\lambda x)(t) = \lambda x(t)$)). Then if $x_1, \dots, x_n$ is a sequence in V and there are points $t_1, \dots, t_n$ in S so that*

$$x_i(t_j) = 0 \quad (i \neq j) \quad \textit{or} \quad 1 \quad (i = j),$$

the sequence $x_1, \dots, x_n$ is linearly independent.

The proof is trivial. If a linear combination $\sum_{i=1}^n \lambda_i x_i$ vanishes, then evaluation at t_j shows that $\lambda_j = 0$.

Examples of dual bases: We calculate the dual bases to

- $(1,1),(1,0)$ for $\mathbf{R}^2$;
- the canonical basis $(1,t,\dots,t^n)$ for $\mathrm{Pol}\,(n)$.

(1) Let $x_1=(1,1)$, $x_2=(1,0)$ and let the dual basis be (f_1,f_2) where $f_1=(\xi_1^1,\xi_2^1)$, $f_2=(\xi_1^2,\xi_2^2)$. Then we have the four equations

$$\begin{array}{rclclcrclcl} f_1(x_1) &=& \xi_1^1+\xi_2^1 &=& 1 & \quad & f_2(x_1) &=& \xi_1^2+\xi_2^2 &=& 0\\ f_1(x_2) &=& \xi_1^1 &=& 0 & & f_2(x_2) &=& \xi_1^2 &=& 1 \end{array}$$

with solutions $f_1=(0,1)$, $f_2=(1,-1)$.
(2) Let f_i be the functional

$$p\mapsto\frac{p^{(i)}(0)}{i!}.$$

The of course, $f_i(t^j)=1$ if $i=j$ and 0 otherwise. Hence (f_i) is the dual basis and if $p\in\mathrm{Pol}\,(n)$, then its expansion

$$p=\sum_{i=0}^n f_i(p)t^i=\sum_{i=0}^n\frac{p^{(i)}(0)}{i!}t^i$$

with respect to the natural basis is the formal Taylor expansion of p.

We now investigate the behaviour of the dual basis under coordinate transformations. Let (x_i) resp. (x'_j) be bases for V with dual bases (f_i) and (f'_j). Let $T=[t_{ij}]$ be the transfer matrix from (x_i) to (x'_j) i.e. we have the equations

$$x'_j=\sum_{i=1}^n t_{ij}x_i.$$

Then we know that the coordinate representations

$$x=\sum_{i=1}^n\lambda_i x_i=\sum_{j=1}^n\lambda'_j x'_j$$

of x with respect to the two bases are related by the equations

$$\lambda_i=\sum_{j=1}^n t_{ij}\lambda'_j\quad \lambda'_i=\sum_{j=1}^n\tilde t_{ij}\lambda_j$$

where the matrix $[\tilde t_{ij}]$ is the inverse of T. Remembering that $\lambda_i=f_i(x)$ we can write these equations in the forms

$$f_i(x)=\sum_{j=1}^n t_{ij}f'_j(x)\quad\text{resp.}\quad f'_i(x)=\sum_{j=1}^n\tilde t_{ij}f_j(x).$$

Since these hold for any $x \in V$ we have

$$f_i = \sum_{j=1}^{n} t_{ij} f'_j \quad \text{resp.} \quad f'_i = \sum_{j=1}^{n} \tilde{t}_{ij} f_j.$$

Comparing these with the defining formula

$$f'_j = \sum_{i=1}^{n} s_{ij} f_i$$

for the transfer matrix S from (f_i) to (f'_j) we see that $S = (T^t)^{-1}$. Thus we have proved

Proposition 39 *If T is the transfer matrix from (x_i) to (x'_j), then $(T^t)^{-1}$ is the transfer matrix from (f_i) to (f'_j).*

We now consider duality for mappings. Suppose that $f : V \to W$ is a linear mapping. We define the transposed mapping f^t (which maps the dual W^* of W into V^*) as follows: if $g \in W^*$, then $f^t(g)$ is defined in the natural way as the composition $g \circ f$ i.e. we have the equation

$$f^t(g) : x \mapsto g(f(x)) \quad \text{or} \quad f^t(g)(x) = g(f(x)).$$

As the notation suggests, this is the coordinate-free version of the transpose of a matrix:

Proposition 40 *If $(x_1, \dots, x_n)$ resp. $(y_1, \dots, y_m)$ are bases for V and W resp. and $f : V \to W$ is a linear mapping with matrix $A = [a_{ij}]$, then the matrix of f^t with respect to the dual bases $(g_1, \dots, g_m)$ and $(f_1, \dots, f_n)$ is A^t, the transpose of A.*

PROOF. The matrix A is determined by the fact that f maps $\sum_{j=1}^n \lambda_j x_j$ into $\sum_{j=1}^n (\sum_{i=1}^m a_{ij}\lambda_j) y_i$ or, in terms of the f_j's and g_i's,

$$\sum_{j=1}^{n} a_{ij} f_j(x) = g_i(f(x)) = f^t g_i(x)$$

(the latter equation by the definition of f^t).

Since this holds for each x we have

$$f^t(g_i) = \sum_{j=1}^{n} a_{ij} f_j$$

and if we compare this with the defining relation

$$f^t(g_i) = \sum_{j=1}^{n} b_{ji} f_j$$

for the matrix B of f^t, we see that $B = A^t$.

The bidual: If V is a vector space, we can form the dual of its dual space i.e. the space $(V^*)^*$ which we denote by V^{**}. As we have already seen, the vector space V is isomorphic to its dual space V^* and hence also to its bidual. However, there is an essential difference between the two cases. The first was dependent on an (arbitrary) choice of basis for V. We shall now show how to define a natural isomorphism from V onto V^{**} which is independent of any additional structure of V.

Definition: If V is a vector space, we construct a mapping i_V from V into V^{**} by defining the form $i_V(x)$ $(x \in V)$ as follows:

$$i_V(x)(f) = f(x) \quad (f \in V^*).$$

It is easy to see that i_V is a linear injection. It is surjective since the dimensions of V and V^{**} coincide.

We now turn to duality for subspaces: let M be a subspace of V. Then

$$M^o = \{f \in V^* : f(x) = 0 \quad \text{for all} \quad x \in M\}$$

is a subspace of V^*—called **the annihilator** of M (in V^*). Similarly, if N is a subspace of V^*, then

$$N_o = \{x \in V : f(x) = 0 \quad \text{for all} \quad f \in N\}$$

is the annihilator of N in V. Notice that this is just $i_V^{-1}(N^o)$ where N^o is the annihilator of N in V^{**}, the dual of V^*. Hence there is a perfect symmetry in the relationship between M and M^o resp. N and N_o. Because of this fact, it will often suffice to prove just half of the statements of some of the next results.

Proposition 41 *If $M \subset V$ and $N \subset V^*$, then*

$$\dim M + \dim M^o = \dim V = \dim V^* = \dim N + \dim N_o.$$

Proof. Choose a basis (x_i) for V so that $(x_1, \dots, x_r)$ is one for M. Let (f_i) be the dual basis. Then it is clear that $M^o = [f_{r+1}, \dots, f_n]$ (cf. the calculation above) from which the first half of the equation follows. The second follows from the symmetry mentioned above.

