Matrix algebra for beginners, Part II
linear transformations, eigenvectors and eigenvalues

Jeremy Gunawardena

Department of Systems Biology
Harvard Medical School
200 Longwood Avenue, Cambridge, MA 02115, USA
jeremy@hms.harvard.edu

February 10, 2006

Contents

1 Introduction

In Part I we introduced matrices as rectangular arrays of numbers and we motivated this in terms of solving linear equations. Although this is an important application, matrices also arise in geometry, particularly in studying certain kinds of geometric transformations. Although it takes a little work to understand this connection, it is worth doing so because many aspects of matrix algebra that seem mysterious at first sight become clearer and more intuitive when seen from a geometrical perspective.

2 Vector spaces and linear transformations

A *vector* is a line segment with a specified direction, usually indicated by an arrow at one end. Vectors arise in physics as mathematical representations of quantities like force and velocity which have both magnitude and direction. If we fix a point as the *origin* , then the collection of vectors which originate from this point form a *vector space* . To have something concrete in mind, we can think of a 2 dimensional space, or plane; everything we say holds equally well in any number of dimensions. (Indeed, vector spaces are the mathematical way in which the idea of "dimension" can be made precise.) Notice that, once we have fixed an origin, each point in the space corresponds, in a one-to-one fashion, to a vector from the origin to that point. Vector spaces provide another way of thinking about geometry.

What makes vectors interesting (unlike points in the plane) is that they can be added together. You should be familiar with the *parallelogram law* for forces from elementary physics: if two forces are exerted on an object, then the resultant force is given by the vector to the far corner of the parallelogram bounded by the vectors of the two forces, as in Figure 1. A similar rule holds for velocities and this becomes vital if you ever have to navigate a plane or a boat. Your actual velocity is the parallelogram sum of the velocity determined by your speed and heading and the velocity of the wind or the sea. If you have ever watched an aircraft in flight from the window of another one, then its curious crab-like motion in a direction different to where it is pointing is explained by the parallelogram law applied to the velocities of the two aircraft.

Addition of vectors obeys all the rules that we are familiar with for addition of numbers. For instance, each vector has an inverse vector, which is given by the line segment of the same length but pointing in the opposite direction. If you add a vector to its inverse using the parallelogram law then you get a vector of zero length, the zero vector, corresponding to the zero for addition. You can also multiply a vector by a number; the effect is merely to multiply the magnitude of the vector by that number, while keeping its direction fixed. This kind of multiplication is called *scalar multiplication* and it is easy to check that it obeys all the obvious rules. Let us use $\mathbf{x}$, with bold face, to denote vectors, $\mathbf{x}+\mathbf{y}$ to denote addition of vectors and $\lambda\mathbf{x}$ to denote scalar multiplication by the number λ. Then, for instance,

$$(\lambda\mu)\mathbf{x} = \lambda(\mu\mathbf{x}) \qquad \lambda(\mathbf{x}+\mathbf{y}) = \lambda\mathbf{x} + \lambda\mathbf{y}\ .$$

Figure 7 gives a complete set of rules for addition and scalar multiplication.

If $\mathbf{x}$ and $\mathbf{y}$ point in the same direction, then we can always find a unique number λ, such that $\mathbf{y} = \lambda\mathbf{x}$. In effect, any non-zero vector, like $\mathbf{x}$, defines a unit length along the line corresponding to its direction. Any other vector in that direction must have a length which is some multiple of that unit length. Vectors provide a way of marking out space.

This is all very well but what has it got to do with matrices? Matrices represent certain kinds of transformations on vector spaces. Specifically, they correspond to *linear transformations*. (A transformation is just another word for a function. For historical reasons, the word "transformation" is often preferred when dealing with functions on geometric objects and we will follow that usage here.) For instance, let us consider a rotation about the origin through some angle. Each vector from

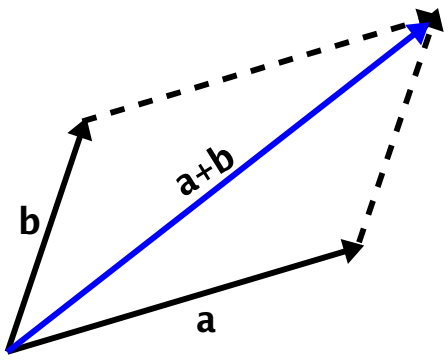

Figure 1: Parallelogram rule for addition of vectors. The sum of vectors a and b, originating from a common point, is the vector from that common point to the far corner of the parallelogram bounded by a and b.

