MA266: Multilinear Algebra

Christian Böhning
Based on notes of Adam Thomas, Derek Holt and David Loeffler

Term 2, 2024

Contents

0 Review of some MA106 material

In this section, we'll recall some ideas from the first year MA106 Linear Algebra module. This will just be a brief reminder; for detailed statements and proofs, go back to your MA106 notes.

0.1 Fields

Recall that a *field* is a number system where we know how to do all of the basic arithmetic operations: we can add, subtract, multiply and divide (as long as we're not trying to divide by zero).

Definition 0.1.1. A field is a non-empty set K together with two operations (maps from $K \times K$ to K) addition, denoted by $+$, and multiplication, denoted by $\cdot$ (or just juxtaposition), satisfying the following axioms:

1. $a + b = b + a$ for all $a, b \in K$;
2. there exists an element $0 \in K$ such that $a + 0 = a$ for all $a \in K$;
3. $(a + b) + c = a + (b + c)$ for all $a, b, c \in K$;
4. there exists an element $-a \in K$ such that $a + (-a) = 0$ for all $a \in K$;
5. $a \cdot b = b \cdot a$;
6. there exists an element $1 \in K$, $1 \neq 0$, such that $1 \cdot a = a$ for all $a \in K$;
7. $(a \cdot b) \cdot c = a \cdot (b \cdot c)$ for all $a, b, c \in K$;
8. there exists an element $a^{-1} \in K$ such that $a \cdot a^{-1} = 1$ for all $0 \neq a \in K$;
9. $a \cdot (b + c) = (a \cdot b) + (a \cdot c)$ for all $a, b, c \in K$.

Examples.

- A non-example is $\mathbb{Z}$, the integers. Here we can add, subtract, and multiply, but we can't always divide without jumping out of $\mathbb{Z}$ into some bigger world. That is to say that Axiom 8 would fail: there are no multiplicative inverses of any integer apart from 1 and -1.
- The real numbers $\mathbb{R}$ and the complex numbers $\mathbb{C}$ are fields, and these are perhaps the most familiar ones.
- The rational numbers $\mathbb{Q}$ are also a field.
- A more subtle example: if p is a prime number, the integers mod p are a field, written as $\mathbb{Z}/p\mathbb{Z}$ or $\mathbb{F}_p$.

There are lots of fields out there, and the reason we take the axiomatic approach is that we know that everything we prove will be applicable to any field we like, as long as we've only used the field axioms in our proofs (rather than any specific properties of the fields we happen to most like). We don't have to know all the fields in existence and check that our proofs are valid for each one separately.

0.2 Vector spaces

Let K be a field[1]. A *vector space* over K is a non-empty set V together with two extra pieces of structure. Firstly, it has to have a notion of *addition*: we need to know what $v + w$ means if v and w are in V. Secondly, it has to have a notion of *scalar multiplication*: we need to know what λv means if v is in V and λ is in K. These have to satisfy some axioms, for which I'm going to refer you again to your MA106 notes.

Definition 0.2.1. A vector space V over a field K is a set V with two operations. The first is addition, a map from $V \times V$ to V satisfying Axioms 1 to 4 in the definition of a field. The second operation is scalar multiplication, a map from $K \times V$ to V denoted by juxtaposition or $\cdot$, satisfying the following axioms:

1. $\alpha(u + v) = \alpha u + \alpha v$ for all $u, v \in V$, $\alpha \in K$;
2. $(\alpha + \beta)v = \alpha v + \beta v$ for all $v \in V$, $\alpha, \beta \in K$;
3. $(\alpha \cdot \beta)v = \alpha(\beta v)$ for all $v \in V$, $\alpha, \beta \in K$;
4. $1 \cdot v = v$ for all $v \in V$.

A *basis* of a vector space is a subset $B \subset V$ such that every $v \in V$ can be written *uniquely* as a finite linear combination of elements of B,

$$v = \lambda_1 b_1 + \cdots + \lambda_n b_n,$$

for some $n \in \mathbb{N}$ and some $\lambda_1, \ldots, \lambda_n \in K$. So for each $v \in V$, we can do this in one and only one way. Another way of saying this is that B is a linearly independent set which spans V, which is the definition you had in MA106. We say V is *finite-dimensional* if there is a finite basis of V. You saw last year that if V has one basis which is finite, then every basis of V is finite, and they all have the same cardinality; and we define the *dimension* of V to be this number which is the number of elements in any basis of V.

Examples. Let $K = \mathbb{R}$.

- The space of polynomials in x with coefficients in $\mathbb{R}$ is certainly a vector space over $\mathbb{R}$; but it's not finite-dimensional (rather obviously).
- For any $d \in \mathbb{N}$, the set $\mathbb{R}^d$ of column vectors with d real entries is a vector space over $\mathbb{R}$ (which, not surprisingly, has dimension d).
- The set
$$\left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \mathbb{R}^3 : x_1 + x_2 + x_3 = 0 \right\}$$
is a vector space over $\mathbb{R}$ if we define vector addition and scalar multiplication componentwise as usual.

The third example above is an interesting one because there's no "natural choice" of basis. It certainly has bases, e.g. the set

$$\left\{ \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \right\},$$

[1] It's conventional to use K as the letter to denote a field; the K stands for the German word "Körper".

but there's no reason why that's better than any other one. This is one of the reasons why we need to worry about the choice of basis – if you want to tell someone else all the wonderful things you've found out about this vector space, you might get into a total muddle if you insisted on using one particular basis and they preferred another different one.

The following lemma (which will be required in the proof of one of our main theorems) is straightforward from the material in MA106 - the proof is left as an exercise to check you are comfortable with such material.

Lemma 0.2.2. *Suppose that U is an m-dimensional subspace of an n-dimensional vector space V and $\mathbf{w}_1, \dots, \mathbf{w}_{n-m}$ extend a basis of U to a basis of V. Then the equation*

$$\alpha_1 \mathbf{w}_1 + \cdots + \alpha_{n-m} \mathbf{w}_{n-m} + \mathbf{u} = \mathbf{0}, \ \textit{where} \ \mathbf{u} \in U, \tag{1}$$

only has the solution $\alpha_i = 0$ for all $1 \le i \le n - m$ and $\mathbf{u} = 0$.

0.3 Linear maps

If V and W are vector spaces (over the same field K), then a *linear map* from V to W is a map $T : V \to W$ which "respects the vector space structures". That is, we know two things that we can do with vectors in a vector space – add them, and multiply them by scalars; and a linear map is a map where adding or scalar-multiplying on the V side, then applying the map T, is the same as applying the map T, then adding or multiplying on the W side. Formally, for T to be a linear map means that we must have

$$T(v_1 + v_2) = T(v_1) + T(v_2) \quad \forall\, v_1, v_2 \in V$$

and

$$T(\lambda v_1) = \lambda T(v_1) \quad \forall\, \lambda \in K, v_1 \in V.$$

Example 1. Let V and W be vector spaces over K. Then $T : V \to W$ defined by $T(v) = 0_W = 0$ for all $v \in V$ is a linear map, called the zero linear map. Furthermore, we have $S : V \to V$ defined by $S(v) = v$ for all $v \in V$ is a linear map, called the identity linear map.

Example 2. Let $V = \mathbb{R}^3$ and $W = \mathbb{R}^2$. Then the following maps $T : V \to W$ are linear.

1. $T\left(\begin{pmatrix} a \\ b \\ c \end{pmatrix}\right) = \begin{pmatrix} a \\ b \end{pmatrix}$;
2. $T\left(\begin{pmatrix} a \\ b \\ c \end{pmatrix}\right) = \begin{pmatrix} b \\ 0 \end{pmatrix}$;
3. $T\left(\begin{pmatrix} a \\ b \\ c \end{pmatrix}\right) = \begin{pmatrix} a+b \\ b+c \end{pmatrix}$.

Whereas, you should check that $T\left(\begin{pmatrix} a \\ b \\ c \end{pmatrix}\right) = \begin{pmatrix} a^2 \\ b \end{pmatrix}$ is NOT a linear map.

0.4 The matrix of a linear map with respect to a choice of (ordered) bases

Let V and W be vector spaces over a field K. Let $T : V \to W$ be a linear map, where $\dim(V) = n$, $\dim(W) = m$. Choose a basis $\mathbf{e}_1, \ldots, \mathbf{e}_n$ of V and a basis $\mathbf{f}_1, \ldots, \mathbf{f}_m$ of W. Note that formally what we are doing here is choosing *ordered* bases- above we defined a basis of a vector space to be simply a *subset* without any preferred ordering, but here we actually make a choice of two ordered sets of bases, $\mathbf{E} = (\mathbf{e}_1, \ldots, \mathbf{e}_n)$ and $\mathbf{F} = (\mathbf{f}_1, \ldots, \mathbf{f}_m)$, the ordering being encoded in the choice of indices.

Now, for $1 \leq j \leq n$, $T(\mathbf{e}_j) \in W$, so $T(\mathbf{e}_j)$ can be written uniquely as a linear combination of $\mathbf{f}_1, \ldots, \mathbf{f}_m$. Let

$$\begin{aligned} T(\mathbf{e}_1) &= \alpha_{11}\mathbf{f}_1 + \alpha_{21}\mathbf{f}_2 + \cdots + \alpha_{m1}\mathbf{f}_m \\ T(\mathbf{e}_2) &= \alpha_{12}\mathbf{f}_1 + \alpha_{22}\mathbf{f}_2 + \cdots + \alpha_{m2}\mathbf{f}_m \\ &\vdots \\ T(\mathbf{e}_n) &= \alpha_{1n}\mathbf{f}_1 + \alpha_{2n}\mathbf{f}_2 + \cdots + \alpha_{mn}\mathbf{f}_m \end{aligned}$$

where the coefficients $\alpha_{ij} \in K$ (for $1 \leq i \leq m$, $1 \leq j \leq n$) are uniquely determined.

The coefficients α_{ij} form an $m \times n$ matrix

$$A = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \ldots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \ldots & \alpha_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{m1} & \alpha_{m2} & \ldots & \alpha_{mn} \end{pmatrix}$$

over K. Then A is called the matrix of the linear map T with respect to the chosen bases of V and W. Note that the columns of A are the images $T(\mathbf{e}_1), \ldots, T(\mathbf{e}_n)$ of the basis vectors of V represented as column vectors with respect to the basis $\mathbf{f}_1, \ldots, \mathbf{f}_m$ of W.

It was shown in MA106 that T is uniquely determined by A, and so there is a one-one correspondence between linear maps $T : V \to W$ and $m \times n$ matrices over K, which depends on the choice of ordered bases of V and W.

For $\mathbf{v} \in V$, we can write $\mathbf{v}$ uniquely as a linear combination of the basis vectors $\mathbf{e}_i$; that is, $\mathbf{v} = x_1\mathbf{e}_1 + \cdots + x_n\mathbf{e}_n$, where the x_i are uniquely determined by $\mathbf{v}$ and the basis $\mathbf{e}_i$. We shall call x_i the *coordinates* of $\mathbf{v}$ with respect to the basis $\mathbf{e}_1, \ldots, \mathbf{e}_n$. We associate the column vector

$$\underline{\mathbf{v}} = \begin{pmatrix} x_1 \\ x_2 \\ . \\ . \\ x_n \end{pmatrix} \in K^{n,1},$$

to $\mathbf{v}$, where $K^{n,1}$ denotes the space of $n \times 1$-column vectors with entries in K. Notice that $\underline{\mathbf{v}}$ depends on the chosen basis $\mathbf{E}$ so a notation such as $\mathbf{v}_{\mathbf{E}}$ or $\underline{\mathbf{v}}_{\mathbf{E}}$ would possibly be better, but also heavier, so we stick with $\underline{\mathbf{v}}$ and assume you bear in mind that $\underline{\mathbf{v}}$ not only depends on $\mathbf{v}$ but also on $\mathbf{E}$.

It was proved in MA106 that if A is the matrix of the linear map T, then for $\mathbf{v} \in V$, we have $T(\mathbf{v}) = \mathbf{w}$ if and only if $A\underline{\mathbf{v}} = \underline{\mathbf{w}}$, where $\underline{\mathbf{w}} \in K^{m,1}$ is the column vector associated with $\mathbf{w} \in W$.

Example. We can write down the matrices for the linear maps in Example 2, using the standard bases for V and W: the standard basis of $\mathbb{R}^n$ is $e_1, \ldots, e_n$ where e_i is the column vector with a 1 in the ith row and all other entries 0 (so it's the $n \times 1$ matrix defined by $\alpha_{j,1} = 1$ if $j = i$ and $\alpha_{j,i} = 0$ otherwise).

1. We calculate that $T(e_1) = e_1 = 1 \cdot e_1 + 0 \cdot e_2$, $T(e_2) = e_2 = 0 \cdot e_1 + 1 \cdot e_2$ and $T(e_3) = 0 = 0 \cdot e_1 + 0 \cdot e_2$ (OK, this could be confusing so we could denote the standard basis for W by f_1, f_2). The matrix is thus
$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}.$$
2. We skip the details but the matrix is
$$A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$
3. This time $T(e_1) = e_1$, $T(e_2) = e_1 + e_2$ and $T(e_3) = e_2$ and so the matrix is
$$A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}.$$

0.5 Change of basis

Let V be a vector space of dimension n over a field K, and let $\mathbf{e}_1, \ldots, \mathbf{e}_n$ and $\mathbf{e}'_1, \ldots, \mathbf{e}'_n$ be two bases of V (ordered of course). Then there is an invertible $n \times n$ matrix $P = (p_{ij})$ such that

$$\mathbf{e}'_j = \sum_{i=1}^{n} p_{ij}\mathbf{e}_i \quad \text{for } 1 \leq j \leq n. \qquad (*)$$

Note that the columns of P are the new basis vectors $\mathbf{e}'_i$ written as column vectors in the old basis vectors $\mathbf{e}_i$. (Recall also that P is the matrix of the identity map $V \to V$ using basis $\mathbf{e}'_1, \ldots, \mathbf{e}'_n$ in the domain and basis $\mathbf{e}_1, \ldots, \mathbf{e}_n$ in the codomain.)

Often, but not always, the original basis $\mathbf{e}_1, \ldots, \mathbf{e}_n$ will be the standard basis of K^n.

Example. Let $V = \mathbb{R}^3$, $\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$, $\mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$, $\mathbf{e}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ (the standard basis) and $\mathbf{e}'_1 = \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}$, $\mathbf{e}'_2 = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}$, $\mathbf{e}'_3 = \begin{pmatrix} -1 \\ 0 \\ 0 \end{pmatrix}$. Then

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

The following result was proved in MA106.

Proposition 0.5.1. *With the above notation, let* $\mathbf{v} \in V$*, and let* $\underline{\mathbf{v}}$ *and* $\underline{\mathbf{v}}'$ *denote the column vectors associated with* $\mathbf{v}$ *when we use the bases* $\mathbf{e}_1, \ldots, \mathbf{e}_n$ *and* $\mathbf{e}'_1, \ldots, \mathbf{e}'_n$*, respectively. Then* $P\underline{\mathbf{v}}' = \underline{\mathbf{v}}$*.*

So, in the example above, if we take $\mathbf{v} = \begin{pmatrix} 1 \\ -2 \\ 4 \end{pmatrix}$, then we have $\mathbf{v} = \mathbf{e}_1 - 2\mathbf{e}_2 + 4\mathbf{e}_3$ (obviously); so the coordinates of $\mathbf{v}$ in the basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ are $\underline{\mathbf{v}} = \begin{pmatrix} 1 \\ -2 \\ 4 \end{pmatrix}$.

On the other hand, we also have $\mathbf{v} = 2\mathbf{e}'_1 - 2\mathbf{e}'_2 - 3\mathbf{e}'_3$, so the coordinates of $\mathbf{v}$ in the basis $\{\mathbf{e}'_1, \mathbf{e}'_2, \mathbf{e}'_3\}$ are

$$\underline{\mathbf{v}}' = \begin{pmatrix} 2 \\ -2 \\ -3 \end{pmatrix},$$

and you can check that

$$P\underline{\mathbf{v}}' = \begin{pmatrix} 0 & 1 & -1 \\ 1 & 2 & 0 \\ 2 & 0 & 0 \end{pmatrix} \begin{pmatrix} 2 \\ -2 \\ -3 \end{pmatrix} = \begin{pmatrix} 1 \\ -2 \\ 4 \end{pmatrix} = \underline{\mathbf{v}},$$

just as Proposition 0.5.1 says.

This equation $P\underline{\mathbf{v}}' = \underline{\mathbf{v}}$ describes the change of coordinates associated with our basis change. If we want to compute the new coordinates from the old ones, we need to use the inverse matrix: $\underline{\mathbf{v}}' = P^{-1}\underline{\mathbf{v}}$. Thus, to enable calculations in the new basis we need both matrices P and P^{-1}. We'll be using this relationship over and over again, so make sure you're happy with it!

Which matrix, P or P^{-1} should be called the *basis change matrix* or *transition matrix* from the original basis $\mathbf{e}_1, \ldots, \mathbf{e}_n$ to the new basis $\mathbf{e}'_1, \ldots, \mathbf{e}'_n$?

Well, the books are split on this. As a historic quirk, the *basis change matrix* in Algebra-1 was always P and the *basis change matrix* in Linear Algebra was P^{-1} since around 2011. We continue with this noble tradition of calling P the *basis change matrix* because, otherwise, we risk introducing typos throughout the text.

Now let $T : V \to W$, $\mathbf{e}_i$, $\mathbf{f}_i$ and A be as in Subsection 0.4 above, and choose new bases $\mathbf{e}'_1, \ldots, \mathbf{e}'_n$ of V and $\mathbf{f}'_1, \ldots, \mathbf{f}'_m$ of W. Then

$$T(\mathbf{e}'_j) = \sum_{i=1}^{m} \beta_{ij}\mathbf{f}'_i \text{ for } 1 \leq j \leq n,$$

where $B = (\beta_{ij})$ is the $m \times n$ matrix of T with respect to the bases $\{\mathbf{e}'_i\}$ and $\{\mathbf{f}'_i\}$ of V and W. Let the $n \times n$ matrix $P = (p_{ij})$ be the basis change matrix for the original basis $\{\mathbf{e}_i\}$ and new basis $\{\mathbf{e}'_i\}$, and let the $m \times m$ matrix $Q = (q_{ij})$ be the basis change matrix for original basis $\{\mathbf{f}_i\}$ and new basis $\{\mathbf{f}'_i\}$. The following theorem was proved in MA106:

Theorem 0.5.2. *With the above notation, we have* $AP = QB$*, or equivalently* $B = Q^{-1}AP$*.*

In most of the applications in this module we will have $V = W$ $(= K^n)$, $\{\mathbf{e}_i\} = \{\mathbf{f}_i\}$, and $\{\mathbf{e}'_i\} = \{\mathbf{f}'_i\}$. So $P = Q$, and hence $B = P^{-1}AP$.

You may have noticed that the above is a bit messy, and it can be difficult to remember the definitions of P and Q (and to distinguish them from their inverses). Experience shows that students (and lecturers) have trouble with this. So here is what I hope is a better and more transparent way to think about change of basis in vector spaces and the way it affects representing matrices for linear maps:

First, we saw in the preceding section, that given:

1. a linear map $T : V \to W$, $\dim(V) = n$, $\dim(W) = m$;
2. ordered bases $\mathbf{E} = (\mathbf{e}_1, \ldots, \mathbf{e}_n)$ and $\mathbf{F} = (\mathbf{f}_1, \ldots, \mathbf{f}_m)$ of V and W;

we can associate to T an $m \times n$-matrix in $K^{m\times n}$ representing the linear map T with respect to the chosen ordered bases. Let's do our book-keeping neatly and try to keep track of all the data involved in our notation: let's denote this matrix temporarily by

$$\mathcal{M}(T)^{\mathbf{F}}_{\mathbf{E}}.$$

Note that the lower index $\mathbf{E}$ remembers the basis in the source V, the upper index $\mathbf{F}$ remembers the basis in the target, and $\mathcal{M}$ just stands for matrix. Of course that's a notational monstrosity, but you will see that for the purpose of explaining base change , it is very convenient. Indeed, choosing different ordered bases for V and W,

$$\mathbf{E}' = (\mathbf{e}'_1, \ldots, \mathbf{e}'_n) \text{ and } \mathbf{F}' = (\mathbf{f}'_1, \ldots, \mathbf{f}'_m)$$

the problem we want to address is: how are the matrices

$$A = \mathcal{M}(T)^{\mathbf{F}}_{\mathbf{E}} \text{ and } B = \mathcal{M}(T)^{\mathbf{F}'}_{\mathbf{E}'}$$

related? The answer to this is very easy if you remember from MA106 that matrix multiplication is compatible with composition of linear maps in the following sense: suppose

$$\begin{array}{ccccc} U & \xrightarrow{R} & V & \xrightarrow{S} & W \\ \mathbf{A} & & \mathbf{B} & & \mathbf{C} \end{array}$$

is a diagram of vector spaces and linear maps, and $\mathbf{A}, \mathbf{B}, \mathbf{C}$ are ordered bases in U, V, W. Then we have *the very basic fact* that

$$\mathcal{M}^{\mathbf{C}}_{\mathbf{A}}(S \circ R) = \mathcal{M}^{\mathbf{C}}_{\mathbf{B}}(S) \cdot \mathcal{M}^{\mathbf{B}}_{\mathbf{A}}(R).$$

Don't be intimidated by the formula and take a second to think about how natural this is! If we form the composite map $S \circ R$ and pass to the matrix representing it with respect to the given ordered bases, we can also get it by matrix-multiplying the matrices for S and R with respect to the chosen ordered bases! Now back to our problem above: consider the sequence of linear maps between vector spaces together with choices of ordered bases:

$$\begin{array}{ccccccc} V & \xrightarrow{\mathrm{id}_V} & V & \xrightarrow{T} & W & \xrightarrow{\mathrm{id}_W} & W \\ \mathbf{E}' & & \mathbf{E} & & \mathbf{F} & & \mathbf{F}' \end{array}$$

Applying the preceding basic fact gives

$$\mathcal{M}(T)^{\mathbf{F}'}_{\mathbf{E}'} = \mathcal{M}(\mathrm{id}_W)^{\mathbf{F}'}_{\mathbf{F}} \cdot \mathcal{M}(T)^{\mathbf{F}}_{\mathbf{E}} \cdot \mathcal{M}(\mathrm{id}_V)^{\mathbf{E}}_{\mathbf{E}'}.$$

Or, putting

$$P := \mathcal{M}(\mathrm{id}_V)^{\mathbf{E}}_{\mathbf{E}'}, \quad Q := \mathcal{M}(\mathrm{id}_W)^{\mathbf{F}}_{\mathbf{F}'}$$

and noticing that

$$\mathcal{M}(\mathrm{id}_W)^{\mathbf{F}'}_{\mathbf{F}} = (\mathcal{M}(\mathrm{id}_W)^{\mathbf{F}}_{\mathbf{F}'})^{-1}$$

we get

$$B = Q^{-1}AP$$

which *proves* Theorem 1.5.2, *but* also gives us a means to *remember* the right definitions of P and Q (which is important because that is the vital information and this is precisely the information students and lecturer always tend to forget): for example, $P = \mathcal{M}(\mathrm{id}_V)^{\mathbf{E}}_{\mathbf{E}'}$ is the matrix whose columns are the basis vectors $\mathbf{e}'_i$ written in the old basis $\mathbf{E}$ with basis vectors $\mathbf{e}_i$. You don't have to remember the entire discussion preceding Theorem 1.5.2 anymore (which is necessary to understand what the theorem says): it's all encoded in the notation! I hope you will never forget this base change formula again.

1 The Jordan Canonical Form

1.1 Introduction

Throughout this section V will be a vector space of dimension n over a field K, $T : V \to V$ will be a linear map, and A will be the matrix of T with respect to a fixed basis $\mathbf{e}_1, \ldots, \mathbf{e}_n$ of V (the same in the source and target V). Our aim is to find a new basis $\mathbf{e}'_1, \ldots, \mathbf{e}'_n$ for V, such that the matrix of T with respect to the new basis is as simple as possible. Equivalently (by Theorem 0.5.2), we want to find an invertible matrix P (the associated basis change matrix) such that $P^{-1}AP$ is as simple as possible.

(Recall that if B is a matrix which can be written in the form $B = P^{-1}AP$, we say B is *similar* to A. So a third way of saying the above is that we want to find a matrix that's similar to A, but which is as nice as possible.)

One particularly simple form of a matrix is a diagonal matrix. So we'd really rather like it if every matrix was similar to a diagonal matrix. But this won't work: we saw in MA106 that the matrix $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$, for example, is not similar to a diagonal matrix. (We say this matrix is not *diagonalizable*.)

The point of this section of the module is to show that although we can't always get to a diagonal matrix, we can get pretty close (at least if K is $\mathbb{C}$). Under this assumption, it can be proved that A is always similar to a matrix B of a certain type (called the *Jordan canonical form* or sometimes *Jordan normal form* of the matrix), which is not far off being diagonal: its only non-zero entries are on the diagonal or just above it.

1.2 Eigenvalues and eigenvectors

We start by summarising some of what we know from MA106 which is going to be relevant to us here.

If we can find some $0 \neq \mathbf{v} \in V$ and $\lambda \in K$ such that $T\mathbf{v} = \lambda\mathbf{v}$, or equivalently $A\underline{\mathbf{v}} = \lambda\underline{\mathbf{v}}$, then λ is an *eigenvalue*, and $\mathbf{v}$ a corresponding *eigenvector* of T (or of A).

From MA106, you have a theorem that tells you when a matrix is diagonalizable:

Proposition 1.2.1. *Let $T : V \to V$ be a linear map. Then the matrix of T is diagonal with respect to some basis of V if and only if V has a basis consisting of eigenvectors of T.*

This is a nice theorem, but it is also more or less a tautology, and it doesn't tell you how you might find such a basis! But there's one case where it's easy, as another theorem from MA106 tells us:

Proposition 1.2.2. *Let $\lambda_1, \ldots, \lambda_r$ be distinct eigenvalues of $T : V \to V$, and let $\mathbf{v}_1, \ldots, \mathbf{v}_r$ be corresponding eigenvectors. (So $T(\mathbf{v}_i) = \lambda_i\mathbf{v}_i$ for $1 \leq i \leq r$.) Then $\mathbf{v}_1, \ldots, \mathbf{v}_r$ are linearly independent.*

Corollary 1.2.3. *If the linear map $T : V \to V$ (or equivalently the $n \times n$ matrix A) has n distinct eigenvalues, where $n = \dim(V)$, then T (or A) is diagonalizable.*

1.3 The minimal polynomial

The minimal polynomial, while arguably not the most important player in the spectral theory of endomorphisms, derives its importance from the fact that it can be used to detect diagonalisability and also classifies nilpotent transformations, and we'll start with it to get off the ground.

If $A \in K^{n,n}$ is a square $n \times n$ matrix over K, and $p \in K[x]$ is a polynomial, then we can make sense of $p(A)$: we just calculate the powers of A in the usual way, and then plug them into the formula defining p, interpreting the constant term as a multiple of I_n.

For instance, if $K = \mathbb{Q}$, $p = 2x^2 - \frac{3}{2}x + 11$, and $A = \begin{pmatrix} 2 & 3 \\ 0 & 1 \end{pmatrix}$, then $A^2 = \begin{pmatrix} 4 & 9 \\ 0 & 1 \end{pmatrix}$, and

$$\begin{aligned} p(A) &= 2\begin{pmatrix} 4 & 9 \\ 0 & 1 \end{pmatrix} - \frac{3}{2}\begin{pmatrix} 2 & 3 \\ 0 & 1 \end{pmatrix} + 11\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 16 & 27/2 \\ 0 & 23/2 \end{pmatrix}. \end{aligned}$$

Warning. Notice that this is in general of course *not* the same as the matrix $\begin{pmatrix} p(2) & p(3) \\ p(0) & p(1) \end{pmatrix}$.

Theorem 1.3.1. *Let $A \in K^{n,n}$. Then there is some non-zero polynomial $p \in K[x]$ of degree at most n^2 such that $p(A)$ is the $n \times n$ zero matrix $\mathbf{0}_n$.*

Proof. The key thing to observe is that $K^{n,n}$, the space of $n \times n$ matrices over K, is itself a vector space over K. Its dimension is n^2.

Let's consider the set $\{I_n, A, A^2, \ldots, A^{n^2}\} \subset K^{n,n}$. Since this is a set of $n^2 + 1$ vectors in an n^2-dimensional vector space, there is a nontrivial linear dependency relation between them. That is, we can find constants $\lambda_0, \lambda_1, \ldots, \lambda_{n^2}$, not all zero, such that

$$\lambda_0 I_n + \cdots + \lambda_{n^2} A^{n^2} = \mathbf{0}_n.$$

Now we define the polynomial $p = \lambda_0 + \lambda_1 x + \cdots + \lambda_{n^2} x^{n^2}$. This isn't zero, and its degree is at most n^2. (It might be less, since λ_{n^2} might be 0.) Then that's it! □

Is there a way of finding a unique polynomial (of minimal degree) that A satisfies? To answer that question, we'll have to think a little bit about arithmetic in $K[x]$.

Note that we can do "division" with polynomials, a bit like with integers. We can divide one polynomial p (with $p \neq 0$) into another polynomial q and get a remainder with degree less than p. For example, if $q = x^5 - 3$, $p = x^2 + x + 1$, then we find $q = sp + r$ with $s = x^3 - x^2 + 1$ and $r = -x - 4$.

If the remainder is 0, so $q = sp$ for some s, we say "p divides q" and write this relation as $p \mid q$.

Finally, a polynomial with coefficients in a field K is called *monic* if the coefficient of the highest power of x is 1. So, for example, $x^3 - 2x^2 + x + 11$ is monic, but $2x^2 - x - 1$ is not.

Theorem 1.3.2. *Let A be an $n \times n$ matrix over K representing the linear map $T : V \to V$. Then*

(i) There is a unique monic non-zero polynomial $p(x)$ with minimal degree and coefficients in K such that $p(A) = 0$.

(ii) If $q(x)$ is any polynomial with $q(A) = 0$, then $p \mid q$.

Proof. (i) If we have any polynomial $p(x)$ with $p(A) = 0$, then we can make p monic by multiplying it by a constant. By Theorem 1.3.1, there exists such a $p(x)$, so there exists one of minimal degree. If we had two distinct monic polynomials $p_1(x)$, $p_2(x)$ of the same minimal degree with $p_1(A) = p_2(A) = 0$, then $p = p_1 - p_2$ would be a non-zero polynomial of smaller degree with $p(A) = 0$, contradicting the minimality of the degree, so p is unique.

(ii) Let $p(x)$ be the minimal monic polynomial in (i) and suppose that $q(A) = 0$. As we saw above, we can write $q = sp + r$ where r has smaller degree than p. If r is non-zero, then $r(A) = q(A) - s(A)p(A) = 0$ contradicting the minimality of p, so $r = 0$ and $p \mid q$. □

Definition 1.3.3. The unique monic non-zero polynomial $\mu_A(x)$ of minimal degree with $\mu_A(A) = 0$ is called the *minimal polynomial* of A.

We know that for $p \in K[x]$, $p(T) = \mathbf{0}_V$ if and only if $p(A) = \mathbf{0}_n$; so μ_A is also the unique monic polynomial of minimal degree such that $\mu_A(T) = 0$ (the minimal polynomial of T.) In particular, since similar matrices A and B represent the same linear map T, and their minimal polynomial is the same as that of T, we have

Proposition 1.3.4. *Similar matrices have the same minimal polynomial.*

By Theorem 1.3.1 and Theorem 1.3.2 (ii), we have

Corollary 1.3.5. *The minimal polynomial of an $n \times n$ matrix A has degree at most n^2.*

(In the next section, we'll see that we can do much better than this.)

Example. If D is a diagonal matrix, say

$$D = \begin{pmatrix} d_{11} & & \\ & \ddots & \\ & & d_{nn} \end{pmatrix},$$

then for any polynomial p we see that $p(D)$ is the diagonal matrix with entries

$$\begin{pmatrix} p(d_{11}) & & \\ & \ddots & \\ & & p(d_{nn}) \end{pmatrix}.$$

Hence $p(D) = 0$ if and only if $p(d_{ii}) = 0$ for all i. So for instance if

$$D = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 2 \end{pmatrix},$$

the minimal polynomial of D is the smallest-degree polynomial which has 2 and 3 as roots, which is clearly $\mu_D(x) = (x-2)(x-3) = x^2 - 5x + 6$.

We can generalize this example as follows

Proposition 1.3.6. *Let D be any diagonal matrix and let $\{\delta_1, \dots, \delta_r\}$ be the set of diagonal entries of D (i.e. without any repetitions, so the values $\delta_1, \dots, \delta_r$ are all different). Then we have*

$$\mu_D(x) = (x - \delta_1)(x - \delta_2) \dots (x - \delta_r).$$

Proof. As in the example, we have $p(D) = 0$ if and only if $p(\delta_i) = 0$ for all $i \in \{1, \dots, r\}$. The smallest-degree monic polynomial vanishing at these points is clearly the polynomial above. □

Corollary 1.3.7. *If A is any diagonalizable matrix, then $\mu_A(x)$ is a product of distinct linear factors.*

Proof. Clear from Proposition 1.3.6 and Proposition 1.3.4. □

Remark. *We'll see later in the course that this is a necessary and sufficient condition: A is diagonalizable* ***if and only if*** *$\mu_A(x)$ is a product of distinct linear factors. But we don't have enough tools to prove this theorem yet – be patient!*

1.4 The Cayley–Hamilton theorem

Theorem 1.4.1 (Cayley–Hamiton)**.** *Let $c_A(x)$ be the characteristic polynomial of the $n \times n$ matrix A over an arbitrary field K. Then $c_A(A) = \mathbf{0}$.*

Proof. Let's agree to drop the various subscripts and bold zeroes – it'll be obvious from context when we mean a zero matrix, zero vector, zero linear map, etc.