Corollar 7 *If M and N are as above, then $M = (M^o)_o$ and$N = (N_o)^o$.*

PROOF. It is clear that $M \subset (M^o)_o$. To verify equality, we count dimensions:

$$\begin{aligned}\dim(M^o)_0 &= \dim V - \dim M^o \\ &= \dim V - (\dim V - \dim M) \\ &= \dim M.\end{aligned}$$

Proposition 42 *If $f : V \to W$ is a linear mapping, then $Ker\, f^t = (Im\, f)^o$, $Im\,(F^t) = (Ker\, f)^o$, $Ker\, f = (Im\, f^t)_o$, $Im\, f = (Ker\, f^t)_o$.*

PROOF. In fact, we only have to prove one of these result, say the first one $\operatorname{Ker} f^t = (\operatorname{Im} f)^o$. This follows from the following chain of equivalences:

$$\begin{array}{lll} g \in \operatorname{Ker} f^t & \text{if and only if} & f^t(g) = 0 \\ & \text{if and only if} & (f^t(g))(x) = 0 \quad \text{for} \quad x \in V \\ & \text{if and only if} & g(f(x)) = 0 \quad \text{for} \quad x \in V \\ & \text{if and only if} & g \in (\operatorname{Im} f)^o. \end{array}$$

In order to obtain the other three we proceed as follows: if we take annihilators of both sides we get

$$(\operatorname{Ker} f^t)_o = ((\operatorname{Im})^o)_o = \operatorname{Im} f$$

which is the fourth equation.

The remaining two are obtained by exchanging the roles of f, V, W with those of f^t, V^* and W^* in the two that we have just proved.

If we apply this result to the linear mapping defined by a matrix, we obtain the following criterium for the solvability of a system of linear equations.

Proposition 43 *Let A be an $m \times n$ matrix. Then the equation $AX = Y$ has a solution if and only if $Y^t Z = 0$ for each solution Z of the homogeneous equation $A^t Z = 0$.*

Product and quotient spaces: If V_1 and V_2 are vector space, then so is $V = V_1 \times V_2$ under the operations

$$\begin{aligned}(x, y) + (x_1, y_1) &= (x + x_1, y + y_1) \\ \lambda(x, y) &= (\lambda x, \lambda y)\end{aligned}$$

as is easily checked. The mappings $x \mapsto (x, o)$ and $y \mapsto (0, y)$ are isomorphisms from V_1 resp. V_2 onto the subspaces

$$\tilde{V_1} = \{(x, 0) : x \in V_1\} \quad \text{resp.} \quad \tilde{V_2} = \{(0, y) : y \in V_2\}$$

and $V = \tilde{V_1} \oplus \tilde{V_2}$. Hence $V_1 \times V_2$ is sometimes called the **external direct sum** of V_1 and V_2.

It is easily checked that the dual $(V_1 \times V_2)^*$ of such a product is naturally isomorphic to $V_1^* \times V_2^*$ where a pair (f, g) in the latter space defines the linear form

$$(x, y) \mapsto f(x) + g(y).$$

We now introduce a construction which is in some sense dual to that of taking subspaces and which can sometimes be used in a similar way to reduce dimension. Suppose that V_1 is a subspace of the vector space V. We introduce an equivalence relation $\sim$ on V as follows:

$$x \sim y \quad \text{if and only} \quad x - y \in V_1.$$

(i.e. we are reducing V_1 and all the affine subspaces parallel to it to points). V/V_1 is, by definition, the corresponding set of equivalence classes $\{[x] : x \in V\}$ where $[x] = \{y : y \sim x\}$. V/V_1 is a vector space in its own right, where we define the operations by the equations

$$[x] + [y] = [x + y]$$
$$\lambda[x] = [\lambda x]$$

and the mapping $\pi : V \to V/V_1$ which maps x onto $[x]$ is linear and surjective. Further we have the following characteristic property:

Proposition 44 *Let f be a linear mapping from V into a vector space W which vanishes on V_1. Then there is a linear mapping $\tilde{f} : V/V_1 \to W$ which is such that $f = \tilde{f} \circ \pi$.*

If we apply this to the case where $W = \mathbf{R}$, we see that the dual space of V/V_1 is naturally isomorphic to the polar V_1^o of V_1 in V^*. From this it follows that the dimension of $(V/V_1)^*$ and hence of V/V_1 is

$$\dim V - \dim V_1.$$

Exercises: 1) Calculate the dual basis to the basis

$$(1, 1, 1), (1, 1, -1), (1, -1, -1)$$

for $\mathbf{R}^3$.
2) Calculate the coordinates of the functional $p \mapsto \int_0^1 p(t)\,dt$ on $\mathrm{Pol}\,(n)$ with respect to the basis (f_{t_i}) where (t_i) is a sequence of distinct points in $[0, 1]$ and $f_{t_i}(p) = p(t_i)$.

3) Let $(x_1, \dots, x_n)$ be a basis for the vector space V with dual basis $(f_1, \dots, f_n)$. Show that the set

$$(x_1, x_2 - \lambda_2 x_1, \dots, x_n - \lambda_n x_1)$$

is a basis and that

$$(f_1 + \lambda_2 f_2 + \dots \lambda_n f_n, f_2 + \lambda_3 f_3 + \dots + \lambda_n f_n, \dots, f_n)$$

is the corresponding dual basis.
4) Find the dual basis to the basis

$$(1, t-a, \dots, (t-a)^n)$$

for $\mathrm{Pol}\,(n)$.
5) Let $t_0, \dots, t_n$ be distinct points of $[0,1]$. Show that the linear forms $f_i : x \to x(t_i)$ form a basis for the dual of $\mathrm{Pol}\,(n)$. What is the dual basis?
6) Let V_1 be a subspace of a vector space V and let $f : V_1 \to W$ be a linear mapping. Show that there is a linear mapping $\tilde{f} : V \to W$ which extends f. Show that if S is a **subset** of V and $f : S \to W$ an arbitrary mapping, then there is an extension of f to a linear mapping $\tilde{f}$ from V into W if and only if whenever $\sum_{i=1}^n \lambda_i x_i = 0$ (for $x_1, \dots, x_n \in S$), then $\sum_{i=1}^n \lambda_i f(x_i) = 0$.
7) Let f be the linear form

$$x \mapsto \int_0^1 x(t)\,dt$$

on $\mathrm{Pol}\,(n)$. Calculate $D^t f$ where D^t is the transpose of the differentiation operator D.
8) Let V_1 and V_2 be subspaces of a vector space V. Show that

$$(V_1 + V_2)^o = V_1^o \cap V_2^o \quad (V_1 \cap V_2)^o = V_1^o + V_2^o.$$

9) Let P be a projection onto the subspace V_1 of V. Show that P^t is also a projection and determine its range.
10) Show that if V and W are vector spaces, then $L(V,W)$ is naturally isomorphic to the dual of $L(V,W)$ under the bilinear mapping

$$(f,g) \mapsto \mathrm{tr}\,(g \circ f).$$

(i.e. the mapping

$$g \mapsto (f \mapsto \mathrm{tr}\,(g \circ f))$$

is an isomorphism from $L(W,V)$ onto $L(V,W)^*$).
11) Show that if f is a linear functional on the vector space M_n so that $f(\mathrm{Id}) = n$ and $f(AB) = f(BA)$ for each pair A, B, then f is the trace functional (i.e. $f(A) = \mathrm{tr}\,A$ for each A).