the origin is transformed into another vector from the origin as a result of the rotation. You should be able to convince yourself quite easily that if you rotate the sume of two vectors, you get the same result as if you first rotated the vectors and then added them. Basically, the whole parallelogram rotates along with the vectors. Similarly, if you scalar multiply a rotated vector by a fixed number you get the same result as if you first scalar multiplied and then rotated. What happens if you translate a vector by moving it in a certain direction and by a certain amount? This would also translate the addition parallelogram. However, it is not a linear transformation because it does not preserve the origin of vectors. (It is an example of what is called an *affine transformation* , which is almost, but not quite, linear.) Transformations which preserve addition and scalar multiplication are called *linear transformations* . More formally, if F denotes a transformation from vectors to vectors, so that $F(\mathbf{x})$ denotes the vector to which $\mathbf{x}$ is transformed, then a linear transformation is one which satisfies

$$F(\mathbf{x}+\mathbf{y}) = F(\mathbf{x}) + F(\mathbf{y}) \qquad F(\lambda\mathbf{x}) = \lambda F(\mathbf{x})\,. \tag{1}$$

We have just seen that rotations are linear transformations. So also are reflections . One easy way to see this is to note that a reflection in 2 dimensions corresponds to a rotation in 3 dimensions through 180 degrees about the axis given by the line of the mirror. Another type of linear transformation is a *dilation* , which corresponds to scalar multiplication: $D_\alpha(\mathbf{x}) = \alpha\mathbf{x}$. It is easy to check using the formulae in Figure 7 and (1) that D_α is a linear transformation. Translating a vector, $\mathbf{x}$, in a certain direction and by a certain amount, is the same as forming the vector sum $\mathbf{x}+\mathbf{v}$, where $\mathbf{v}$ is a vector from the origin in the required direction having the required length. You should be able to check that the transformation given by $\mathbf{x} \to \mathbf{x}+\mathbf{v}$ does not satisfy either of the formulae in (1).

3 Bases and matrices

We need one more idea to see where matrices enter the picture. It goes back to the French philosopher and mathematician René Descartes in the 17th century. Descartes showed how geometry can be turned into algebra by using a *system of coordinates* . Descartes did not know about vector spaces and linear transformations but we can readily translate his ideas into this context. We choose two vectors (for a 2 dimensional vector space) as *basis vectors* , or, in Descartes' language, as the axes of the coordinate system. Descartes took his axes to be at right angles to each other, which is what we still do whenever we draw a graph. Axes which are at right angles define a *Cartesian* coordinate system . However, there is no necessity to do this and our basis vectors can be at any angle to each other, provided only that the angle is not zero. It is not much use to have coincident axes (see the comments below on bases in higher dimensions). Let us call the basis vectors $\mathbf{b}_1$ and $\mathbf{b}_2$. It is important to remember that their choice can be arbitrary, provided only that $\mathbf{b}_1 \neq \lambda\mathbf{b}_2$. Although

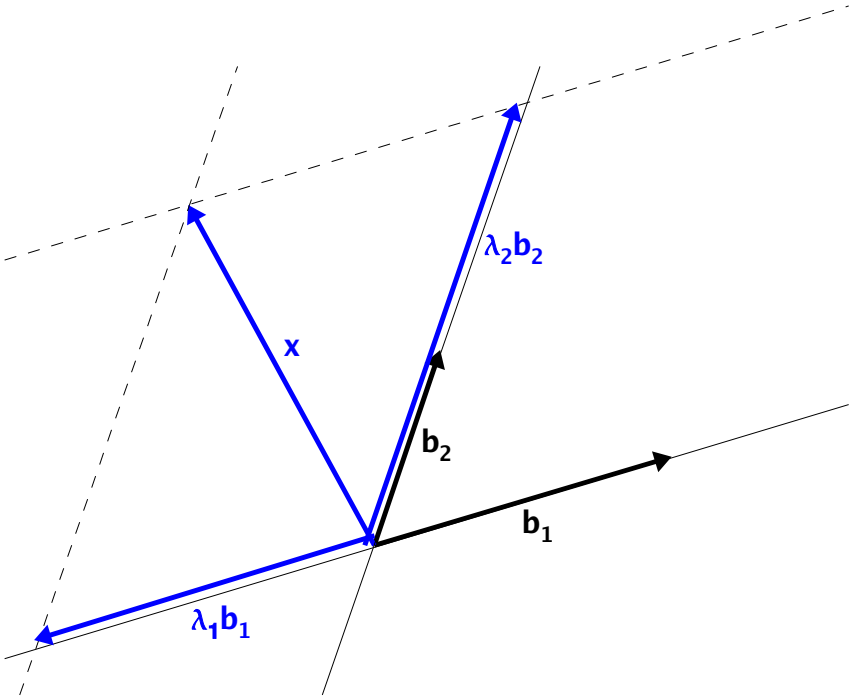

Figure 2: Resolving the vector $\mathbf{x}$ into its components with respect to the basis $\mathbf{b}_1$ and $\mathbf{b}_2$. The component vectors are offset slightly from the origin to make them easier to distinguish from the basis vectors.

the order of the basis vectors is not important for them to constitute a basis, it is important for some of the things we want to do below, so it is best to think of $\mathbf{b}_1$, $\mathbf{b}_2$ as a different basis to $\mathbf{b}_2, \mathbf{b}_1$.