Recall from MA106 that, if B is any $n \times n$ matrix, the "adjugate matrix" of B is another matrix $\operatorname{adj}(B)$ which was constructed along the way to constructing the inverse of B. The entries of $\operatorname{adj}(B)$ are the "cofactors" of B: the (i, j) entry of B is $(-1)^{i+j} c_{ji}$ (note the transposition of indices here!), where $c_{ji} = \det(B_{ji})$, B_{ji} being the $(n-1) \times (n-1)$ matrix obtained by deleting the j-th row and the i-th column of B. The key property of $\operatorname{adj}(B)$ is that it satisfies

$$B \operatorname{adj}(B) = \operatorname{adj}(B) B = (\det B) I_n.$$

(Notice that if B is invertible, this just says that $\operatorname{adj}(B) = (\det B) B^{-1}$, but the adjugate matrix still makes sense even if B is not invertible.)

Let's apply this to the matrix $B = A - xI_n$. By definition, $\det(B)$ is the characteristic polynomial $c_A(x)$, so

$$\operatorname{adj}(A - xI_n)(A - xI_n) = c_A(x) I_n. \tag{2}$$

Now we use the following statement whose proof is obvious: suppose $P(x) = \sum_j P_j x^j$ and $Q(x) = \sum_k Q_k x^k$ are two polynomials in the indeterminate x with *matrix coefficients*; so P_j and Q_k are $n \times n$ matrices. Then the product of P and Q is $R(x) = \sum_l R_l x^l$ with

$$R_l = \sum_{j+k=l} P_j Q_k.$$

Then if an $n \times n$ matrix M commutes with all the coefficients of Q we have $R(M) = P(M)Q(M)$. We now apply this observation with

$$P(x) = \operatorname{adj}(A - xI_n), \quad Q(x) = A - xI_n, \quad M = A.$$

Since $Q(A) = 0$, we get $c_A(A) = 0$. □

Corollary 1.4.2. *For any $A \in K^{n,n}$, we have $\mu_A \mid c_A$, and in particular $\deg(\mu_A) \leq n$.*

Example. Let D be the diagonal matrix $\begin{pmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 2 \end{pmatrix}$ from the previous example. We saw above that $\mu_A(x) = (x-2)(x-3)$. However, it's easy to see that

$$c_A(x) = \begin{vmatrix} 3-x & 0 & 0 \\ 0 & 3-x & 0 \\ 0 & 0 & 2-x \end{vmatrix} = -(x-2)(x-3)^2.$$

How NOT to prove the Cayley–Hamilton theorem It is very tempting to try and prove the Cayley–Hamilton theorem as follows: we know that

$$c_A(x) = \det(A - xI_n),$$

so shouldn't we have

$$c_A(A) = \det(A - AI_n) = \det(A - A) = \det(0) = 0?$$

This is **wrong**. In fact, $c_A(A)$ is a matrix, and $\det(A - AI_n)$ is an element of K, so they are not even objects of the same type in general.

1.5 Calculating the minimal polynomial

We will present two methods for this.

Method 1 ("top down"; always never works in practice; it only works well if a benign lecturer or some other benevolent power reveals to you the factorisation of the characteristic polynomial into irreducibles).

Lemma 1.5.1. *Let λ be any eigenvalue of A. Then $\mu_A(\lambda) = 0$.*

Proof. Let $\underline{\mathbf{v}} \in K^{n,1}$ be an eigenvector corresponding to λ. Then $A^n\underline{\mathbf{v}} = \lambda^n\underline{\mathbf{v}}$, and hence for any polynomial $p \in K[x]$, we have

$$p(A)\underline{\mathbf{v}} = p(\lambda)\underline{\mathbf{v}}.$$

We know that $\mu_A(A)\underline{\mathbf{v}} = 0$, since $\mu_A(A)$ is the zero matrix. Hence $\mu_A(\lambda)\underline{\mathbf{v}} = 0$, and since $\underline{\mathbf{v}} \neq 0$ and $\mu_A(\lambda)$ is an element of K (not a matrix!), this can only happen if $\mu_A(\lambda) = 0$. □

This lemma, together with Cayley–Hamilton, give us very, very few possibilities for μ_A. Let's look at an example.

Example. Take $K = \mathbb{C}$ and let

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

This is rather large, but it has a fair few zeros, so you can calculate its characteristic polynomial fairly quickly by hand and find out that

$$c_A(x) = x^4 - 11x^3 + 45x^2 - 81x + 54.$$

Some trial and error shows that 2 is a root of this, and we find that

$$c_A(x) = (x-2)(x^3 - 9x^2 + 27x - 27) = (x-2)(x-3)^3.$$

So $\mu_A(x)$ divides $(x-2)(x-3)^3$. On the other hand, the eigenvalues of A are the roots of $c_A(x)$, namely $\{2,3\}$; and we know that μ_A must have each of these as roots. So the only possibilities for $\mu_A(x)$ are:

$$\mu_A(x) \in \left\{ \begin{matrix} (x-2)(x-3), \\ (x-2)(x-3)^2, \\ (x-2)(x-3)^3. \end{matrix} \right\}.$$

Some slightly tedious calculation shows that $(A-2)(A-3)$ isn't zero, and nor is $(A-2)(A-3)^2$, and so it must be the case that $(x-2)(x-3)^3$ is the minimal polynomial of A.

Method 2 ("bottom up"; this works well, also for large matrices)

This is based on

Lemma 1.5.2. *Let $T\colon V \to V$ be a linear map of an n-dimensional vector space V over K to itself, and suppose $W_1, \ldots, W_k$ are finitely many T-invariant subspaces spanning V. In other words, we require $T(W_i) \subset W_i$ and*

$$V = W_1 + \cdots + W_k$$

(but the sum doesn't have to be direct). Let $\mu_i(x)$ be the minimal polynomial of $T|_{W_i}$. Then

$$\mu_T(x) = \text{l.c.m.}\{\mu_1, \ldots, \mu_k\}.$$

In words: the minimal polynomial of T is the least common multiple of the minimal polynomials of the $T|_{W_i}$, $i = 1, \ldots, k$.

Proof. First we will show that setting

$$f(x) = \text{l.c.m.}\{\mu_1, \ldots, \mu_k\}$$

we have that $\mu_T(x)$ divides $f(x)$. Indeed, if $v \in W_i$, then writing $f(x) = g_i(x)\mu_i(x)$ we calculate

$$f(T)v = g_i(T)\mu_i(T)v = g_i(T|_{W_i})\mu_i(T|_{W_i})v = 0$$

since $\mu_i(T\mid_{W_i}) = 0$. Since this argument is valid for any i and the W_i's span V, we conclude that $f(T)$ annihilates all of V hence is the zero linear map on V. Thus $f(x)$ is divisible by $\mu_T(x)$.

But $f(x)$ also divides $\mu_T(x)$: indeed, $\mu_T(T) = 0$, and hence also $\mu_T(T\mid_{W_i}) = 0$ for any i. Hence, $\mu_T(x)$ is divisible by any $\mu_i(x)$, and consequently by their least common multiple, too.

Since both $f(x)$ and $\mu_T(x)$ are monic, they must be equal. □

The preceding Lemma allows us to come up with a sensible algorithm to compute the minimal polynomial of T:

Algorithm:

Pick any $v \neq 0$ in V and set

$$W = \text{span}\{v,\ T(v),\ T^2(v), \dots\}.$$

By definition, W is T-invariant. Now let d be the minimal positive integer such that

$$v,\ T(v), \dots, T^d(v)$$

are linearly dependent. In particular,

$$v,\ T(v), \dots, T^{d-1}(v)$$

are linearly independent, and if $p(x)$ is any polynomial of degree $\leq d-1$, $p(T)v$ will never be zero: hence the minimal polynomial $\mu_{T|_W}(x)$ has degree $\geq d$. There is a nontrivial linear dependency relation of the form

$$T^d(v) + c_{d-1}T^{d-1}(v) + \cdots + c_1 T(v) + c_0 v = 0.$$

Consider the polynomial

$$x^d + c_{d-1}x^{d-1} + \cdots + c_1 x + c_0.$$

We claim this must be $\mu_{T|_W}(x)$: indeed, it is monic, $\mu_{T|_W}(T\mid_W)$ annihilates W, and $\mu_{T|_W}(x)$ is of smallest possible degree d with this property. Therefore we have computed $\mu_{T|_W}(x)$, and we can set

$$W_1 := W,\ \mu_1(x) := \mu_{T|_W}(x).$$

If $W_1 \neq V$, pick a vector v' not in W_1 and repeat the preceding procedure, leading to a T-invariant subspace W_2 such that the span of W_1 and W_2 will be strictly larger than W_1. Since V is finite-dimensional, after finitely many steps, we compute in this way $W_1, \dots, W_k$ and polynomials $\mu_1(x), \dots, \mu_k(x)$ satisfying the conditions in Lemma 1.5.2. Since computing a least common multiple presents no problem (use the Euclidean algorithm for polynomials repeatedly), we are done.

1.6 Jordan chains and Jordan blocks

We'll now consider some special vectors attached to our matrix, which satisfy a condition a little like eigenvectors (but weaker). These will be the stepping-stones towards the Jordan canonical form.

As always let $T\colon V \to V$ be a linear self-map of an n-dimensional K-vector space. In particular, the vector space could be $K^{n,1}$ and then the linear map would be given by some matrix A. Note that choosing an ordered basis in V amounts to fixing a bijective linear map $\beta\colon K^{n,1} \to V$ and $\beta^{-1} \circ T \circ \beta$ is then a linear map from $K^{n,1}$ to itself given by the matrix A of T with respect to that chosen ordered basis. In the following, we work with T or A depending on the situation- it comes down to the same thing in every instance.

Definition 1.6.1. A non-zero vector $\mathbf{v} \in V$ such that $(T - \lambda \mathrm{id}_V)^i \mathbf{v} = 0$, for some $i > 0$, is called a *generalised eigenvector* of T with respect to the eigenvalue λ.

Note that, for fixed $i > 0$,

$$N_i(T, \lambda) := \{\, \mathbf{v} \in V \mid (T - \lambda \mathrm{id}_V)^i \mathbf{v} = 0 \,\}$$

is the nullspace of $(T - \lambda \mathrm{id}_V)^i$, and is called the *generalised eigenspace of index i* of T with respect to λ.

The generalised eigenspace of index 1 is just called the *eigenspace* of T w.r.t. λ; it consists of the eigenvectors of T w.r.t. λ, together with the zero vector. We sometimes also consider the *full generalised eigenspace* of T w.r.t. λ, which is the set of all generalised eigenvectors together with the zero vector; this is the union of the generalised eigenspaces of index i over all $i \in \mathbb{N}$.

We can arrange generalised eigenvectors into "chains":

Definition 1.6.2. A *Jordan chain of length k* is a sequence of non-zero vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k \in V$ that satisfies

$$T\mathbf{v}_1 = \lambda \mathbf{v}_1, \;\; T\mathbf{v}_i = \lambda \mathbf{v}_i + \mathbf{v}_{i-1}, \;\; 2 \le i \le k,$$

for some eigenvalue λ of T.

Equivalently, $(T - \lambda \mathrm{id}_V)\mathbf{v}_1 = \mathbf{0}$ and $(T - \lambda \mathrm{id}_V)\mathbf{v}_i = \mathbf{v}_{i-1}$ for $2 \le i \le k$, so $(T - \lambda \mathrm{id}_V)^i \mathbf{v}_i = \mathbf{0}$ for $1 \le i \le k$. Thus all of the vectors in a Jordan chain are generalised eigenvectors, and $\mathbf{v}_i$ lies in the generalised eigenspace of index i.

Lemma 1.6.3. *Let $\mathbf{v}_1, \ldots, \mathbf{v}_k \in V$ be a Jordan chain of length k for eigenvalue λ of T. Then $\mathbf{v}_1, \ldots, \mathbf{v}_k$ are linearly independent.*

Proof. Exercise. □

For example, take $K = \mathbb{C}$ and consider the matrix

$$A = \begin{pmatrix} 3 & 1 & 0 \\ 0 & 3 & 1 \\ 0 & 0 & 3 \end{pmatrix}.$$

We see that, for $\{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3\}$ the standard basis of $\mathbb{C}^{3,1}$, we have $A\mathbf{b}_1 = 3\mathbf{b}_1$, $A\mathbf{b}_2 = 3\mathbf{b}_2 + \mathbf{b}_1$, $A\mathbf{b}_3 = 3\mathbf{b}_3 + \mathbf{b}_2$, so $\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3$ is a Jordan chain of length 3 for the eigenvalue 3 of A. The generalised eigenspaces of index 1, 2, and 3 are respectively $\langle \mathbf{b}_1 \rangle$, $\langle \mathbf{b}_1, \mathbf{b}_2 \rangle$, and $\langle \mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3 \rangle$.

Note that this isn't the only possible Jordan chain. Obviously, $\{17\mathbf{b}_1, 17\mathbf{b}_2, 17\mathbf{b}_3\}$ would be a Jordan chain; but there are more devious possibilities – you can check that $\{\mathbf{b}_1, \mathbf{b}_1 + \mathbf{b}_2, \mathbf{b}_2 + \mathbf{b}_3\}$ is a Jordan chain, so there can be several Jordan chains with the same first vector. On the other hand, two Jordan chains with the same *last* vector are the same and in particular have the same length.

What are the generalised eigenspaces here? The only eigenvalue is 3. For this eigenvalue, the generalised eigenspace of index 1 is $\langle \mathbf{b}_1 \rangle$ (the linear span of $\mathbf{b}_1$); the generalised eigenspace of index 2 is $\langle \mathbf{b}_1, \mathbf{b}_2 \rangle$; and the generalised eigenspace of index 3 is the whole space $\langle \mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3 \rangle$. So the dimensions are $(1, 2, 3)$.

Definition 1.6.4. We define the *Jordan block* of degree k with eigenvalue λ to be the $k \times k$ matrix $J_{\lambda,k}$ whose entries are given by

$$\gamma_{ij} = \begin{cases} \lambda & \text{if } j = i \\ 1 & \text{if } j = i + 1 \\ 0 & \text{otherwise.} \end{cases}$$

So, for example,

$$J_{1,2} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \qquad J_{4i-7,3} = \begin{pmatrix} 4i-7 & 1 & 0 \\ 0 & 4i-7 & 1 \\ 0 & 0 & 4i-7 \end{pmatrix}, \qquad \text{and} \qquad J_{0,4} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

are Jordan blocks.

It should be clear that the matrix of T with respect to the basis $\mathbf{v}_1, \ldots, \mathbf{v}_n$ of V is a Jordan block of degree n if and only if $\mathbf{v}_1, \ldots, \mathbf{v}_n$ is a Jordan chain for T.

Note that the minimal polynomial of $J_{\lambda,k}$ is equal to $(x - \lambda)^k$, and the characteristic polynomial is $(\lambda - x)^k$.

Warning. Some authors put the 1's below rather than above the main diagonal in a Jordan block. This corresponds to writing the Jordan chain in reverse order. This is an arbitrary choice but in this course we stick to our convention - when you read other notes/books be careful to check which convention they use.

1.7 Jordan bases and the Jordan canonical form

Definition 1.7.1. A *Jordan basis* for T is a basis of V consisting of one or more Jordan chains strung together.

Such a basis will look like

$$w_{11}, \ldots, w_{1k_1}, w_{21}, \ldots, w_{2k_2}, \ldots, w_{s1}, \ldots, w_{sk_s},$$

where, for $1 \leq i \leq s$, $w_{i1}, \ldots, w_{ik_i}$ is a Jordan chain (for some eigenvalue λ_i).

We denote the $m \times n$ matrix in which all entries are 0 by $\mathbf{0}_{m,n}$. If A is an $m \times m$ matrix and B an $n \times n$ matrix, then we denote the $(m+n) \times (m+n)$ matrix with block form

$$\left(\begin{array}{c|c} A & \mathbf{0}_{m,n} \\ \hline \mathbf{0}_{n,m} & B \end{array}\right),$$

by $A \oplus B$, the *direct sum* of A and B. For example

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

It's clear that the matrix of T with respect to a Jordan basis is the direct sum $J_{\lambda_1,k_1} \oplus J_{\lambda_2,k_2} \oplus \cdots \oplus J_{\lambda_s,k_s}$ of the corresponding Jordan blocks.

The following lemma is left as an exercise.

Lemma 1.7.2. *Suppose that $M = A \oplus B$. Then the characteristic polynomial $c_M(x)$ is the product of $c_A(x)$ and $c_B(x)$, and the minimal polynomial $\mu_M(x)$ is the lowest common multiple of $\mu_A(x)$ and $\mu_B(x)$.*

It is now time for us to state the main theorem of this section, which says that if K is the complex numbers $\mathbb{C}$, then Jordan bases exist.

Theorem 1.7.3. *Let $T\colon V \to V$ be a linear self-map of an n-dimensional complex vector space V. Then there exists a Jordan basis for T. In particular, any $n \times n$ matrix A over $\mathbb{C}$ is similar to a matrix J which is a direct sum of Jordan blocks. The Jordan blocks occurring in J are uniquely determined by A. This J is said to be the* Jordan canonical form *(JCF) or sometimes* Jordan normal form *of A.*

In fact, the uniqueness follows from the following more precise statement: Let λ be an eigenvalue of a matrix $A \in \mathbb{C}^{n,n}$, and let J be the JCF of A. Then

(i) The number of Jordan blocks of J with eigenvalue λ is equal to $\mathrm{nullity}(A - \lambda I_n)$.

(ii) More generally, for $i > 0$, the number of Jordan blocks of J with eigenvalue λ and degree at least i is equal to $\mathrm{nullity}((A - \lambda I_n)^i) - \mathrm{nullity}((A - \lambda I_n)^{i-1})$.

Remark. *The only reason we need $K = \mathbb{C}$ in this theorem is to ensure that T (or A) has at least one eigenvalue. If $K = \mathbb{R}$ (or $\mathbb{Q}$), we'd run into trouble with $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$; this matrix has* no *eigenvalues, since $c_A(x) = x^2 + 1$ has no roots in K. So it certainly has no Jordan chains. The theorem is valid more generally for any field K which is such that any non-constant polynomial in $K[x]$ has a root in K (one calls such fields algebraically closed; there are many more of them out there than just $\mathbb{C}$).*

Proof. EXISTENCE:

We proceed by induction on $n = \dim(V)$. The case $n = 1$ is clear.

We are looking for a vector space of dimension less than n, related to T to apply our inductive hypothesis to. Let λ be an eigenvalue of T and set $S := T - \lambda I_V$. Then we let $U = \mathrm{im}(S)$ and $m = \dim(U)$. Using the Rank-Nullity Theorem we see that $m = \mathrm{rank}(S) = n - \mathrm{nullity}(S) < n$, because there exists at least one eigenvector of T for λ, which lies in the nullspace of $S = T - \lambda I_V$. For $\mathbf{u} \in U$, we have $\mathbf{u} = S(\mathbf{v})$ for some $\mathbf{v} \in V$, and hence $T(\mathbf{u}) = TS(\mathbf{v}) = ST(\mathbf{v}) \in \mathrm{im}(S) = U$. Note that $TS = ST$ because $T(T - \lambda I_V) = T^2 - T\lambda I_V = T^2 - \lambda I_V T = (T - \lambda I_V)T$. So T maps U to U and thus T restricts to a linear map $T_U : U \to U$. Since $m < n$, we can apply our inductive hypothesis to T_U to deduce that U has a basis

$\mathbf{e}_1, \ldots, \mathbf{e}_m$, which is a disjoint union of Jordan chains for T_U (for all eigenvalues of T_U).

It is our job to show how to extend this Jordan basis of U to one of V. We do this in two stages. Firstly, let $\mathbf{v}_1, \ldots, \mathbf{v}_k$ be one of the l disjoint Jordan chains for eigenvalue λ for T_U (where l could be 0), so we have $T(\mathbf{v}_1) = T_U(\mathbf{v}_1) = \lambda \mathbf{v}_1,\ T(\mathbf{v}_i) = T_U(\mathbf{v}_i) = \lambda \mathbf{v}_i + \mathbf{v}_{i-1},\ 2 \leq i \leq k$. Now, since $\mathbf{v}_k \in U = \operatorname{im} S = \operatorname{im}(T - \lambda I_V)$, we can find $\mathbf{v}_{k+1} \in V$ with $T(\mathbf{v}_{k+1}) = \lambda \mathbf{v}_{k+1} + \mathbf{v}_k$, thereby extending the chain by an extra vector of V.

We do this for each of the l disjoint chains for λ and so at this point we have adjoined l new vectors to the basis. Let us call these new vectors $\mathbf{w}_1, \ldots, \mathbf{w}_l$.

For the second stage, observe that the first vector in each of the l chains lies in the eigenspace of T_U for λ. We know that the dimension of the eigenspace of T for λ is the dimension of the nullspace of S, which is $n - m$. So we can adjoin $(n-m) - l$ (which could be 0) further eigenvectors of T to the l that we have already to complete a basis of the nullspace of $(T - \lambda I_V)$. Let us call these $(n-m) - l$ new vectors $\mathbf{w}_{l+1}, \ldots, \mathbf{w}_{n-m}$. They are adjoined to our basis of V in the second stage. They each form a Jordan chain of length 1 (since they are not in the image of $S = T - \lambda I_V$), so we now have a collection of n vectors which form a disjoint union of Jordan chains.

To complete the proof, we need to show that these n vectors form a basis of V, for which it is enough to show that they are linearly independent.

Suppose that

$$\alpha_1 \mathbf{w}_1 + \cdots + \alpha_{n-m} \mathbf{w}_{n-m} + \mathbf{x} = \mathbf{0}, \text{ where } \mathbf{x} = \beta_1 \mathbf{e}_1 + \ldots + \beta_m \mathbf{e}_m\,, \tag{3}$$

a linear combination of the basis vectors $\mathbf{e}_1, \ldots, \mathbf{e}_m$ of U. We now apply S to both sides of this equation, recalling that $S(\mathbf{w}_{l+i}) = 0$ for $i \geq 1$, by definition.

$$\alpha_1 S(\mathbf{w}_1) + \cdots + \alpha_l S(\mathbf{w}_l) + S(\mathbf{x}) = \mathbf{0}. \tag{4}$$

By the construction of the $\mathbf{w}_i$, each of the $S(\mathbf{w}_i)$ for $1 \leq i \leq l$ is the last member of one of the l Jordan chains for T_U. Let this set of l vectors e_j be $L = \{j \mid \mathbf{e}_j = S(\mathbf{w}_i) \text{ for some } 1 \leq i \leq 1\}$. Now examine the last term

$$S(\mathbf{x}) = (T - \lambda I_n)(\mathbf{x}) = (T_U - \lambda I_m)(\mathbf{x}) = \beta_1 (T_U - \lambda I_m)(\mathbf{e}_1) + \cdots + \beta_m (T_U - \lambda I_m)(\mathbf{e}_m)\,.$$

Each $(T_U - \lambda I_m)(\mathbf{e}_j)$ is a linear combination of the basis vectors of U from the subset

$$\{\mathbf{e}_1, \ldots, \mathbf{e}_m\} \setminus \{\mathbf{e}_j \mid j \in L\}.$$

Indeed, this follows because after application of S we must have 'moved' down our Jordan chains for T_U. It now follows from the linear independence of the basis $\mathbf{e}_1, \ldots \mathbf{e}_m$, that $\alpha_i = 0$ for all $1 \leq i \leq l$.

So Equation (4) is now just

$$S(\mathbf{x}) = \mathbf{0}\,,$$

and so $\mathbf{x}$ is in the eigenspace of T_U for the eigenvalue λ. Equation (3) looks like

$$\alpha_{l+1} \mathbf{w}_{l+1} + \cdots + \alpha_{n-m} \mathbf{w}_{n-m} + \mathbf{x} = \mathbf{0}. \tag{5}$$

By construction, $\mathbf{w}_{l+1}, \ldots, \mathbf{w}_{n-m}$ extend a basis of the eigenspace of T_U to a basis of the eigenspace of T for λ. Lemma 0.2.2 now applies (to the eigenspace of T), yielding $\alpha_i = 0$ for $l + 1 \leq i \leq n - m$ and $\mathbf{x} = 0$. Since $\mathbf{e}_1, \ldots, \mathbf{e}_m$ is a basis for U, we must have all $\beta_j = 0$, which completes the proof.

UNIQUENESS: The corresponding generalised eigenspaces of A and J have the same dimensions, so we may assume WLOG that $A = J$. So A is a direct sum of several Jordan blocks $J_{\lambda_1,k_1} \oplus \cdots \oplus J_{\lambda_s,k_s}$.

However, it's easy to see that the dimension of the generalised λ-eigenspace of index i of a direct sum $A \oplus B$ is the sum of the dimensions of the generalised λ eigenspaces of index i of A and of B. Hence it suffices to prove the theorem for a single Jordan block $J_{\lambda,k}$.

But we know that $(J_{\lambda,k} - \lambda I_k)^i$ has a single diagonal line of ones i places above the diagonal, for $i < k$, and is 0 for $i \geq k$. Hence the dimension of its kernel is i for $0 \leq i \leq k$ and k for $i \geq k$. This clearly implies the theorem when A is a single Jordan block, and hence for any A.

□

Theorem 1.7.4 (Consequences of the JCF). *Let $A \in \mathbb{C}^{n,n}$, and $\{\lambda_1, \ldots, \lambda_r\}$ be the set of eigenvalues of A.*

(i) The characteristic polynomial of A is

$$(-1)^n \prod_{i=1}^{r} (x - \lambda_i)^{a_i},$$

where a_i is the sum of the degrees of the Jordan blocks of A of eigenvalue λ_i.

(ii) The minimal polynomial of A is

$$\prod_{i=1}^{r} (x - \lambda_i)^{b_i},$$

where b_i is the largest among the degrees of the Jordan blocks of A of eigenvalue λ_i.

(iii) A is diagonalizable if and only if $\mu_A(x)$ has no repeated factors.

Proof. We know that the characteristic and minimal polynomials of A and J, its JCF, are the same. So the first two parts follow from applying Lemma 1.7.2 (multiple times) to J. For the last part, notice that if A is diagonalizable, the JCF of A is just the diagonal form of A; since the JCF is unique, it follows that A is diagonalizable if and only if every Jordan block for A has size 1, so all of the numbers b_i are 1. □

1.8 The JCF when n=2 and 3

When $n = 2$ and $n = 3$, the JCF can be deduced just from the minimal and characteristic polynomials. Let us consider these cases.

When $n = 2$, we have either two distinct eigenvalues λ_1, λ_2, or a single repeated eigenvalue λ_1. If the eigenvalues are distinct, then by Corollary 1.2.3 A is diagonalizable and the JCF is the diagonal matrix $J_{\lambda_1,1} \oplus J_{\lambda_2,1}$.

Example 3. $A = \begin{pmatrix} 1 & 4 \\ 1 & 1 \end{pmatrix}$. We calculate $c_A(x) = x^2 - 2x - 3 = (x-3)(x+1)$, so there are two distinct eigenvalues, 3 and -1. Associated eigenvectors are $\begin{pmatrix} 2 \\ 1 \end{pmatrix}$ and $\begin{pmatrix} -2 \\ 1 \end{pmatrix}$, so we put $P = \begin{pmatrix} 2 & -2 \\ 1 & 1 \end{pmatrix}$ and then $P^{-1}AP = \begin{pmatrix} 3 & 0 \\ 0 & -1 \end{pmatrix}$.

If the eigenvalues are equal, then there are two possible JCFs, $J_{\lambda_1,1} \oplus J_{\lambda_1,1}$, which is a scalar matrix, and $J_{\lambda_1,2}$. The minimal polynomial is respectively $(x - \lambda_1)$ and $(x - \lambda_1)^2$ in these two cases. In fact, these cases can be distinguished without any calculation whatsoever, because in the first case A is a scalar multiple of the identity, and in particular A is already in JCF.

In the second case, a Jordan basis consists of a single Jordan chain of length 2. To find such a chain, let $\mathbf{v}_2$ be any vector for which $(A - \lambda_1 I_2)\mathbf{v}_2 \neq \mathbf{0}$ and let $\mathbf{v}_1 = (A - \lambda_1 I_2)\mathbf{v}_2$. (Note that, in practice, it is often easier to find the vectors in a Jordan chain in reverse order.)

Example 4. $A = \begin{pmatrix} 1 & 4 \\ -1 & -3 \end{pmatrix}$. We have $c_A(x) = x^2 + 2x + 1 = (x+1)^2$, so there is a single eigenvalue -1 with multiplicity 2. Since the first column of $A + I_2$ is non-zero, we can choose $\mathbf{v}_2 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\mathbf{v}_1 = (A + I_2)\mathbf{v}_2 = \begin{pmatrix} 2 \\ -1 \end{pmatrix}$, so $P = \begin{pmatrix} 2 & 1 \\ -1 & 0 \end{pmatrix}$ and $P^{-1}AP = \begin{pmatrix} -1 & 1 \\ 0 & -1 \end{pmatrix}$.

Now let $n = 3$. If there are three distinct eigenvalues, then A is diagonalizable.

Suppose that there are two distinct eigenvalues, so one has multiplicity 2, and the other has multiplicity 1. Let the eigenvalues be $\lambda_1, \lambda_1, \lambda_2$, with $\lambda_1 \neq \lambda_2$. Then there are two possible JCFs for A, $J_{\lambda_1,1} \oplus J_{\lambda_1,1} \oplus J_{\lambda_2,1}$ and $J_{\lambda_1,2} \oplus J_{\lambda_2,1}$, and the minimal polynomial is $(x - \lambda_1)(x - \lambda_2)$ in the first case and $(x - \lambda_1)^2(x - \lambda_2)$ in the second.

In the first case, a Jordan basis is a union of three Jordan chains of length 1, each of which consists of an eigenvector of A.

Example 5. $A = \begin{pmatrix} 2 & 0 & 0 \\ 1 & 5 & 2 \\ -2 & -6 & -2 \end{pmatrix}$. Then

$$c_A(x) = (2-x)[(5-x)(-2-x)+12] = (2-x)(x^2-3x+2) = (2-x)^2(1-x).$$

We know from the theory above that the minimal polynomial must be $(x - 2)(x-1)$ or $(x-2)^2(x-1)$. We can decide which simply by calculating $(A - 2I_3)(A - I_3)$ to test whether or not it is 0. We have

$$A - 2I_3 = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 3 & 2 \\ -2 & -6 & -4 \end{pmatrix}, \quad A - I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 4 & 2 \\ -2 & -6 & -3 \end{pmatrix},$$

and the product of these two matrices is 0, so $\mu_A = (x-2)(x-1)$.

The eigenvectors $\mathbf{v}$ for $\lambda_1 = 2$ satisfy $(A - 2I_3)\mathbf{v} = \mathbf{0}$, and we must find two linearly independent solutions; for example we can take $\mathbf{v}_1 = \begin{pmatrix} 0 \\ 2 \\ -3 \end{pmatrix}$, $\mathbf{v}_2 =$

$\begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$. An eigenvector for the eigenvalue 1 is $\mathbf{v}_3 = \begin{pmatrix} 0 \\ 1 \\ -2 \end{pmatrix}$, so we can choose

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

and then $P^{-1}AP$ is diagonal with entries $2, 2, 1$.

In the second case, there are two Jordan chains, one for λ_1 of length 2, and one for λ_2 of length 1. For the first chain, we need to find a vector $\mathbf{v}_2$ with $(A - \lambda_1 I_3)^2\mathbf{v}_2 = \mathbf{0}$ but $(A - \lambda_1 I_3)\mathbf{v}_2 \neq \mathbf{0}$, and then the chain is $\mathbf{v}_1 = (A - \lambda_1 I_3)\mathbf{v}_2, \mathbf{v}_2$. For the second chain, we simply need an eigenvector for λ_2.

Example 6. $A = \begin{pmatrix} 3 & 2 & 1 \\ 0 & 3 & 1 \\ -1 & -4 & -1 \end{pmatrix}$. Then

$$c_A(x) = (3-x)[(3-x)(-1-x)+4] - 2 + (3-x) = -x^3 + 5x^2 - 8x + 4 = (2-x)^2(1-x),$$

as in Example 3. We have

$$A - 2I_3 = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 1 \\ -1 & -4 & -3 \end{pmatrix}, \quad (A-2I_3)^2 = \begin{pmatrix} 0 & 0 & 0 \\ -1 & -3 & -2 \\ 2 & 6 & 4 \end{pmatrix}, \quad (A - I_3) = \begin{pmatrix} 2 & 2 & 1 \\ 0 & 2 & 1 \\ -1 & -4 & -2 \end{pmatrix}.$$

and we can check that $(A - 2I_3)(A - I_3)$ is non-zero, so we must have $\mu_A = (x-2)^2(x-1)$.