5.2 Duality in euclidean spaces

As we have seen, any vector space V is isomorphic to its dual space. In the special case where $V = \mathbf{R}^n$ we used the particular isomorphism $y \mapsto f_y$ where f_y is the linear form $x \mapsto \sum_i \xi_i \eta_i$. In this case we see the special role of the scalar product and this suggests the following result:

Proposition 45 *Let V be an euclidean space with scalar product $(\ |\ )$. then the mapping $\tau : y \to f_y$ where $f_y(x) = (x|y)$ is an isomorphism from V onto V^*.*

PROOF. τ is clearly linear and injective. It is surjective since $\dim V = \dim V^*$.

Using this fact, the duality theory for euclidean spaces can be given a more symmetric form. We illustrate this by discussing briefly the concept of covariant and contravariant vectors (cf. Chapter II.9). If (x_i) is a basis, consider the dual basis (f_i) for V^*. Then if we define x^i to be $\tau^{-1}(f_i)$, (x^i) is of course a basis for V. Hence each $x \in V$ has two representations,namely

$$x = \sum f_i(x) x_i = \sum_i (x|x^i) x_i$$

which is called the **contravariant representation** and

$$x = \sum_i (x|x_i) x^i$$

the **covariant representation**.

Note that a basis (x_i) is orthonormal exactly when it coincides with its dual basis. Then the two representations for x are the same.

We have been guilty of an abuse of notation by denoting the adjoint of an operator between euclidean spaces studied in Chapter VIII and the adjoint introduced here by the same symbol. This is justified by the fact that they coincide up to the identification of the spaces with their duals *via* the mapping τ.

5.3 Multilinear mappings

In this and the following section, we shall consider the concepts of multilinear mappings and tensors. In fact, these are just two aspects of the same mathematical phenomenon—the difference in language having arisen during their historical development. We begin with the concept of a multilinear mapping:

Definition: Let $V_1, \dots, V_n$ and W be vector spaces. A multilinear mapping from $V_1 \times \dots \times V_n$ into W is a mapping

$$f : V_1 \times \dots \times V_n \to W$$

so that

$$f(x_1, \dots, \lambda x_i + \mu x_i', \dots, x_n) = \lambda f(x_1, \dots, x_i, \dots, x_n) + \mu f(x_1, \dots, x_i', \dots, x_n)$$

for each choice of $i, x_1, \dots, x_i, x_i', \dots, x_n, \lambda, \mu$ (i.e. T is linear in each variable separately). Note that the linear mappings correspond to the special case $n = 1$.

We denote the space of all such mappings by $L^n(V_1, \dots, V_n; W)$. If $V_1 = \dots = V_n$ (which will almost always be the case in applications), we write $L^n(V; W)$. If the range space is $\mathbf{R}$, we denote it by $L^n(V_1, \dots, V_n)$ resp. $L^n(V)$. In particular, the space $L^1(V)$ is just the dual V^*.

In order to simplify the presentation, we shall, for the present, confine ourselves to the simple case $L^2(V_1, V_2)$ of bilinear forms on the product of *two* vector spaces. In this case, the defining conditions can be combined in the form

$$f(\lambda_1 x_1 + \lambda_2 x_2, \mu_1 y_1 + \mu_2 y_2) = \sum_{i,j=1,2} \lambda_i \mu_j f(x_i, y_j).$$

More generally, we have

$$f(\sum_{i-1}^{n} \lambda_i x_i, \sum_{j=1}^{m} \mu_j y_j) = \sum_{i,j} \lambda_i \mu_j f(x_i, y_j)$$

for any m. We have already met several examples of bilinear forms. For example, the scalar product on a euclidean space and the bilinear form $\sum_{i,j} a_{ij} \xi_i \eta_j$ associated with a conic section. In fact, the typical bilinear form on $\mathbf{R}^n$ can be written as

$$f(x, y) = \sum_{i,j} a_{ij} \xi_i \eta_j$$

for suitable coefficients $[a_{ij}]$. For if we put $a_{ij} = f(e_i, e_j)$ then

$$f(x, y) = f(\sum_{i=1}^{n} \xi_i e_i, \sum_{j-1}^{n} \eta_j e_j) = \sum_{i,j} a_{ij} \xi_i \eta_j.$$

In matrix notation this can be conveniently written in the form $X^t AY$ where X and Y are the column matrices

$$\begin{bmatrix} \xi_1 \\ \vdots \\ \xi_n \end{bmatrix} \qquad \begin{bmatrix} \eta_1 \\ \vdots \\ \eta_n \end{bmatrix}.$$

Just as in the case of the representation of linear operators by matrices, this is completely general and so if V_1 and V_2 are spaces with bases $(x_1, \dots, x_m)$ resp. $(y_1, \dots, y_n)$ and if $f \in L^2(V_1, V_2)$, then $A = [a_{ij}]$ where $a_{ij} = f(x_i, y_j)$ is called the **matrix** of f with respect to these bases and we have the representation

$$f(x, y) = \sum_{i,j} a_{ij} \lambda_i \mu_j$$

where $x = \sum_i \lambda_i x_i$ and $y = \sum_j \mu_j x_j$. We can express this fact in a more abstract way as follows. Suppose that $f \in V_1^*$ and $g \in V_2^*$. Then we define a linear functional $f \otimes g$ on $V_1 \times V_2$ as follows:

$$f \otimes f : (x, y) \mapsto f(x)g(y).$$

Proposition 46 *If (f_i) and (g_i) are the dual bases of V_1 and V_2, then $(f_i \otimes g_j)$ is a basis for $L^2(V_1, V_2)$. Hence the dimension of the latter space is* $\dim V_1 \cdot \dim V_2$.

PROOF. The argument above shows that these elements span $L^2(V_1, V_2)$. On the other hand, $f_i \otimes g_j(x_k, y_l)$ vanishes unless $i = k$ and $j = l$ in which case its value is one. Hence the set is linearly independent by the Lemma above.

We have thus seen that both linear mappings and bilinear forms are representable by matrices. However, it is important to note that the formula for the change in the representing matrices induced by new coordinate systems is different in each case as we shall now see. For suppose that we introduce new bases $(x_1', \dots, x_m')$ resp. $(y_1', \dots, y_n')$ in the above situation with transfer matrices $S = [s_{ij}]$ and $T = [t_{kl}]$ i.e.

$$x_j' = \sum_i s_{ij} x_i \quad y_j' = \sum_k t_{kl} y_k.$$

Now if A' is the matrix of f with respect to the new bases, then

$$\begin{aligned} a_{ji}' &= f(x_j', y_l') \\ &= f(\sum_i s_{ij} x_i, \sum_k t_{kl} y_k) \\ &= \sum_i \sum_k s_{ij} a_{ik} t_{kl} \end{aligned}$$

which is the (j,l)-th element S^tAT. Thus we have the formula

$$A' = S^tAT$$

for the new matrix which should be compared with that for the change in the matrix of a linear mapping.