Now, if we have any vector $\mathbf{x}$ then we can use the parallelogram rule in reverse to project it into the two basis vectors, as shown in Figure 2. This gives us two *component vectors* lying in the direction of the basis vectors. Each component is some scalar multiple of the corresponding basis vector. In other words, there are scalars λ_1 and λ_2, which are often called the *scalar components* of $\mathbf{x}$, such that

$$\mathbf{x} = \lambda_1 \mathbf{b}_1 + \lambda_2 \mathbf{b}_2 \,. \tag{2}$$

All we have done, in effect, is to work out the coordinates of the point at the end of $\mathbf{x}$ in the coordinate system defined by $\mathbf{b}_1$ and $\mathbf{b}_2$. Descartes would still recognise this, despite all the fancy new language.

We have to be a little careful to define a basis correctly in higher dimensions. The vectors $\mathbf{b}_1$, $\mathbf{b}_2$, $\cdots$, $\mathbf{b}_n$ form a basis set in a vector space if, and only if, each vector in the space can be represented **uniquely** as a sum of scalar multiples of the basis vectors, as in (2). There are two requirements here. The first is that you need enough basis vectors to represent every vector in the space. So one vector would be insufficient in the plane. The second is the you need to choose them in such a way that each vector has a **unique** representation. If you use 3 vectors for a basis in the plane, then some vectors can be represented in the basis in more than one way (why?). In particular, the size of a basis set equals the dimension of the vector space. Indeed, that is how "dimension" is mathematically defined: it is the size of a maximal basis set. However, it is not the case that just because you have n vectors in n dimensions, they form a basis. In two dimensions, it is no good choosing two vectors in the same direction. In 3 dimensions, it is no good choosing three vectors in such a way that one of them (and hence each of them) lies in the plane defined by the other two; and so on.

Once we have chosen a basis, $\mathbf{b}_1$, $\mathbf{b}_2$, $\cdots$, $\mathbf{b}_n$, each vector in the vector space can be represented by

n numbers in an unique way, just as we did in (2) for vectors in the plane:

$$\mathbf{x} = \lambda_1 \mathbf{b}_1 + \lambda_2 \mathbf{b}_2 + \cdots + \lambda_n \mathbf{b}_n \,. \tag{3}$$

We can represent these numbers in a $n \times 1$ matrix

$$\begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{pmatrix} .$$

(You might wonder why we do not use a $1 \times n$ matrix, which is easier to write down. We could have done but it would have forced other choices subsequently, which would feel more awkward.) The conversion of vectors into $n \times 1$ matrices, with respect to a chosen basis, is fundamental to what we are doing here. We are going to use square brackets, $[\mathbf{x}]$, to denote conversion from vectors to $n \times 1$ matrices, so that

$$[\mathbf{x}] = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{pmatrix} , \tag{4}$$

It is very important to remember that this conversion depends on the choice of basis. If it is necessary to keep track of which basis we are using then we will decorate the square brackets with a subscript, $[\mathbf{x}]_B$, but, on the general basis of minimising notation, we will drop the subscript when the basis is clear from the context.

What are the $n \times 1$ matrices corresponding to the basis vectors $\mathbf{b}_1, \mathbf{b}_2, \cdots, \mathbf{b}_n$ themselves? It should be obvious that they are just

$$\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} , \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix} , \cdots , \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} . \tag{5}$$

In other words, they are the the columns of the $n \times n$ identity matrix. You should appreciate the significance of this below.

It is quite easy to check that the matrix of a sum of two vectors is the sum of the matrices of the two vectors and that scalar multiplication of vectors corresponds in a similar way to scalar multiplication of matrices:

$$[\mathbf{x} + \mathbf{y}] = [\mathbf{x}] + [\mathbf{y}] \qquad [\lambda \mathbf{x}] = \lambda [\mathbf{x}] \,. \tag{6}$$

In other words, forming the $n \times 1$ matrix of a vector with respect to a basis is itself a linear transformation from vectors to $n \times 1$ matrices. The set of $n \times 1$ matrices forms a vector space, obeying all the rules in Figure 7.

What does a linear transformation look like, if we have chosen a basis? If we apply a linear transformation, F, to the expression for $\mathbf{x}$ in (3) and use the rules in (1), we find that

$$F(\mathbf{x}) = \lambda_1 F(\mathbf{b}_1) + \lambda_2 F(\mathbf{b}_2) + \cdots + \lambda_n F(\mathbf{b}_n) \,.$$

In other words, if we know what the basis vectors transform into (ie: $F(\mathbf{b}_i)$ then we can work out what x transforms into, using just the components of $\mathbf{x}$ (ie: λ_i) in the same basis. If we unravel all this, we discover that we are using matrices.

So what is the matrix? Before describing it, let us agree to use the same square bracket notation to denote the matrix corresponding to a transformation, $[F]$. If we are in n-dimensional vector space

then this matrix will turn out to be an $n \times n$ matrix. **It is very important to remember that the matrix depends on the choice of basis.** As above, we will use a subscript, $[F]_B$, to denote the basis if we need to. Well, $[F]$ is just the $n \times n$ matrix whose columns are given by the $n \times 1$ matrices of the transformations of the basis vectors.