For the Jordan chain of length 2, we need a vector with $(A - 2I_3)^2\mathbf{v}_2 = \mathbf{0}$ but $(A - 2I_3)\mathbf{v}_2 \neq \mathbf{0}$, and we can choose $\mathbf{v}_2 = \begin{pmatrix} 2 \\ 0 \\ -1 \end{pmatrix}$. Then $\mathbf{v}_1 = (A - 2I_3)\mathbf{v}_2 = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$. An eigenvector for the eigenvalue 1 is $\mathbf{v}_3 = \begin{pmatrix} 0 \\ 1 \\ -2 \end{pmatrix}$, so we can choose

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

and then

$$P^{-1}AP = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Finally, suppose that there is a single eigenvalue, λ_1, so $c_A = (\lambda_1 - x)^3$. There are three possible JCFs for A, namely $J_{\lambda_1,1} \oplus J_{\lambda_1,1} \oplus J_{\lambda_1,1}$, $J_{\lambda_1,2} \oplus J_{\lambda_1,1}$, and $J_{\lambda_1,3}$, and the minimal polynomials in the three cases are $(x - \lambda_1)$, $(x - \lambda_1)^2$, and $(x - \lambda_1)^3$, respectively.

In the first case, J is a scalar matrix, and $A = PJP^{-1} = J$, so this is recognisable immediately.

In the second case, there are two Jordan chains, one of length 2 and one of length 1. For the first, we choose $\mathbf{v}_2$ with $(A - \lambda_1 I_3)\mathbf{v}_2 \neq 0$, and let $\mathbf{v}_1 = (A - \lambda_1 I_3)\mathbf{v}_2$. (This case is easier than the case illustrated in Example 4, because we have $(A - \lambda_1 I_3)^2\mathbf{v} = \mathbf{0}$ for all $\mathbf{v} \in \mathbb{C}^{3,1}$.) For the second Jordan chain, we choose $\mathbf{v}_3$ to be an eigenvector for λ_1 such that $\mathbf{v}_1$ and $\mathbf{v}_3$ are linearly independent.

Example 7. $A = \begin{pmatrix} 0 & 2 & 1 \\ -1 & -3 & -1 \\ 1 & 2 & 0 \end{pmatrix}$. Then

$$c_A(x) = -x[(3+x)x+2] - 2(x+1) - 2 + (3+x) = -x^3 - 3x^2 - 3x - 1 = -(1+x)^3.$$

We have

$$A + I_3 = \begin{pmatrix} 1 & 2 & 1 \\ -1 & -2 & -1 \\ 1 & 2 & 1 \end{pmatrix},$$

and we can check that $(A + I_3)^2 = 0$. The first column of $A + I_3$ is non-zero, so $(A + I_3)\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \neq \mathbf{0}$, and we can choose $\mathbf{v}_2 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ and $\mathbf{v}_1 = (A + I_3)\mathbf{v}_2 = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$. For $\mathbf{v}_3$ we need to choose a vector which is not a multiple of $\mathbf{v}_1$ such that $(A + I_3)\mathbf{v}_3 = \mathbf{0}$, and we can choose $\mathbf{v}_3 = \begin{pmatrix} 0 \\ 1 \\ -2 \end{pmatrix}$. So we have

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

and then

$$P^{-1}AP = \begin{pmatrix} -1 & 1 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

In the third case, there is a single Jordan chain, and we choose $\mathbf{v}_3$ such that $(A - \lambda_1 I_3)^2\mathbf{v}_3 \neq 0$, $\mathbf{v}_2 = (A - \lambda_1 I_3)\mathbf{v}_3$, $\mathbf{v}_1 = (A - \lambda_1 I_3)^2\mathbf{v}_3$.

Example 8. $A = \begin{pmatrix} 0 & 1 & 0 \\ -1 & -1 & 1 \\ 1 & 0 & -2 \end{pmatrix}$. Then

$$c_A(x) = -x[(2+x)(1+x)] - (2+x) + 1 = -(1+x)^3.$$

We have

$$A + I_3 = \begin{pmatrix} 1 & 1 & 0 \\ -1 & 0 & 1 \\ 1 & 0 & -1 \end{pmatrix}, \quad (A + I_3)^2 = \begin{pmatrix} 0 & 1 & 1 \\ 0 & -1 & -1 \\ 0 & 1 & 1 \end{pmatrix},$$

so $(A + I_3)^2 \neq 0$ and $\mu_A = (x+1)^3$. For $\mathbf{v}_3$, we need a vector that is not in the nullspace of $(A + I_3)^2$. Since the second column, which is the image of $\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$

is non-zero, we can choose $\mathbf{v}_3 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$, and then $\mathbf{v}_2 = (A + I_3)\mathbf{v}_3 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ and $\mathbf{v}_1 = (A + I_3)\mathbf{v}_2 = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$. So we have

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

and then

$$P^{-1}AP = \begin{pmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & -1 \end{pmatrix}.$$

1.9 Examples for $n \geq 4$

In the examples above, we could tell what the sizes of the Jordan blocks were for each eigenvalue from the dimensions of the eigenspaces, since the dimension of the eigenspace for each eigenvalue λ is the number of blocks for that eigenvalue. This doesn't work for $n = 4$: for instance, the matrices

$$A_1 = J_{\lambda,2} \oplus J_{\lambda,2}$$

and

$$A_2 = J_{\lambda,3} \oplus J_{\lambda,1}$$

both have only one eigenvalue (λ) with the eigenspace being of dimension 2.

(Knowing the minimal polynomial helps, but it's a bit of a pain to calculate – generally the easiest way to find the minimal polynomial is to calculate the JCF first! Worse still, it still doesn't uniquely determine the JCF in large dimensions, since

$$A_3 = J_{\lambda,3} \oplus J_{\lambda,3} \oplus J_{\lambda,1}$$

and

$$A_4 = J_{\lambda,3} \oplus J_{\lambda,2} \oplus J_{\lambda,2}$$

have the same minimal polynomial, the same characteristic polynomial, and the same number of blocks.)

In general, we can compute the JCF from the dimensions of the generalised eigenspaces. Notice that the matrices A_1 and A_2 can be distinguished by looking at the dimensions of their generalised eigenspaces: the generalised eigenspace for λ of index 2 has dimension 4 for A_1 (it's the whole space) but dimension only 3 for A_2.

Example 9. $A = \begin{pmatrix} -1 & -3 & -1 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 3 & 1 & -1 \end{pmatrix}$. Then $c_A(x) = (-1-x)^2(2-x)^2$, so there are two eigenvalues $-1, 2$, both with multiplicity 2. There are four possibilities for the JCF (one or two blocks for each of the two eigenvalues). We could determine the JCF by computing the minimal polynomial μ_A but it is probably easier to compute the nullities of the eigenspaces and use the second part of Theorem 1.7.3. We have

$$A + I_4 = \begin{pmatrix} 0 & -3 & -1 & 0 \\ 0 & 3 & 1 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 3 & 1 & 0 \end{pmatrix}, \quad (A - 2I_4) = \begin{pmatrix} -3 & -3 & -1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 3 & 1 & -3 \end{pmatrix},$$

$$(A - 2I_4)^2 = \begin{pmatrix} 9 & 9 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & -9 & 0 & 9 \end{pmatrix}.$$

The rank of $A + I_4$ is clearly 2, so its nullity is also 2, and hence there are two Jordan blocks with eigenvalue -1. The three non-zero rows of $(A - 2I_4)$ are linearly independent, so its rank is 3, hence its nullity 1, so there is just one Jordan block with eigenvalue 2, and the JCF of A is $J_{-1,1} \oplus J_{-1,1} \oplus J_{2,2}$.

For the two Jordan chains of length 1 for eigenvalue -1, we just need two linearly independent eigenvectors, and the obvious choice is $\mathbf{v}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$, $\mathbf{v}_2 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$. For the Jordan chain $\mathbf{v}_3, \mathbf{v}_4$ for eigenvalue 2, we need to choose $\mathbf{v}_4$ in the nullspace of $(A - 2I_4)^2$ but not in the nullspace of $A - 2I_4$. (This is why we calculated $(A - 2I_4)^2$.) An obvious choice here is $\mathbf{v}_4 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}$, and then $\mathbf{v}_3 = \begin{pmatrix} -1 \\ 1 \\ 0 \\ 1 \end{pmatrix}$, and to transform A to JCF, we put

$$P = \begin{pmatrix} 1 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{pmatrix}, \quad P^{-1} = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & -1 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \quad P^{-1}AP = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 2 \end{pmatrix}.$$

Example 10. $A = \begin{pmatrix} -2 & 0 & 0 & 0 \\ 0 & -2 & 1 & 0 \\ 0 & 0 & -2 & 0 \\ 1 & 0 & -2 & -2 \end{pmatrix}$. Then $c_A(x) = (-2 - x)^4$, so there is a single eigenvalue -2 with multiplicity 4. We find $(A + 2I_4) = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & -2 & 0 \end{pmatrix}$, and $(A + 2I_4)^2 = 0$, so $\mu_A = (x + 2)^2$, and the JCF of A could be $J_{-2,2} \oplus J_{-2,2}$ or $J_{-2,2} \oplus J_{-2,1} \oplus J_{-2,1}$.

To decide which case holds, we calculate the nullity of $A + 2I_4$ which, by Theorem 1.7.3, is equal to the number of Jordan blocks with eigenvalue -2. Since $A + 2I_4$ has just two non-zero rows, which are distinct, its rank is clearly 2, so its nullity is $4 - 2 = 2$, and hence the JCF of A is $J_{-2,2} \oplus J_{-2,2}$.

A Jordan basis consists of a union of two Jordan chains, which we will call $\mathbf{v}_1, \mathbf{v}_2$, and $\mathbf{v}_3, \mathbf{v}_4$, where $\mathbf{v}_1$ and $\mathbf{v}_3$ are eigenvectors and $\mathbf{v}_2$ and $\mathbf{v}_4$ are generalised eigenvectors of index 2. To find such chains, it is probably easiest to find $\mathbf{v}_2$ and $\mathbf{v}_4$ first and then to calculate $\mathbf{v}_1 = (A + 2I_4)\mathbf{v}_2$ and $\mathbf{v}_3 = (A + 2I_4)\mathbf{v}_4$.

Although it is not hard to find $\mathbf{v}_2$ and $\mathbf{v}_4$ in practice, we have to be careful, because they need to be chosen so that no linear combination of them lies in the nullspace of $(A + 2I_4)$. In fact, since this nullspace is spanned by the second and fourth standard basis vectors, the obvious choice is $\mathbf{v}_2 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$, $\mathbf{v}_4 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}$, and then $\mathbf{v}_1 = (A + 2I_4)\mathbf{v}_2 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$, $\mathbf{v}_3 = (A + 2I_4)\mathbf{v}_4 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ -2 \end{pmatrix}$, so to transform A to JCF, we put

$$P = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & -2 & 0 \end{pmatrix}, \quad P^{-1} = \begin{pmatrix} 0 & 2 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \quad P^{-1}AP = \begin{pmatrix} -2 & 1 & 0 & 0 \\ 0 & -2 & 0 & 0 \\ 0 & 0 & -2 & 1 \\ 0 & 0 & 0 & -2 \end{pmatrix}.$$

1.10 An algorithm to compute the Jordan canonical form in general (brute force)

Whereas the examples above explain some shortcuts, tricks and computational recipes to compute, given a matrix $A \in \mathbb{C}^{n,n}$, a Jordan canonical form J for A as well as a matrix P (invertible) such that $J = P^{-1}AP$, it may also be useful to know how this can be done systematically, *provided* we know all the eigenvalues, $\lambda_1, \dots, \lambda_s$, say, of A.

Algorithm:

Step 1: Compute J. This amounts to knowing, for a given eigenvalue λ, the number of Jordan blocks of degree/size i in J. By Theorem 1.7.3, (ii), this number is

$$\begin{aligned}(\dim N_i(A,\lambda) - \dim N_{i-1}(A,\lambda)) - (\dim N_{i+1}(A,\lambda) - \dim N_i(A,\lambda)) \\ = 2\dim N_i(A,\lambda) - \dim N_{i-1}(A,\lambda) - \dim N_{i+1}(A,\lambda).\end{aligned}$$

So the computation of J is no problem then.

Step 2: Compute P. You can proceed as follows: pick an eigenvalue λ. Now suppose

$$N_1 \geq N_2 \geq \cdots \geq N_r$$

are the sizes of the Jordan blocks with eigenvalue λ (repeats among the N_i allowed if there are several blocks of the same size; we order them according to decreasing size for definiteness). Then pick a vector $v_{1,1} \in V$ with

$$(A - \lambda I_n)^{N_1} v_{1,1} = 0, \; (A - \lambda I_n)^{N_1 - 1} v_{1,1} \neq 0$$

(note that this amounts to solving several systems of linear equations ultimately-we leave the details of how to accomplish this step to you). Then put

$$v_{1,2} := (A - \lambda I_n) v_{1,1}, \; v_{1,3} := (A - \lambda I_n)^2 v_{1,1}, \dots, v_{1,N_1} := (A - \lambda I_n)^{N_1 - 1} v_{1,1}.$$

Note that $(v_{1,N_1}, \dots, v_{1,1})$ is then a Jordan chain. If $r = 1$, we are done, else we choose a vector $v_{2,1} \in V$ with

$$(A - \lambda I_n)^{N_2} v_{2,1} = 0,\ (A - \lambda I_n)^{N_2 - 1} v_{2,1} \notin \langle v_{1,1}, \dots, v_{1,N_1} \rangle.$$

So note that the second condition has become more restrictive: we want that $(A - \lambda I_n)^{N_2-1} v_{2,1}$ is not just nonzero, but not in the span $V_1 := \langle v_{1,1}, \dots, v_{1,N_1} \rangle$ of the first bunch of basis vectors. Equivalently, we want it to be nonzero in the quotient V/V_1, for those of you who know what quotient vector spaces are (which isn't required). We then put

$$v_{2,2} := (A - \lambda I_n) v_{2,1},\ v_{2,3} := (A - \lambda I_n)^2 v_{2,1}, \dots, v_{2,N_2} := (A - \lambda I_n)^{N_2 - 1} v_{2,1}.$$

Then by construction $(v_{2,N_2}, \dots, v_{2,1})$ is a Jordan chain, and $v_{1,1}, \dots, v_{1,N_1}, v_{2,1}, \dots, v_{2,N_2}$ are linearly independent (for those who know quotient spaces, an easy way to check this is to notice that $v_{2,N_2}, \dots, v_{2,1}$ are a Jordan chain in V/V_1). If $r = 2$, we are done, otherwise we continue in the same fashion: pick $v_{3,1} \in V$ with

$$(A - \lambda I_n)^{N_3} v_{3,1} = 0,\ (A - \lambda I_n)^{N_3 - 1} v_{3,1} \notin \langle v_{1,1}, \dots, v_{1,N_1}, v_{2,1}, \dots, v_{2,N_2} \rangle,$$

and now you should see what the pattern to continue is. Finally you end up with vectors

$$v_{1,1}, \dots, v_{1,N_1}, v_{2,1}, \dots, v_{2,N_2}, \dots, v_{r,1}, \dots, v_{r,N_r} \in \mathbb{C}^n.$$

Listing these in reverse order gives us the first $N_1 + \dots + N_r$ columns of P. Now we repeat the same procedure for the remaining eigenvalues of A other than λ, adding a bunch of columns to P at each step in this way. That gives us the desired base change matrix P.

1.11 Grand finale

At this point we would like to take a step back and formulate the **basic facts of the spectral theory of matrices** we have obtained so far in a way that is both easy to remember and convenient to use in many applications. We use the more standard $\mathbb{C}^{n \times n}$ for $\mathbb{C}^{n,n}$ and $\mathbb{C}^n$ for $\mathbb{C}^{n,1}$ below.

Theorem 1.11.1. *Let $A \in \mathbb{C}^{n \times n}$ be a square matrix with complex entries, $p \in \mathbb{C}[x]$ any polynomial. Then if λ is an eigenvalue of A, $p(\lambda)$ is an eigenvalue of $p(A)$, and any eigenvalue of $p(A)$ is of this form.*

Theorem 1.11.2. *For $A \in \mathbb{C}^{n \times n}$ let*

$$N_i(A, \lambda)$$

be the null-space of $(A - \lambda I_n)^i$, so non-zero elements in $N_i(A, \lambda)$ are generalised eigenvectors of A w.r.t. λ of index i (index 1 being genuine eigenvectors). Then every vector in $\mathbb{C}^n$ can be written as a sum of eigenvectors of A, genuine or generalised.

This follows immediately from Theorem 1.7.3.

Theorem 1.11.3. *(i) Suppose $A, B \in \mathbb{C}^{n\times n}$ are similar, in the sense that there exists an invertible $n \times n$ matrix S with $B = S^{-1}AS$. Then A and B have the same set of eigenvalues:*

$$\lambda_1 = \mu_1, \ldots, \lambda_k = \mu_k$$

(here the λ's are the eigenvalues for A, the μ's the ones for B), and in addition we have

$$(*) \qquad \dim N_i(A, \lambda_j) = \dim N_i(B, \mu_j)$$

for all i, j.

(ii) Conversely, if $A, B \in \mathbb{C}^{n\times n}$ have the same eigenvalues $\lambda_1 = \mu_1, \ldots, \lambda_k = \mu_k$ as above, and $(*)$ holds for all i and j, then A and B are similar.

Whereas (i) is obvious, (ii) follows from the uniqueness part of Theorem 1.7.3.

The three results above are the basic results of spectral theory, in some sense even more basic than the Jordan canonical form itself. Also clearly

$$N_1(A, \lambda) \subset N_2(A, \lambda) \subset N_3(A, \lambda) \subset \ldots$$

and denoting by $d(\lambda)$ the smallest index from which these spaces are equal to each other (the index of the eigenvalue λ), we have: if $\lambda_1, \ldots, \lambda_k$ are the distinct eigenvalues of A, we have for the minimal polynomial

$$\mu_A(x) = \prod_{i=1}^{k} (x - \lambda_i)^{d(\lambda_i)}.$$

This is just Theorem 1.7.4, (ii) together with Theorem 1.7.3.

2 Functions of matrices

2.1 Powers of matrices

The theory of Jordan canonical form we developed can be used to compute powers of matrices efficiently. Suppose we need to compute A^n for large n where $A = \begin{pmatrix} -2 & 0 & 0 & 0 \\ 0 & -2 & 1 & 0 \\ 0 & 0 & -2 & 0 \\ 1 & 0 & -2 & -2 \end{pmatrix}$ is the matrix from Example 10 in 1.9.

There are two practical ways of computing A^n by hand for a general matrix A and a very large n. The first one involves the JCF of A.

If $J = P^{-1}AP$ is the JCF of A then it is sufficient to compute J^n because of the telescoping product:

$$A^n = (PJP^{-1})^n = PJP^{-1}PJP^{-1}P \dots JP^{-1} = PJ^nP^{-1}.$$

How do we work out what J^n is? Firstly, we need to convince ourselves that

$$(B \oplus C)^n = B^n \oplus C^n$$

for square matrices B, C. We leave this as an exercise in understanding the multiplication of direct sums of matrices (it might help to look at some small examples!) and we have already required this when thinking about the minimal polynomial of direct sums of matrices. Clearly, it extends to the direct sum of any finite number of square matrices.

So we are left to consider what the power of an individual Jordan block is. Again a small example will help us:

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}^2 = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}, \dots, \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}^n = \begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix}.$$

The eigenvalue being 1 hides things a little so let's do a slightly more complicated example.

$$\begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}^2 = \begin{pmatrix} 4 & 4 & 1 \\ 0 & 4 & 4 \\ 0 & 0 & 4 \end{pmatrix}, \quad \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}^3 = \begin{pmatrix} 8 & 12 & 6 \\ 0 & 8 & 12 \\ 0 & 0 & 8 \end{pmatrix}.$$

At this point you should be willing to believe the following formula, which is left as an exercise (use induction!) to prove.

$$J_{\lambda,k}^n = \begin{pmatrix} \lambda^n & n\lambda^{n-1} & \dots & \binom{n}{k-2}\lambda^{n-k+2} & \binom{n}{k-1}\lambda^{n-k+1} \\ 0 & \lambda^n & \dots & \binom{n}{k-3}\lambda^{n-k+3} & \binom{n}{k-2}\lambda^{n-k+2} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & \lambda^n & n\lambda^{n-1} \\ 0 & 0 & \dots & 0 & \lambda^n \end{pmatrix} \qquad (6)$$

where $\binom{n}{t} = \frac{n!}{(n-t)!t!}$ is the choose-function (or binomial coefficient), interpreted as $\binom{n}{t} = 0$ whenever $t > n$.

Let us apply it to the matrix A above:

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

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

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

The second method of computing A^n uses Lagrange's interpolation polynomial. It is less labour intensive and more suitable for pen-and-paper calculations.

Suppose $\psi(M) = 0$ for a polynomial $\psi(z)$, in practice we will choose $\psi(z)$ to be either the minimal or characteristic polynomial. Dividing with remainder gives $z^n = q(z)\psi(z) + h(z)$, and we conclude that

$$A^n = q(A)\psi(A) + h(A) = h(A).$$

Division with remainder may appear problematic[2] for large n but there is a shortcut. If we know the roots of $\psi(z)$, say $\alpha_1, \ldots, \alpha_k$ with their multiplicities $m_1, \ldots, m_k$, then $h(z)$ can be found by solving the system of simultaneous equations in coefficients of $h(z)$:

$$f^{(t)}(\alpha_j) = h^{(t)}(\alpha_j),\ 1 \leq j \leq k,\ 0 \leq t < m_j$$

where $f(z) = z^n$ and $f^{(t)}$ is the t-th derivative of f with respect to z. In other words, $h(z)$ is what is known as Lagrange's interpolation polynomial for the function z^n at the roots of $\psi(z)$. Note that we only ever need to take $h(z)$ to be a polynomial of degree $m_1 + \cdots + m_k - 1$.

Let's use this to find A^n again for A as above. We know the minimal polynomial $\mu_A(z) = (z+2)^2$. Given $\mu_A(z)$ is degree 2 we can take the Lagrange interpolation of z^n at the roots of $(z+2)^2$ to be $h(z) = \alpha z + \beta$. To determine α and β we have to solve

$$\begin{cases} (-2)^n & = & h(-2) & = & -2\alpha + \beta \\ n(-2)^{n-1} & = & h'(-2) & = & \alpha \end{cases}$$

Solving them gives $\alpha = n(-2)^{n-1}$ and $\beta = (1-n)(-2)^n$. It follows that

$$A^n = n(-2)^{n-1}A + (1-n)(-2)^n I = \begin{pmatrix} (-2)^n & 0 & 0 & 0 \\ 0 & (-2)^n & n(-2)^{n-1} & 0 \\ 0 & 0 & (-2)^n & 0 \\ n(-2)^{n-1} & 0 & n(-2)^n & (-2)^n \end{pmatrix}.$$

[2] Try to divide z^{2022} by $z^2 + z + 1$ without reading any further.

2.2 Applications to difference equations

Let us consider *an initial value problem* for an *autonomous* system with discrete time:

$$\mathbf{x}(n+1) = A\mathbf{x}(n),\ n \in \mathbb{N},\ \mathbf{x}(0) = w.$$

Here $\mathbf{x}(n) \in K^m$ is a sequence of vectors in a vector space over a field K. One thinks of $\mathbf{x}(n)$ as a state of the system at time n. The initial state is $\mathbf{x}(0) = w$. The $n \times n$-matrix A with coefficients in K describes the evolution of the system. The adjective *autonomous* means that the evolution equation does not change with the time[3].

It takes longer to formulate this problem than to solve it. The solution is straightforward:

$$\mathbf{x}(n) = A\mathbf{x}(n-1) = A^2\mathbf{x}(n-2) = \ldots = A^n\mathbf{x}(0) = A^n w. \tag{7}$$

As a working example, let us consider the Fibonacci numbers:

$$F_0 = 0,\ F_1 = 1 \text{ and } F_n = F_{n-1} + F_{n-2}\ (n \geq 2).$$

The recursion relations for them turn into

$$\begin{pmatrix} F_n \\ F_{n+1} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} F_{n-1} \\ F_n \end{pmatrix}$$

so that (7) immediately yields a general solution

$$\begin{pmatrix} F_n \\ F_{n+1} \end{pmatrix} = A^n \begin{pmatrix} 0 \\ 1 \end{pmatrix} \text{ where } A = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}. \tag{8}$$

We compute the characteristic polynomial of A to be $c_A(z) = z^2 - z - 1$. Its discriminant is 5. The roots of $c_A(z)$ are the golden ratio $\lambda = (1+\sqrt{5})/2$ and $1 - \lambda = (1-\sqrt{5})/2$. It is useful to observe that

$$2\lambda - 1 = \sqrt{5} \quad \text{and} \quad \lambda(1-\lambda) = -1.$$

Let us introduce the number $\mu_n = \lambda^n - (1-\lambda)^n$. Suppose the Lagrange interpolation of z^n at the roots of $z^2 - z - 1$ is $h(z) = \alpha z + \beta$. The condition on the coefficients is given by

$$\begin{cases} \lambda^n & = & h(\lambda) & = & \alpha\lambda + \beta \\ (1-\lambda)^n & = & h(1-\lambda) & = & \alpha(1-\lambda) + \beta \end{cases}$$

Solving them gives

$$\alpha = \mu_n/\sqrt{5} \text{ and } \beta = \mu_{n-1}/\sqrt{5}\,.$$

It follows that

$$A^n = \alpha A + \beta = \mu_n/\sqrt{5}A + \mu_{n-1}/\sqrt{5}I_2 = \begin{pmatrix} \mu_{n-1}/\sqrt{5} & \mu_n/\sqrt{5} \\ \mu_n/\sqrt{5} & (\mu_n + \mu_{n-1})/\sqrt{5} \end{pmatrix}.$$

Equation (8) immediately implies that

$$F_n = \mu_n/\sqrt{5} \text{ and } A^n = \begin{pmatrix} F_{n-1} & F_n \\ F_n & F_{n+1} \end{pmatrix}.$$

If we try and do this for more complicated difference equations, we could meet matrices which aren't diagonalisable. Here's an example (taken from the book by Kaye and Wilson, §14.11), done using Jordan canonical form.

[3] A nonautonomous system would be described by $\mathbf{x}(n+1) = A(n)\mathbf{x}(n)$ here.

Example. Let x_n, y_n, z_n be sequences of complex numbers satisfying

$$\begin{cases} x_{n+1} &= 3x_n + z_n, \\ y_{n+1} &= -x_n + y_n - z_n, \\ z_{n+1} &= y_n + 2z_n. \end{cases}$$

with $x_0 = y_0 = z_0 = 1$.

We can write this as

$$\mathbf{v}_{n+1} = \begin{pmatrix} 3 & 0 & 1 \\ -1 & 1 & -1 \\ 0 & 1 & 2 \end{pmatrix} \mathbf{v}_n.$$

So we have

$$\mathbf{v}_n = A^n \mathbf{v}_0 = A^n \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

where A is the 3×3 matrix above.

We find that the JCF of A is $J = P^{-1}AP$ where

$$J = J_{2,3} = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}, \quad P = \begin{pmatrix} 1 & 1 & 1 \\ 0 & -1 & 0 \\ -1 & 0 & 0 \end{pmatrix}.$$

The formula for the entries of J^k for J a Jordan block tells us that

$$\begin{aligned} J^n &= \begin{pmatrix} 2^n & n2^{n-1} & \binom{n}{2} 2^{n-2} \\ 0 & 2^n & n2^{n-1} \\ 0 & 0 & 2^n \end{pmatrix} \\ &= 2^n \begin{pmatrix} 1 & \frac{1}{2}n & \frac{1}{4}\binom{n}{2} \\ 0 & 1 & \frac{1}{2}n \\ 0 & 0 & 1 \end{pmatrix} \end{aligned}$$

We therefore have

$$\begin{aligned} A^n &= PJ^nP^{-1} \\ &= 2^n \begin{pmatrix} 1 & 1 & 1 \\ 0 & -1 & 0 \\ -1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & \frac{1}{2}n & \frac{1}{4}\binom{n}{2} \\ 0 & 1 & \frac{1}{2}n \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 & -1 \\ 0 & -1 & 0 \\ 1 & 1 & 1 \end{pmatrix} \\ &= 2^n \begin{pmatrix} 1 & 1 + \frac{1}{2}n & 1 + \frac{1}{2}n + \frac{1}{4}\binom{n}{2} \\ 0 & -1 & -\frac{1}{2}n \\ -1 & -\frac{1}{2}n & -\frac{1}{4}\binom{n}{2} \end{pmatrix} \begin{pmatrix} 0 & 0 & -1 \\ 0 & -1 & 0 \\ 1 & 1 & 1 \end{pmatrix} \\ &= 2^n \begin{pmatrix} 1 + \frac{1}{2}n + \frac{1}{4}\binom{n}{2} & \frac{1}{4}\binom{n}{2} & \frac{1}{2}n + \frac{1}{4}\binom{n}{2} \\ -\frac{1}{2}n & 1 - \frac{1}{2}n & -\frac{1}{2}n \\ -\frac{1}{4}\binom{n}{2} & \frac{1}{2}n - \frac{1}{4}\binom{n}{2} & 1 - \frac{1}{4}\binom{n}{2} \end{pmatrix} \end{aligned}$$

Finally, we obtain

$$A^n \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = 2^n \begin{pmatrix} 1 + n + \frac{3}{4}\binom{n}{2} \\ 1 - \frac{3}{2}n \\ 1 + \frac{1}{2}n - \frac{3}{4}\binom{n}{2} \end{pmatrix}$$

or equivalently, using the fact that $\binom{n}{2} = \frac{n(n-1)}{2}$,

$$\begin{cases} x_n &= 2^n(\frac{3}{8}n^2 + \frac{5}{8}n + 1), \\ y_n &= 2^n(1 - \frac{3}{2}n), \\ z_n &= 2^n(-\frac{3}{8}n^2 + \frac{7}{8}n + 1). \end{cases}$$

2.3 Motivation: Systems of Differential Equations

Suppose we want to expand our repertoire and solve a system of first-order simultaneous differential equations, say

$$\frac{\mathrm{d}a}{\mathrm{d}t} = 3a - 4b + 8c,$$
$$\frac{\mathrm{d}b}{\mathrm{d}t} = a - c,$$
$$\frac{\mathrm{d}c}{\mathrm{d}t} = a + b + c.$$

These are common in the Differential Equations course last year. Let's write the system in a different form. We consider $\mathbf{v}(t) = \begin{pmatrix} a(t) \\ b(t) \\ c(t) \end{pmatrix}$, a vector-valued function of time, and write the above system as

$$\frac{d\mathbf{v}}{dt} = A\mathbf{v}$$

where A is the matrix

$$\begin{pmatrix} 3 & -4 & 8 \\ 1 & 0 & -1 \\ 1 & 1 & 1 \end{pmatrix}.$$

"Aha!" we say. "We know the solution is $\mathbf{v}(t) = e^{tA}\mathbf{v}(0)$!" But then we pause, and say "Hang on, what does e^{tA} actually mean?" In the next section, we'll use what we now know about special forms of matrices to define e^{tA}, and other functions of a matrix, in a sensible way that will make this actually work; and having got our definition, we'll work out how to calculate with it.

2.4 Definition of a function of a matrix

Suppose we have a "nice" one variable complex-valued function $f(z)$. What is $f(A)$? In general, there is no natural answer. We had one for $f(z) = z^n$ in Section 2.1 and we choose to generalise this to define $f(A)$ using the Jordan canonical form of A as follows. Let $J = P^{-1}AP$ with $J = J_{\lambda_1,k_1} \oplus \cdots \oplus J_{\lambda_t,k_t}$ being the JCF of A. We define

$$f(A) = Pf(J)P^{-1}, \text{ where } f(J) = f(J_{\lambda_1,k_1}) \oplus \cdots \oplus f(J_{\lambda_t,k_t}),$$

and

$$f(J_{\lambda,k}) = \begin{pmatrix} f(\lambda) & f'(\lambda) & \dots & f^{[k-2]}(\lambda) & f^{[k-1]}(\lambda) \\ 0 & f(\lambda) & \dots & f^{[k-3]}(\lambda) & f^{[k-2]}(\lambda) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & f(\lambda) & f'(\lambda) \\ 0 & 0 & \dots & 0 & f(\lambda) \end{pmatrix}. \quad (9)$$

The notation $f^{[k]}(z)$ is known as the divided power derivative and defined as

$$f^{[k]}(z) := \frac{1}{k!} f^{(k)}(z).$$

So $f^{[1]} = f', f^{[2]} = \frac{1}{2}f'', f^{[3]} = \frac{1}{6}f'''$, etc. As you might imagine, deciding exactly what a "nice" function is, and whether this is definition is sensible for functions defined by power series etc. is more analysis than it is algebra. Thus, in this course we will ignore such issues. We are mainly interested in the exponential of a matrix. Taylor's series at zero of the exponential function is $\sum_{k=0}^{\infty} \frac{z^k}{k!}$ and so we might think that the following equation should be true.

$$e^A = I_n + A + \frac{A^2}{2} + \frac{A^3}{6} + \cdots = \sum_{k=0}^{\infty} \frac{A^k}{k!}. \tag{10}$$

It is indeed true, i.e. this coincides with our definition of $e^A = f(A)$ where f is the standard exponential function. Note, however, that not everything we know about the exponential function of complex numbers is true when we apply it to matrices. For example, it is *not* true that $e^{B+C} = e^B e^C$ for general matrices B and C; you may wish to find an explicit example.