In the particular case where $V_1 = V_2 = V$ and we use the *same* basis for each space, the above equation takes on the form

$$A' = S^tAS.$$

It is instructive to verify this formula with the use of coordinates. In matrix notation we have

$$f(x,y) = X^tAX = (X')^tA'(Y')$$

where $X,,Y,X',Y'$ are the column matrices composed of the coordinates of x and y with respect to the corresponding matrices. Now we know that $X = SX'$ and $Y = TY'$ and if we substitute this in the formula we get

$$f(x,y) = (SX')^tA(TY') = (X')^t(S^tAT)(Y')$$

as required.

We can distinguish two particularly important classes of bilinear forms f on the product $V \times V$ of a vector space with itself. $f \in L^2(V)$ is said to be

- **symmetric** if $f(x,y) = f(y,x)(x,y \in V)$;
- **alternating** if $f(x,y) = -f(y,x)(x,y \in V)$.

If f has the coordinate representation

$$\sum_{i,j} a_{ij} f_i \otimes f_j$$

with respect to the basis $(x_1,\dots,x_n)$, then f is symmetric (resp. alternating) if and only if $A = A^t$ (resp. $A = -A^t$). (For $a_{ij} = f(x_i,x_j) = \pm f(x_j,x_i) = \pm a_{ji}$ according as f is symmetric or alternating).

Symmetric forms with representations

$$f = \sum_{i=1}^{p} f_i \otimes f_i - \sum_{i=p+1}^{p+q} f_i \otimes f_i$$

are particularly transparent. On $\mathbf{R}^n$ these are the forms

$$f(x,y) = \xi_1\eta_1 + \cdots + \xi_p\eta_p - (\xi_{p+1}\eta_{p+1} + \cdots + \xi_{p+q}\eta_{p+q}).$$

The central result on symmetric forms is the following:

Proposition 47 *Let f be a symmetric bilinear form on V. Then there is a basis (x_i) of V and integers p, q with $p + q \leq n$ so that*

$$f = \sum_{i=1}^{p} f_i \otimes f_i - \sum_{i=p+1}^{p+q} f_i \otimes f_i.$$

Before proving this result, we restate it as one on matrices:

Proposition 48 *Let A be a symmetric $n \times n$ matrix. Then there is an invertible $n \times n$ matrix S so that*

$$S^t AS = diag\,(1, \ldots, 1, -1, \ldots, -1, 0, \ldots, 0).$$

Proof. We prove the matrix form of the result. First note that the result on the diagonalisation of symmetric operators on euclidean space provides a unitary matrix U so that $U^t AU = \mathrm{diag}\,(\lambda - 1, \ldots, \lambda_n)$ where the λ_i are the eigenvalues and can be ordered so that the first p (say) are positive, the next q are negative and the rest zero. Now put

$$T = \mathrm{diag}\,(\frac{1}{\sqrt{\lambda_1}}, \ldots, \frac{1}{\sqrt{\lambda_p}}, \frac{1}{\sqrt{-\lambda_{p+1}}}, \ldots, \frac{1}{\sqrt{-\lambda_{p+q}}}, 1, \ldots, 1).$$

Then if $S = UT$,

$$S^t AS = \begin{bmatrix} I_p & 0 & 0 \\ 0 & -I_q & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

We now turn to the signs involved in the canonical form

$$\begin{bmatrix} I_p & 0 & 0 \\ 0 & -I_q & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Although there is a wide range of choice of the diagonalising matrix S, it turns out that the arrangement of signs is invariant:

Proposition 49 *Proposition (Sylvester's law of inertia) Suppose that the symmetric form f has the representations*

$$\begin{bmatrix} I_p & 0 & 0 \\ 0 & -I_q & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad \textit{resp.} \quad \begin{bmatrix} I_{p'} & 0 & 0 \\ 0 & -I_{q'} & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

with respect to bases (x_i) resp. (x'_j). Then $p = p'$ and $q = q'$.

PROOF. First note that $p+q$ and $p'+q'$ are both equal to the rank of the corresponding matrices and so are equal. Now put

$$M_p = [x_1, \dots, x_p] \qquad M_{p'} = [x'_1, \dots, x'_{p'}]$$

$$N_q = [x_{p+1}, \dots, x_n] \qquad N_{q'} = [x'_{p'+1}, \dots, x'_n]$$

and note that $f(x,x) > 0$ for $x \in M_p \setminus \{0\}$ (resp. $M_{p'} \setminus \{0\}$) and $f(x,x) \leq 0$ for $x \in N_q$ resp. $N_{q'}$.

Hence $M_{p'} \cap N_q = \{0\} = M_p \cap N_{q'}$. Suppose that $p \neq p'$, say $p' > p$. Then

$$\begin{aligned}\dim(M_{p'} \cap N_q) &= \dim M_{p'} + \dim N_q - \dim(M_{p'} + N_q) \\ &\geq p' + (n-p) - n \\ &= p' - p > 0\end{aligned}$$

which is a contradiction. Then $p = p'$ and so $q = q'$.

The above proof of the diagonalisation of a symmetric bilinear form is rather artificial in that it introduces in an arbitrary way the inner product on $\mathbf{R}^n$. We give a direct, coordinate-free proof as follows: suppose that ϕ is such a form. We shall show, by induction on the dimension of V, how to construct a basis $(x_1, \dots, x_n)$ with respect to which the matrix of ϕ is

$$\begin{bmatrix} I_{p'} & 0 & 0 \\ 0 & -I_{q'} & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

for suitable p, q.

PROOF. We choose a vector $\tilde{x}_1$ with $\phi(\tilde{x}_1, \tilde{x}_1) \neq 0$. (If there is no such vector then the form ϕ vanishes and the result is trivially true). Now let V_1 be the linear span $[\tilde{x}_1]$ and put

$$V_2 = \{y \in V : \phi(\tilde{x}_1, y) = 0\}.$$

It is clear that $V_1 \cap V_2 = \{0\}$ and the expansion

$$y = \frac{\phi(\tilde{x}_1, y)}{\phi(\tilde{x}_1, \tilde{x}_1}\tilde{x}_1 + z$$

where $z = \dfrac{\phi(\tilde{x}_1, y)}{\phi(\tilde{x}_1, \tilde{x}_1}\tilde{x}_1$ (from which it follows that $z \in V_2$) shows that $V = V_1 \otimes V_2$. By the induction hypothesis, there is a suitable basis $(x_2, \dots, x_n)$ for V_2. We define x_1 to be $\dfrac{\tilde{x}_1}{\sqrt{\phi(\tilde{x}_1, \tilde{x}_1)}}$ if $\phi(\tilde{x}_1, \tilde{x}_1) > 0$ and $\dfrac{\tilde{x}_1}{\sqrt{-\phi(\tilde{x}_1, \tilde{x}_1)}}$ otherwise. Then the basis $(x_1, \dots, x_n)$ has the required properties.

A symmetric bilinear form ϕ is said to be **non-singular** if whenever $x \in V$ is such that $\phi(x, y) = 0$ for each $y \in V$, then x vanishes. The reader can check that this is equivalent to the fact that the rank of the matrix of ϕ is equal to the dimension of V (i.e. $p + q = n$). In this case, just as in the special case of a scalar product, the mapping

$$\tau : x \mapsto (y \mapsto \phi(x, y))$$

is an isomorphism from V onto its dual space V^*. The classical example of a non positive definite form which is non-singular, is the mapping

$$\phi : (x, y) \mapsto \xi_1\eta_1 - \xi_2\eta_2$$

on $\mathbf{R}^2$.