More formally, let us suppose that we have a basis, $B = \mathbf{b}_1, \mathbf{b}_2, \cdots, \mathbf{b}_n$, in an n-dimensional vector space and suppose that F is any linear transformation acting on this space. For each basis element $\mathbf{b}_i$, we can form the $n \times 1$ matrix $[F(\mathbf{b}_i)]$, just as in (4). Then, $[F]$ is the matrix whose i-th column is $[F(\mathbf{b}_i)]$:

$$\boxed{[F] = ([F(\mathbf{b}_1)] \mid [F(\mathbf{b}_2)] \mid \cdots \mid [F(\mathbf{b}_n)])} \tag{7}$$

4 Examples—rotations and reflections

What is the matrix of the dilation, D_α? It should be easy to see that this is αI_n, where I_n is the $n \times n$ identity matrix. Let us do something a bit more challenging. What is the matrix of a rotation through θ degrees counterclockwise, with respect to a Cartesian basis? For simplicity, we will assume that the two basis vectors, $\mathbf{b}_1$ and $\mathbf{b}_2$ have the same length, as shown in Figure 3(A). We will use the notation, $|\mathbf{x}|$ to indicate the length of a vector, so that $|\mathbf{b}_1| = |\mathbf{b}_2|$. It follows from elementary trigonometry that the components of the rotated $\mathbf{b}_1$ have lengths $|\mathbf{b}_1| \cos\theta$ and $|\mathbf{b}_1| \sin\theta$. If we express each component as a scalar multiple of $\mathbf{b}_1$ and $\mathbf{b}_2$, respectively, then the scalar components of the rotated $\mathbf{b}_1$ are $\cos\theta$ and $\sin\theta$, respectively. (Notice that we have used the fact that $|\mathbf{b}_1| = |\mathbf{b}_2|$.) Similarly, the scalar components of $\mathbf{b}_2$ are $-\sin\theta$ and $\cos\theta$. It follows that the matrix of the rotation, with respect to this basis, is

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} . \tag{8}$$

Just to check this, a rotation through $\theta = \pi$ should be the same as a dilation by -1. Since $\cos\pi = -1$ and $\sin\pi = 0$, that is exactly what we get (the identity matrix multiplied by -1).

You might have noticed that (8) is a matrix of the same form as those which represented complex numbers, which we studied in §4 of Part I. What is the reason for that?

What about a reflection, with respect to the same basis? Figure 3(B) shows a reflection through a mirror line at an angle θ to the $\mathbf{b}_1$ axis. You should be able to work through the elementary trigonometry to show that the matrix of the reflection, with respect to this basis, is

$$\begin{pmatrix} \cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta \end{pmatrix} . \tag{9}$$

To check this, let us take $\theta = \pi/2$. In this case vectors on the $\mathbf{b}_1$ axis are reflected through the origin, while vectors on the $\mathbf{b}_2$ axis are left fixed. Is that what the matrix does?

5 Isomorphism between linear transformations and matrices

The basic property that we have been working towards through all these preliminaries is something that you should now be able to check from the definitions. If F is any linear transformation, then

$$[F(\mathbf{x})] = [F].[\mathbf{x}] . \tag{10}$$

In other words, the $n \times 1$ matrix of $F(\mathbf{x})$ is given by the product of the $n \times n$ matrix of F and $n \times 1$ matrix of $\mathbf{x}$, all the matrices being taken with respect to the same basis B.

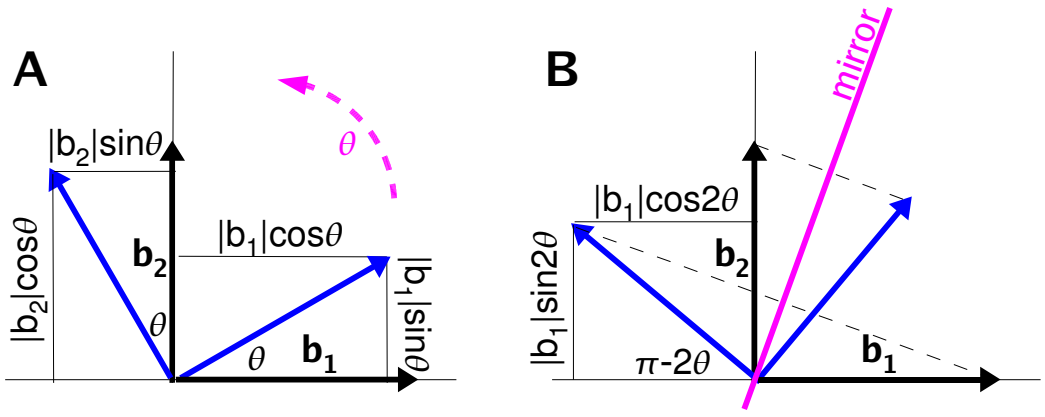

Figure 3: Two linear transformations, considered with respect to the Cartesian basis, $\mathbf{b}_1, \mathbf{b}_2$, with basis vectors of equal length. **(A)** Counterclockwise rotation through an angle θ. **(B)** Reflection in the mirror at an angle θ to the $\mathbf{b}_1$ axis. The lengths of the line segments are given as absolute lengths, with $|\mathbf{b}_1|$ and $|\mathbf{b}_2|$ denoting the lengths of the basis vectors.