Let's start by calculating e^A for a matrix A.

Example 11. Consider $A = \begin{pmatrix} 1 & 4 \\ 1 & 1 \end{pmatrix}$. This was Example 3 from Section 1.8 above, and we saw that $P^{-1}AP = J$ where

$$P = \begin{pmatrix} 2 & -2 \\ 1 & 1 \end{pmatrix}, \quad J = \begin{pmatrix} 3 & 0 \\ 0 & -1 \end{pmatrix}.$$

Hence

$$\begin{aligned} e^A &= Pe^J P^{-1} \\ &= \begin{pmatrix} 2 & -2 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} e^3 & 0 \\ 0 & e^{-1} \end{pmatrix} \begin{pmatrix} 2 & -2 \\ 1 & 1 \end{pmatrix}^{-1} \\ &= \frac{1}{4} \begin{pmatrix} 2 & -2 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} e^3 & 0 \\ 0 & e^{-1} \end{pmatrix} \begin{pmatrix} 1 & 2 \\ -1 & 2 \end{pmatrix} \\ &= \begin{pmatrix} \frac{1}{2}e^3 + \frac{1}{2}e^{-1} & e^3 - e^{-1} \\ \frac{1}{4}e^3 - \frac{1}{4}e^{-1} & \frac{1}{2}e^3 + \frac{1}{2}e^{-1} \end{pmatrix}. \end{aligned}$$

Let's see another way to calculate e^A. We can again use Lagrange's interpolation method, which is often easier in practice.

Example 12. We compute e^A for the matrix A from Example 10, Section 1.9, using Lagrange interpolation. Suppose that $h(z) = \alpha z + \beta$ is the interpolation of e^z at the roots of $\mu_A(z) = (z+2)^2$. The condition on the coefficients is given by

$$\begin{cases} e^{-2} &= h(-2) &= -2\alpha + \beta \\ e^{-2} &= h'(-2) &= \alpha \end{cases}$$

Solving them gives $\alpha = e^{-2}$ and $\beta = 3e^{-2}$. It follows that

$$e^A = h(A) = e^{-2}A + 3e^{-2}I = \begin{pmatrix} e^{-2} & 0 & 0 & 0 \\ 0 & e^{-2} & e^{-2} & 0 \\ 0 & 0 & e^{-2} & 0 \\ e^{-2} & 0 & -2e^{-2} & e^{-2} \end{pmatrix}.$$

Our motivation for defining the exponential of a matrix was to find e^{tA} so let's do that in the next example. It is important to note that t here should be seen as a constant when we differentiate $f(z) = e^{zt}$. So $f^{[1]}(z) = te^{zt}$, $f^{[2]}(z) = \frac{1}{2}t^2e^{zt}$, etc.

Example 13. Let
$$A = \begin{pmatrix} 1 & 0 & -3 \\ 1 & -1 & -6 \\ -1 & 2 & 5 \end{pmatrix}.$$

Using the methods of the last chapter we can check that its JCF is $J = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ and the basis change matrix P such that $J = P^{-1}AP$ is given by $P = \begin{pmatrix} 3 & 0 & 2 \\ 3 & 1 & 1 \\ -1 & -1 & 0 \end{pmatrix}$.

Applying the argument above, we see that $e^{tA} = Pe^{tJ}P^{-1}$ where
$$e^{tJ} = \begin{pmatrix} e^{2t} & te^{2t} & 0 \\ 0 & e^{2t} & 0 \\ 0 & 0 & e^t \end{pmatrix}.$$

We can now calculate e^{tA} explicitly by doing the matrix multiplication to get the entries of $Pe^{Jt}P^{-1}$, as we did in the 2×2 example above.

It looks messy. Do we really want to write it down here?

Well, let us not do it. In a pen-and-paper calculation, except a few cases (for example, diagonal matrices) it is simpler to use Lagrange's interpolation.

Example 14. Let us consider a harmonic oscillator described by the equation $y''(t) + y(t) = 0$. The general solution $y(t) = \alpha \sin(t) + \beta \cos(t)$ is well known. Let us obtain it using matrix exponents. Setting
$$x(t) = \begin{pmatrix} y(t) \\ y'(t) \end{pmatrix}, \ A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$
the harmonic oscillator becomes the initial value problem with a solution $x(t) = e^{tA}x(0)$. The eigenvalues of A are i and $-i$. Interpolating e^{tz} at these values of z gives the following condition on $h(z) = \alpha z + \beta$
$$\begin{cases} e^{ti} & = & h(i) & = & \alpha i + \beta \\ e^{-ti} & = & h(-i) & = & -\alpha i + \beta \end{cases}$$

Solving them gives $\alpha = (e^{ti} - e^{-ti})/2i = \sin(t)$ and $\beta = (e^{ti} + e^{-ti})/2 = \cos(t)$. It follows that
$$e^{tA} = \sin(t)A + \cos(t)I_2 = \begin{pmatrix} \cos(t) & \sin(t) \\ -\sin(t) & \cos(t) \end{pmatrix}$$
and so
$$x(t) = \begin{pmatrix} \cos(t)y(0) + \sin(t)y'(0) \\ -\sin(t)y(0) + \cos(t)y'(0) \end{pmatrix}.$$

The final solution is thus $y(t) = \cos(t)y(0) + \sin(t)y^{'}(0)$.

Example 15. Let us consider a system of differential equations

$$\begin{cases} y_1' &= y_1 - 3y_3 \\ y_2' &= y_1 - y_2 - 6y_3 \\ y_3' &= -y_1 + 2y_2 + 5y_3 \end{cases}, \text{with the initial condition} \begin{cases} y_1(0) &= 1 \\ y_2(0) &= 1 \\ y_3(0) &= 0 \end{cases}$$

Using matrices

$$x(t) = \begin{pmatrix} y_1(t) \\ y_2(t) \\ y_3(t) \end{pmatrix}, \ w = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \ A = \begin{pmatrix} 1 & 0 & -3 \\ 1 & -1 & -6 \\ -1 & 2 & 5 \end{pmatrix},$$

it becomes an initial value problem. The characteristic polynomial is $c_A(z) = -z^3 + 5z^2 - 8z + 4 = (1-z)(2-z)^2$. We need to interpolate e^{tz} at 1 and 2 by $h(z) = \alpha z^2 + \beta z + \gamma$. At the multiple root 2 we need to interpolate up to order 2 that involves tracking the derivative $(e^{tz})' = te^{tz}$:

$$\begin{cases} e^t &= h(1) &= \alpha + \beta + \gamma \\ e^{2t} &= h(2) &= 4\alpha + 2\beta + \gamma \\ te^{2t} &= h'(2) &= 4\alpha + \beta \end{cases}$$

Solving, $\alpha = (t-1)e^{2t} + e^t$, $\beta = (4-3t)e^{2t} - 4e^t$, $\gamma = (2t-3)e^{2t} + 4e^t$. It follows that

$$e^{tA} = e^{2t} \begin{pmatrix} 3t-3 & -6t+6 & -9t+6 \\ 3t-2 & -6t+4 & -9t+3 \\ -t & 2t & 3t+1 \end{pmatrix} + e^t \begin{pmatrix} 4 & -6 & -6 \\ 2 & -3 & -3 \\ 0 & 0 & 0 \end{pmatrix}$$

and

$$x(t) = \begin{pmatrix} y_1(t) \\ y_2(t) \\ y_3(t) \end{pmatrix} = e^{tA} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} (3-3t)e^{2t} - 2e^t \\ (2-3t)e^{2t} - e^t \\ te^{2t} \end{pmatrix}.$$

3 Bilinear Maps and Quadratic Forms

We'll now introduce another, rather different kind of object you can define for vector spaces: a *bilinear map*. These are a bit different from linear maps: rather than being machines that take a vector and spit out another vector, they take two vectors as input and spit out a number.

3.1 Bilinear maps: definitions

Let V and W be vector spaces over a field K.

Definition 3.1.1. A *bilinear map* on V and W is a map $\tau : V \times W \to K$ such that

(i) $\tau(\alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2, \mathbf{w}) = \alpha_1 \tau(\mathbf{v}_1, \mathbf{w}) + \alpha_2 \tau(\mathbf{v}_2, \mathbf{w})$; and

(ii) $\tau(\mathbf{v}, \alpha_1 \mathbf{w}_1 + \alpha_2 \mathbf{w}_2) = \alpha_1 \tau(\mathbf{v}, \mathbf{w}_1) + \alpha_2 \tau(\mathbf{v}, \mathbf{w}_2)$

for all $\mathbf{v}, \mathbf{v}_1, \mathbf{v}_2 \in V$, $\mathbf{w}, \mathbf{w}_1, \mathbf{w}_2 \in W$, and $\alpha_1, \alpha_2 \in K$.

So $\tau(\mathbf{v}, \mathbf{w})$ is linear in $\mathbf{v}$ for each $\mathbf{w}$, and linear in $\mathbf{w}$ for each $\mathbf{v}$ – linear in two different ways, hence the term "bilinear".

Clearly if we fix bases of V and W, a bilinear map will be determined by what it does to the basis vectors. Choose a basis $\mathbf{e}_1, \dots, \mathbf{e}_n$ of V and a basis $\mathbf{f}_1, \dots, \mathbf{f}_m$ of W.

Let $\tau : V \times W \to K$ be a bilinear map, and let $\alpha_{ij} = \tau(\mathbf{e}_i, \mathbf{f}_j)$, for $1 \le i \le n$, $1 \le j \le m$. Then the $n \times m$ matrix $A = (\alpha_{ij})$ is defined to be the matrix of τ with respect to the bases $\mathbf{e}_1, \dots, \mathbf{e}_n$ and $\mathbf{f}_1, \dots, \mathbf{f}_m$ of V and W.

For $\mathbf{v} \in V$, $\mathbf{w} \in W$, let $\mathbf{v} = x_1 \mathbf{e}_1 + \cdots + x_n \mathbf{e}_n$ and $\mathbf{w} = y_1 \mathbf{f}_1 + \cdots + y_m \mathbf{f}_m$, so the coordinates of $\mathbf{v}$ and $\mathbf{w}$ with respect to our bases are

$$\underline{\mathbf{v}} = \begin{pmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ x_n \end{pmatrix} \in K^{n,1}, \quad \text{and} \quad \underline{\mathbf{w}} = \begin{pmatrix} y_1 \\ y_2 \\ \cdot \\ \cdot \\ y_m \end{pmatrix} \in K^{m,1}.$$

Then, by using the equations (i) and (ii) above, we get

$$\tau(\mathbf{v}, \mathbf{w}) = \sum_{i=1}^{n} \sum_{j=1}^{m} x_i \, \tau(\mathbf{e}_i, \mathbf{f}_j) \, y_j = \sum_{i=1}^{n} \sum_{j=1}^{m} x_i \, \alpha_{ij} \, y_j = \underline{\mathbf{v}}^{\mathrm{T}} A \underline{\mathbf{w}}. \tag{2.1}$$

So once we've fixed bases of V and W, every bilinear map on V and W corresponds to an $n \times m$ matrix, and conversely every matrix determines a bilinear map.

For example, let $V = W = \mathbb{R}^2$ and use the natural basis of V. Suppose that $A = \begin{pmatrix} 1 & -1 \\ 2 & 0 \end{pmatrix}$. Then

$$\tau((x_1, x_2), (y_1, y_2)) = (x_1 \; x_2) \begin{pmatrix} 1 & -1 \\ 2 & 0 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = x_1 y_1 - x_1 y_2 + 2 x_2 y_1.$$

3.2 Bilinear maps: change of basis

We retain the notation of the previous section, so τ is a bilinear map on V and W, and A is the matrix of τ with respect to some bases $\mathbf{e}_1, \dots, \mathbf{e}_n$ of V and $\mathbf{f}_1, \dots, \mathbf{f}_m$ of W.

As in §1.5 of the course, suppose that we choose new bases $\mathbf{e}'_1, \dots, \mathbf{e}'_n$ of V and $\mathbf{f}'_1, \dots, \mathbf{f}'_m$ of W, and let P and Q be the associated basis change matrices. Let B be the matrix of τ with respect to these new bases.

Let $\mathbf{v}$ be any vector in V. Then we know (from Proposition 0.5.1) that if $\underline{\mathbf{v}} \in K^{n,1}$ is the column vector of coordinates of $\mathbf{v}$ with respect to the old basis $\mathbf{e}_1, \dots, \mathbf{e}_n$, and $\underline{\mathbf{v}}'$ the coordinates of $\mathbf{v}$ in the new basis $\mathbf{e}'_1, \dots, \mathbf{e}'_n$, then we have $P\underline{\mathbf{v}}' = \underline{\mathbf{v}}$. Similarly, for any $\mathbf{w} \in W$, the coordinates $\underline{\mathbf{w}}$ and $\underline{\mathbf{w}}'$ of $\mathbf{w}$ with respect to the old and new bases of W are related by $Q\underline{\mathbf{w}}' = \underline{\mathbf{w}}$.

We know that we have

$$\underline{\mathbf{v}}^{\mathrm{T}} A \underline{\mathbf{w}} = \tau(\mathbf{v}, \mathbf{w}) = (\underline{\mathbf{v}}')^{\mathrm{T}} B \underline{\mathbf{w}}'.$$

Substituting in the formulae from Proposition 0.5.1, we have

$$\begin{aligned}(\underline{\mathbf{v}}')^{\mathrm{T}} B \underline{\mathbf{w}}' &= (P\underline{\mathbf{v}}')^{\mathrm{T}} A (Q\underline{\mathbf{w}}') \\ &= (\underline{\mathbf{v}}')^{\mathrm{T}}\, P^{\mathrm{T}} A Q\, \underline{\mathbf{w}}'.\end{aligned}$$

Since this relation must hold for all $\underline{\mathbf{v}}' \in K^{n,1}$ and $\underline{\mathbf{w}}' \in K^{m,1}$, the two matrices in the middle must be equal (exercise!): that is, we have $B = P^{\mathrm{T}} A Q$. So we've proven:

Theorem 3.2.1. *Let A be the matrix of the bilinear map $\tau : V \times W \to K$ with respect to the bases $\mathbf{e}_1, \dots, \mathbf{e}_n$ and $\mathbf{f}_1, \dots, \mathbf{f}_m$ of V and W, and let B be its matrix with respect to the bases $\mathbf{e}'_1, \dots, \mathbf{e}'_n$ and $\mathbf{f}'_1, \dots, \mathbf{f}'_m$ of V and W. Let P and Q be the basis change matrices, as defined above. Then $B = P^{\mathrm{T}} A Q$.*

Compare this result with Theorem 0.5.2.

We shall be concerned from now on only with the case where $V = W$. A bilinear map $\tau : V \times V \to K$ is called a *bilinear form* on V. Theorem 3.2.1 then becomes:

Theorem 3.2.2. *Let A be the matrix of the bilinear form τ on V with respect to the basis $\mathbf{e}_1, \dots, \mathbf{e}_n$ of V, and let B be its matrix with respect to the basis $\mathbf{e}'_1, \dots, \mathbf{e}'_n$ of V. Let P the basis change matrix with original basis $\{\mathbf{e}_i\}$ and new basis $\{\mathbf{e}'_i\}$. Then $B = P^{\mathrm{T}} A P$.*

Let's give a name to this relation between matrices:

Definition 3.2.3. Two matrices A and B are called *congruent* if there exists an invertible matrix P with $B = P^{\mathrm{T}} A P$.

So congruent matrices represent the same bilinear form in different bases. Notice that congruence is very different from similarity; if τ is a bilinear form on V and T is a linear operator on V, it might be the case that τ and T have the same matrix A in some specific basis of V, but that doesn't mean that they have the same matrix in any other basis – they inhabit different worlds.

So, in the example at the end of Subsection 3.1, if we choose the new basis $\mathbf{e}'_1 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$, $\mathbf{e}'_2 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ then $P = \begin{pmatrix} 1 & 1 \\ -1 & 0 \end{pmatrix}$, $P^{\mathrm{T}}AP = \begin{pmatrix} 0 & -1 \\ 2 & 1 \end{pmatrix}$, and

$$\tau((y'_1\mathbf{e}'_1 + y'_2\mathbf{e}'_2, x'_1\mathbf{e}'_1 + x'_2\mathbf{e}'_2)) = -y'_1x'_2 + 2y'_2x'_1 + y'_2x'_2.$$

Since P is an invertible matrix, P^{T} is also invertible (its inverse is $(P^{-1})^{\mathrm{T}}$), and so the matrices $P^{\mathrm{T}}AP$ and A are "equivalent matrices" in the sense of MA106, and hence have the same rank.

The rank of the bilinear form τ is defined to be the rank of its matrix A. So we have just shown that the rank of τ is a well-defined property of τ, not depending on the choice of basis we've used.

In fact we can say a little more. It's clear that a vector $\underline{\mathbf{v}} \in K^{n,1}$ is zero if and only if $\underline{\mathbf{v}}^{\mathrm{T}}\underline{\mathbf{w}} = 0$ for all vectors $\underline{\mathbf{w}} \in K^{n,1}$. Since

$$\tau(\mathbf{v}, \mathbf{w}) = \underline{\mathbf{v}}^{\mathrm{T}}A\underline{\mathbf{w}},$$

the kernel of A is equal to the space

$$\{\mathbf{v} \in V : \tau(\mathbf{w}, \mathbf{v}) = 0 \;\forall \mathbf{w} \in V\}$$

(the *right radical* of τ) and the kernel of A^{T} is equal to the space

$$\{\mathbf{v} \in V : \tau(\mathbf{v}, \mathbf{w}) = 0 \;\forall \mathbf{w} \in V\}$$

(the *left radical*). Since A^{T} and A have the same rank, the left and right radicals both have dimension $n - r$, where r is the rank of τ. In particular, the rank of τ is n if and only if the left and right radicals are zero. If this occurs, we'll say τ is *nondegenerate*; so τ is nondegenerate if and only if its matrix (in any basis) is nonsingular.

You could be forgiven for expecting that we were about to launch into a long study of how to choose, given a bilinear form τ on V, the "best" basis for V which makes the matrix of τ as nice as possible. We are *not* going to do this, because although it's a very natural question to ask, it's *extremely* hard! Instead, we'll restrict ourselves to a special kind of bilinear form where life is much easier, which covers most of the bilinear forms that come up in "real life".

Definition 3.2.4. We say bilinear form τ on V is *symmetric* if $\tau(\mathbf{w}, \mathbf{v}) = \tau(\mathbf{v}, \mathbf{w})$ for all $\mathbf{v}, \mathbf{w} \in V$.

We say τ is *antisymmetric* (or sometimes *alternating*) if $\tau(\mathbf{v}, \mathbf{v}) = 0$ for all $\mathbf{v} \in V$.

The antisymmetry condition implies for all $\mathbf{v}, \mathbf{w} \in V$

$$\tau(\mathbf{v} + \mathbf{w}, \mathbf{v} + \mathbf{w}) = \tau(\mathbf{v}, \mathbf{w}) + \tau(\mathbf{w}, \mathbf{v}) = 0$$

hence for all $\mathbf{v}, \mathbf{w} \in V$

$$\tau(\mathbf{v}, \mathbf{w}) = -\tau(\mathbf{w}, \mathbf{v}).$$

If $2 \neq 0$ in K, the condition $\tau(\mathbf{v}, \mathbf{w}) = -\tau(\mathbf{w}, \mathbf{v})$ implies antisymmetry (take $\mathbf{v} = \mathbf{w}$, but you need to be able to divide by 2).

An $n \times n$ matrix A is called symmetric if $A^{\mathrm{T}} = A$, and anti-symmetric if $A^{\mathrm{T}} = -A$ and it has zeros along the diagonal. We then clearly have:

Proposition 3.2.5. *The bilinear form τ is symmetric or anti-symmetric if and only if its matrix (with respect to any basis) is symmetric or anti-symmetric.*

The best known example of a symmetric form is when $V = \mathbb{R}^n$, and τ is defined by

$$\tau\left(\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}\right) = x_1y_1 + x_2y_2 + \cdots + x_ny_n.$$

This form has matrix equal to the identity matrix I_n with respect to the standard basis of $\mathbb{R}^n$. Geometrically, it is equal to the normal scalar product $\tau(\mathbf{v}, \mathbf{w}) = |\mathbf{v}||\mathbf{w}|\cos\theta$, where θ is the angle between the vectors $\mathbf{v}$ and $\mathbf{w}$.

On the other hand, the form on $\mathbb{R}^2$ defined by $\tau\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}\right) = x_1y_2 - x_2y_1$ is anti-symmetric. This has matrix $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$.

Proposition 3.2.6. *Suppose that $2 \neq 0$ in K. Then any bilinear form τ can be written uniquely as $\tau_1 + \tau_2$ where τ_1 is symmetric and τ_2 is antisymmetric.*

Proof. We just put $\tau_1(\mathbf{v}, \mathbf{w}) = \frac{1}{2}(\tau(\mathbf{v}, \mathbf{w}) + \tau(\mathbf{w}, \mathbf{v}))$ and $\tau_2(\mathbf{v}, \mathbf{w}) = \frac{1}{2}(\tau(\mathbf{v}, \mathbf{w}) - \tau(\mathbf{w}, \mathbf{v}))$. It's clear that τ_1 is symmetric and τ_2 is antisymmetric.

Moreover, given any other such expression $\tau = \tau_1' + \tau_2'$, we have

$$\begin{aligned} \tau_1(\mathbf{v}, \mathbf{w}) &= \frac{\tau_1'(\mathbf{v}, \mathbf{w}) + \tau_1'(\mathbf{w}, \mathbf{v}) + \tau_2'(\mathbf{v}, \mathbf{w}) + \tau_2'(\mathbf{w}, \mathbf{v})}{2} \\ &= \frac{\tau_1'(\mathbf{v}, \mathbf{w}) + \tau_1'(\mathbf{v}, \mathbf{w}) + \tau_2'(\mathbf{v}, \mathbf{w}) - \tau_2'(\mathbf{v}, \mathbf{w})}{2} \end{aligned}$$

from the symmetry and antisymmetry of τ_1' and τ_2'. The last two terms cancel each other and we just have

$$= \frac{2\tau_1'(\mathbf{v}, \mathbf{w})}{2} = \tau_1'(\mathbf{v}, \mathbf{w}).$$

So $\tau_1 = \tau_1'$, and so $\tau_2 = \tau - \tau_1 = \tau - \tau_1' = \tau_2'$, so the decomposition is unique. □

(Notice that $\frac{1}{2}$ has to exist in K for all this to make sense!)

3.3 Quadratic forms

Definition 3.3.1. Let V be a vector space over the field K. Then a *quadratic form* on V is a function $q : V \to K$ that satisfies that

$$q(\lambda\mathbf{v}) = \lambda^2 q(\mathbf{v}), \quad \forall\, \mathbf{v} \in V, \lambda \in K$$

and that

$$(*) \quad \tau_q(\mathbf{v}_1, \mathbf{v}_2) := q(\mathbf{v}_1 + \mathbf{v}_2) - q(\mathbf{v}_1) - q(\mathbf{v}_2)$$

is a symmetric bilinear form on V.

As we can see from the definition, symmetric bilinear forms and quadratic forms are closely related. Indeed, given a bilinear form τ we can define a quadratic form by

$$q_\tau(\mathbf{v}) := \tau(\mathbf{v}, \mathbf{v}).$$

Moreover, given a quadratic form, (*) above gives us a symmetric bilinear form. These processes are *almost* inverse to each other: indeed, one can easily compute that starting with a quadratic form q and bilinear form τ

$$q_{\tau_q} = 2q, \quad \tau_{q_\tau} = 2\tau.$$

So as long as $2 \neq 0$ in our K, quadratic forms and bilinear forms correspond to each other in a one-to-one way if we make the associations

$$q \mapsto \frac{1}{2}\tau_q, \quad \tau \mapsto q_\tau.$$

If $2 = 0$ in K (e.g. in $\mathbb{F}_2 = \mathbb{Z}/2\mathbb{Z}$, but there are again lots of other examples of such fields) this correspondence breaks down: indeed, in that case there are quadratic forms that are *not* of the form $\tau(-,-)$ for any symmetric bilinear form τ on V; e.g. let $V = \mathbb{F}_2^2$, the space of pairs (x_1, x_2) with $x_i \in \mathbb{F}_2$. We would certainly like to be able to call

$$q((x_1, x_2)) = x_1 x_2$$

a quadratic form on V. On the other hand, a general symmetric bilinear form on V looks like

$$\tau((x_1, x_2), (y_1, y_2)) = ax_1y_1 + bx_1y_2 + bx_2y_1 + cx_2y_2$$

so that putting $(x_1, x_2) = (y_1, y_2)$ we only get quadratic forms that a sums of squares.

There is an important and highly developed theory of quadratic forms also when $2 = 0$ in K (exposed in for example the books by Merkurjev-Karpenko-Elman or Kneser on quadratic forms), but the normal forms for them are a bit different from the case when $2 \neq 0$ and though the theory is not actually harder it divides naturally according to whether $2 = 0$ or $2 \neq 0$ in K. So from now on till the rest of this Chapter we make the:

Assumption: In our field K, we have that $2 = 1 + 1$ is not equal to 0.

Let $\mathbf{e}_1, \ldots, \mathbf{e}_n$ be a basis of V. Recall that the coordinates of $\mathbf{v}$ with respect to this basis are defined to be the field elements x_i such that $\mathbf{v} = \sum_{i=1}^n x_i \mathbf{e}_i$.

Let $A = (\alpha_{ij})$ be the matrix of a symmetric bilinear form τ with respect to this basis. We will also call A the matrix of $q = q_\tau$ with respect to this basis. Then A is symmetric because τ is, and by Equation (2.1) of Subsection 3.1, we have

$$q(\mathbf{v}) = \underline{\mathbf{v}}^{\mathrm{T}} A \underline{\mathbf{v}} = \sum_{i=1}^n \sum_{j=1}^n x_i \alpha_{ij} x_j = \sum_{i=1}^n \alpha_{ii} x_i^2 + 2 \sum_{i=1}^n \sum_{j=1}^{i-1} \alpha_{ij} x_i x_j. \tag{3.1}$$

Conversely, if we are given a quadratic form as in the right hand side of Equation (3.1), then it is easy to write down its matrix A. For example, if $n = 3$ and $q(\mathbf{v}) = 3x^2 + y^2 - 2z^2 + 4xy - xz$, then $A = \begin{pmatrix} 3 & 2 & -1/2 \\ 2 & 1 & 0 \\ -1/2 & 0 & -2 \end{pmatrix}$.

3.4 Nice bases for quadratic forms

We'll now show how to choose a basis for V which makes a given symmetric bilinear form (or, equivalently, quadratic form) "as nice as possible". This will turn out to be much easier than the corresponding problem for linear operators.

Theorem 3.4.1. *Let V be a vector space of dimension n equipped with a symmetric bilinear form τ (or, equivalently, a quadratic form q).*

Then there is a basis $\mathbf{b}_1, \dots, \mathbf{b}_n$ of V, and constants $\beta_1, \dots, \beta_n$, such that

$$\tau(\mathbf{b}_i, \mathbf{b}_j) = \begin{cases} \beta_i & \text{if } j = i \\ 0 & \text{if } j \neq i \end{cases}.$$

Equivalently,

- *given any symmetric matrix A, there is an invertible matrix P such that $P^{\mathrm{T}} A P$ is a diagonal matrix (i.e. A is congruent to a diagonal matrix);*
- *given any quadratic form q on a vector space V, there is a basis $\mathbf{b}_1, \dots, \mathbf{b}_n$ of V and constants $\beta_1, \dots, \beta_n$ such that*

$$q(x_1\mathbf{b}_1 + \dots + x_n\mathbf{b}_n) = \beta_1 x_1^2 + \dots + \beta_n x_n^2.$$

Proof. We shall prove this by induction on $n = \dim V$. If $n = 0$ then there is nothing to prove, so let's assume that $n \geq 1$.

If τ is zero, then again there is nothing to prove, so we may assume that $\tau \neq 0$. Then the associated quadratic form q is not zero either, so there is a vector $\mathbf{v} \in V$ such that $q(\mathbf{v}, \mathbf{v}) \neq 0$. Let $\mathbf{b}_1 = \mathbf{v}$ and let $\beta_1 = q(\mathbf{v})$.

Consider the linear map $V \to K$ given by $\mathbf{w} \mapsto \tau(\mathbf{w}, \mathbf{v})$. This is not the zero map, so its image has rank 1; so its kernel W has rank $n - 1$. Moreover, this $(n-1)$-dimensional subspace doesn't contain $\mathbf{b}_1 = \mathbf{v}$.

By the induction hypothesis, we can find a basis $\mathbf{b}_2, \dots, \mathbf{b}_n$ for the kernel such that $\tau(\mathbf{b}_i, \mathbf{b}_j) = 0$ for all $2 \leq i < j \leq n$; and all of these vectors lie in the space W, so we also have $\tau(\mathbf{b}_1, \mathbf{b}_j) = 0$ for all $2 \leq j \leq n$. Since $\mathbf{b}_1 \notin W$, it follows that $\mathbf{b}_1, \dots, \mathbf{b}_n$ is a basis of V, so we're done. □

Finding the good basis: The above proof is quite short and slick, and gives us very little help if we explicitly want to find the diagonalizing basis. So let's unravel what's going on a bit more explicitly. We'll see in a moment that what's going on is very closely related to "completing the square" in school algebra.

So let's say we have a quadratic form q. As usual, let $B = (\beta_{ij})$ be the matrix of q with respect to some arbitrary basis $\mathbf{b}_1, \dots, \mathbf{b}_n$. We'll modify the basis $\mathbf{b}_i$ step-by-step in order to eventually get it into the nice form the theorem predicts.

Step 1: Arrange that $q(\mathbf{b}_1) \neq 0$. Here there are various cases to consider.

- If $\beta_{11} \neq 0$, then we're done: this means that $q(\mathbf{b}_1) \neq 0$, so we don't need to do anything.

- If $\beta_{11} = 0$, but $\beta_{ii} \neq 0$ for some $i > 1$, then we just interchange $\mathbf{b}_1$ and $\mathbf{b}_i$ in our basis.
- If $\beta_{ii} = 0$ for all i, but there is some i and j such that $\beta_{ij} \neq 0$, then we replace $\mathbf{b}_i$ with $\mathbf{b}_i + \mathbf{b}_j$; since
$$q(\mathbf{b}_i + \mathbf{b}_j) = q(\mathbf{b}_i) + q(\mathbf{b}_j) + 2\tau(\mathbf{b}_i, \mathbf{b}_j) = 2\beta_{ij},$$
after making this change we have $q(\mathbf{b}_i) \neq 0$, so we're reduced to one of the two previous cases.
- If $\beta_{ij} = 0$ for all i and j, we can stop: the matrix of q is zero, so it's certainly diagonal.

Step 2: Modify $\mathbf{b}_2, \dots, \mathbf{b}_n$ to make them orthogonal to $\mathbf{b}_1$. Suppose we've done Step 1, but we haven't stopped, so β_{11} is now non-zero. We want to arrange that $\tau(\mathbf{b}_1, \mathbf{b}_i)$ is 0 for all $i > 1$. To do this, we just replace $\mathbf{b}_i$ with
$$\mathbf{b}_i - \frac{\beta_{1i}}{\beta_{11}}\mathbf{b}_1.$$
This works because
$$\tau(\mathbf{b}_1, \mathbf{b}_i - \frac{\beta_{1i}}{\beta_{11}}\mathbf{b}_1) = \tau(\mathbf{b}_1, \mathbf{b}_i) - \frac{\beta_{1i}}{\beta_{11}}\tau(\mathbf{b}_1, \mathbf{b}_1) = \beta_{1i} - \frac{\beta_{1i}}{\beta_{11}}\beta_{11} = 0.$$
This is where the relation to "completing the square" comes in. We've changed our basis by the matrix
$$P = \begin{pmatrix} 1 & -\frac{\beta_{12}}{\beta_{11}} & \cdots & -\frac{\beta_{1n}}{\beta_{11}} \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{pmatrix}$$
so the coordinates of a vector $\mathbf{v} \in V$ change by the inverse of this, which is just
$$P^{-1} = \begin{pmatrix} 1 & \frac{\beta_{12}}{\beta_{11}} & \cdots & \frac{\beta_{1n}}{\beta_{11}} \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{pmatrix}$$
This corresponds to writing
$$q(x_1\mathbf{b}_1 + \cdots + x_n\mathbf{b}_n) = \beta_{11}x_1^2 + 2\beta_{12}x_1x_2 + \cdots + 2\beta_{1n}x_1x_n + C$$
where C doesn't involve x_1 at all, and writing this as
$$\beta_{11}\left(x_1 + \frac{\beta_{12}}{\beta_{11}}x_2 + \cdots + \frac{\beta_{1n}}{\beta_{11}}x_n\right)^2 + C'$$
where C' also doesn't involve x_1. Then our change of basis changes the coordinates so the whole bracketed term becomes the first coordinate of $\mathbf{v}$; we've eliminated "cross terms" involving x_1 and one of the other variables.

Step 3: Induct on n. Now we've managed to engineer a basis $\mathbf{b}_1, \dots, \mathbf{b}_n$ such that the matrix $B = \beta_{ij}$ of q looks like

$$\begin{pmatrix} \beta_{11} & 0 & \dots & 0 \\ 0 & ? & \dots & ? \\ \vdots & \vdots & \ddots & \vdots \\ 0 & ? & \dots & ? \end{pmatrix}$$

So we can now repeat the process with V replaced by the $(n-1)$-dimensional vector space W spanned by $\mathbf{b}_2, \dots, \mathbf{b}_n$. We can mess around as much as we like with the vectors $\mathbf{b}_2, \dots, \mathbf{b}_n$ without breaking the fact that they pair to zero with $\mathbf{b}_1$, since this is true of any vector in W. So we go back to step 1 but with a smaller n, and keep going until we either have an 0-dimensional space or a zero form, in which case we can safely stop.