Most of the results above can be carried over to the space $L(V_1, \dots, V_r; W)$ of multilinear mappings from $V_1 \times \dots \times V_r$ into W. We content ourselves with the remark that if we have bases $(x_1^1, \dots, x_{n_1}^1), \dots (x_1^r, \dots, x_{n_r}^r)$ for $V_1, \dots, V_r$ with the corresponding dual bases and $(y_1, \dots, y_p)$ for W, then the set

$$(f_{i_1}^1 \otimes \dots \otimes f_{i_r}^r \otimes y_j : 1 \leq i_1 \leq n_1, \dots, 1 \leq i_r \leq n_r, 1 \leq j \leq p)$$

is a basis for $L(V_1, \dots, V_r; W)$ where for $f_i \in V_1^*, \dots, f_r \in V_r^*$, and $y \in W$, $f_1 \otimes \dots \otimes f_r \otimes y$ is the mapping

$$(x_1, \dots, x_r) \mapsto f_1(x_1) \dots f_r(x_r) y.$$

In particular, the dimension of the latter space is

$$(\dim V_1) \dots (\dim V_r)(\dim W).$$

Example: Calculate the coordinates of

$$f : (x, y) \mapsto 2\xi_1\eta_1 - \xi_1\eta_2 + \xi_2\eta_1 - \xi_2\eta_2$$

with respect to the basis $(1, 0), (1, 1)$.
Solution: The dual basis is $f_1 = (1, -1)$, $f_2 = (0, 1)$. Then $f = \sum a_{ij} f_i \otimes f_j$ where $a_{ij} = f(x_i, x_j)$ with $x_1 = (1, 0)$, $x_2 = (1, 1)$. Thus $a_{11} = 2$, $a_{21} = 1$, $a_{21} = 3$, $a_{22} = 1$ i.e.

$$f = 2f_1 \otimes f_1 + f_1 \otimes f_2 + 3f_2 \otimes f_1 + f_2 \otimes f_2.$$

Exercises: 1) Reduce the following forms on $\mathbf{R}^3$ to their canonical forms:

- $(\xi_1, \xi_2, \xi_3) \mapsto \xi_1\xi_2 + \xi_2\xi_3 + \xi_3\xi_2$;
- $(\xi_1, \xi_2, \xi_3) \mapsto \xi_2\eta 1 + \xi_1\eta_2 + 2\xi_2\eta_2 + 2\xi_2\eta_3 + 2\xi_3\eta_2 + 5\xi_3\eta_3$.

2) Find the matrices of the bilinear forms

$$(x, y) \mapsto \int_0^1 x(t)y(t)\,dt$$
$$(x, y) \mapsto x(0)y(0)$$
$$(x, y) \mapsto x(0)y'(0)$$

on Pol (n).
3) Let f be a symmetric bilinear form on V and ϕ be the mapping $x \mapsto f(x, x)$. Show that

- $f(x, y) = \frac{1}{4}(\phi(x + y) - \phi(x - y))$ (V real);
- $f(x, y) = \frac{1}{4}(\phi(x+y) - \phi(x-y) + i\phi(x+iy) - i\phi(x-iy))$ (V complex).

(This example shows how we can recover a symmetric 2-form from the quadratic form it generates i.e. its values on the diagonal).
4) Let $x_1, \dots, x_{n-1}$ be elements of $\mathbf{R}^n$. Show that there exists a unique element y of $\mathbf{R}^n$ so that$(x|y) = \det X$ for each $x \in \mathbf{R}^n$ where X is the matrix with rows $x_1, x_2, \dots, x_{n-1}, x$. If we denote this y by

$$x_1 \times x_2 \times \cdots \times x_{n-1}$$

show that this **cross-product** is linear in each variable x_i (i.e. it is an $(n-1)$-linear mapping from $\mathbf{R}^n \times \cdots \times \mathbf{R}^n$ into $\mathbf{R}^n$). (When $n = 3$, this coincides with the classical vector product studied in Chapter II).
5) Two spaces V and W with symmetric bilinear forms ϕ and ψ are said to be **isometric** if there is a vector space isomorphism $f : V \to W$ so that

$$\psi(f(x), f(y)) = \phi(x, y) \quad (x, y \in V).$$

Show that this is the case if and only if the dimensions of V and W respectively the rank and signatures of ϕ and ψ coincide.
6) Let A be a symmetric, invertible $n \times n$ matrix. Show that the quadratic form on $\mathbf{R}^n$ induced by A^{-1} is

$$(x, y) \mapsto -\frac{1}{\det A} \begin{bmatrix} 0 & \xi_1 & \dots & \xi_n \\ \eta_1 & & & \\ \vdots & & A & \\ \eta_n & & \dots & \end{bmatrix}.$$

7) Let ϕ be a symmetric bilinear form on the vector space V. Show that V has a direct sum representation

$$V = V_+ \oplus V_- \oplus V_0$$

where

$$\begin{aligned} V_0 &= \{x : \phi(x,x) = 0\} \\ V_+ &= \{x : \phi(x,x) > 0\} \cup \{0\} \\ V_- &= \{x : \phi(x,x) < 0\} \cup \{0\}. \end{aligned}$$

8) Let $(\,|\,)$ be a symmetric bilinear form on the vector space V and put

$$N = \{x \in V : (x|y) = 0 \quad \text{for each} \quad y \in V\}.$$

Show that one can define a bilinear form $(\,|\,)_1$ on V/N so that $(\pi(x)|\pi(y))_1 = (x|y)$ for each x, y. Show that this form is non-singular. What is $\dim V/N$? (π denotes the natural mapping $x \mapsto [x]$ from V onto the quotient space).
9) What is the rank of the bilinear form with matrix

$$\begin{bmatrix} a & 1 & 1 & \dots & 1 \\ 1 & a & 1 & \dots & 1 \\ \vdots & & & & \vdots \\ 1 & 1 & 1 & \dots & a \end{bmatrix}?$$

10) Show that a multilinear form $f \in L^r(V)$ is alternating if and only if $f(x_1, \dots, x_r) = 0$ whenever two of the x_i coincide.

Show that every bilinear form on V has a unique representation $f = f_s + f_a$ as a sum of a symmetric and an alternating form.
11) $\mathbf{R}^2$, together with the bilinear form

$$(x, y) \mapsto \xi_1\eta_1 - \xi_2\eta_2$$

is often called the **hyperbolic plane**. Show that if V is a vector space with a non-singular inner product and there is a vector x with $(x|x) = 0$, then V contains a two dimensional subspace which is isometric to the hyperbolic plane.
12) Suppose that ϕ and ψ are bilinear forms on a vector space V so that

$$\{x : \phi(x,x) = 0\} \cap \{x : \psi(x,x) = 0\} = \{0\}.$$

Show that there is a basis for V with respect to which both ϕ and ψ have upper triangular matrices. Deduce that if ϕ and ψ are symmetric, there is a basis for which both are diagonal.