Formula (10) further tells us that $[-]$ converts the key algebraic operation on transformations (namely, composition of transformations) into the product of the corresponding matrices. To see this, let F and G be two linear transformations. The composition of F and G is defined by $(FG)(\mathbf{x}) = F(G(\mathbf{x}))$: first apply G, then apply F. Let us see what this looks like in matrix terms. Using (10), first applied to F and then applied to G, we see that

$$[(FG)(\mathbf{x})] = [F(G(\mathbf{x}))] = [F].[G(\mathbf{x})] = [F].([G].[\mathbf{x}]) = ([F].[G]).[\mathbf{x}] \, .$$

(Notice that we had to use the associativity of the matrix product in the last step but only for the case in which the third matrix is $n \times 1$.) On the other hand, using (10) applied to FG, we know that

$$[(FG)(\mathbf{x})] = [FG].[\mathbf{x}] \, .$$

Hence, since these equations hold for any vector $\mathbf{x}$,

$$[FG] = [F][G] \, . \tag{11}$$

The matrix of the composition of two transformations is the product of the matrices of the individual transformations.

This may all seem rather intricate but the basic point is quite simple. On the one hand we have vectors and linear transformations of vectors. If we choose a basis in the vector space, then that defines a function $[-]$, which converts vectors to $n \times 1$ matrices and linear transformations into $n \times n$ matrices. These conversions are one-to-one correspondences (you should convince yourself of this) which preserve the corresponding algebraic operations. They are hence *isomorphisms* , in the sense explained in Part I. The corresponding structures, along with their algebraic operations, are effectively equivalent and we can as well work with $n \times 1$ matrices as with vectors and with $n \times n$ matrices as with linear transformations.

6 Geometric interpretation of products and determinants

Now that we know that linear transformations and vector spaces are isomorphic to $n \times n$ matrices and $n \times 1$ matrices, respectively, a number of properties fall into place. First, we know that composition

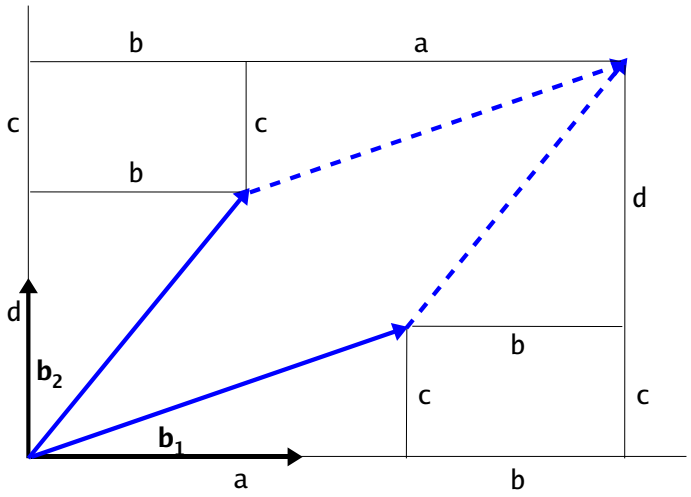

Figure 4: Calculating the area of the parallelogram bounded by the transforms of the Cartesian basis vectors $\mathbf{b}_1$ and $\mathbf{b}_2$ after applying the matrix in (12). The lengths of the line segments are given relative to the lengths of the corresponding basis vectors.

of transformations is associative: $F(GH) = (FG)H$. This is a basic property of functions acting on sets, which has nothing to do with the specific functions—linear transformations—and specific sets—vector spaces—being used here. It follows from (11) that the product of matrices must also be associative. This is not quite a proof of associativity, because we had to use a special case of associativity to prove (11), but it provides the intuition for why matrix product has to be associative. Second, it becomes obvious that we cannot expect the matrix product to be commutative, since composition of functions is certainly not commutative. For instance, if we consider the two functions $x \to x + 1$ and $x \to x^2$ on numbers, then $(x + 1)^2 \neq x^2 + 1$.

The determinant has a particularly nice interpretation in terms of linear transformations: it corresponds to the change in volume caused by the transformation. Let us see this in 2 dimensions. To keep things simple, let us choose a Cartesian basis, $\mathbf{b}_1, \mathbf{b}_2$, so that $\mathbf{b}_1$ and $\mathbf{b}_2$ are at right angles to each other, as in Figure 4. Pick an arbitrary matrix,

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \tag{12}$$

and consider it as a linear transformation acting on this vector space. The basis vectors are transformed into vectors whose components in the basis are given by the columns of A:

$$\begin{pmatrix} a \\ c \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} b \\ d \end{pmatrix} . \tag{13}$$