Example. Let $V = \mathbb{R}^3$ and $q\left(\begin{pmatrix} x \\ y \\ z \end{pmatrix}\right) = xy + 3yz - 5xz$, so the matrix of q with respect to the standard basis $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ is

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

Since we have only 3 variables, it's much less work to call them x, y, z than x_1, x_2, x_3. When we change the variables, we will write x_1, y_1, z_1 and so on. We still proceed as in the previous proof and you need to read the proof first! We will use $\overset{\heartsuit}{=}$ for the equalities that need no checking (they are for information purposes only).

First change of basis: All the diagonal entries of A are zero, so we're in Case 3 of Step 1 of the proof above. But α_{12} is $1/2$, which isn't zero; so we replace $\mathbf{e}_1$ with $\mathbf{e}_1 + \mathbf{e}_2$. That is, we work in the basis

$$\mathbf{b}_1 := \mathbf{e}_1 + \mathbf{e}_2,\ \mathbf{b}_2 := \mathbf{e}_2,\ \mathbf{b}_3 := \mathbf{e}_3\,.$$

Thus the basis change matrix from $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ to $\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3$ is

$$P = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{so that} \quad \begin{pmatrix} x \\ y \\ z \end{pmatrix} \overset{\heartsuit}{=} P \begin{pmatrix} x_1 \\ y_1 \\ Z/1 \end{pmatrix}$$

where $\begin{pmatrix} x_1 \\ y_1 \\ Z/1 \end{pmatrix}$ is the coordinate expression in the new basis (remember, P takes new coordinates to old). And we have

$$q(x_1\mathbf{b}_1 + y_1\mathbf{b}_2 + z_1\mathbf{b}_3) = q\begin{pmatrix} x_1 \\ x_1 + y_1 \\ Z/1 \end{pmatrix} =$$

$$= x_1(x_1 + y_1) + 3(x_1 + y_1)z_1 - 5x_1z_1 = x_1^2 + x_1y_1 - 2x_1z_1 + 3y_1z_1,$$

so the matrix of q in the basis $\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3$ is

$$B = \begin{pmatrix} 1 & 1/2 & -1 \\ 1/2 & 0 & 3/2 \\ -1 & 3/2 & 0 \end{pmatrix} \overset{\heartsuit}{=} P^T A P.$$

Second change of basis: Now we can use Step 2 of the proof to clear the entries in the first row and column by modifying $\mathbf{b}_2$ and $\mathbf{b}_3$, this is the "completing the square" step. As specified in Step 2 of the proof, we introduce a new basis $\mathbf{b}'$ as follows

$$\mathbf{b}'_1 := \mathbf{b}_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \ \mathbf{b}'_2 := \mathbf{b}_2 - \frac{1}{2}\mathbf{b}_1 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} - \frac{1}{2}\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} -1/2 \\ 1/2 \\ 0 \end{pmatrix},$$

$$\mathbf{b}'_3 := \mathbf{b}_3 - (-1)\mathbf{b}_1 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

So the basis change matrix from $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ to $\mathbf{b}'_1, \mathbf{b}'_2, \mathbf{b}'_3$ is

$$P' = \begin{pmatrix} 1 & -1/2 & 1 \\ 1 & 1/2 & 1 \\ 0 & 0 & 1 \end{pmatrix} \overset{\heartsuit}{=} PQ \ \text{ where } Q = \begin{pmatrix} 1 & -1/2 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

This corresponds to writing

$$\begin{aligned} \left[x_1^2 + x_1 y_1 - 2x_1 z_1\right] + 3y_1 z_1 &= \left[(x_1 + \frac{1}{2}y_1 - z_1)^2 - \frac{1}{4}y_1^2 - z_1^2 + y_1 z_1\right] + 3y_1 z_1 \\ &= (x_1 + \frac{1}{2}y_1 - z_1)^2 - \frac{1}{4}y_1^2 + 4y_1 z_1 - z_1^2. \end{aligned}$$

In the new basis $x_2\mathbf{b}'_1 + y_2\mathbf{b}'_2 + z_2\mathbf{b}'_3 = (x_2 - \frac{1}{2}y_2 + z_2)\mathbf{b}_1 + (x_2 + \frac{1}{2}y_2 + z_2)\mathbf{b}_2 + z_2\mathbf{b}_3$, which tells us that

$$q(x_2\mathbf{b}'_1 + y_2\mathbf{b}'_2 + z_2\mathbf{b}'_3) = x_2^2 - \frac{1}{4}y_2^2 + 4y_2 z_2 - z_2^2.$$

so the matrix of q with respect to the $\mathbf{b}'$ basis is

$$B' = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1/4 & 2 \\ 0 & 2 & -1 \end{pmatrix} \overset{\heartsuit}{=} Q^T B Q \overset{\heartsuit}{=} (P')^T A P'.$$

Third change of basis: Now we are in Step 3 of the proof, concentrating on the bottom right 2×2 block. We must change the second and third basis vectors. Any subsequent changes of basis we make will keep the first basis vector unchanged. We have

$$q(y_2\mathbf{b}'_2 + z_2\mathbf{b}'_3) = -\frac{1}{4}y_2^2 + 4y_2 z_2 - z_2^2,$$

the "leftover terms" of the bottom right corner. This is a 2-variable quadratic form.

Since $q(\mathbf{b}'_2) = -1/4 \neq 0$, we don't need to do anything for Step 1 of the proof. Using Step 2 of the proof, we replace $\mathbf{b}'_1, \mathbf{b}'_2, \mathbf{b}'_3$ by another new basis $\mathbf{b}''$:

$$\mathbf{b}''_1 := \mathbf{b}'_1,\ \mathbf{b}''_2 := \mathbf{b}'_2,\ \mathbf{b}''_3 := \mathbf{b}'_3 - \frac{2}{-1/4}\mathbf{b}'_2 = \begin{pmatrix}1\\1\\1\end{pmatrix} + 8\begin{pmatrix}-1/2\\1/2\\0\end{pmatrix} = \begin{pmatrix}-3\\5\\1\end{pmatrix}.$$

So the change of basis matrix from $\mathbf{e}$ to $\mathbf{b}''$ is

$$P'' = \begin{pmatrix}1 & -1/2 & -3\\1 & 1/2 & 5\\0 & 0 & 1\end{pmatrix} \overset{\heartsuit}{=} P'Q' \text{ where } Q' = \begin{pmatrix}1 & 0 & 0\\0 & 1 & 8\\0 & 0 & 1\end{pmatrix}.$$

This corresponds, of course, to the completing-the-square operation

$$-\frac{1}{4}y_2^2 + 4y_2z_2 - z_2^2 = -\frac{1}{4}(y_2 - 8z_2)^2 + 15z_2^2.$$

So the matrix of q is now

$$B'' = \begin{pmatrix}1 & 0 & 0\\0 & -1/4 & 0\\0 & 0 & 15\end{pmatrix} \overset{\heartsuit}{=} (Q')^T B' Q' \overset{\heartsuit}{=} (P'')^T A P''.$$

This is diagonal, so we're done: the matrix of q in the basis $\mathbf{b}''_1, \mathbf{b}''_2, \mathbf{b}''_3$ is the diagonal matrix B''.

Notice that the choice of "good" basis, and the resulting "good" matrix, are extremely far from unique. For instance, in the example above we could have replaced $\mathbf{b}''_2$ with $2\mathbf{b}''_2$ to get the (perhaps nicer) matrix

$$\begin{pmatrix}1 & 0 & 0\\0 & -1 & 0\\0 & 0 & 15\end{pmatrix}.$$

In the case $K = \mathbb{C}$, we can do even better. After reducing q to the form $q(\mathbf{v}) = \sum_{i=1}^n \alpha_{ii}x_i^2$, we can permute the coordinates if necessary to get $\alpha_{ii} \neq 0$ for $1 \leq i \leq r$ and $\alpha_{ii} = 0$ for $r+1 \leq i \leq n$, where $r = \operatorname{rank}(q)$. We can then make a further change of coordinates $x'_i = \sqrt{\alpha_{ii}}x_i$ $(1 \leq i \leq r)$, giving $q(\mathbf{v}) = \sum_{i=1}^r (x'_i)^2$. Hence we have proved:

Proposition 3.4.2. *A quadratic form q over $\mathbb{C}$ has the form $q(\mathbf{v}) = \sum_{i=1}^r x_i^2$ with respect to a suitable basis, where $r = \operatorname{rank}(q)$.*

Equivalently, given a symmetric matrix $A \in \mathbb{C}^{n,n}$, there is an invertible matrix $P \in \mathbb{C}^{n,n}$ such that $P^{\mathrm{T}}AP = B$, where $B = (\beta_{ij})$ is a diagonal matrix with $\beta_{ii} = 1$ for $1 \leq i \leq r$, $\beta_{ii} = 0$ for $r+1 \leq i \leq n$, and $r = \operatorname{rank}(A)$.

In particular, up to changes of basis, a quadratic form on $\mathbb{C}^n$ is uniquely determined by its rank. We say the rank is the only *invariant* of a quadratic form over $\mathbb{C}$.

When $K = \mathbb{R}$, we cannot take square roots of negative numbers, but we can replace each positive α_i by 1 and each negative α_i by -1 to get:

Proposition 3.4.3 (Sylvester's Theorem)**.** *A quadratic form q over $\mathbb{R}$ has the form $q(\mathbf{v}) = \sum_{i=1}^{t} x_i^2 - \sum_{i=1}^{u} x_{t+i}^2$ with respect to a suitable basis, where $t + u = \operatorname{rank}(q)$.*

Equivalently, given a symmetric matrix $A \in \mathbb{R}^{n,n}$, there is an invertible matrix $P \in \mathbb{R}^{n,n}$ such that $P^{\mathrm{T}}AP = B$, where $B = (\beta_{ij})$ is a diagonal matrix with $\beta_{ii} = 1$ for $1 \le i \le t$, $\beta_{ii} = -1$ for $t+1 \le i \le t+u$, and $\beta_{ii} = 0$ for $t+u+1 \le i \le n$, and $t + u = \operatorname{rank}(A)$.

We shall now prove that the numbers t and u of positive and negative terms are invariants of q. The pair of integers (t, u) is called the *signature* of q.

Theorem 3.4.4 (Sylvester's Law of Inertia)**.** *Suppose that q is a quadratic form on the vector space V over $\mathbb{R}$, and that $\mathbf{e}_1, \ldots, \mathbf{e}_n$ and $\mathbf{e}'_1, \ldots, \mathbf{e}'_n$ are two bases of V such that*

$$q(x_1\mathbf{e}_1 + \cdots + x_n\mathbf{e}_n) = \sum_{i=1}^{t} x_i^2 - \sum_{i=1}^{u} x_{t+i}^2$$

and

$$q(x_1\mathbf{e}'_1 + \cdots + x_n\mathbf{e}'_n) = \sum_{i=1}^{t'} x_i^2 - \sum_{i=1}^{u'} x_{t'+i}^2.$$

Then $t = t'$ and $u = u'$.

Proof. We know that $t + u = t' + u' = \operatorname{rank}(q)$, so it is enough to prove that $t = t'$. Suppose not; by symmetry we may suppose that $t > t'$.

Let V_1 be the span of $\mathbf{e}_1, \ldots, \mathbf{e}_t$, and let V_2 be the span of $\mathbf{e}'_{t'+1}, \ldots, \mathbf{e}'_n$. Then for any non-zero $\mathbf{v} \in V_1$ we have $q(\mathbf{v}) > 0$; while for any $\mathbf{w} \in V_2$ we have $q(\mathbf{w}) \le 0$. So there cannot be any non-zero $\mathbf{v} \in V_1 \cap V_2$.

On the other hand, we have $\dim(V_1) = t$ and $\dim(V_2) = n - t'$. It was proved in MA106 that

$$\dim(V_1 + V_2) = \dim(V_1) + \dim(V_2) - \dim(V_1 \cap V_2),$$

so

$$\dim(V_1 \cap V_2) = t + (n - t') - \dim(V_1 + V_2) = (t - t') + n - \dim(V_1 + V_2) > 0.$$

The last inequality follows from our assumption on $t - t'$ and the fact $V_1 + V_2$ is a subspace of V and thus has dimension at most n. Since we have shown that $V_1 \cap V_2 = \{0\}$, this is a contradiction, which completes the proof. □

Remark. *Notice that any non-zero $x \in \mathbb{R}$ is either equal to a square, or -1 times a square, but not both. This property is shared by the finite field $\mathbb{F}_7$ of integers mod 7, so any quadratic form over $\mathbb{F}_7$ can be written as a diagonal matrix with only 0's, 1's and -1's down the diagonal (i.e. Sylvester's Theorem holds over $\mathbb{F}_7$). But Sylvester's law of inertia isn't valid in $\mathbb{F}_7$: in fact, we have*

$$\begin{pmatrix} 2 & 3 \\ 4 & 2 \end{pmatrix}^{\mathrm{T}} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 2 & 3 \\ 4 & 2 \end{pmatrix} = \begin{pmatrix} 20 & 14 \\ 14 & 20 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix},$$

so the same form has signature $(2, 0)$ and $(0, 2)$! The proof breaks down because there's no good notion of a "positive" element of $\mathbb{F}_7$, so a sum of non-zero squares can be zero (the easiest example is $1^2 + 2^2 + 3^2 = 0$). So Sylvester's law of inertia is really using something quite special about $\mathbb{R}$.

3.5 Euclidean spaces, orthonormal bases and the Gram–Schmidt process

In this section, we're going to suppose $K = \mathbb{R}$. As usual, we let V be an n-dimensional vector space over K, and we let q be a quadratic form on V, with associated symmetric bilinear form τ.

Definition 3.5.1. The quadratic form q is said to be *positive definite* if $q(\mathbf{v}) > 0$ for all $0 \neq \mathbf{v} \in V$.

It is clear that this is the case if and only if $t = n$ and $u = 0$ in Proposition 3.4.3; that is, if q has signature $(n, 0)$.

The associated symmetric bilinear form τ is also called positive definite when q is.

Definition 3.5.2. A vector space V over $\mathbb{R}$ together with a positive definite symmetric bilinear form τ is called a *Euclidean space*.

In this case, Proposition 3.4.3 says that there is a basis $\{\mathbf{e}_i\}$ of V with respect to which $\tau(\mathbf{e}_i, \mathbf{e}_j) = \delta_{ij}$, where

$$\delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j. \end{cases}$$

(so the matrix A of q is the identity matrix I_n.) We call a basis of a Euclidean space V with this property an *orthonormal basis* of V. We call a basis *orthogonal* if the matrix of τ is diagonal (with diagonal entries not necessarily equal to 1).

(More generally, any set $\mathbf{v}_1, \dots, \mathbf{v}_r$ of vectors in V, not necessarily a basis, will be said to be *orthonormal* if $\tau(\mathbf{v}_i, \mathbf{v}_j) = \delta_{ij}$ for $1 \leq i, j \leq r$. Same for orthogonal.)

We shall assume from now on that V is a Euclidean space, and that we have chosen an orthonormal basis $\mathbf{e}_1, \dots, \mathbf{e}_n$. Then τ corresponds to the standard dot product and we shall write $\mathbf{v} \cdot \mathbf{w}$ instead of $\tau(\mathbf{v}, \mathbf{w})$.

Note that $\mathbf{v} \cdot \mathbf{w} = \underline{\mathbf{v}}^{\mathrm{T}} \underline{\mathbf{w}}$ where, as usual, $\underline{\mathbf{v}}$ and $\underline{\mathbf{w}}$ are the column vectors associated with $\mathbf{v}$ and $\mathbf{w}$.

For $\mathbf{v} \in V$, define $|\mathbf{v}| = \sqrt{\mathbf{v} \cdot \mathbf{v}}$. Then $|\mathbf{v}|$ is the length of $\mathbf{v}$. Hence the length, and also the cosine $\mathbf{v} \cdot \mathbf{w} / (|\mathbf{v}||\mathbf{w}|)$ of the angle between two vectors can be defined in terms of the scalar product. Thus a set of vectors is orthonormal if the vectors all have length 1 and are at right angles to each other.

The following theorem tells us that we can modify every given basis of V to an orthonormal basis in a controlled way.

Theorem 3.5.3 (Gram-Schmidt process/orthonormalisation procedure)**.** *Let V be a euclidean space of dimension n, and suppose that $\mathbf{g}_1, \dots, \mathbf{g}_n$ is a basis of V. Then there exists an orthonormal basis $\mathbf{f}_1, \dots, \mathbf{f}_n$ of V with the property that for all $1 \leq i \leq n$*

$$\operatorname{span}\{\mathbf{f}_1, \dots, \mathbf{f}_i\} = \operatorname{span}\{\mathbf{g}_1, \dots, \mathbf{g}_i\}.$$

More precisely, it suffices to put

$$\mathbf{f}_1 := \frac{\mathbf{g}_1}{|\mathbf{g}_1|}$$

and then inductively, supposing that $\mathbf{f}_1, \ldots, \mathbf{f}_i$ have already been computed, we set

$$\mathbf{f}'_{i+1} := \mathbf{g}_{i+1} - \sum_{\alpha=1}^{i} (\mathbf{f}_\alpha \cdot \mathbf{g}_{i+1}) \mathbf{f}_\alpha$$

$$\mathbf{f}_{i+1} = \frac{\mathbf{f}'_{i+1}}{|\mathbf{f}'_{i+1}|}.$$

Moreover, note that this means that the basis change matrix $\mathcal{M}(\mathrm{id}_V)^{(\mathbf{g}_1,\ldots,\mathbf{g}_n)}_{(\mathbf{f}_1,\ldots,\mathbf{f}_n)}$ is upper triangular.

Proof. In fact, the statement of the Theorem already contains most of the ideas for the proof- we just have to check the algorithm does indeed what we claim it does. Indeed, the statement about spans follows directly from the construction, and all we have to check is that $\mathbf{f}_1, \ldots, \mathbf{f}_n$ is orthonormal. That all vectors have length 1 is obvious by construction. So it suffices to check that $\mathbf{f}'_{i+1}$ is orthogonal (=has dot product zero) with each of $\mathbf{f}_1, \ldots, \mathbf{f}_i$ for all $i = 1, \ldots n-1$. That's how we have constructed/defined $\mathbf{f}'_{i+1}$: for $j \leq i$

$$\mathbf{f}_j \cdot \mathbf{f}'_{i+1} := \mathbf{f}_j \cdot \mathbf{g}_{i+1} - \sum_{\alpha=1}^{i} (\mathbf{f}_\alpha \cdot \mathbf{g}_{i+1})(\mathbf{f}_j \cdot \mathbf{f}_\alpha) = \mathbf{f}_j \cdot \mathbf{g}_{i+1} - \mathbf{f}_j \cdot \mathbf{g}_{i+1} = 0.$$

□

Note that as an immediate corollary of the Gram-Schmidt process we obtain that if for some r with $0 \leq r \leq n$, $\mathbf{f}_1, \ldots, \mathbf{f}_r$ are vectors in V such that

$$\mathbf{f}_i \cdot \mathbf{f}_j = \delta_{ij} \quad \text{for} \quad 1 \leq i, j \leq r. \tag{$*$}$$

Then $\mathbf{f}_1, \ldots, \mathbf{f}_r$ can be extended to an orthonormal basis $\mathbf{f}_1, \ldots, \mathbf{f}_n$ of V. Indeed, just extend $\mathbf{f}_1, \ldots, \mathbf{f}_r$ in some way to a (not necessarily orthonormal) basis

$$\mathbf{f}_1, \ldots, \mathbf{f}_r, \mathbf{f}'_{r+1}, \ldots, \mathbf{f}'_n$$

and run the Gram-Schmidt orthonormalisation procedure above for this set of vectors.

Example. Let $V = \mathbb{R}^3$ with the standard dot product. It is straightforward to check that $\begin{pmatrix}1\\-1\\1\end{pmatrix}, \begin{pmatrix}1\\0\\1\end{pmatrix}, \begin{pmatrix}1\\1\\2\end{pmatrix}$ is a basis for V but it is not orthonormal. Let's use the Gram-Schmidt process to fix that. Thus here $g_1 = \begin{pmatrix}1\\-1\\1\end{pmatrix}$, $g_2 = \begin{pmatrix}1\\0\\1\end{pmatrix}$ and $g_3 = \begin{pmatrix}1\\1\\2\end{pmatrix}$.

Then $f'_1 := g_1$ and so $f_1 = f'_1/|f'_1| = \frac{1}{\sqrt{3}} \begin{pmatrix}1\\-1\\1\end{pmatrix}$,

$f'_2 := g_2 - (f_1 \cdot g_2) f_1 = g_2 - \frac{2}{\sqrt{3}} f_1 = \frac{1}{3} \begin{pmatrix}1\\2\\1\end{pmatrix}$ and so $f_2 = \frac{1}{\sqrt{6}} \begin{pmatrix}1\\2\\1\end{pmatrix}$,

$f_3' := g_3 - (f_1 \cdot g_3)f_1 - (f_2 \cdot g_3)f_2 = g_3 - \frac{2}{\sqrt{3}}f_1 - \frac{5}{\sqrt{6}}f_2 = \frac{1}{2}\begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$ and so $f_3 = \frac{1}{\sqrt{2}}\begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$.

thus we have now got an orthonormal basis f_1, f_2, f_3 (always good to check this at the end!).

3.6 Orthogonal transformations

Definition 3.6.1. A linear map $T{:}V \to V$ is said to be *orthogonal* if it preserves the scalar product on V. That is, if $T(\mathbf{v}) \cdot T(\mathbf{w}) = \mathbf{v} \cdot \mathbf{w}$ for all $\mathbf{v}, \mathbf{w} \in V$.

Since length and angle can be defined in terms of the scalar product, an orthogonal linear map preserves distance and angle. In $\mathbb{R}^2$, for example, an orthogonal map is either a rotation about the origin, or a reflection about a line through the origin.

If A is the matrix of T (with respect to some orthonormal basis), then $T(\mathbf{v}) = A\underline{\mathbf{v}}$ and so

$$T(\mathbf{v}) \cdot T(\mathbf{w}) = \underline{\mathbf{v}}^{\mathrm{T}} A^{\mathrm{T}} A \underline{\mathbf{w}}.$$

Hence T is orthogonal (the right hand side equals $\mathbf{v} \cdot \mathbf{w}$) if and only if $A^{\mathrm{T}}A = I_n$, or equivalently if $A^{\mathrm{T}} = A^{-1}$.

Definition 3.6.2. An $n \times n$ matrix is called *orthogonal* if $A^{\mathrm{T}}A = I_n$.

So we have proved:

Proposition 3.6.3. *A linear map $T : V \to V$ is orthogonal if and only if its matrix A (with respect to an orthonormal basis of V) is orthogonal.*

Incidentally, the fact that $A^{\mathrm{T}}A = I_n$ tells us that A (and hence T) is invertible, so $\det(A)$ is non-zero. In fact we can do a little better than that:

Proposition 3.6.4. *An orthogonal matrix has determinant ± 1.*

Proof. We have $A^{\mathrm{T}}A = I_n$, so $\det(A^{\mathrm{T}}A) = \det(I_n) = 1$.

On the other hand, $\det(A^{\mathrm{T}}A) = \det(A^{\mathrm{T}})\det(A) = (\det A)^2$. So $(\det A)^2 = 1$, implying that $\det A = \pm 1$. □

Example. For any $\theta \in \mathbb{R}$, let $A = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$. (This represents a anticlockwise rotation through an angle θ.) Then it is easily checked that $A^{\mathrm{T}}A = AA^{\mathrm{T}} = I_2$.

One can check that every orthogonal 2×2 matrix with determinant $+1$ is a rotation by some angle θ, and similarly that any orthogonal 2×2 matrix of det -1 is a reflection in some line through the origin. In higher dimensions the taxonomy of orthogonal matrices is a bit more complicated – we'll revisit this in a later section of the course.

Notice that the columns of A are mutually orthogonal vectors of length 1, and the same applies to the rows of A. Let $\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_n$ be the columns of the matrix A. As we observed in §1, $\mathbf{c}_i$ is equal to the column vector representing $T(\mathbf{e}_i)$. In other words, if $T(\mathbf{e}_i) = \mathbf{f}_i$, say, then $\underline{\mathbf{f}_i} = \mathbf{c}_i$.

Since the (i, j)-th entry of $A^{\mathrm{T}}A$ is $\mathbf{c}_i^{\mathrm{T}}\mathbf{c}_j = \mathbf{f}_i \cdot \mathbf{f}_j$, we see that T and A are orthogonal if and only if

$$\mathbf{f}_i \cdot \mathbf{f}_i = 1 \text{ and } \mathbf{f}_i \cdot \mathbf{f}_j = 0 \ (i \neq j), \ \ 1 \leq i, j \leq n. \qquad (*)$$

By Proposition 3.6.4, an orthogonal linear map is invertible, so $T(\mathbf{e}_i)$ $(1 \leq i \leq n)$ forms a basis of V, and we have:

Proposition 3.6.5. *A linear map T is orthogonal if and only if $T(\mathbf{e}_1), \ldots, T(\mathbf{e}_n)$ is an orthonormal basis of V.*

The Gram-Schmidt process readily gives:

Proposition 3.6.6 (QR decomposition)**.** *Let A be any $n \times n$ real matrix. Then we can write $A = QR$ where Q is orthogonal and R is upper-triangular.*

Proof. The proof when A is invertible goes as follows. Let $\mathbf{E}$ be the standard basis of $\mathbb{R}^n$, $\mathbf{G}$ the basis $\mathbf{g}_1, \ldots, \mathbf{g}_n$ given by the columns of A, and $\mathbf{F}$ be the orthonormal basis $\mathbf{f}_1, \ldots, \mathbf{f}_n$ from the Gram-Schmidt process applied to $\mathbf{G}$. Then by definition $A = \mathcal{M}(\mathrm{id}_V)^{\mathbf{E}}_{\mathbf{G}}$ and thus

$$A = \mathcal{M}(\mathrm{id}_V)^{\mathbf{E}}_{\mathbf{G}} = \mathcal{M}(\mathrm{id}_V)^{\mathbf{E}}_{\mathbf{F}}\mathcal{M}(\mathrm{id}_V)^{\mathbf{F}}_{\mathbf{G}}.$$

Since $\mathbf{F}$ is orthonormal, $Q := \mathcal{M}(\mathrm{id}_V)^{\mathbf{E}}_{\mathbf{F}}$ is orthogonal, and since we know by Theorem 3.5.3 that $\mathcal{M}(\mathrm{id}_V)^{\mathbf{G}}_{\mathbf{F}}$ is upper triangular, and

$$\mathcal{M}(\mathrm{id}_V)^{\mathbf{F}}_{\mathbf{G}} = (\mathcal{M}(\mathrm{id}_V)^{\mathbf{G}}_{\mathbf{F}})^{-1}$$

is also upper triangular as the inverse of an invertible upper triangular matrix, we can put $R := \mathcal{M}(\mathrm{id}_V)^{\mathbf{F}}_{\mathbf{G}}$ and are done in the case when A is invertible.

To deal with the case when A isn't invertible (so the columns of A no longer form a basis) we can do the following. We first show that any matrix A can be written as $A = BR'$ where B is invertible and R' is upper triangular; then writing $B = QR$ we have $A = QRR'$, and RR' is also upper-triangular. We leave the details to you, as we won't need the result for noninvertible A. □

Example. Consider the matrix

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

We have $\det(A) = 10$, so A is non-singular. Let $\mathbf{g}_1, \mathbf{g}_2, \mathbf{g}_3$ be the columns of A.

Then $|\mathbf{g}_1| = \sqrt{5}$, so

$$\mathbf{f}_1 = \frac{\mathbf{g}_1}{\sqrt{5}} = \begin{pmatrix} -1/\sqrt{5} \\ 2/\sqrt{5} \\ 0 \end{pmatrix}.$$

For the next step, we take $\mathbf{f}_2' = \mathbf{g}_2 - (\mathbf{f}_1 \cdot \mathbf{g}_2)\mathbf{f}_1 = \mathbf{g}_2$, since $\mathbf{f}_1 \cdot \mathbf{g}_2 = 0$. So

$$\mathbf{f}_2 = \frac{\mathbf{g}_2}{|\mathbf{g}_2|} = \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix}.$$

For the final step, we take the vector

$$\mathbf{f}_3' = \mathbf{g}_3 - (\mathbf{f}_1 \cdot \mathbf{g}_3)\mathbf{f}_1 - (\mathbf{f}_2 \cdot \mathbf{g}_3)\mathbf{f}_2.$$

We have

$$\mathbf{f}_1 \cdot \mathbf{g}_3 = \begin{pmatrix} -1/\sqrt{5} \\ 2/\sqrt{5} \\ 0 \end{pmatrix} \cdot \begin{pmatrix} -2 \\ -1 \\ -2 \end{pmatrix} = 0, \quad \mathbf{f}_2 \cdot \mathbf{g}_3 = \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix} \cdot \begin{pmatrix} -2 \\ -1 \\ -2 \end{pmatrix} = 2.$$

So $\mathbf{f}_3' = \mathbf{g}_3 - 2\mathbf{f}_2 = \begin{pmatrix} -2 \\ -1 \\ 0 \end{pmatrix}$. We have $|\mathbf{f}_3'| = \sqrt{5}$ again, so

$$\mathbf{f}_3 = \frac{\mathbf{f}_3'}{\sqrt{5}} = \begin{pmatrix} -2/\sqrt{5} \\ -1/\sqrt{5} \\ 0 \end{pmatrix}.$$

Thus Q is the matrix whose columns are $\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3$, that is

$$Q = \begin{pmatrix} -1/\sqrt{5} & 0 & -2/\sqrt{5} \\ 2/\sqrt{5} & 0 & -1/\sqrt{5} \\ 0 & -1 & 0 \end{pmatrix}.$$

and we have

$$\mathbf{g}_1 = \sqrt{5}\mathbf{f}_1, \quad \mathbf{g}_2 = 2\mathbf{f}_2, \quad \mathbf{g}_3 = 2\mathbf{f}_2 + \sqrt{5}\mathbf{f}_3$$

so $A = QR$ where

$$R = \begin{pmatrix} \sqrt{5} & 0 & 0 \\ 0 & 2 & 2 \\ 0 & 0 & \sqrt{5} \end{pmatrix}.$$

The QR decomposition theorem is a very important technique in numerical calculations. For example, if you know a QR decomposition of an invertible matrix A and you want to solve a linear system of equations $A\mathbf{x} = \mathbf{b}$, that's easy: just solve $QR\mathbf{x} = \mathbf{b}$, or equivalently

$$R\mathbf{x} = Q^T\mathbf{b}$$

(since R is upper triangular, this can be quickly done substituting backwards).

3.7 Nice orthonormal bases

If T is any linear map, then $(\mathbf{v}, \mathbf{w}) \mapsto (T\mathbf{v}) \cdot \mathbf{w}$ is a bilinear form; so there must be some linear map S such that

$$(T\mathbf{v}) \cdot \mathbf{w} = \mathbf{v} \cdot (S\mathbf{w}) \qquad (*)$$

for all $\mathbf{v}$ and $\mathbf{w}$.

Definition 3.7.1. If $T : V \to V$ is a linear map on a Euclidean space V, then the unique linear map S such that $(*)$ holds is called the *adjoint* of T. We write this as T^*.

When talking about adjoints, people sometimes prefer to call linear maps linear *operators*. That is because adjoints are particularly important in functional analysis, where the linear maps can be pretty complicated, so people initially were afraid of them and chose a complicated name ("operator" instead of "map") to reflect their fear.

If we have chosen an orthonormal basis, then the matrix of T^* is just the transpose of the matrix of T. It follows from this that a linear operator is orthogonal if and only if $T^* = T^{-1}$; one can also prove this directly from the definition.

We say T is *selfadjoint* if $T^* = T$, or equivalently if the bilinear form $\tau(\mathbf{v}, \mathbf{w}) = T\mathbf{v} \cdot \mathbf{w}$ is symmetric. Notice that 'selfadjointness', like 'orthogonalness', is something that only makes sense for linear operators on Euclidean spaces; it doesn't make sense to ask if a linear operator on a general vector space is selfadjoint. It should be clear that T is selfadjoint if and only if its matrix in an orthonormal basis of V is a symmetric matrix.

So if V is a Euclidean space of dimension n, the following problems are all actually the same:

- given a quadratic form q on V, find an orthonormal basis of V making the matrix of q as nice as possible;
- given a selfadjoint linear operator T on V, find an orthonormal basis of V making the matrix of T as nice as possible;
- given an $n \times n$ symmetric real matrix A, find an orthogonal matrix P such that $P^{\mathrm{T}}AP$ is as nice as possible.

First, we'll warm up by proving a proposition which we'll need in proving the main result solving these equivalent problems.

Proposition 3.7.2. *Let A be an $n \times n$ real symmetric matrix. Then A has an eigenvalue in $\mathbb{R}$, and all complex eigenvalues of A lie in $\mathbb{R}$.*

Proof. (To simplify the notation, we will write just $\mathbf{v}$ for a column vector $\underline{\mathbf{v}}$ in this proof.)