13) Let $A = [a_{ij}]$ be an $n \times n$ symmetric matrix and letA_k denote the submatrix

$$\begin{bmatrix} a_{11} & \dots & a_{1k} \\ \vdots & & \vdots \\ a_{k1} & \dots & a_{kk} \end{bmatrix}.$$

Show that if the determinants of each of the A_k are non-zero, then the corresponding quadratic form $Q(x) = (Ax|x)$ can be written in the form

$$\sum_{k=1}^{n} \frac{\det A_k}{\det A_{k-1}} \eta_k^2$$

where $\eta_k = \xi_k + \sum_{j=k+1}^{n} b_{jk}\xi_j$ for suitable b_{ik}.

Deduce that A is positive definite if and only if each $\det A_k$ is positive. Can you give a corresponding characterisation of positive semi-definiteness?

14) Show that if f is a symmetric mapping in $L^r(V;W)$, then

$$f(x_1, \dots, x_r) = \frac{1}{r!2^r} \sum \epsilon_1 \dots \epsilon_r f(\epsilon_1 x_1 + \dots + \epsilon_r x_r)$$

the sum being taken over all choices (ϵ_i) of sign (i.e. each ϵ_i is either 1 or -1—there being 2^r summands).

15) Let ϕ be an alternating bilinear form on a vector space V. Show that V has a basis so that the matrix of the form is

$$\begin{bmatrix} 0 & I_r & 0 \\ -I_r & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

16) Let ϕ be as above and suppose that the rank of ϕ is n. Then it follows from the above that n is even i.e. of the form $2k$ for some k and V has a basis so that the matrix of ϕ is

$$J = \begin{bmatrix} 0 & I_k \\ -I_k & 0 \end{bmatrix}.$$

A space V with such a ϕ is called a **symplectic** vector space. A linear mapping f on such a space is called **symplectic** if $\phi(f(x), f(y)) = \phi(x, y)$. Show that this is the case if and only if $A^t JA = A$ where J is as above and A is the matrix of f with respect to this basis. Deduce that A has determinant 1 and so an isomorphism. Show that if λ is an eigenvalue of f, then so are $\frac{1}{\lambda}$, $\bar{\lambda}$ and $\frac{1}{\bar{\lambda}}$.

5.4 Tensors

We now turn to tensor products. In fact, these are also multilinear mappings which are now defined on the *dual* space. However, because of the symmetry between a vector space and its dual this is of purely notational significance.

Definition: If V_1 and V_2 are vector space, the **tensor product** $V_1 \otimes V_2$ of V_1 and V_2 is the space $L^2(V_1^*, V_2^*)$ of bilinear forms on the dual spaces V_1^* and V_2^*. If $x \in V_1$, $y \in V_2$, $x \otimes y$ is the form

$$(f, g) \mapsto f(x)f(y)$$

on $V_1^* \times V_2^*$. We can then translate some previous results as follows:

- the mapping $(x, y) \mapsto x \otimes y$ is bilinear from $V_1 \times V_2$ into $V_1 \otimes V_2$. In other words we multiply out tensors in the usual way:

$$(\sum_i \lambda_i x_i) \otimes (\sum_j \mu_j y_j) = \sum_{i,j} \lambda_i \mu_j x_i \otimes y_j;$$

- if (x_i) resp. (y_j) is a basis for V_1 resp. V_2, then

$$(x_i \otimes y_j : 1 \leq i \leq m, 1 \leq j \leq n)$$

  is a basis for $V_1 \otimes V_2$ and so each $z \in V_1 \otimes V_2$ has a representation $z = \sum_{i,j} a_{ij} x_i \otimes y_j$ where $a_{ij} = z(f_i, g_j)$.

Once again, this last statement implies that every tensor is described by a matrix. Of course the transformation laws for the matrix of a tensor are again different from those that we have met earlier and in fact we have the formula

$$A' = S^{-1} A (T^{-1})^t$$

where A is the matrix of z with respect to (x_i) and (y_j), A' is the matrix with respect to (x_i') and (z_j') and S and T are the corresponding transfer matrices.

Every tensor $z \in V_1 \otimes V_2$ is thus representable as a linear combination of so-called **simple tensors** i.e. those of the form $x \otimes y$ ($x \in V_1$, $y \in V_2$) (stated more abstractly, the image of $V_1 \times V_2$ in $V_1 \otimes V_2$ spans the latter). Not every tensor is simple. This can be perhaps most easily verified as follows: if $x = \sum_i \lambda_i x_i$ and $y = \sum_j \mu_j y_j$, then the matrix of $x \otimes y$ is $[\lambda_i \mu_j]$ and this has rank 1. Hence if the matrix of a tensor has rank more than one, it is not a simple tensor.

Tensor products of linear mappings: Suppose now that we have linear mappings $f \in L(V_1, W_1)$ and $g \in L(V_2, W_2)$. Then we can define a linear mapping $f \otimes g$ from $V_1 \otimes V_2$ into $W_1 \otimes W_2$ by putting

$$f \otimes g(\sum_i x_i \otimes y_i) = \sum_i f(x_i) \otimes g(y_i).$$

Higher order tensors: Similarly, we can define $V_1 \otimes \cdots \otimes V_r$ to be the space $L^r(V_1^*, \dots, V_r^*)$ of multilinear mappings on $V_1^* \times \cdots \times V_r^*$. Then if $(x_1^1, \dots, x_{n_1}^1), \dots, (x_1^r, \dots, x_{n_r}^r)$ are bases for $V_1, \dots, V_r$, the family

$$(x_{i_1}^1 \otimes \cdots \otimes x_{i_r}^r : 1 \leq i_1 \leq n_1, \dots, 1 \leq i_r \leq n_r)$$

is a basis for the tensor product and so every tensor has a representation

$$\sum_{1 \leq i_1, \dots, i_r \leq n} t_{i_1 \dots i_r} x_{i_1}^1 \otimes \cdots \otimes x_{i_r}^r.$$

In practice one is almost always interested in tensor products of the form

$$V \otimes V \otimes \cdots \otimes V \otimes V^* \otimes \cdots \otimes V^*$$

with p copies of V and q copies of the dual. We denote this space by $\bigotimes_p^q V$—its elements are called $(p+q)$ tensors on V. They are **contravariant of degree** p and **covariant of degree** q.

Thus the notation $\bigotimes_p^q V$ is just a new one for the space

$$L^{p+q}(V^*, \dots, V^*, V, \dots, V)$$

of multilinear forms on $V^* \times \cdots \times V^* \times V \times \cdots \times V$. (Strictly speaking, the last q spaces should be V^{**}'s but we are tacitly identifying V with V^{**}).

We now bring a useful list of operations on tensors. In doing so, we shall specify them only on simple tensors. This means that they will be defined on typical basis elements. They can then be extended by linearity or multilinearity to arbitrary tensors.