What has the transformation done to areas? The transformation squishes the parallelogram bounded by $\mathbf{b}_1$ and $\mathbf{b}_2$ into the parallelogram bounded by the transformed vectors, as shown in Figure 4. We can work out the area of this parallelogram, using the lengths of the line segments shown in Figure 4 and the well-known formulae for the areas of a triangle and a rectangle. (Contrary to everyday practice, mathematical areas are signed quantities. The analogy you should make is with the "area under a curve" in integration. This area has a sign depending on the order of the limits of integration. Integration is the technical method by which the concept of area is actually defined.) The lengths are given as multiples of the lengths of the basis vectors. The area of the transformed parallelogram is

$$(a|\mathbf{b}_1| + b|\mathbf{b}_1|)(c|\mathbf{b}_2| + d|\mathbf{b}_2|) - 2b|\mathbf{b}_1|c|\mathbf{b}_2| - a|\mathbf{b}_1|c|\mathbf{b}_2| - b|\mathbf{b}_1|d|\mathbf{b}_2| = (ad - bc)|\mathbf{b}_1||\mathbf{b}_2| .$$

The determinant appears again! This formula holds more generally: with respect to any basis, the determinant of a matrix is the volume of the parallelipiped formed by the vectors corresponding to the columns of the matrix, divided by the volume of the parallelipiped formed by the basis vectors themselves.

Suppose we apply a second matrix B to the picture in Figure 4. The first matrix A squishes the volume by $\det A$, with respect to the original Cartesian basis, $\mathbf{b}_1, \mathbf{b}_2$. The second matrix further squishes the volume by $\det B$, with respect to the basis vectors transformed by A. These are the vectors whose corresponding 2×1 matrices with respect to the original basis are given by the columns of A. (We have to be a little careful here to make sure that these transformed basis vectors actually form a basis. However, the only reason this might not happen is if $\det A = 0$. Why? This is a special case that can be treated separately.) The net squish will be $(\det A)(\det B)$, with respect to the original Cartesian basis. It follows that $\det AB = (\det A)(\det B)$. Hopefully, this gives some intuition for why the determinant is a multiplicative function, which is one of the more surprising formulae in matrix algebra.

7 Change of basis, similarity and eigenvectors

It is nice to have more intuition for the determinant but if that is all there was to it, it would not have been worth going through all this material on vector spaces and linear transformations. The real reason for explaining this geometrical perspective is to guide you towards a different way of thinking about matrices, in which the matrix itself becomes less important than the linear transformation which it defines. This is quite a profound change and it is fundamental to appreciating some of the deeper properties of matrices.

Linear transformations exist independently of any choice of basis. For instance, we can define rotations, reflections and dilations without reference to any basis. If we choose a basis, then we can represent any linear transformation by a matrix, as we saw in (7). If we change the basis, then the matrix will change but the linear transformation will remain the same. Perhaps we should treat different matrices, which represent the same linear transformation with respect to different bases, as, in some sense, equivalent. If the particular matrix we have is complicated, perhaps we can find a simpler "equivalent" matrix, which represents the same linear transformation but with respect to a different basis. Notice, that it would have been hard to have even thought about this without knowing about vectors and linear transformations.

To see what this might mean, we need to understand how a matrix changes when we change the basis of the vector space. Let us suppose we have an n-dimensional vector space, V, and a linear transformation, $F : V \to V$. Let us choose two basis sets, $B = \mathbf{b}_1, \mathbf{b}_2, \cdots, \mathbf{b}_n$, and $C = \mathbf{c}_1, \mathbf{c}_2, \cdots, \mathbf{c}_n$. What is the relationship between $[F]_B$ and $[F]_C$?

The easy way to work this out is shown in Figure 5. Let us denote V with the basis B by V_B and V with the basis C by V_C. These are the same vector space but they have different bases. If we consider F as a transformation on V_B: $F : V_B \to V_B$, then we can use (7) to determine $[F]_B$. Similarly, if we consider F as a transformation on V_C: $F : V_C \to V_C$, we can determine $[F]_C$. To link these together, we need only interpose the identity transformation on V (ie: the transformation that leaves each vector fixed) but take it from V_B to V_C and from V_C to V_B. At the level of vectors, it is obvious that $IFI = F$, but if we determine the matrices with respect to the bases shown in Figure 5 we find that

$$[F]_B = [I]_{BC}[F]_C[I]_{CB} \,, \tag{14}$$

where we have used an obvious notation to indicate that the matrix of I should be taken with respect to different bases on each side of the transformation.

We are in a slightly different situation here to the one considered in (7) because we have a different basis on each side of the transformation. However, exactly the same rule as in (7) works for this more

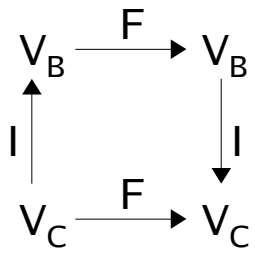

Figure 5: The linear transformation $F : V \to V$ can be viewed with respect to either basis B or basis C. The identity transformation, $I : V \to V$, enables changing between the two bases. From the viewpoint of vectors, $IFI = F$, while from the viewpoint of matrices we get (14).

general situation. The columns of the matrix $[I]_{BC}$ are given by the components of the transformed vectors, with respect to the new basis C. Since the identity transformation does not do anything, this amounts to determining $[\mathbf{b}_i]_C$. Hence,

$$[I]_{BC} = ([\mathbf{b}_1]_C \mid [\mathbf{b}_2]_C \mid \cdots \mid [\mathbf{b}_n]_C) \ ,$$

and a similar formula holds for $[I]_{CB}$ with the roles of B and C interchanged. However, it is simpler to note that if we go from V_B to V_C with the identity transformation and then from V_C to V_B with the identity transformation, we not only do not change the vector, we also do not change the basis. It follows that the two matrices are each other's inverse: $[I]_{CB} = [I]_{BC}^{-1}$.