The characteristic equation $\det(A - xI_n) = 0$ is a polynomial equation of degree n in x, and since $\mathbb{C}$ is an algebraically closed field, it certainly has a root $\lambda \in \mathbb{C}$, which is an eigenvalue for A if we regard A as a matrix over $\mathbb{C}$. We shall prove that any such λ lies in $\mathbb{R}$, which will prove the proposition.

For a column vector $\mathbf{v}$ or matrix B over $\mathbb{C}$, we denote by $\overline{\mathbf{v}}$ or $\overline{B}$ the result of replacing all entries of $\mathbf{v}$ or B by their complex conjugates. Since the entries of A lie in $\mathbb{R}$, we have $\overline{A} = A$.

Let $\mathbf{v}$ be a complex eigenvector associated with λ. Then

$$A\mathbf{v} = \lambda\mathbf{v} \tag{1}$$

so,taking complex conjugates and using $\overline{A} = A$, we get

$$A\overline{\mathbf{v}} = \overline{\lambda}\overline{\mathbf{v}}. \tag{2}$$

Transposing (1) and using $A^{\mathrm{T}} = A$ gives

$$\mathbf{v}^{\mathrm{T}} A = \lambda \mathbf{v}^{\mathrm{T}}, \tag{3}$$

so by (2) and (3) we have

$$\lambda \mathbf{v}^{\mathrm{T}} \overline{\mathbf{v}} = \mathbf{v}^{\mathrm{T}} A \overline{\mathbf{v}} = \overline{\lambda} \mathbf{v}^{\mathrm{T}} \overline{\mathbf{v}}.$$

But if $\mathbf{v} = (\alpha_1, \alpha_2, \ldots, \alpha_n)^{\mathrm{T}}$, then $\mathbf{v}^{\mathrm{T}} \overline{\mathbf{v}} = \alpha_1 \overline{\alpha_1} + \cdots + \alpha_n \overline{\alpha_n}$, which is a non-zero real number (eigenvectors are non-zero by definition). Thus $\lambda = \overline{\lambda}$, so $\lambda \in \mathbb{R}$. □

Now let's prove the main theorem of this section.

Theorem 3.7.3. *Let V be a Euclidean space of dimension n. Then:*

- *Given any quadratic form q on V, there is an orthonormal basis $\mathbf{f}_1, \ldots, \mathbf{f}_n$ of V and constants $\alpha_1, \ldots, \alpha_n$, uniquely determined up to reordering, such that*

  $$q(x_1 \mathbf{f}_1 + \cdots + x_n \mathbf{f}_n) = \sum_{i=1}^{n} \alpha_i (x_i)^2$$

  for all $x_1, \ldots, x_n \in \mathbb{R}$.

- *Given any linear operator $T : V \to V$ which is selfadjoint, there is an orthonormal basis $\mathbf{f}_1, \ldots, \mathbf{f}_n$ of V consisting of eigenvectors of T.*

- *Given any $n \times n$ real symmetric matrix A, there is an orthogonal matrix P such that $P^{\mathrm{T}} A P = P^{-1} A P$ is a diagonal matrix.*

Proof. We've already seen that these three statements are equivalent to each other, so we can prove whichever one of them we like. Notice that in the second and third forms of the statement, it's clear that the diagonal matrix we obtain is similar to the original one; that tells us that in the first statement the constants $\alpha_1, \ldots, \alpha_n$ are uniquely determined (possibly up to re-ordering).

We'll prove the second statement using induction on $n = \dim V$. If $n = 0$ there is nothing to prove, so let's assume the proposition holds for $n - 1$.

Let T be our linear operator. By Proposition 3.7.2, T has an eigenvalue in $\mathbb{R}$. Let $\mathbf{v}$ be a corresponding eigenvector in V. Then $\mathbf{f}_1 = \mathbf{v}/|\mathbf{v}|$ is also an eigenvector, and $|\mathbf{f}_1| = 1$. Let α_1 be the corresponding eigenvalue.

We consider the space $W = \{\mathbf{w} \in V : \mathbf{w} \cdot \mathbf{f}_1 = 0\}$. Since W is the kernel of a surjective linear map

$$V \longrightarrow \mathbb{R}, \quad \mathbf{v} \mapsto \mathbf{v} \cdot \mathbf{f}_1\,,$$

it is a subspace of V of dimension $n - 1$. We claim that T maps W into itself. So suppose $\mathbf{w} \in W$; we want to show that $T(\mathbf{w}) \in W$ also.

We have

$$T(\mathbf{w}) \cdot \mathbf{f}_1 = \mathbf{w} \cdot T(\mathbf{f}_1)$$

since T is selfadjoint. But we know that $T(\mathbf{f}_1) = \alpha_1 \mathbf{f}_1$, so it follows that

$$T(\mathbf{w}) \cdot \mathbf{f}_1 = \alpha_1 (\mathbf{w} \cdot \mathbf{f}_1) = 0,$$

since $\mathbf{w} \in W$ so $\mathbf{w} \cdot \mathbf{f}_1 = 0$.

So T maps W into itself. Moreover, W is a euclidean space of dimension $n-1$, so we may apply the induction hypothesis to the restriction of T to W. This gives us an orthonormal basis $\mathbf{f}_2, \ldots, \mathbf{f}_n$ of W consisting of eigenvectors of T. By definition of W, $\mathbf{f}_1$ is orthogonal to $\mathbf{f}_2, \ldots, \mathbf{f}_n$ and it follows that $\mathbf{f}_1, \ldots, \mathbf{f}_n$ is an orthonormal basis of V, consisting of eigenvectors of T. □

Although it is not used in the proof of the theorem above, the following proposition is useful when calculating examples. It helps us to write down more vectors in the final orthonormal basis immediately, without having to use Theorem 3.5.3 repeatedly.

Proposition 3.7.4. *Let A be a real symmetric matrix, and let λ_1, λ_2 be two distinct eigenvalues of A, with corresponding eigenvectors $\mathbf{v}_1$, $\mathbf{v}_2$. Then $\mathbf{v}_1 \cdot \mathbf{v}_2 = 0$.*

Proof. (As in Proposition 3.7.2, we will write $\mathbf{v}$ rather than $\underline{\mathbf{v}}$ for a column vector in this proof. So $\mathbf{v}_1 \cdot \mathbf{v}_2$ is the same as $\mathbf{v}_1^{\mathrm{T}}\mathbf{v}_2$.) We have

$$A\mathbf{v}_1 = \lambda_1 \mathbf{v}_1, \tag{1}$$
$$A\mathbf{v}_2 = \lambda_2 \mathbf{v}_2. \tag{2}$$

The trick is now to look at the expression $\mathbf{v}_1^{\mathrm{T}} A \mathbf{v}_2$. On the one hand, by (2) we have

$$\mathbf{v}_1^{\mathrm{T}} A \mathbf{v}_2 = \mathbf{v}_1 \cdot (A\mathbf{v}_2) = \mathbf{v}_1^{\mathrm{T}}(\lambda_2 \mathbf{v}_2) = \lambda_2(\mathbf{v}_1 \cdot \mathbf{v}_2). \tag{3}$$

On the other hand, $A^{\mathrm{T}} = A$, so $\mathbf{v}_1^{\mathrm{T}} A = \mathbf{v}_1^{\mathrm{T}} A^{\mathrm{T}} = (A\mathbf{v}_1)^{\mathrm{T}}$, so using (1) we have

$$\mathbf{v}_1^{\mathrm{T}} A \mathbf{v}_2 = (A\mathbf{v}_1)^{\mathrm{T}} \mathbf{v}_2 = (\lambda_1 \mathbf{v}_1^{\mathrm{T}})\mathbf{v}_2 = \lambda_1(\mathbf{v}_1 \cdot \mathbf{v}_2). \tag{4}$$

Comparing (3) and (4), we have $(\lambda_2 - \lambda_1)(\mathbf{v}_1 \cdot \mathbf{v}_2) = 0$. Since $\lambda_2 - \lambda_1 \neq 0$ by assumption, we have $\mathbf{v}_1^{\mathrm{T}}\mathbf{v}_2 = 0$. □

Example 16. Let $n = 2$ and let A be the symmetric matrix $A = \begin{pmatrix} 1 & 3 \\ 3 & 1 \end{pmatrix}$. Then

$$\det(A - xI_2) = (1-x)^2 - 9 = x^2 - 2x - 8 = (x-4)(x+2),$$

so the eigenvalues of A are 4 and -2. Solving $A\mathbf{v} = \lambda\mathbf{v}$ for $\lambda = 4$ and -2, we find corresponding eigenvectors $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ and $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$. Proposition 3.7.4 tells us that these vectors are orthogonal to each other (which we can of course check directly!), so if we divide them by their lengths to give vectors of length 1, giving $\begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}$ and $\begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{-1}{\sqrt{2}} \end{pmatrix}$ then we get an orthonormal basis consisting of eigenvectors of A, which is what we want. The corresponding basis change matrix P has these vectors as columns, so $P = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \end{pmatrix}$, and we can check that $P^{\mathrm{T}}P = I_2$ (i.e. P is orthogonal) and that $P^{\mathrm{T}}AP = \begin{pmatrix} 4 & 0 \\ 0 & -2 \end{pmatrix}$.

Example 17. Let's do an example of the "quadratic form" version of the above theorem. Let $n = 3$ and

$$q(\mathbf{v}) = 3x^2 + 6y^2 + 3z^2 - 4xy - 4yz + 2xz,$$
$$\text{so} \quad A = \begin{pmatrix} 3 & -2 & 1 \\ -2 & 6 & -2 \\ 1 & -2 & 3 \end{pmatrix}.$$

Then, expanding by the first row,

$$\begin{aligned} \det(A - xI_3) &= (3-x)(6-x)(3-x) - 4(3-x) - 4(3-x) + 4 + 4 - (6-x) \\ &= -x^3 + 12x^2 - 36x + 32 = (2-x)(x-8)(x-2), \end{aligned}$$

so the eigenvalues are 2 (repeated) and 8. For the eigenvalue 8, if we solve $A\mathbf{v} = 8\mathbf{v}$ then we find a solution $\mathbf{v} = \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}$. Since 2 is a repeated eigenvalue, we need two corresponding eigenvectors, which must be orthogonal to each other. The equations $A\mathbf{v} = 2\mathbf{v}$ all reduce to $a - 2b + c = 0$, and so any vector $\begin{pmatrix} a \\ b \\ c \end{pmatrix}$ satisfying this equation is an eigenvector for $\lambda = 2$. By Proposition 3.7.4 these eigenvectors will all be orthogonal to the eigenvector for $\lambda = 8$, but we will have to choose them orthogonal to each other. We can choose the first one arbitrarily, so let's choose $\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$. We now need another solution that is orthogonal to this. In other words, we want a, b and c not all zero satisfying $a - 2b + c = 0$ and $a - c = 0$, and $a = b = c = 1$ is a solution. So we now have a basis $\begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ of three mutually orthogonal eigenvectors. To get an orthonormal basis, we just need to divide by their lengths, which are, respectively, $\sqrt{6}$, $\sqrt{2}$, and $\sqrt{3}$, and then the basis change matrix P has these vectors as columns, so

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

It can then be checked that $P^{\mathrm{T}}P = I_3$ and that $P^{\mathrm{T}}AP$ is the diagonal matrix with entries 8, 2, 2. So if $\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3$ is this basis, we have

$$q(x\mathbf{f}_1 + y\mathbf{f}_2 + z\mathbf{f}_3) = 8x^2 + 2y^2 + 2z^2.$$

3.8 Quadratic forms in geometry

3.8.1 Reduction of the general second degree equation

The general equation of a second degree polynomial in n variables $x_1, \ldots, x_n$ is

$$\sum_{i=1}^{n} \alpha_i x_i^2 + \sum_{i=1}^{n} \sum_{j=1}^{i-1} \alpha_{ij} x_i x_j + \sum_{i=1}^{n} \beta_i x_i + \gamma = 0. \quad (\dagger)$$

For fixed values of the α's, β's and γ, this defines a *quadric* curve or surface or threefold or... in general $(n-1)$-fold, in n-dimensional euclidean space. To study the possible shapes thus defined, we first simplify this equation by applying coordinate changes resulting from isometries (rigid motions) of $\mathbb{R}^n$; that is, transformations that preserve distance and angle.

By Theorem 3.7.3, we can apply an orthogonal basis change (that is, an isometry of $\mathbb{R}^n$ that fixes the origin) which has the effect of eliminating the terms $\alpha_{ij}x_ix_j$ in the above sum. To carry out this step we consider the

$$\sum_{i=1}^{n} \alpha_i x_i^2 + \sum_{i=1}^{n}\sum_{j=1}^{i-1} \alpha_{ij}x_ix_j$$

term and, when making the orthogonal change of coordinates, we then have to consider its impact on the terms in $\sum_{i=1}^{n} \beta_i x_i$.

For example, suppose we have $x^2 + xy + y^2 + x = 0$. Then $x^2 + xy + y^2$ is the quadratic form associated to the bilinear form with matrix

$$\begin{pmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & 1 \end{pmatrix}$$

with eigenvalues $3/2$ and $1/2$ with associated normalised eigenvectors

$$\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

and indeed, in terms of the new coordinates,

$$x' = \frac{1}{\sqrt{2}}(x+y), \quad y' = \frac{1}{\sqrt{2}}(x-y)$$

we get

$$\begin{aligned} x^2 + xy + y^2 &= \left(\frac{1}{\sqrt{2}}(x'+y')\right)^2 + \left(\frac{1}{\sqrt{2}}(x'+y')\right)\left(\frac{1}{\sqrt{2}}(x'-y')\right) + \left(\frac{1}{\sqrt{2}}(x'-y')\right)^2 \\ &= \frac{3}{2}(x')^2 + \frac{1}{2}(y')^2. \end{aligned}$$

Note that this base change is orthogonal (we can't just complete the square here writing $x^2 + xy + y^2 = (x + (1/2)y)^2 + (3/4)y^2$ because this will **not** give an orthogonal base change!)

Now, whenever $\alpha_i \neq 0$, we can replace x_i by $x_i - \beta_i/(2\alpha_i)$, and thereby eliminate the term $\beta_i x_i$ from the equation. This transformation is just a translation, which is also an isometry.

For example, suppose we have $x^2 - 3x = 0$. Then we are completing the square again, but this time in one variable. So $x^2 - 3x = 0$ is just $(x - \frac{3}{2})^2 - \frac{9}{4} = 0$ and we use $x_1 = x - \frac{3}{2}$ to write it as $x_1^2 - \frac{9}{4}$.

If $\alpha_i = 0$, then we cannot eliminate the term $\beta_i x_i$. Let us permute the coordinates such that $\alpha_i \neq 0$ for $1 \leq i \leq r$, and $\beta_i \neq 0$ for $r+1 \leq i \leq r+s$.

If $s > 1$, we want to leave the x_i alone for $1 \leq i \leq r$ but replace $\sum_{i=1}^{s} \beta_{r+i}x_{r+i}$ by $\beta x'_{r+1}$. We put

$$x'_{r+1} := \frac{1}{\sqrt{\sum_{i=1}^{s} \beta_{r+i}^2}} \sum_{i=1}^{s} \beta_{r+i}x_{r+i}, \quad \beta = \sqrt{\sum_{i=1}^{s} \beta_{r+i}^2}.$$

Then

$$x_1, \ldots, x_r, x'_{r+1}$$

are orthonormal (with respect to the standard inner product $\sum_i a_i b_i$ between $\sum_i a_i x_i$, $\sum_i b_i x_i$) and

$$x_1, \ldots, x_r, x'_{r+1}, x_{r+2}, \ldots, x_n$$

are a basis which we can make orthonormal by running the Gram-Schmidt procedure in Theorem 3.5.3 on it (this corresponds to an orthogonal base change on the e_i, too, since the transpose and inverse of an orthogonal matrix are orthogonal). By abuse of notation (or using dynamical names for the variables), we again denote the resulting new coordinates by $x_1, \ldots, x_n$.

So we have reduced our equation to at most one non-zero β_i; either there are no linear terms at all, or there is just $\beta_{r+1}x_{r+1}$. Dividing through by a constant we can choose β_{r+1} to be -1 for convenience.

Finally, if there is a linear term, we can then perform the translation that replaces x_{r+1} by $x_{r+1} + \gamma$, and thereby eliminate the constant γ. When there is no linear term then we divide the equation through by a constant, to assume that γ is 0 or -1 and we put γ on the right hand side for convenience.

We have proved the following theorem:

Theorem 3.8.1. *By rigid motions of euclidean space, we can transform the set defined by the general second degree equation* (†) *into the set defined by an equation having one of the following three forms:*

$$\sum_{i=1}^{r} \alpha_i x_i^2 = 0,$$

$$\sum_{i=1}^{r} \alpha_i x_i^2 = 1,$$

$$\sum_{i=1}^{r} \alpha_i x_i^2 - x_{r+1} = 0.$$

Here $0 \leq r \leq n$ *and* $\alpha_1, \ldots, \alpha_r$ *are non-zero constants, and in the third case* $r < n$.

We shall assume that $r \neq 0$, because otherwise we have a linear equation. The sets defined by the first two types of equation are called *central quadrics* because they have central symmetry; i.e. if a vector $\mathbf{v}$ satisfies the equation, then so does $-\mathbf{v}$.

We shall now consider the types of curves and surfaces that can arise in the familiar cases $n = 2$ and $n = 3$. These different types correspond to whether the α_i are positive, negative or zero, and whether $\gamma = 0$ or 1.

We shall use x, y, z instead of x_1, x_2, x_3, and $\pm\alpha$, $\pm\beta$, $\pm\gamma$ instead of α_1, α_2, α_3, assuming also that α, β, γ are all strictly positive. When the coefficient of the right hand side is 0, we will divide through by -1 at will. For example, Case (i) in the next section contains both $\alpha x^2 = 0$ and $-\alpha x^2 = 0$, which of course need not be counted twice. Moreover, if swapping the names of x and y (whilst swapping the arbitrary positive real numbers α and β) gives the same equation, we will only consider it once. For example, we do this for the list in the next section by only listing Case (vii) once ($-\alpha x^2 + \beta y^2 = 1$ is also in this case).

3.8.2 The case $n = 2$

When $n = 2$ we have the following possibilities.

(i) $\alpha x^2 = 0$. This just defines the line $x = 0$ (the y-axis).

(ii) $\alpha x^2 = 1$. This defines the two parallel lines $x = \pm 1/\sqrt{\alpha}$.

(iii) $-\alpha x^2 = 1$. This is the empty set!

(iv) $\alpha x^2 + \beta y^2 = 0$. The single point $(0, 0)$.

(v) $\alpha x^2 - \beta y^2 = 0$. This defines two straight lines $y = \pm\sqrt{\alpha/\beta}\, x$, which intersect at $(0, 0)$.

(vi) $\alpha x^2 + \beta y^2 = 1$. An ellipse.

(vii) $\alpha x^2 - \beta y^2 = 1$. A hyperbola.

(viii) $-\alpha x^2 - \beta y^2 = 1$. The empty set again.

(ix) $\alpha x^2 - y = 0$. A parabola.

3.8.3 The case $n = 3$

When $n = 3$, we still get the nine possibilities (i) – (ix) that we had in the case $n = 2$, but now they must be regarded as equations in the three variables x, y, z that happen not to involve z.

So, in Case (i), we now get the plane $x = 0$, in Case (ii) we get two parallel planes $x = \pm 1/\sqrt{\alpha}$, in Case (iv) we get the line $x = y = 0$ (the z-axis), in Case (v) two intersecting planes $y = \pm\sqrt{\alpha/\beta} x$, and in Cases (vi), (vii) and (ix), we get, respectively, elliptical, hyperbolic and parabolic cylinders.

The remaining cases involve all of x, y and z. We omit $-\alpha x^2 - \beta y^2 - \gamma z^2 = 1$, which is empty.

(x) $\alpha x^2 + \beta y^2 + \gamma z^2 = 0$. The single point $(0, 0, 0)$.

(xi) $\alpha x^2 + \beta y^2 - \gamma z^2 = 0$. See Fig. 1.

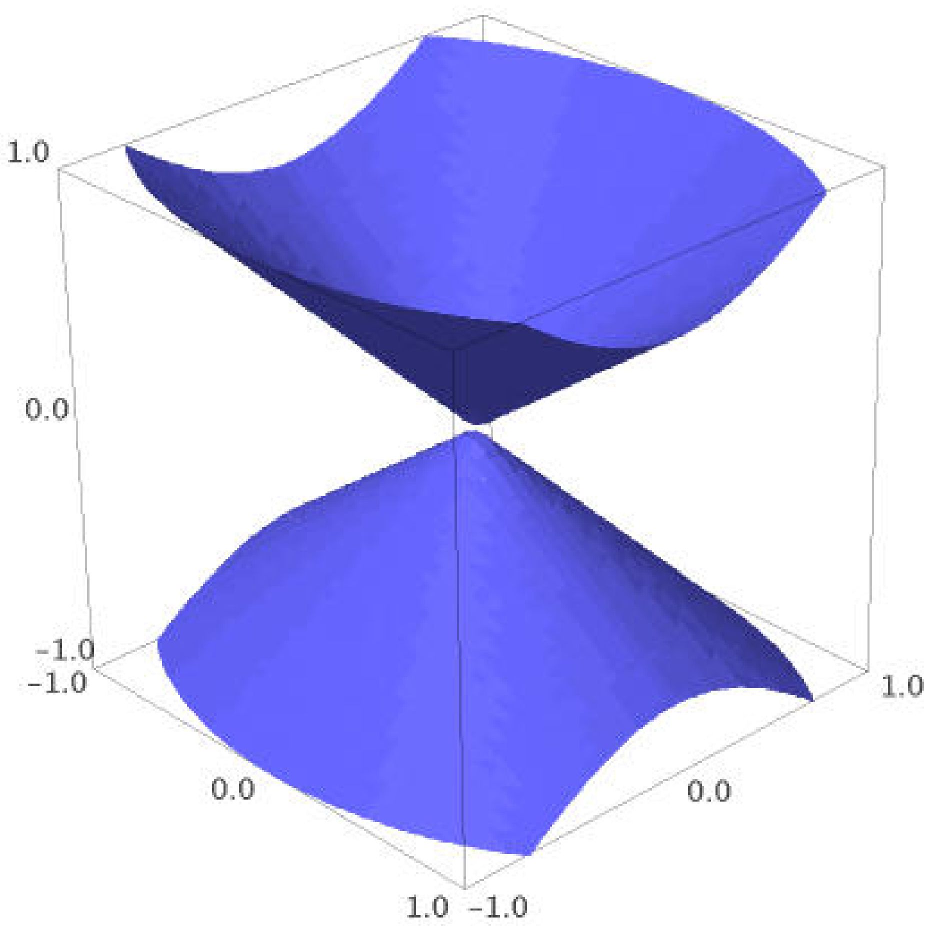

Figure 1: $\frac{1}{2}x^2 + y^2 - z^2 = 0$

This is an elliptical cone. The cross sections parallel to the xy-plane are ellipses of the form $\alpha x^2 + \beta y^2 = c$, whereas the cross sections parallel to the other coordinate planes are generally hyperbolas. Notice also that if a particular point (a, b, c) is on the surface, then so is $t(a, b, c)$ for any $t \in \mathbb{R}$. In other words, the surface contains the straight line through the origin and any of its points. Such lines are called *generators*. When each point of a 3-dimensional surface lies on one or more generators, it is possible to make a model of the surface with straight lengths of wire or string.

(xii) $\alpha x^2 + \beta y^2 + \gamma z^2 = 1$. An ellipsoid. See Fig. 2.

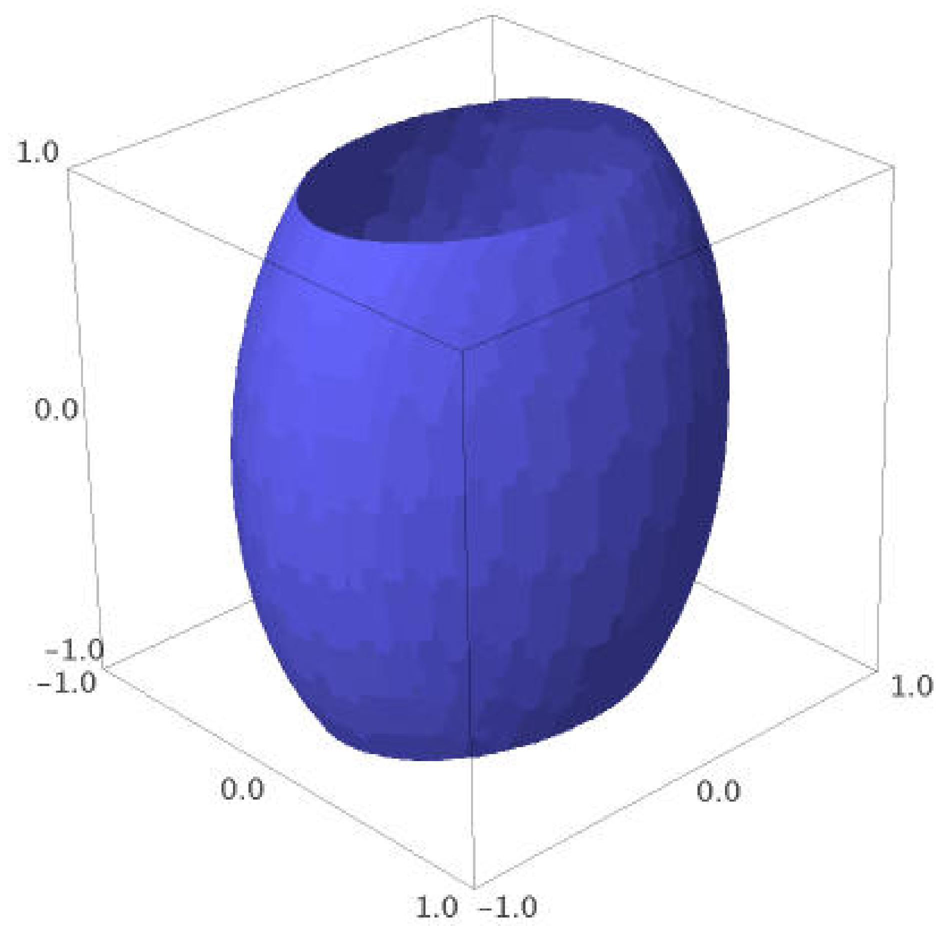

Figure 2: $2x^2 + y^2 + \frac{1}{2}z^2 = 1$

This is a “squashed sphere”. It is bounded, and hence clearly has no generators. Notice that if α, β, and γ are distinct, it has only the finite group of symmetries given by reflections in x, y and z, but if some two of the coefficients coincide, it picks up an infinite group of rotation symmetries.

(xiii) $\alpha x^2 + \beta y^2 - \gamma z^2 = 1$. A hyperboloid. See Fig. 3.

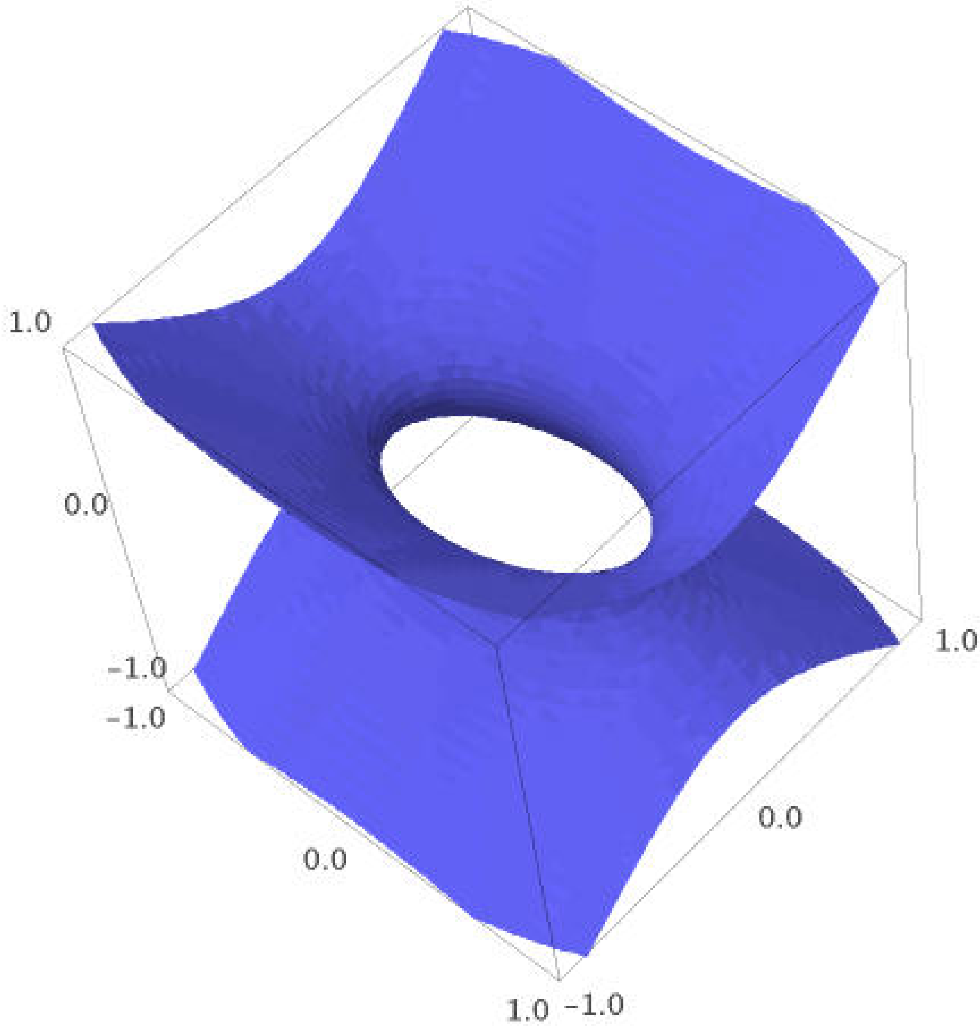

Figure 3: $3x^2 + 8y^2 - 8z^2 = 1$

There are two types of 3-dimensional hyperboloids. This one is connected, and is known as a *hyperboloid of one sheet*. Any cross-section in the xy direction will be an ellipse, and these get larger as z grows (notice the hole in the middle in the picture). Although it is not immediately obvious, each point of this surface lies on exactly two generators; that is, lines that lie entirely on the surface. For each $\lambda \in \mathbb{R}$, the line defined by the pair of equations

$$\sqrt{\alpha}\, x - \sqrt{\gamma}\, z = \lambda(1 - \sqrt{\beta}\, y); \qquad \lambda(\sqrt{\alpha}\, x + \sqrt{\gamma}\, z) = 1 + \sqrt{\beta}\, y.$$

lies entirely on the surface; to see this, just multiply the two equations together. The same applies to the lines defined by the pairs of equations

$$\sqrt{\beta}\, y - \sqrt{\gamma}\, z = \mu(1 - \sqrt{\alpha}\, x); \qquad \mu(\sqrt{\beta}\, y + \sqrt{\gamma}\, z) = 1 + \sqrt{\alpha}\, x.$$

It can be shown that each point on the surface lies on exactly one of the lines in each of these two families.

There is a photo at http://home.cc.umanitoba.ca/~gunderso/model_photos/misc/hyperboloid_of_one_sheet.jpg depicting a rather nice wooden model of a hyperboloid of one sheet, which gives a good idea how these lines sit inside the surface.

(xiv) $\alpha x^2 - \beta y^2 - \gamma z^2 = 1$. Another kind of hyperboloid. See Fig. 4.