Multiplication: We can multiply a tensor of degree (p, q) with one of degree (p_1, q_1) to obtain one of degree $(p + p_1, q + q_1)$. More precisely, there is a bilinear mapping m from $\bigotimes_q^p V \times \bigotimes_{q_1}^{p_1} V$ into $\bigotimes_{q+q_1}^{p+p_1} V$ whereby

$$m(x_1 \otimes \cdots \otimes x_p \otimes f_1 \otimes \cdots \otimes f_q, x_1' \otimes \cdots \otimes x_{p_1}' \otimes f_1' \otimes \cdots \otimes f_{q_1}')$$

is

$$(x_1 \otimes \cdots \otimes x_p \otimes x_1' \otimes \cdots \otimes x_{p_1}' \otimes f_1 \otimes \cdots \otimes f_{q_1}').$$

Contraction: We can reduce a tensor of order (p, q) to one of degree $(p - 1, q - 1)$ by applying a covariant component to a contravariant one. More precisely, there is a linear mapping c from $\bigotimes_q^p V$ into $\bigotimes_{q-1}^{p-1} V$ where

$$c(x_1 \otimes \cdots \otimes x_p \otimes f_1 \otimes \dots f_q) = f_1(x_p) x_1 \otimes \dots x_{p-1} \otimes f_2 \otimes \cdots \otimes f_q.$$

Raising or lowering an index: In the case where V is a euclidean space, we can apply the operator τ or its inverse to a component of a coordinate to change it from a covariant one to a contravariant one or *vice versa*. For example we can map $\bigotimes_q^p V$ into $\bigotimes_{q+1}^{p-1} V$ as follows:

$$x_1 \otimes \dots \xi_p \otimes f_1 \otimes \cdots \otimes f_q \mapsto x_1 \otimes \cdots \otimes x_{p-1} \otimes \tau x_p \otimes f_1 \otimes \cdots \otimes f_q.$$

We continue with some brief remarks on the standard notation for tensors. V is a vector space with basis $(e_1, \dots, e_n)$ and we denote by $(f_1, \dots, f_n)$ the dual basis. Then we write $(e^1, \dots, e^n)$ for the corresponding basis for V, identified with V^* by way of a scalar product (i.e. $e^i = \tau^{-1}(f_i)$). Then we have bases

- (e_{ij}) for $V \otimes V$ where $e_{ij} = e_i \otimes e_j$ and a typical tensor has a representation $z = \sum_{ij} \xi^{ij} e_{ij}$ where $\xi^{ij} = f_i \otimes f_j(z) = (e^{ij}|z)$ with $e^{ij} = e^i \otimes e^j$).
- (e_i^j) for $V \otimes V^*$ where $e_i^j = e_i \otimes e^j$ and a typical tensor z has the representation $\sum_{i,j} \xi_j^i e_i^j$ where $\xi_j^i = (z|e_j^i)$.

In the general tensor space $\bigotimes_q^p V$ we have a basis

$$(e_{i_1 \dots i_p}^{j_1 \dots j_q})$$

where the typical element is

$$e_{i_1} \otimes \cdots \otimes e_{i_p} \otimes e^{j_1} \otimes \cdots \otimes e^{j_q}$$

and the tensor z has the representation

$$\sum \xi_{j_1 \dots j_q}^{i_1 \dots i_p} e_{i_1 \dots i_p}^{j_1 \dots j_q}.$$

Example: Let f be the operator induced by the 2×2 matrix $A = \begin{bmatrix} 0 & 2 \\ 1 & -1 \end{bmatrix}$ resp. by

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

. We calculate the matrix of $f \otimes g$ with respect to the basis (y_1, y_2, y_3, y_4) where

$$y_1 = e_1 \otimes e_1 \quad y_2 = e_1 \otimes e_2 \quad y_3 = e_2 \otimes e_1 \quad y_4 = e_4 \otimes e_4.$$

Then

$$\begin{aligned} f(y_1) &= -y_3 + 2y_4 \\ f(y_2) &= y_3 \\ f(y_3) &= -2y_1 + 4y_2 + y_3 - 2y_4 \\ f(y_4) &= 2y_1 - y_3. \end{aligned}$$

Hence the required matrix is

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

We can write this in the suggestive form

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

from which the following general pattern should be clear.

Proposition 50 *If A is the matrix of the linear mapping f and B that of g, then the matrix of $f \otimes f$ is*

$$\begin{bmatrix} a_{11}B & a_{12}B & \dots & a_{1n}B \\ \vdots & & & \vdots \\ a_{n1}B & a_{2n}B & \dots & a_{nn}B \end{bmatrix}.$$

(This matrix is called the **Kronecker product** *of A and B—written $A \otimes B$).*

Note that the basis used to define the matrix of the tensor product is that one which is obtained by considering the array $(x_i \otimes y_n)$ of tensor products and numbering it by reading along the successive rows n the customary manner.

Using these matrix representation one can check the following results:

- if f and g are injective, then so is $f \otimes g$;

- $r(f \otimes g) = r(f)r(g)$;
- $\mathrm{tr}\,(f \otimes g) = \mathrm{tr}\, f \cdot \mathrm{tr}\, g$;
- $\det(f \otimes g) = (\det f)^m (\det g)^n$ where m is the dimension of the space on which f acts, resp. n that of g.

One proves these results by choosing bases for which the matrices have a simple form and then examining the Kronecker product. For example, consider 3) and 4). We choose bases so that f and g are in Jordan form. Then it is clear that the Kronecker product is upper triangle (try it out for small matrices) and the elements in the diagonal are products of the form $\lambda_i \mu_j$ where λ_i is an eigenvalue of f and μ_j one of g. The formulae 3) and 4) follow by taking the sum resp. the product and counting how often the various eigenvalues occur.

If the underlying spaces to be tensored are euclidean, then the same is true of the tensor product space. For example, if V_1 and V_2 are euclidean, then the mapping

$$(x \otimes y | x_1 \otimes y_1) \mapsto (x|x_1)(y|y_1)$$

can be extended to a scalar product on $V_1 \otimes V_2$. Note that the latter is defined so as to ensure that if (x_i) is an orthonormal basis for V_1 and (y_j) one for V_2, then $(x_i \otimes y_j)$ is also orthonormal. In this context, we have the natural formula

$$(f \otimes g)^t = f^t \otimes g^t$$

relating the adjoint of $f \otimes g$ with those of f and g. This easily implies that the tensor product of two self-adjoint mappings is itself self-adjoint. The same holds for normal mappings. Also we have the formula

$$(f \otimes g)^\dagger = f^\dagger \otimes g^\dagger$$

for the Moore-Penrose inverse of a tensor product.

Example: Consider the matrix

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

Then its Moore-Penrose inverse can be easily calculated by noting that it is the Kronecker product of the matrices

$$\begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 1 & 0 & 1 \\ 1 & 3 & -1 \end{bmatrix}$$

whose Moore-Penrose inverses can be readily computed.

Tensor products can be used to solve certain types of matrix equation and we continue this section with some remarks on this theme. Firstly we note that if p is a polynomial in two variables, say

$$p(s,t) = \sum_{i,j} a_{ij} s^i t^j,$$

and A and B are $n \times n$ matrices, then we can define a new operator $p(A, B)$ by means of the formula

$$p(A,B) = \sum_{i,j} a_{i,j} A^i \otimes B^j$$

(Warning: this is not the matrix obtained by substituting A and B for s and t resp.—the latter is an $n \times n$ matrix whereas the matrix above is $n^2 \times n^2$).