Let us summarise what we have learned. If A is any matrix, then the representation of the same linear transformation with respect to any other basis looks like TAT^{-1}, where T corresponds to the change of basis matrix, $T = I_{BC}$. Note that there is nothing special about change of basis matrices, other than that they are always invertible (so that $\det T \neq 0$): given any basis, then any invertible matrix is the change of basis matrix to some other basis. We say that two matrices A and B are *similar* to each other if there exists some invertible matrix T such that $B = TAT^{-1}$. It is easy to check that this defines an *equivalence relation* on the set of all matrices, so that matrices can be partitioned into disjoint subsets, within which all the matrices are similar to each other. It is fundamental to the new way that we want to think about matrices that some of their most important properties are properties of the underlying linear transformation and therefore invariant up to similarity. For instance, this should be true of the determinant since, as we saw in §6, it can be defined in terms of the underlying linear transformation. Indeed, since it is a multiplicative function, so that $\det(AB) = \det(A)\det(B)$, it follows that $\det TAT^{-1} = \det AT^{-1}T = \det A$.

There is one issue that we need to clarify. The relationship between vectors and $n \times 1$ matrices or between linear transformations and $n \times n$ matrices depends on choosing a basis. But which basis? There is no "canonical" choice of basis in most vector spaces. For the most part, it does not matter which basis is chosen. However, this flexibility is actually rather awkward when we want to write down a specific vector or linear transformation. It is convenient therefore to fix upon a particular vector space in which there is a "special" basis and to work from now on in that vector space and with that basis. In n dimensions the vector space we use is the set of $n \times 1$ matrices, which we will denote V_n. This is a perfectly good vector space—it satisfies all the rules laid out in Figure 7—but we do lose the appreciation of geometry that we have with lines and planes. However, V_n has a preferred basis: the columns of the $n \times n$ identity matrix, as in (5). If we have some other n-dimensional vector space, W, and have chosen a particular basis, $B = \mathbf{b}_1, \cdots, \mathbf{b}_n$, then the transformation $[-]$ defines an isomorphism between W and V_n. We can always translate any calculation from V_n to W in much the same way that we moved back and forth between two different bases B and C in Figure 5. So, from now on, vectors will be column vectors and linear transformations acting on

vectors will be matrices multiplying column vectors. We will also go back to using matrix notation, with lower case $x, y, v, \cdots$ to denote vectors.

This leads us back to the questions that we started with. Given a matrix A, can we find a matrix similar to it, TAT^{-1}, which is simpler or in some way easier to deal with? The simplest matrices are diagonal matrices. So the first question we shall ask is, when is a matrix similar to a diagonal matrix? Or, more formally, when is a matrix *diagonalisable* ? This amounts to asking whether we can find a matrix T such that

$$AT^{-1} = T^{-1}L\,, \tag{15}$$

where L is a diagonal matrix. Let us disentangle this equation on a column by column basis. Let λ_j be the entry on the j-th column of L. You should be clear that (15) is satisfied if, and only if,

$$A(T^{-1})_{\cdot j} = \lambda_i (T^{-1})_{\cdot j}\,.$$

In other words, the columns of T^{-1} must represent vectors whose directions are left fixed by A. This leads to one of the most fundamental definitions in matrix algebra.

The non-zero vector x is said to be an *eigenvector* of matrix A for the *eigenvalue* λ, if and only if,

$$Ax = \lambda x\,. \tag{16}$$

(We exclude the zero vector for the obvious reason that it would otherwise be an eigenvector for any eigenvalue.) We have just worked out that for a matrix A to be diagonalisable, it needs to have a basis of eigenvectors. If we look at the underlying linear transformation in this new basis, then the new matrix will be diagonal. The eigenvectors provide the columns of the change of basis matrix that takes us from the canonical basis to the new basis of eigenvectors.

Notice that the eigenvector for a given eigenvalue is only defined up to a scalar multiple. If x is an eigenvector for the eigenvalue λ then so too is αx, where $\alpha \neq 0$. Eigenvectors specify direction, not magnitude. Accordingly when we say that two eigenvectors are "distinct", we explicitly disallow the possibility that one is a scalar multiple of the other: distinct eigenvectors must point in different directions.