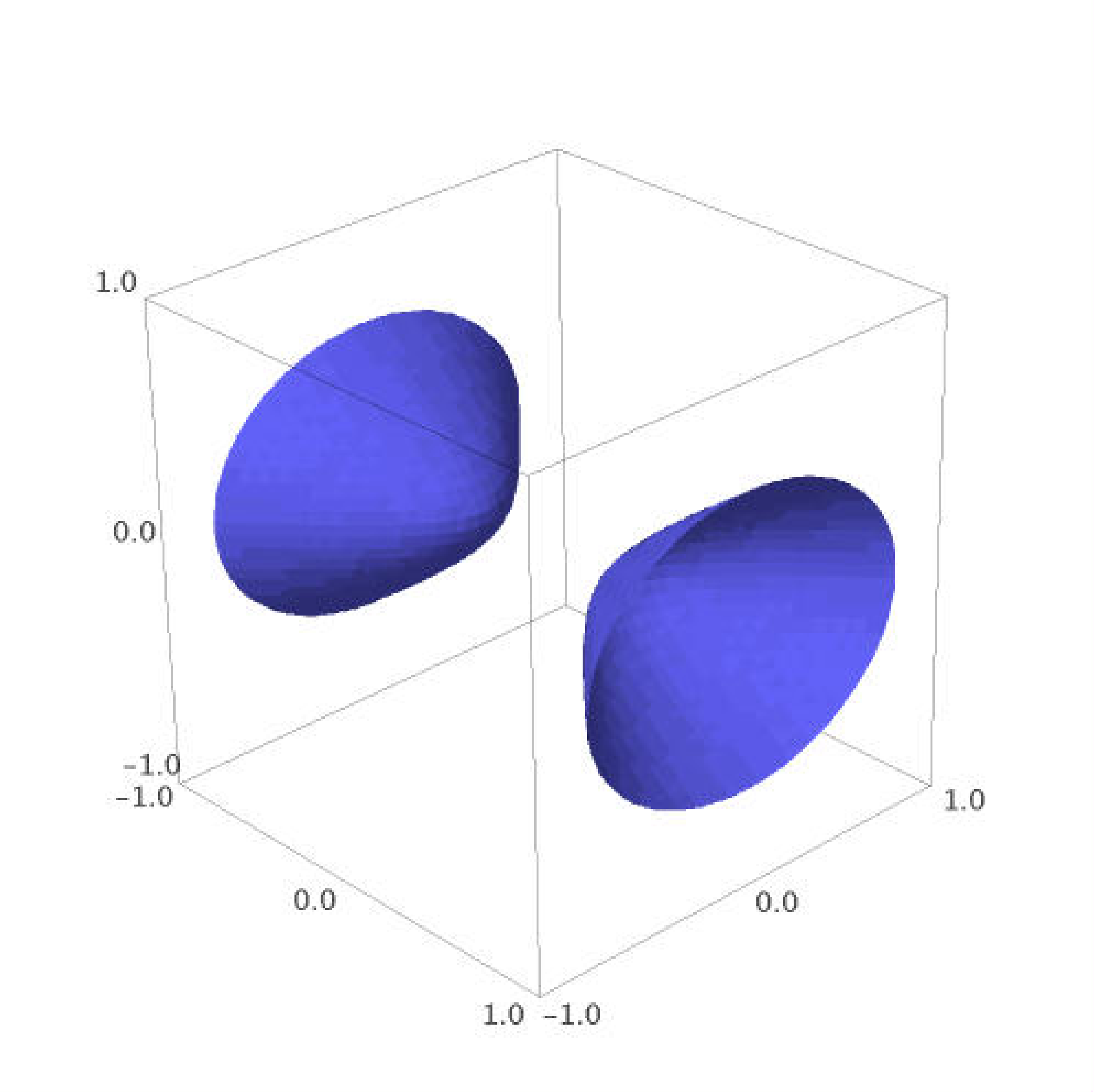

Figure 4: $8x^2 - 12y^2 - 20z^2 = 1$

This one has two connected components and is called a *hyperboloid of two sheets*. It does not have generators.

(xv) $\alpha x^2 + \beta y^2 - z = 0$. An elliptical paraboloid. See Fig. 5.

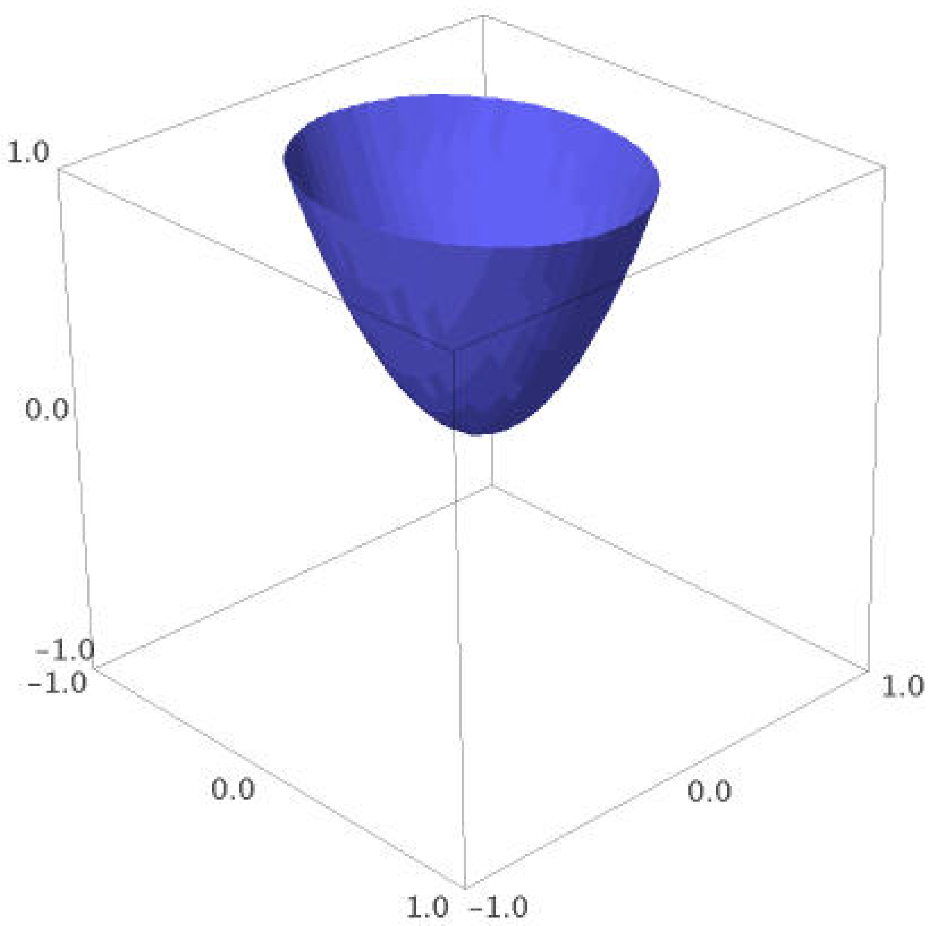

Figure 5: $2x^2 + 3y^2 - z = 0$

Cross-sections of this surface parallel to the xy plane are ellipses, while cross-sections in the yz and xz directions are parabolas. It can be regarded as the limit of a family of hyperboloids of two sheets, where one "cap" remains at the origin and the other recedes to infinity.

(xvi) $\alpha x^2 - \beta y^2 - z = 0$. A hyperbolic paraboloid (a rather elegant saddle shape). See Fig. 6.

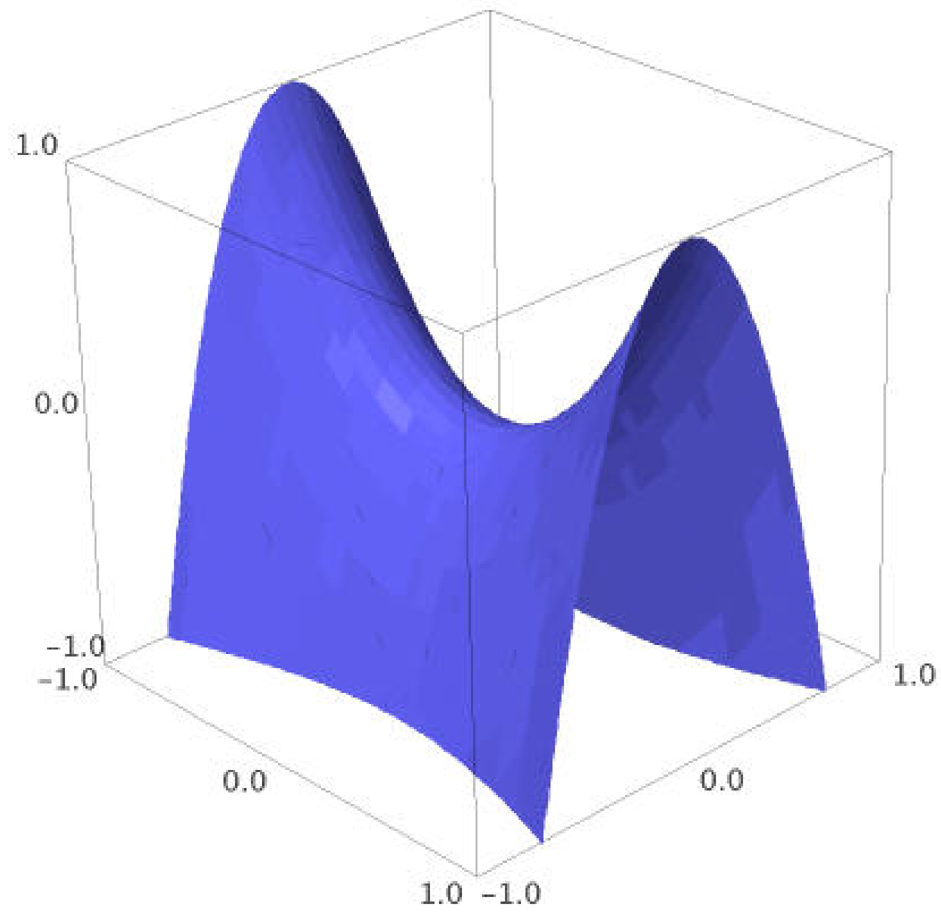

Figure 6: $x^2 - 4y^2 - z = 0$

As in the case of the hyperboloid of one sheet, there are two generators passing through each point of this surface, one from each of the following two families of lines:

$$\lambda(\sqrt{\alpha}\,x - \sqrt{\beta}\,y) = z; \qquad \sqrt{\alpha}\,x + \sqrt{\beta}\,y = \lambda.$$
$$\mu(\sqrt{\alpha}\,x + \sqrt{\beta}\,y) = z; \qquad \sqrt{\alpha}\,x - \sqrt{\beta}\,y = \mu.$$

Just as the elliptical paraboloid was a limiting case of a hyperboloid of two sheets, so the hyperbolic paraboloid is a limiting case of a hyperboloid of one sheet: you can imagine gradually deforming the hyperboloid of one sheet so the elliptical hole in the middle becomes bigger and bigger, and the result is the hyperbolic paraboloid.

3.9 Singular value decomposition

In this section we want to study what linear maps $T : V \to W$ between Euclidean spaces look like? From MA106 Linear Algebra we know that we can choose bases in V and W such that the matrix of T in Smith normal form is $\left(\begin{array}{c|c} I_n & 0 \\ \hline 0 & 0 \end{array}\right)$ where n is the rank of T. This answer is unsatisfactory in our case because it does not take the Euclidean geometry of V and W into account. In other words, we want to choose orthonormal bases, not just any bases. This leads us to the singular value decomposition, SVD for short.

Notation: We will see various diagonal matrices in the following so we will use the shorthand $\text{diag}(d_1, \ldots, d_n)$ for an $n \times n$ diagonal matrix with diagonal entries $d_1, \ldots, d_n$.

Theorem 3.9.1 (SVD for linear maps). *Suppose $T : V \to W$ is a linear map of rank n between Euclidean spaces. Then there exist unique positive numbers $\gamma_1 \geq \gamma_2 \geq \ldots \geq \gamma_n > 0$, called the singular values of T, and orthonormal bases of V and W such that the matrix of T with respect to these bases is*

$$\left(\begin{array}{c|c} D & 0 \\ \hline 0 & 0 \end{array}\right) \quad \textit{where} \quad D = \textit{diag}(\gamma_1, \ldots, \gamma_n).$$

In fact, the γ's are nothing but the positive square-roots of the nonzero eigenvalues of T^*T, each one appearing as many times as the dimension of the corresponding eigenspace, where T^* is the adjoint of T. Here by adjoint we mean the unique linear map $T^*\colon W \to V$ such that

$$\langle T\mathbf{v}, \mathbf{w}\rangle_W = \langle \mathbf{v}, T^*\mathbf{w}\rangle_V$$

where $\langle\cdot,\cdot\rangle_V$ and $\langle\cdot,\cdot\rangle_W$ are the inner products on V and W (we will also denote them just by a dot if there is no risk of confusion).

Proof. We will consider a new symmetric bilinear form on V defined as follows.

$$\mathbf{u} \star \mathbf{v} := T(\mathbf{u}) \cdot T(\mathbf{v}) = \mathbf{u} \cdot T^*T(\mathbf{v}) \ .$$

Note that $\mathbf{v} \star \mathbf{v} = T(\mathbf{v}) \cdot T(\mathbf{v}) \geq 0$; we call such a bilinear form *positive semidefinite* (note that it need not be positive definite because T can have a non-zero kernel). By Theorem 3.7.3, there exist unique constants $\alpha_1 \geq \ldots \geq \alpha_m$ (eigenvalues of the matrix of the $\star$ bilinear form) and an orthonormal basis $\mathbf{e}_1, \ldots, \mathbf{e}_m$ of V such that the bilinear form $\star$ is given by $\text{diag}(\alpha_1, \ldots, \alpha_m)$ in this basis. Since $\star$ is positive semidefinite we see that all α_i are non-negative. Suppose $\alpha_k > 0$ is the last positive eigenvalue, that is, $\alpha_{k+1} = \cdots = \alpha_m = 0$.

The kernel of T^*T is equal to the kernel of T (they are the same subspace of V) because

$$T(\mathbf{v}) \cdot T(\mathbf{v}) = \mathbf{v} \cdot (T^*T)(\mathbf{v}).$$

and hence $T(\mathbf{e}_{k+1}) = \cdots = T(\mathbf{e}_m) = 0$. Moreover, $T(\mathbf{e}_1), \ldots, T(\mathbf{e}_k)$ form an orthogonal set of vectors in W. It follows that k is the rank of T since a set of orthogonal vectors is linearly independent. Thus, $k = n$. We define $\gamma_i := \sqrt{\alpha_i}$ for all $i \leq k$.

We now use these image vectors $T(\mathbf{e}_i)$ to build an orthonormal basis of W. Since $T(\mathbf{e}_i) \cdot T(\mathbf{e}_i) = \mathbf{e}_i \star \mathbf{e}_i = \alpha_i$, we know that $|T(\mathbf{e}_i)| = \sqrt{\alpha_i} = \gamma_i$. Let $\mathbf{f}_i := \frac{T(\mathbf{e}_i)}{\gamma_i}$ for all $i \leq n$. We can then extend this orthonormal set of vectors to an orthonormal basis of W by the Gram-Schmidt process (Theorem 3.5.3). Since $T(\mathbf{e}_i) = \gamma_i \mathbf{f}_i$ for $i \leq n$ and $T(\mathbf{e}_j) = 0$ for $j > n$, the matrix of T with respect to these bases has the required form.

It remains to prove the uniqueness of the singular values. Suppose we have orthonormal bases $\mathbf{e}'_1, \ldots, \mathbf{e}'_m$ of V and $\mathbf{f}'_1, \ldots, \mathbf{f}'_s$ of W, in which T is represented by a matrix $\left(\begin{array}{c|c} B & 0 \\ \hline 0 & 0 \end{array}\right)$ where $B = \operatorname{diag}(\beta_1, \ldots, \beta_t)$ with $\beta_1 \geq \ldots \geq \beta_t > 0$. Put $\beta_i = 0$ for $i > t$. Then $\mathbf{e}'_i \star \mathbf{e}'_j = \beta_i \mathbf{f}'_i \cdot \beta_j \mathbf{f}'_j = \delta_{ij}\beta_i^2$. Thus, $\operatorname{diag}(\beta_1^2, \ldots, \beta_m^2)$ is the matrix of the bilinear form $\star$ in the basis $\mathbf{e}'_1, \ldots, \mathbf{e}'_m$. Uniqueness in Theorem 3.7.3 implies the uniqueness of the singular values. □

Before we proceed with some examples, all on the standard euclidean spaces $\mathbb{R}^n$, let us restate the SVD for matrices:

Corollary 3.9.2 (SVD for matrices)**.** *Given any real $k \times m$ matrix A, there exist unique singular values $\gamma_1 \geq \gamma_2 \geq \ldots \geq \gamma_n > 0$ and (non-unique) orthogonal matrices P and Q such that*

$$\left(\begin{array}{c|c} D & 0 \\ \hline 0 & 0 \end{array}\right) = P^T A Q \text{ where } D = \operatorname{diag}(\gamma_1, \ldots, \gamma_n).$$

Equivalently, we say the SVD of A is

$$A = P \left(\begin{array}{c|c} D & 0 \\ \hline 0 & 0 \end{array}\right) Q^T \text{ where } D = \operatorname{diag}(\gamma_1, \ldots, \gamma_n).$$

Here the γ's are the positive square roots of the nonzero eigenvalues of $A^T A$.

Example. Consider a linear map $\mathbb{R}^2 \to \mathbb{R}^2$, given by the symmetric matrix $A = \begin{pmatrix} 1 & 3 \\ 3 & 1 \end{pmatrix}$, in the example from Section 3.7. There we found the orthogonal matrix $P = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \end{pmatrix}$ with $P^T A P = \begin{pmatrix} 4 & 0 \\ 0 & -2 \end{pmatrix}$. This is not the SVD of A because the diagonal matrix contains a negative entry. To get to the SVD we just need to pick different bases for the domain and the range: the columns $\mathbf{c}_1$, $\mathbf{c}_2$ can still be a basis of the domain, while the basis of the range could become $\mathbf{c}_1$, $-\mathbf{c}_2$. This is the SVD:

$$P = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}, \quad Q = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \end{pmatrix}, \quad P^T A Q = \begin{pmatrix} 4 & 0 \\ 0 & 2 \end{pmatrix}.$$

The same method works for any symmetric matrix: the SVD is just orthogonal diagonalisation with additional care needed for signs. If the matrix is not symmetric, we need to follow the proof of Theorem 3.9.1 during the calculation.

Example. Consider a linear map $\mathbb{R}^3 \to \mathbb{R}^2$, given by $A = \begin{pmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{pmatrix}$. Since $\mathbf{x} \star \mathbf{y} = A\mathbf{x} \cdot A\mathbf{y} = (A\mathbf{x})^T A\mathbf{y} = \mathbf{x}^T (A^T A)\mathbf{y}$, the matrix of the bilinear form $\star$ in the standard basis is

$$A^T A = \begin{pmatrix} 4 & 8 \\ 11 & 7 \\ 14 & -2 \end{pmatrix} \begin{pmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{pmatrix} = \begin{pmatrix} 80 & 100 & 40 \\ 100 & 170 & 140 \\ 40 & 140 & 200 \end{pmatrix}.$$

The eigenvalues of this matrix are 360, 90 and 0. Hence the singular values of A are

$$\gamma_1 = \sqrt{360} = 6\sqrt{10} \geq \gamma_2 = \sqrt{90} = 3\sqrt{10}\,.$$

At this stage we are assured of the existence of orthogonal matrices P and Q such that
$$P^{\mathrm{T}}AQ = \begin{pmatrix} 6\sqrt{10} & 0 & 0 \\ 0 & 3\sqrt{10} & 0 \end{pmatrix}.$$
To find such orthogonal matrices we first need to find an orthonormal basis of eigenvectors of A^TA. Since the eigenvalues are distinct on this occasion we only need to find an eigenvector for each eigenvalue and normalise it so it has length 1. This leads to:
$$\mathbf{e}_1 = \begin{pmatrix} 1/3 \\ 2/3 \\ 2/3 \end{pmatrix}, \ \mathbf{e}_2 = \begin{pmatrix} -2/3 \\ -1/3 \\ 2/3 \end{pmatrix}, \ \mathbf{e}_3 = \begin{pmatrix} 2/3 \\ -2/3 \\ 1/3 \end{pmatrix}.$$
These make up Q. Then we need to find the images of these vectors under A divided by the corresponding singular value (so only the eigenvectors for the non-zero eigenvalues of A^TA):
$$\mathbf{f}_1 = \frac{1}{6\sqrt{10}}A\mathbf{e}_1 = \begin{pmatrix} 3/\sqrt{10} \\ 1/\sqrt{10} \end{pmatrix}, \ \mathbf{f}_2 = \frac{1}{3\sqrt{10}}A\mathbf{e}_2 = \begin{pmatrix} 1/\sqrt{10} \\ -3/\sqrt{10} \end{pmatrix}.$$
The proof says we need to extend this to a basis of W, which is easy here because we already have two vectors and so we don't need anymore for a basis of $\mathbb{R}^2$. Hence, the orthogonal matrices are
$$P = \begin{pmatrix} 3/\sqrt{10} & 1/\sqrt{10} \\ 1/\sqrt{10} & -3/\sqrt{10} \end{pmatrix}, \ Q = \begin{pmatrix} 1/3 & -2/3 & 2/3 \\ 2/3 & -1/3 & -2/3 \\ 2/3 & 2/3 & 1/3 \end{pmatrix}.$$

3.10 The complex story

The results in Subsection 3.7 applied only to vector spaces over the real numbers $\mathbb{R}$. There are corresponding results for spaces over the complex numbers $\mathbb{C}$, which we shall summarize here. We only include one proof, although the others are similar and analogous to those for spaces over $\mathbb{R}$.

3.10.1 Sesquilinear forms

The key thing that made everything work over $\mathbb{R}$ was the fact that if $x_1, \ldots, x_n$ are real numbers, and $x_1^2 + \cdots + x_n^2 = 0$, then all the x_i are zero. This doesn't work over $\mathbb{C}$: take $x_1 = 1$ and $x_2 = i$. But we do have something similar if we bring *complex conjugation* into play. As usual, for $z \in \mathbb{C}$, we let $\bar{z}$ denote the complex conjugate of z. Then if $z_1\bar{z}_1 + \cdots + z_n\bar{z}_n = 0$, each z_i must be zero. So we need to "put bars on half of our formulae". Notice that there was a hint of this in the proof of Proposition 3.7.2.

We'll do this as follows.

Definition 3.10.1. A *sesquilinear form* on a complex vector space V is a function $\tau : V \times V \to \mathbb{C}$ such that
$$\tau(\mathbf{v}, a_1\mathbf{w}_1 + a_2\mathbf{w}_2) = a_1\tau(\mathbf{v}, \mathbf{w}_1) + a_2\tau(\mathbf{v}, \mathbf{w}_2)$$

(as before), but

$$\tau(a_1\mathbf{v}_1 + a_2\mathbf{v}_2, \mathbf{w}) = \bar{a}_1\tau(\mathbf{v}_1, \mathbf{w}) + \bar{a}_2\tau(\mathbf{v}_2, \mathbf{w}),$$

for all vectors v_1, v_2, v, w_1, w_2, w and all $a_1, a_2 \in \mathbb{C}$.

We say such a form is *hermitian symmetric* if

$$\tau(\mathbf{w}, \mathbf{v}) = \overline{\tau(\mathbf{v}, \mathbf{w})}.$$

The word "sesquilinear" literally means "one-and-a-half-times-linear" from its Latin meaning – it's linear in the second argument, but only halfway there in the first argument! We'll often abbreviate "hermitian-symmetric sesquilinear form" to just "hermitian form".

We can represent these by matrices in a similar way to bilinear forms. If τ is a sesquilinear form, and $\mathbf{e}_1, \ldots, \mathbf{e}_n$ is a basis of V, we define the matrix of τ to be the matrix A whose i, j entry is $\tau(\mathbf{e}_i, \mathbf{e}_j)$. Then we have

$$\tau(\mathbf{v}, \mathbf{w}) = \overline{(\underline{\mathbf{v}}^{\mathrm{T}})}A\underline{\mathbf{w}}$$

where $\underline{\mathbf{v}}$ and $\underline{\mathbf{w}}$ are the coordinates of $\mathbf{v}$ and $\mathbf{w}$ as usual. We'll shorten this to $\underline{\mathbf{v}}^* A\underline{\mathbf{w}}$, where the $*$ denotes "conjugate transpose". The condition to be hermitian symmetric translates to the relation $a_{ji} = \overline{a_{ij}}$, so τ is hermitian if and only if A satisfies $A^* = A$.

We have a version here of Sylvester's two theorems (Proposition 3.4.3 and Theorem 3.4.4):

Theorem 3.10.2. *If τ is a hermitian form on a complex vector space V, there is a basis of V in which the matrix of τ is given by*

$$\left(\begin{array}{c|c|c} I_t & & \\ \hline & -I_u & \\ \hline & & 0 \end{array}\right)$$

for some uniquely determined *integers t and u.*

As in the real case, we call the pair (t, u) the *signature* of τ, and we say τ is *positive definite* if its signature is $(n, 0)$ (if V is an n-dimensional space). In this case, the theorem tells us that there is a basis of V in which the matrix of τ is the identity, and in such a basis we have

$$\tau(\mathbf{v}, \mathbf{v}) = \sum_{i=1}^{n} |v_i|^2$$

where $v_1, \ldots, v_n$ are the coordinates of $\mathbf{v}$. Hence $\tau(\mathbf{v}, \mathbf{v}) > 0$ for all non-zero $\mathbf{v} \in V$.

Just as we defined a euclidean space to be a real vector space with a choice of positive definite bilinear form, we have a similar definition here:

Definition 3.10.3. A *Hilbert space* is a finite-dimensional complex vector space endowed with a choice of positive-definite hermitian-symmetric sesquilinear form.

These are the complex analogues of euclidean spaces. If V is a Hilbert space, we write $\mathbf{v} \cdot \mathbf{w}$ for the sesquilinear form on V, and we refer to it as an *inner product*. For any Hilbert space, we can always find a basis $\mathbf{e}_1, \ldots, \mathbf{e}_n$ of V such that $\mathbf{e}_i \cdot \mathbf{e}_j = \delta_{ij}$ (an orthonormal basis). Then we can write the inner product matrix-wise as

$$\mathbf{v} \cdot \mathbf{w} = \underline{\mathbf{v}}^* \underline{\mathbf{w}},$$

where $\underline{\mathbf{v}}$ and $\underline{\mathbf{w}}$ are the coordinates of $\mathbf{v}$ and $\mathbf{w}$ and $\underline{\mathbf{v}}^* = \overline{\underline{\mathbf{v}}}^{\mathrm{T}}$ as before.

The canonical example of a Hilbert space is $\mathbb{C}^n$, with the standard inner product given by

$$\mathbf{v} \cdot \mathbf{w} = \sum_{i=1}^{n} \overline{v_i} w_i,$$

for which the standard basis is obviously orthonormal.

Remark. *Technically, we should say "finite-dimensional Hilbert space". There are lots of interesting infinite-dimensional Hilbert spaces, but we won't say anything about them in this course. (Curiously, one never seems to come across infinite-dimensional euclidean spaces.)*

3.10.2 Operators on Hilbert spaces

In our study of linear operators on euclidean spaces, the idea of the *adjoint* of an operator was important. There's an analogue of it here:

Definition 3.10.4. Let $T : V \to V$ be a linear operator on a Hilbert space V. Then there is a unique linear operator $T^* : V \to V$ (the *hermitian adjoint* of T) such that

$$T(\mathbf{v}) \cdot \mathbf{w} = \mathbf{v} \cdot T^*(\mathbf{w}).$$

It's clear that if A is the matrix of T in an orthonormal basis, then the matrix of T^* is A^*.

Definition 3.10.5. We say that T is

- *selfadjoint* if $T^* = T$,
- *unitary* if $T^* = T^{-1}$,
- *normal* if $T^*T = TT^*$.

Exercise. If T is unitary, then $T(\mathbf{u}) \cdot T(\mathbf{v}) = \mathbf{u} \cdot \mathbf{v}$ for all $\mathbf{u}, \mathbf{v}$ in V.

Using this exercise we can also replicate Proposition 3.6.5 in the complex world. This shows that 'unitary' is the complex analgoue of 'orthogonal'. The proof is entirely similar to that of Proposition 3.6.5 (which comes before the statement).

Proposition 3.10.6. *Let $e_1, \ldots, e_n$ be an orthonormal basis of a Hilbert space V. A linear map T is unitary if and only if $T(e_1), \ldots, T(e_n)$ is an orthonormal basis of V.*

If A is the matrix of T in an orthonormal basis, then it's clear that T is selfadjoint if and only if $A^* = A$ (a hermitian-symmetric matrix), unitary if and only if $A^* = A^{-1}$ (a *unitary matrix*), and normal if and only if $A^*A = AA^*$ (a *normal matrix*). In other words, these properties are preserved under unitary base changes:

Lemma 3.10.7. *If $A \in \mathbb{C}^{n,n}$ is normal (selfadjoint, unitary) and $P \in \mathbb{C}^{n,n}$ is unitary, then P^*AP is normal (selfadjoint, unitary).*

Proof. Let $B = P^*AP$. Using the property $(MN)^* = N^*M^*$, we compute that in the first (normal) case,

$$BB^* = (P^*AP)(P^*AP)^* = P^*APP^*A^*P = P^*AA^*P = P^*A^*AP = (P^*A^*P)(P^*AP) = B^*B.$$

In the second (selfadjoint) case, $B^* = P^*A^*P = P^*AP = B$. In the third (unitary) case, $BB^* = P^*APP^*A^*P = P^*AA^*P = P^*P = I$. □

Notice that if A is unitary and the entries of A are real, then A must be orthogonal, but the definition also includes things like

$$\begin{pmatrix} i & 0 \\ 0 & i \end{pmatrix}.$$

Similarly, a matrix with real entries is hermitian-symmetric if and only if it's symmetric, but

$$\begin{pmatrix} 2 & i \\ -i & 3 \end{pmatrix}$$

is a hermitian-symmetric matrix that's not symmetric.

Both selfadjoint and unitary operators are normal. The generalisation of Theorem 3.7.3 applies to all three types of operators.

Theorem 3.10.8. *The following statements hold for a linear operator $T : V \to V$ on a Hilbert space.*

(i) T is normal if and only if there exists an orthonormal basis of V consisting of eigenvectors of T.

(ii) T is selfadjoint if and only if there exists an orthonormal basis of V consisting of eigenvectors of T with real eigenvalues.

(iii) T is unitary if and only if there exists an orthonormal basis of V consisting of eigenvectors of T with eigenvalues of absolute value 1.

Example. Let $A = \begin{pmatrix} 6 & 2+2i \\ 2-2i & 4 \end{pmatrix}$. Then

$$c_A(x) = (6-x)(4-x) - (2+2i)(2-2i) = x^2 - 10x + 16 = (x-2)(x-8),$$

so the eigenvalues are 2 and 8. Corresponding eigenvectors are $\mathbf{v}_1 = (1+i, -2)^{\mathrm{T}}$ and $\mathbf{v}_2 = (1+i, 1)^{\mathrm{T}}$. We find that $|\mathbf{v}_1|^2 = \mathbf{v}_1^*\mathbf{v}_1 = 6$ and $|\mathbf{v}_2|^2 = 3$, so we divide by their lengths to get an orthonormal basis $\mathbf{v}_1/|\mathbf{v}_1|$, $\mathbf{v}_2/|\mathbf{v}_2|$ of $\mathbb{C}^2$. Then the matrix

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

having this basis as columns is selfadjoint and satisfies $P^*AP = \begin{pmatrix} 2 & 0 \\ 0 & 8 \end{pmatrix}$.

4 Duality, quotients, tensors and all that

In this section we will introduce and discuss at length properties of the dual vector space to a vector space. After that we will turn to some ubiquitous and very useful constructions in multilinear algebra: tensor products, the exterior and symmetric algebra, and several applications.

From this point onwards we will abandon the practice of denoting vectors by lower case boldface letters (to prepare you for real life outside the Warwick UG curriculum since you are grownups now and many text books and research articles do not adhere to that notational practice). We will always be absolutely clear about the meaning of each symbol introduced, so there will be no risk of confusion. Vectors tend to be, as usual, lower case Roman letters such as $v, w, \dots$, and scalars in the ground field K have a penchant to be lower case Greek letter such as $\lambda, \mu, \nu \dots$.

4.1 The dual vector space and quotient spaces

Let V be any vector space over a field K (which need not even be of finite dimension at this point). We consider the set of all linear forms on V, i.e., the set of all linear mappings $l\colon V \to K$, and denote it by V^*. More generally, for any vector space W, we denote by

$$\mathrm{Hom}_K(V, W)$$

the set of all K-linear mappings from V to W; we note that this is a vector space in a natural way if we define addition and scalar multiplication "pointwise" as follows:

$$\forall\, f, g \in \mathrm{Hom}_K(V, W),\ \forall\, v \in V,\ \forall\, \lambda \in K\colon \quad (f+g)(v) := f(v) + g(v),$$
$$(\lambda f)(v) := \lambda f(v).$$

Thus in particular, V^* is again a vector space over K, which we call the *dual vector space* to V. In the remainder of this subsection I will try to convince you that V^* is a really cool and useful thing that can be used to solve many linear algebra problems conceptually and transparently; moreover, duality as a process is used everywhere in mathematics, in representation theory, functional analysis, commutative and homological algebra, topology...

First we need to develop some basic properties of V^*. The first and most obvious is that the construction is, in fancy language, "functorial" with respect to linear maps of vector spaces and reverses all arrows: this means that if you have a linear map

$$f\colon V \to W$$

you get a natural linear map in the other direction between duals:

$$f^*\colon W^* \to V^*$$

by defining

$$\forall l_W \in W^*\, \forall v \in V\colon \quad (f^*(l_W))(v) = (l \circ f)(v) = l(f(v))$$

(this is just "precomposing the given linear form on W with the linear map f"). We call f^* *the linear map dual to* f. As a little exercise you should check that f^* is surjective resp. injective if and only if f is injective resp. surjective.

It is then straightforward and boring to check the following for vector spaces V, W, T (which you should do because you are just learning about duals and need the practice to get a feeling for them):

$$\forall f_1 \in \mathrm{Hom}_K(V,W),\, \forall f_2 \in \mathrm{Hom}_K(W,T)\colon \quad (f_2 \circ f_1)^* = f_1^* \circ f_2^*, (\mathrm{id}_V)^* = \mathrm{id}_V.$$

Here id_V is the identity map from V to V. If you want to intimidate other students learning about this and brag about the range of words you command, you can say that the operation of taking duals defines a contravariant functor from the category of vector spaces to itself (which is what the preceding formulas amount to).

Moreover, $(-)^*$ is compatible with the vector space structure on $\mathrm{Hom}_K(V,W)$ in the sense that

$$\forall f, g \in \mathrm{Hom}_K(V,W),\, \forall \lambda, \mu \in K\colon \quad (\lambda f + \mu g)^* = \lambda f^* + \mu g^*.$$

So far so good. Now assume V is finite dimensional with basis $e_1, \dots, e_n$. Define elements $e_i^* \in V^*$ by

$$e_i^*(e_j) = \delta_{ij}$$

where δ_{ij} is the Kronecker delta symbol- by definition 1 if $i = j$ and 0 otherwise.