The result which we shall require is the following:

Proposition 51 *The eigenvalues of the above matrix are the scalars of the form $p(\lambda_i, \mu_j)$ where $\lambda_i, \dots, \lambda_n$ are the eigenvalues of A and $\mu_1, \dots, \mu_n$ are those of B.*

This is proved by using a basis for which A has Jordan form. The required matrix then has an upper triangular block form with diagonal matrices of type $p(\lambda_i, B)$ i.e. matrices which are obtained by substituting an eigenvalue λ_i of A for s and the matrix B for t. This matrix has eigenvalues $p(\lambda_i, \mu_j)$ $(j_1, \dots, n)$ from which the result follows.

The basis of our application of tensor products is the following simple remark. The space $M_{m,n}$ of $m \times n$ matrices is of course identifiable as a vector space with $\mathbf{R}^{mn}$. In the following we shall do this systematically by associating to an $m \times n$ matrix $X = [X_1 \dots X_n]$ the column vector

$$\tilde{X} = \begin{bmatrix} X_1 \\ \vdots \\ X_n \end{bmatrix}$$

i.e. we place the columns of X on top of each other. The property that we shall require in order to deal with matrix equations is the following:

Proposition 52 *Let A be an $m \times m$ matrix, X an $m \times n$ matrix and B an $n \times n$ matrix. Then we have the equation*

$$\widetilde{AXB} = (B^t \otimes A)\tilde{X}.$$

PROOF. If $X = [X_1 \dots X_n]$ then the j-th column of AXB is $\sum_{k=1}^{n}(b_{jk}A)X_k$ and this is

$$[b_{1j}A \quad b_{2j}A \dots b_{nj}A]\tilde{X}$$

which implies the result.

The following special cases will be useful below.

$$\begin{aligned} \widetilde{AX} &= (I_n \otimes A)X; \\ \widetilde{XB} &= (B^t \otimes I_m)X; \\ \widetilde{AX + XB} &= ((I_n \otimes A) + (B^t \otimes I_m)\tilde{X}. \end{aligned}$$

The above results can be used to tackle equations of the form

$$A_1 X B_1 + \dots + A^r X B^r = C.$$

Here the A's are the given $m \times m$ matrices, the B's are $n \times n$ and V is $m \times n$. X is the unknown. Using the above apparatus, we can rewrite the equation in the form $G\tilde{X} = \tilde{C}$ where $C = \sum_{j=1}^{r} B_j^t \otimes A_j$.

Rather than consider the most general case, we shall confine our attention to one special one which often occurs in applications, namely the equation

$$AX + XB = C.$$

In this case, the matrix G is

$$I_n \otimes A + B^t \otimes I_m.$$

This is just $p(A, B^t)$ where $p(s,t) = s+t$. Hence by our preparatory remarks, the eigenvalues of G are the scalars of the form $\lambda_i + \mu_j$ where the λ's are the eigenvalues of A and the μ_j are those of B.

Hence we have proved the following result:

Proposition 53 *The equation $AX + XB = C$ has a solution for any C if and only if $\lambda_i + \mu_j \neq 0$ for each pair λ_i and μ_j of eigenvalues of A and B respectively.*

For the above condition means that 0 is not an eigenvalue of $\tilde{G}$ i.e. this matrix is invertible.

We can also get information for the case where the above general equation is not always solvable. Consider, for example, the equation

$$AXB = C$$

where we are not assuming that A and B are invertible. Suppose that S and T are generalised inverses for A and B respectively. Then it is clear that $S \otimes T$ is one for $A \otimes B$. If we rewrite the equation $AXB = C$ in the form

$$(A^t \otimes B)\tilde{X} = \tilde{C}$$

then it follows form the general theory of such inverses that it has a solution if and only if we have the equality

$$ASCTB = C.$$

The general solution is then

$$X = SCT + Y - SAYBT$$

whereby Y is arbitrary.

In general, one would choose the Moore-Penrose inverses $A^\dagger$ and $B^\dagger$ for S and T. This gives the solution

$$X = A^\dagger C B^\dagger$$

which is best possible in the usual sense.

Exercises: 1) Let $x_1, \dots, x_n$ be elements of V. Show that

$$x_1 \otimes \cdots \otimes x_n = 0$$

if and only if at least one x_i vanishes. If x, y, x', y' are non-vanishing elements of V show that $x \otimes y = x' \otimes y'$ if and only if there is a $\lambda \in \mathbf{R}$ so that $x = \lambda x'$ and $y = \frac{1}{\lambda} y'$.
2) Let V_1 and V_2 be vector spaces and denote by ϕ the natural bilinear mapping $(x, y) \mapsto x \otimes y$ from $V_1 \times V_2$ into $V_1 \otimes V_2$. Show that for every bilinear mapping $T : V_1 \times V_2 \to W$, there is a unique linear mapping $f : V_1 \otimes V_2 \to W$ so that $T = f \circ \phi$.

Show that this property characterises the tensor product i.e. if U is a vector space and ψ is a bilinear mapping from $V_1 \times V_2$ into U so that the natural analogue of the above property holds, then there is an isomorphism g from U onto $V_1 \otimes V_2$ so that $g \circ \psi = \phi$.
3) Let (x_i) resp. (x_i') be bases for the vector space V and suppose that $A = [a_{ij}]$ is the corresponding transfer matrix. Show that if a p-tensor z has the coordinate representation

$$z = \sum_{1 \leq i_1, \dots, i_p \leq n} t_{i_1, \dots, i_p} f_{i_1} \otimes \cdots \otimes f_{i_p}$$

with respect to the first basis and

$$z = \sum_{1 \leq i_1, \dots, i_p \leq n} t'_{i_1, \dots, i_p} f'_{i_1} \otimes \cdots \otimes f'_{i_p}$$

with respect to the second, then

$$t'_{i_1, \dots, i_p} = \sum_{1 \leq j_1, \dots, j_p \leq n} a_{i_1 j_1} \dots a_{i_p j_p} t_{j_1 \dots j_p}.$$

4) Show that the tensor product

$$\mathrm{Pol}\,(m) \otimes \mathrm{Pol}\,(n)$$

is naturally isomorphic to the vector space of polynomials in two variable, of degree at most m in the first and at most n in the second.
5) What is the trace resp. the determinant of the linear mapping

$$f \mapsto \phi \circ f \circ \psi$$

on $L(V)$ where ϕ and ψ are fixed members of $L(V)$.
6) Consider the linear mapping

$$\Phi : f \mapsto f^t$$

on $L(V)$ where V is Hermitian. Calculate the matrix of f with respect to the usual basis for $L(V)$ derived from an orthonormal basis for V. Is Φ normal resp. unitary with respect to the scalar product

$$(f|g) = \mathrm{tr}\,(g^t f).$$

Calculate the characteristic polynomial resp. the eigenvalues of Φ.
7) Let p and q be polynomials and A and B be matrices with p and q as characteristic polynomials (for example, the companion matrices of p and q). Show that the number $\Delta = \det(A \otimes I - I \otimes B)$ is a resolvent of p and q i.e. has the property that it vanishes if and only if p and q have a common zero.
8) Consider the mapping

$$X \mapsto AXA^t$$

on the vector space of the $n \times n$ symmetric matrices, where A is a given $n \times n$ matrix. Show that the determinant of this mapping is $(\det A)^{n+1}$.