Lemma 4.1.1. *The elements $e_1^*, \dots, e_n^*$ form a basis of V^*.*

We call $e_1^*, \dots, e_n^*$ the dual basis to the basis $e_1, \dots, e_n$. Thus given an ordered basis $\mathbf{E} = (e_1, \dots, e_n)$ in V, the operation $(-)^*$ spits out another ordered basis $\mathbf{E}^* = (e_1^*, \dots, e_n^*)$ in V^* "dual" to the given one.

Proof. We need to check that $e_1^*, \dots, e_n^*$ are linearly independent in V^* and generate V^*. Suppose

$$\lambda_1 e_1^* + \dots + \lambda_n e_n^* = 0$$

is a linear dependency relation in V^* between the e_i^*. Here the λ_i are in K of course. Applying the linear map on the left hand side of the previous displayed equation to e_j yields $\lambda_j = 0$, hence the $e_1^*, \dots, e_n^*$ are linearly independent in V^*.

To show that $e_1^*, \dots, e_n^*$ generate V^* we have to use that V is finite dimensional (otherwise it is not necessarily true by the way). Indeed, let $l \in V^*$ be arbitrary. Then the linear form

$$L := l(e_1)e_1^* + \dots + l(e_n)e_n^*$$

takes the same values on all the e_i, $i = 1, \dots, n$, as l, hence $L = l$ and consequently $e_1^*, \dots, e_n^*$ generate V^*. □

Now suppose we are given two finite-dimensional vector spaces V, W of dimension $\dim V = n$, $\dim W = m$, and let $\mathbf{B} = (b_1, \dots b_n)$ and $\mathbf{C} = (c_1, \dots, c_m)$ be ordered bases in V and W respectively. Consider a linear map

$$f\colon V \to W.$$

We know we can associate to this setup an $m \times n$ matrix with entries in K, representing f with respect to the given bases in source and target; this matrix is

$$M_{\mathbf{B}}^{\mathbf{C}}(f)$$

in our previously used notation. Now a natural question is: what is

$$M_{\mathbf{C}^*}^{\mathbf{B}^*}(f^*)$$

and how is it related to $M_{\mathbf{B}}^{\mathbf{C}}(f)$? It is clear that $M_{\mathbf{C}^*}^{\mathbf{B}^*}(f^*)$ is an $n \times m$ matrix, so a natural guess is it could be the transpose of $M_{\mathbf{B}}^{\mathbf{C}}(f)$. That is indeed the case.

Lemma 4.1.2. *We have*

$$M_{\mathbf{C}^*}^{\mathbf{B}^*}(f^*) = \left(M_{\mathbf{B}}^{\mathbf{C}}(f)\right)^T.$$

Proof. The main point is to pull yourself together and unravel all the symbols systematically and correctly, then the proof is obvious and requires no ideas. Here is how it goes: the first easy observation is that the (i, j)-entry of $M_{\mathbf{B}}^{\mathbf{C}}(f)$ is nothing but

$$c_j^*(f(b_i))$$

whereas the $(j, i)-$entry of $M_{\mathbf{C}^*}^{\mathbf{B}^*}(f^*)$ is

$$(f^*(c_j^*))(b_i),$$

so all we need to do is show that these two are equal. But by definition of f^*

$$(f^*(c_j^*))(b_i) = c_j^*(f(b_i))$$

so we are done. Boom. That's all there is to it. □

As a next step it is natural to wonder how the kernels and images of $f\colon V \to W$ and $f^*\colon W^* \to V^*$ are related. We keep the assumption that V, W are of finite dimension $\dim V = n$ and $\dim W = m$, respectively, and have ordered bases as above.

Proposition 4.1.3. *For $V, W, f\colon V \to W$ as before, let $i\colon \mathrm{Im}(f) \to W$ be the inclusion. Then the dual linear map*

$$i^*\colon W^* \to \mathrm{Im}(f)^*$$

factors over the linear map $W^* \to \mathrm{Im}(f^*)$ induced by f^*, inducing an isomorphism

$$\mathrm{Im}(f^*) \simeq \mathrm{Im}(f)^*.$$

Proof. Again the proof is confusing, but easy once one has managed to unravel what the statement says: suppose l_W is a linear form on W that maps to zero in $\mathrm{Im}(f^*)$ under the linear map $W^* \to \mathrm{Im}(f^*)$ induced by f^*. This just means that $l_W \circ f$ is a linear form on V that is identically zero. But that means l_W restricted to the image of f is identically zero, so l_W is in the kernel of i^*. Therefore i^* factors uniquely over the linear map $W^* \to \mathrm{Im}(f^*)$ induced by f^*, thus giving us a linear map

$$\overline{i^*}\colon \mathrm{Im}(f^*) \to \mathrm{Im}(f)^*.$$

We just need to show that this map is injective and surjective. Concretely, $\overline{i^*}$ is given as follows: write l_V in $\mathrm{Im}(f^*)$ as $l_V = l_W \circ f$ with $l_W \in W^*$, then

$l_W \circ i \in \mathrm{Im}(f)^*$ is $\overline{i^*}(l_V)$. Suppose then that $l_W \circ i$ is zero. That just means that l_W restricted to $\mathrm{Im}(f)$ is zero, so l_V is zero. This shows injectivity. Surjectivity is follows because we can write any element in $\mathrm{Im}(f)^*$ in the form $l_W \circ i$ (extend a linear form on $\mathrm{Im}(f)$ to all of W), and then $l_V = l_W \circ f$ gives a preimage in $\mathrm{Im}(f^*)$ under $\overline{i^*}$ of the element you started with. $\square$

In particular, $\mathrm{Im}(f)$ and $\mathrm{Im}(f^*)$ have the same dimension. Thus:

Corollary 4.1.4. *The ranks of the two matrices*

$$M^{\mathbf{B}^*}_{\mathbf{C}^*}(f^*), \quad M^{\mathbf{C}}_{\mathbf{B}}(f)$$

are equal; in particular, using Lemma 4.1.2, the row rank of any $m \times n$ matrix over K is equal to its column rank.

You will have seen a proof of the last statement in your first linear algebra module, but here the proof falls into our laps basically effortlessly, and it is conceptually much more illuminating.

We can ask what happens if we take duals twice, i.e., pass from V to V^*, then to $(V^*)^*$ etc.

Proposition 4.1.5. *Define a natural linear map*

$$D\colon V \to (V^*)^*$$

as follows: to a vector $v \in V$ the map D associates the linear form on V^* that is given by evaluation of linear forms in v. Then D is an isomorphism if V is finite dimensional.

In the following we write more simply V^{**} for $(V^*)^*$.

Proof. Suppose v is in the kernel of D. That means that given any linear form l on V, $l(v)$ is zero. But this means that v must be zero! (Check this as an exercise if you are not convinced). By Lemma 4.1.1 we know that V and V^{**} have the same dimension, so D is an isomorphism. $\square$

OK, that's all pretty neat, but maybe you're not yet completely sold that the dual space is the perfect jack of all trades device of linear algebra, so let me give you another application.

Example. Consider an n-dimensional K-vector space V together with a nondegenerate symmetric bilinear form $\beta\colon V \times V \to K$. Then $q(v) = \beta(v, v)$ is a quadratic form, and we are interested in the maximum dimension of linear subspaces lying on the quadric $\{v \in V \mid q(v) = 0\}$. This is a very natural geometric problem occurring in various situations. We can get information using duality as follows.

The fact that β is nondegenerate means that the linear map

$$B\colon V \to V^*$$

which sends a vector $v \in V$ to the linear form $B(v) \in V^*$ defined by $B(v)(v') = \beta(v, v')$, $v' \in V$, is an isomorphism. Suppose $L \subset V$ is a linear subspace, and

$p\colon V^* \to L^*$ the surjection induced by the inclusion of L in V by dualising. The kernel of the composite map $p \circ B$ is clearly

$$L^{\perp} = \{v \in V \mid \beta(v, w) = 0 \,\forall\, w \in L\}.$$

Moreover, $p \circ B$ is surjective, therefore, by the rank-nullity theorem/dimension formula for linear maps, we get

$$\dim V = \dim(L) + \dim(L^{\perp}).$$

If q is identically zero on L, this means $L \subset L^{\perp}$. In particular,

$$2 \dim(L) \le \dim V$$

so that

$$\dim(L) \le \left[\frac{\dim V}{2}\right].$$

It is not hard to see that the bound is attained if $K = \mathbb{C}^n$; then we may assume $V = \mathbb{C}^n$ and $q = 0$ is just a sum of squares being zero in suitable coordinates:

$$x_1^2 + x_2^2 + \cdots + x_n^2 = 0.$$

A linear subspace defined by $x_1 = \sqrt{-1}x_2$, $x_3 = \sqrt{-1}x_4, \dots$ will do the job/be of maximum dimension $[n/2]$ in this case. Over other fields the situation can be different, and in fact, the equation $q = 0$ may have only the zero solution at all to begin with.

Here is another application of duals that might even convince the most practically-minded hardliners among you that duals are cool:

Theorem 4.1.6. *Let $I \subset \mathbb{R}$ be a closed interval and let $t_1, \dots, t_n \in I$ be distinct points. Then there exist n (real) numbers $m_1, \dots, m_n$ such that for all (real) polynomials p of degree $\le n-1$ we have*

$$\int_I p(t)\, dt = m_1 p(t_1) + \cdots + m_n p(t_n).$$

Proof. Polynomials p of degree $\le n-1$ form a real vector space $V_{<n}$ of dimension n (a basis would be $1, t, t^2, \dots, t^{n-1}$). Evaluation in t_i defines a linear form l_i on $V_{<n}$, hence an element in $V_{<n}^*$. We claim that these $l_1, \dots, l_n$ are linearly independent. Indeed, if

$$c_1 l_1 + \cdots + c_n l_n = 0$$

where the $c_i \in \mathbb{R}$, is a linear dependency relation in $V_{<n}^*$, then we can apply the linear form on the left hand side to the following polynomials in $V_{<n}$:

$$q_k(t) := \prod_{j \ne k, 1 \le j \le n} (t - t_j)$$

where $k = 1, \dots, n$. The polynomial q_k is nonzero at t_k and zero at all other t_i, so we get that $c_k = 0$. Since this holds for all k, the $l_1, \dots, l_n$ are linearly independent. By Lemma 4.1.1, $V_{<n}^*$ has dimension n, so $l_1, \dots, l_n$ must be a basis. Therefore, any linear form on $V_{<n}$ can be written as a linear combination of the l_i. In particular, this holds for the integral in the statement of the Theorem. Boom. It's as easy as that. □

We now turn to another useful construction, which we will use in a subsequent section, too, *quotient spaces*. We start by asking: given a K-vector space V and a subspace $U \subset V$, is there always a vector space W with a surjective linear map

$$\pi \colon V \to W$$

whose kernel is precisely U? Well, one way to solve this is to dualize the entire problem: if such a thing as we ask for exists, then

$$\pi^* \colon W^* \to V^*$$

will be an injective linear map with the property that the image of W^* in V^* is precisely the kernel of the surjective map $i^* \colon V^* \to U^*$ induced by the inclusion $i \colon U \to V$. In fact, we can then simply let $W^* = \mathrm{Ker}(i^*)$ and define W as

$$W := \mathrm{Ker}(i^*)^*$$

which will have the required property (using the natural isomorphism in Proposition 4.1.5). But that way to solve the problem is a bit cranky, and we mentioned it mainly to emphasise the connection with duals. A nicer way to solve the problem is this: the datum of the subspace U in V induces an equivalence relation on V by viewing v, v' as equivalent if their difference lies in U. In a formula:

$$v \sim_U v' :\iff v - v' \in U.$$

We define W to be the set of equivalence classes. We also denote this by V/U (read V modulo U). For $v \in V$ we denote by $[v] \in V/U$ its equivalence class. The set V/U can be endowed with a vector space structure by defining

$$[v] + [v'] := [v + v'], \quad \lambda[v] := [\lambda v]$$

for $v, v' \in V$, $\lambda \in K$. One uses the fact that U is a subspace to show that vector addition and scalar multiplication are well-defined on V/U, i.e., independent of the choice of representatives for the equivalence classes.

It is possible to characterise V/U by a universal property that is often useful: the quotient vector space V/U of a vector space V by a subspace U is a vector space together with a surjection $\pi \colon V \to V/U$ such that any linear map $f \colon V \to T$ from V to another vector space T with $U \subset \mathrm{Ker}(f)$ factors uniquely over V/U, i.e., there exists a unique linear map $\bar{f} \colon V/U \to T$ such that $f = \pi \circ \bar{f}$.

Proposition 4.1.7. *Let V be a finite-dimensional K-vector space, U a subspace. Then*

$$\dim U + \dim V/U = \dim V.$$

Proof. Rank nullity theorem applied to the canonical projection $\pi \colon V \to V/U$. □

Example. Here is a particularly striking application of quotients that is of immense importance in algebra. We do not give all details since this will be done in lectures on field and Galois theory, and we just want to convey the main idea here.

Suppose K is a field and $p(x)$ some irreducible polynomial in $K[x]$. Very often one wants to construct a field L containing K as a subfield (i.e., an overfield of K) in which $p(x)$ has a root. This is almost effortless using quotient spaces. We consider the set $I_p := \{p(x)q(x) \mid q(x) \in K[x]\} \subset K[x]$ of all polynomials in $K[x]$ that are divisible by $p(x)$. This is obviously a K-vector subspace and we can form the quotient space

$$L := K[x]/I_p.$$

L contains K as a K-subspace. One can define a multiplication in L that turns L even into an overfield of K. Indeed, simply define

$$[r] \cdot [s] := [r \cdot s], \quad r, s \in K[x]$$

where $r \cdot s$ is multiplication in the polynomial ring $K[x]$. It then needs a few checks that this is (a) well-defined and (b) makes L into a field (for the latter you need to use that $p(x)$ was assumed to be irreducible), but basically that is not too difficult. The point is that once you know that L thus defined is a field, the polynomial p obviously has a zero in L: the equivalence class $[x]$ of the variable x!

4.2 Tensors, the exterior and symmetric algebra

First, given two vector spaces U, V we define their *tensor product* $U \otimes V$ (sometimes also denoted by $U \otimes_K V$ if we want to recall the ground field) as follows. Let $F(U, V)$ be the vector space which has the set $U \times V$ as a basis, i.e., the free vector space (over K of course as always) generated by the pairs (u, v) where $u \in U$ and $v \in V$. Let R be the vector subspace of $F(U, V)$ spanned by all elements of the form

$$(u + u', v) - (u, v) - (u', v), \quad (u, v + v') - (u, v) - (u, v'),$$
$$(ru, v) - r(u, v), \quad (u, rv) - r(u, v)$$

where $u, u' \in U$, $v, v' \in V$, $r \in K$.

Definition 4.2.1. The quotient vector space

$$U \otimes V := F(U, V)/R$$

is called the *tensor product* of U and V. The image of $(u, v) \in F(U, V)$ under the projection $F(U, V) \to U \otimes V$ will be denoted by $u \otimes v$. We define the *canonical bilinear mapping*

$$\beta\colon U \times V \to U \otimes V$$

by $\beta(u, v) = u \otimes v$. Being very precise, one should refer to the pair $(U \otimes V, \beta)$ as the tensor product of U and V, but usually people just use the term for $U \otimes V$ with β tacitly understood.

Sometimes one does not need to know the construction of $U \otimes V$ when working with it, but only has to use the following property it enjoys in proofs.

Proposition 4.2.2. *Let W be a vector space with a bilinear mapping $\psi\colon U \times V \to W$. We say that (W, ψ) has the* universal factorisation property *for $U \times V$ if for every vector space S and every bilinear mapping $f\colon U \times V \to S$ there exists a unique linear mapping $g\colon W \to S$ such that $f = g \circ \psi$.*

Then the couple $(U \otimes V, \beta)$ has the universal factorisation property for $U \times V$. If a couple (W, ψ) has the universal factorisation property for $U \times V$, then $(U \otimes V, \beta)$ and (W, ψ) are canonically isomorphic in the sense that there exists a unique isomorphism $\sigma\colon U \otimes V \to W$ such that $\psi = \sigma \circ \beta$.

Proof. Suppose we are given any bilinear mapping $f\colon U \times V \to S$. Since $U \times V$ is a basis of $F(U,V)$ we can extend f to a unique linear mapping $f'\colon F(U,V) \to S$. Now f' vanishes on R since f is bilinear so induces a linear mapping $g\colon U \otimes V \to S$ on the quotient. Clearly, $f = g \circ \beta$ by construction. The uniqueness of such a map g follows from the fact that $\beta(U \times V)$ spans $U \otimes V$, so we have no other choice in defining g.

Now if (W, ψ) is a couple having the universal factorisation property for $U \times V$, then by the universal factorisation property of $(U \otimes V, \beta)$ (resp. of (W, ψ)), there exists a unique linear mapping $\sigma\colon U \otimes V \to W$ (resp. $\tau\colon W \to U \otimes V$) such that $\psi = \sigma \circ \beta$ (resp. $\beta = \tau \circ \psi$). Hence

$$\beta = \tau \circ \sigma \circ \beta, \quad \psi = \sigma \circ \tau \circ \psi.$$

Using the uniqueness of the g in the universal factorisation property, we conclude that $\tau \circ \sigma$ and $\sigma \circ \tau$ are the identity on $U \times V$ and W respectively. □

This universal property of the tensor product can be used to prove a great many formal properties of the tensor product in a way that is almost mechanical once one gets practice with it. All these proofs are boring. So we give one, and you can easily work out the rest for some practice with this.

Proposition 4.2.3. *The tensor product has the following properties.*

(a) *There is a unique isomorphism of $U \otimes V$ onto $V \otimes U$ sending $u \otimes v$ to $v \otimes u$ for all $u \in U, v \in V$*

(b) *There is a unique isomorphism of $K \otimes U$ with U sending $r \otimes u$ to ru for all $r \in K$ and $u \in U$; similarly for $U \otimes K$ and U.*

(c) *There is a unique isomorphism of $(U \otimes V) \otimes W$ onto $U \otimes (V \otimes W)$ sending $(u \otimes v) \otimes w$ to $u \otimes (v \otimes w)$ for all $u \in U, v \in V, w \in W$.*

(d) *Given linear mappings*

$$f_i\colon U_i \to V_i,\ i = 1, 2,$$

there exists a unique linear mapping $f\colon U_1 \otimes U_2 \to V_1 \otimes V_2$ such that

$$f(u_1 \otimes u_2) = f_1(u_1) \otimes f_2(u_2)$$

for all $u_1 \in U_1, u_2 \in U_2$.

(e) *There is a unique isomorphism from $(U_1 \oplus U_1) \otimes V$ onto $(U_1 \otimes V) \oplus (U_1 \otimes V)$ sending $(u_1, u_2) \otimes v$ to $(u_1 \otimes v, u_2 \otimes v)$ for all $u_1 \in U_1$, $u_2 \in U_2$, $v \in V$.*

(f) *If $u_1, \ldots, u_m$ is a basis for U and $v_1, \ldots, v_n$ is a basis for V, then $u_i \otimes v_j$, $i = 1, \ldots, m$, $j = 1, \ldots, n$, is a basis for $U \otimes V$. In particular,* $\dim U \otimes V = \dim U \dim V$.

(g) *Let U^* be the dual vector space to U. Then there is a unique isomorphism g from $U \otimes V$ onto $Hom(U^*, V)$ such that*

$$(g(u \otimes v))(u^*) = \langle u, u^* \rangle v \quad \text{for all } u \in U, v \in V, u^* \in U^*.$$

(h) *There is a unique isomorphism h of $U^* \otimes V^*$ onto $(U \otimes V)^*$ such that*

$$(h(u^* \otimes v^*))(u \otimes v) = \langle u, u^* \rangle \langle v, v^* \rangle$$

for all $u \in U$, $u^* \in U^*$, $v \in V$, $v^* \in V^*$.

Proof. We prove a) and g) just to illustrate the method, and leave the rest as easy exercises.

For a) let $f\colon U \times V \to V \otimes U$ be the bilinear mapping with $f(u,v) = v \otimes u$. By the universal property of the tensor product, there is a unique linear mapping

$$g\colon U \otimes V \to V \otimes U$$

such that $g(u \otimes v) = v \otimes u$. Similarly, there is a unique linear mapping $g'\colon V \otimes U \to U \otimes V$ with $g'(v \otimes u) = u \otimes v$. Clearly, $g' \circ g$ and $g \circ g'$ are the identity transformations.

We now prove g) (using f)). Consider the bilinear mapping $f\colon U \times V \to \operatorname{Hom}(U^*, V)$ given by

$$(f(u,v))(u^*) = \langle u, u^* \rangle v$$

and apply the universal property of the tensor product in Proposition 4.2.2. Thus there exists a unique linear mapping $g\colon U \otimes V \to \operatorname{Hom}(U^*, V)$ such that $(g(u \otimes v))(u^*) = \langle u, u^* \rangle v$. To prove that g is an isomorphism, let $u_1, \ldots, u_m$ be a basis for U, $u_1^*, \ldots, u_m^* \in U^*$ the dual basis, and $v_1, \ldots, v_n$ a basis for V. We show that

$$\{g(u_i \otimes v_j)\colon i = 1, \ldots, m,\ j = 1, \ldots, n\}$$

is a linearly independent set of vectors. Indeed

$$\sum_{i,j} a_{ij} g(u_i \otimes v_j) = 0, \quad a_{ij} \in K$$

gives

$$0 = \sum_{i,j} a_{ij} g(u_i \otimes v_j)(u_k^*) = \sum_j a_{kj} v_j$$

hence all the a_{ij} vanish. Since $\dim U \otimes V = \dim \operatorname{Hom}(U^*, V)$ by f), g is an isomorphism. □

We now consider a vector space V and put $V^{\otimes r} := V \otimes \cdots \otimes V$ (r-times), and set

$$T^\bullet(V) = \bigoplus_{r \geq 0} V^{\otimes r}.$$

If $e_1, \ldots, e_n$ is a basis for V, then

$$\{e_{i_1} \otimes \cdots \otimes e_{i_r}\colon 1 \leq i_1, \ldots, i_r \leq n\}$$

is a basis for $V^{\otimes r}$, applying f) of Proposition 4.2.3 inductively. $T^\bullet(V)$ has more structure than just the structure of a K-vector space (of infinite dimension in general!):

Definition 4.2.4. 1. Let $\mathcal{A}$ be a (not necessarily commutative) ring with unit, and suppose that there is a field K that is a subring of $\mathcal{A}$. Then $\mathcal{A}$ is called a K-algebra; in particular, $\mathcal{A}$ is also a K-vector space.

2. We call $\mathcal{A}$ a graded algebra if there is a direct sum decomposition as a K-vector space

$$\mathcal{A} = \bigoplus_{n=0}^{\infty} \mathcal{A}_n$$

such that $\mathcal{A}_i \cdot \mathcal{A}_j \subset \mathcal{A}_{i+j}$. Elements in $\mathcal{A}_i$ are said to have degree i. So the last condition means that the product of an element of degree i and one of degree j has degree $i + j$.

3. Suppose $I \subset \mathcal{A}$ is a vector subspace that has the additional properties:

$$\mathcal{A} \cdot I \subset I, \quad I \cdot \mathcal{A} \subset I.$$

Then we call I a two-sided ideal in $\mathcal{A}$.

It is a routine check that if $I \subset \mathcal{A}$ is a two-sided ideal, the quotient K-vector space $\mathcal{A}/I$ becomes a K-algebra by defining $[a] \cdot [a'] := [a \cdot a']$.

With this terminology, we can say that $T^\bullet(V)$ is a graded K-algebra, associative, but not commutative, if we define the product

$$(v_1 \otimes \cdots \otimes v_r) \cdot (w_1 \otimes \cdots \otimes w_s) = v_1 \otimes \cdots \otimes v_r \otimes w_1 \otimes \cdots \otimes w_s$$

and extend by K-linearity to all of $T^\bullet(V)$. We call $T^\bullet(V)$ the *tensor algebra* of V.

Definition 4.2.5. Let I be the two-sided ideal of $T^\bullet(V)$ generated by all elements of the form $v \otimes v$ for $v \in V$. The quotient

$$\Lambda^\bullet(V) := T^\bullet(V)/I$$

is called the *exterior algebra* of V.

Similarly, if J denotes the two-sided ideal of $T^\bullet(V)$ generated by all elements of the form $v \otimes w - w \otimes v$ for $v, w \in V$, then

$$\mathrm{Sym}^\bullet(V) = T^\bullet(V)/J$$

is called the *symmetric algebra* of V.

We denote the image of $v_1 \otimes \cdots \otimes v_r$ in $\Lambda^\bullet(V)$ by $v_1 \wedge \cdots \wedge v_r$, and the image in $\mathrm{Sym}^\bullet(V)$ by $v_1 \cdot \ldots \cdot v_r$, or simply $v_1 \ldots v_r$. We will also denote the algebra product in $\Lambda^\bullet(V)$ simply by $\wedge$ and call it the wedge product. Similarly, we denote the algebra product in $\mathrm{Sym}^\bullet(V)$ by a dot or simply by concatenation.

Both the symmetric and exterior algebras inherit a natural grading from the tensor algebra. The r-th graded component $\Lambda^r(V)$ of $\Lambda^\bullet(V)$ (resp. $\mathrm{Sym}^r(V)$ of $\mathrm{Sym}^\bullet(V)$) is called the r-th exterior power of V (resp. r-th symmetric power of V).

In fact, other types of important algebras can be defined in a similar way as quotients of the tensor algebra $T^\bullet(V)$, for example, Clifford algebras. But in fact, the exterior algebra

$$\Lambda^\bullet(V) = \bigoplus_{r=0}^{\infty} \Lambda^r(V)$$

will be most important for us below. We only mentioned the symmetric algebra because it would have weighed too heavily on our conscience if we hadn't- it is so important in other contexts. In fact, it is a good exercise to convince yourself that $\mathrm{Sym}^\bullet(V)$ is simply isomorphic to a polynomial algebra $K[X_1, \ldots, X_n]$ with one variable X_i corresponding to each basis vector e_i of V.

We now turn to the properties of the exterior algebra we will need later. First of all it is clear that for any $v, w \in V$ we have

$$v \wedge v = 0, \quad v \wedge w = -w \wedge v,$$

the first because $v \otimes v$ maps to zero under the quotient map $T^\bullet(V) \to \Lambda^\bullet(V)$, and the second is implied by $(v+w) \wedge (v+w) = 0$. We say the wedge-product is *alternating* or *anti-symmetric*. More generally, this implies that if $\omega \in \Lambda^r(V)$ and $\varphi \in \Lambda^s(V)$, then

$$\omega \wedge \varphi = (-1)^{rs} \varphi \wedge \omega.$$

Proposition 4.2.6. *The exterior powers and exterior algebra have the following properties.*

(a) If $F\colon V \times \cdots \times V \to W$ (r copies of V) is a multilinear alternating mapping of vector spaces (which means $F(v_1, \dots, v_r)$ is linear in each argument separately and zero if two of the v_i are equal), then there is a unique linear map

$$\bar{F}\colon \Lambda^r(V) \to W$$

with $\bar{F}(v_1 \wedge \cdots \wedge v_r) = F(v_1, \dots, v_r)$.

(b) If $\varphi\colon V \to W$ is a linear mapping, there is a unique linear mapping

$$\Lambda^r(\varphi)\colon \Lambda^r(V) \to \Lambda^r(W)$$

with the property

$$\Lambda^r(\varphi)(v_1 \wedge \cdots \wedge v_r) = \varphi(v_1) \wedge \cdots \wedge \varphi(v_r).$$

(c) If $e_1, \dots, e_n$ is a basis of V, then

$$\{e_{i_1} \wedge \cdots \wedge e_{i_r} \colon 1 \le i_1 < \cdots < i_r \le n\}$$

is a basis of $\Lambda^r(V)$. Consequently,

$$\dim \Lambda^r(V) = \binom{n}{r}$$

and $\Lambda^i(V) = 0$ for $i > n$. Moreover, note that $\dim \Lambda^n(V) = 1$.

(d) If $f\colon V \to V$ is an endomorphism, $\dim V = n$, then the induced map

$$\Lambda^n(f)\colon \Lambda^n(V) \to \Lambda^n(V)$$

is multiplication by $\det(f)$.

(e) For an n-dimensional vector space V we have a natural non-degenerate bilinear pairing

$$\Lambda^r(V^*) \times \Lambda^r(V) \to K$$

mapping

$$(v_1^* \wedge \cdots \wedge v_r^*, w_1 \wedge \cdots \wedge w_r) \mapsto \det(v_i^*(w_j))_{1 \le i,j \le r}$$

which induces an isomorphism

$$\Lambda^r(V^*) \simeq (\Lambda^r(V))^*.$$

Proof. For a) notice that repeated application of the universal property of the tensor product furnishes us with a linear map

$$\tilde{F}\colon V^{\otimes r} \to W$$

with $\tilde{F}(v_1 \otimes \cdots \otimes v_r) = F(v_1, \ldots, v_r)$; this factors over

$$\Lambda^r(V) = V^{\otimes r} / \left(V^{\otimes r} \cap I\right)$$

since the ideal I is generated by elements $v \otimes v$ that get mapped to zero since F is alternating.

To prove b) notice that inductive application of Proposition 4.2.3, d), gives an induced mapping

$$\otimes^r \varphi \colon V^{\otimes r} \to W^{\otimes r}$$

and this maps $V^{\otimes r} \cap I$ into the corresponding piece of the ideal we divide out by to get $\Lambda^r W$, so descends to give $\Lambda^r(\varphi)$ as desired.

For c) we first show that $\Lambda^n(V) \simeq K$ via the map induced by the determinant. Indeed, since the elements $v_1 \wedge \cdots \wedge v_r$ generate $\Lambda^r(V)$, it is clear that, if $e_1, \ldots, e_n$ is a basis of V, then

$$\{e_{i_1} \wedge \cdots \wedge e_{i_r} \colon 1 \le i_1 < \cdots < i_r \le n\}$$

is at least a generating set for $\Lambda^r(V)$. In particular, $\Lambda^n(V)$ is at most one-dimensional, and exactly one-dimensional, generated by $e_1 \wedge \cdots \wedge e_n$, if we can show it is nonzero. But by a), the determinant gives a map $\overline{\det} \colon \Lambda^n(V) \to K$ sending $e_1 \wedge \cdots \wedge e_n$ to 1.

Now suppose there was a linear dependence relation between the $e_{i_1} \wedge \cdots \wedge e_{i_r}$:

$$\sum_I a_I e_I = 0$$

where we use multi-index notation $I = (i_1, \ldots, i_r)$, $1 \le i_1 < \cdots < i_r \le n$, $a_I \in K$, $e_I = e_{i_1} \wedge \cdots \wedge e_{i_r}$. For a certain multi-index $J = (j_1, \ldots, j_r)$, let $\bar{J}$ be the complimentary indices to J in $\{1, \ldots, n\}$, increasingly ordered. Then

$$(\sum_I a_I e_I) \wedge e_{\bar{J}} = \pm a_J e_1 \wedge \cdots \wedge e_n = 0.$$

Hence all coefficients a_J are zero, proving c).

The endomorphism $f \colon V \to V$ gives a commutative diagram

$$\begin{array}{ccc} \Lambda^n(V) & \xrightarrow{\Lambda^n(f)} & \Lambda^n(V) \\ \downarrow{\scriptstyle \overline{\det}} & & \downarrow{\scriptstyle \overline{\det}} \\ K & \xrightarrow{\text{mult}(c)} & K \end{array}$$

where the lower horizontal arrow is multiplication by some constant c. We want to show that $c = \det(f)$ and for this it suffices to consider what happens to $\overline{\det}(e_1 \wedge \cdots \wedge e_n) = 1$: this gets mapped to the determinant of the matrix with columns $(f(e_1), \ldots, f(e_n))$, which is $\det(f)$. This proves d).

For e) first note that

$$\det((v_i^*(w_j))_{1 \le i,j \le r}$$

is alternating in both the $v_1^*, \ldots, v_n^*$ and the $w_1, \ldots, w_n$, and multilinear in these sets of variables; hence by an application of a), we get a well-defined map

$$\beta \colon \Lambda^r(V^*) \times \Lambda^r(V) \to K$$

of the type in e). All that remains to prove is that this pairing is nondegenerate, i.e. that for any nonzero $\omega \in \Lambda^r(V)$ there is a $\psi \in \Lambda^r(V^*)$ with $\beta(\psi, \omega) \neq 0$, and vice versa, for any nonzero $\psi' \in \Lambda^r(V^*)$ there is an $\omega' \in \Lambda^r(V)$ with $\beta(\psi', \omega') \neq 0$. We prove the first assertion since the second is then proven completely analogously. If $e_1, \ldots, e_n$ is a basis of V, write in multi-index notation

$$\omega = \sum_I a_I e_I.$$

Since $\omega \neq 0$, there is an $a_J \neq 0$. Then let $\psi = e_J^* = e_{j_1}^* \wedge \cdots \wedge e_{j_r}^*$ where $e_1^*, \ldots, e_n^*$ is the dual basis to $e_1, \ldots, e_n$. We have $\beta(\psi, \omega) = a_J \neq 0$ then. □