8 Eigenvalues and the characteristic equation

When does equation (16) hold? We can rewrite it using the identity matrix as

$$(A - \lambda I)x = 0\,,$$

which implies that the matrix $A - \lambda I$ takes a non-zero vector, x, to the zero vector. Accordingly, $A - \lambda I$ cannot be invertible (why?) and so its determinant must be 0:

$$\det(A - \lambda I) = 0\,. \tag{17}$$

This equation for λ is the *characteristic equation* of A. Having a solution of it is a necessary condition for an eigenvector x to exist, whose eigenvalue is the solution of (17). It is a also a sufficient condition. If we can find a λ that solves (17), then we know that the matrix $A - \lambda I$ is not invertible and that therefore there must exist a non-zero vector x which satisfies (16). This last assertion is not so obvious but it can be shown that one of the columns of the adjugate matrix of $A - \lambda I$ can provide the necessary x. (You might have guessed as much from the formula for the adjugate which we studied in §6 of Part I.)

Let us calculate the characteristic equation in dimension 2 using the same matrix A that we used in (12) to work out volumes. We find that

$$\det(A - \lambda I) = \det \begin{pmatrix} a - \lambda & b \\ c & d - \lambda \end{pmatrix} = (a - \lambda)(d - \lambda) - bc = \lambda^2 - (a + d)\lambda + (ad - bc)\,.$$

For a general n-dimensional matrix, the characteristic equation equates to 0 a polynomial in λ of degree n. We shall denote this *characteristic polynomial* $p_A(\lambda)$ or just $p(\lambda)$ when the matrix is clear from the context. Since we introduced it in the context of diagonalisation, let us make sure that the it is a property of the underlying linear transformation and not just of the matrix A. Using the fact that the determinant is a multiplicative function, we see that

$$\det(TAT^{-1} - \lambda I) = \det T(A - \lambda I)T^{-1} = \det(A - \lambda I)\,,$$

and hence that

$$p_{TAT^{-1}}(\lambda) = p_A(\lambda)\,.$$

In particular, all of the coefficients of the characteristic polynomial are invariant under similarity. We already know this for the constant term of the polynomial, $p_A(0) = \det A$. It is also helpful to single out one of the other coefficients. The *trace* of a matrix, denoted $\mathrm{Tr}A$, is the sum of the terms on the main diagonal. For A in (12), $\mathrm{Tr}A = a + d$. It is not hard to check that the coefficient of λ^{n-1} in $p_A(\lambda)$ is $(-1)^{n+1}\mathrm{Tr}A$. In two dimensions, as we see from the calculation we just did, the trace and determinant completely determine the characteristic equation:

$$p_A(\lambda) = \lambda^2 - \mathrm{Tr}A\lambda + \det A\,. \tag{18}$$

There is one situation in which the characteristic equation and the eigenvalues can be easily determined. If the matrix is upper (or lower) triangular then it should be easy to see that its eigenvalues are precisely the entries on the main diagonal. If these are $d_1, \cdots, d_n$, then the characteristic polynomial already comes to us broken up into linear factors:

$$(\lambda - d_1)(\lambda - d_2)\cdots(\lambda - d_n) = 0\,.$$

In general, the characteristic polynomial is harder to work out. We know, as a consequence of the fundamental theorem of algebra, mentioned above, that a polynomial equation of degree n has exactly n solutions (or "roots" as they are usually called). The problem is that some of these roots may be complex numbers. For instance, you should know how to solve the quadratic polynomial (18). Its roots are

$$\lambda = \frac{\mathrm{Tr}A \pm \sqrt{(\mathrm{Tr}A)^2 - 4\det A}}{2}\,. \tag{19}$$

The quantity under the square root sign is the *discriminant* of the polynomial. For the characteristic polynomial of A, we will denote it $\triangle A$:

$$\triangle A = (\mathrm{Tr}A)^2 - 4\det A\,.$$

If $\triangle A < 0$, then the characteristic polynomial has two complex conjugate roots. In this case, A has no eigenvectors. An example of this is the rotation matrix (8). In this case, $\mathrm{Tr}A = 2\cos\theta$, while $\det A = \cos^2\theta + \sin^2\theta = 1$, so that $\triangle A = 4(\cos^2\theta - 1)$. Evidently, the discriminant is negative except when $\theta = 0$ or $\theta = \pi$. In the former case the rotation is the identity; in the latter case it is the dilation by -1. In all other cases, it is clear from the geometric picture of a rotation that there can be no vector which continues to point in the same direction after rotation. We see from this that rotation matrices, other than the two degenerate ones just mentioned, cannot be diagonalised.

The simplest situation in two dimensions arises when $\triangle A > 0$. In this case the matrix has two distinct real eigenvalues. Each eigenvalue has a corresponding eigenvector and it is not hard to see that these two eigenvectors form a basis (why?). Accordingly, any such matrix can be diagonalised. Reflections are a good example of this. It should be rather easy to see geometrically that for any reflection there are always two vectors (or, rather, directions) which are left fixed by the reflection. One is the line of the mirror itself and the other is the line at right angles to it. It should also be clear that the eigenvalue of the former is 1, while that of the latter must be -1. Can we deduce this

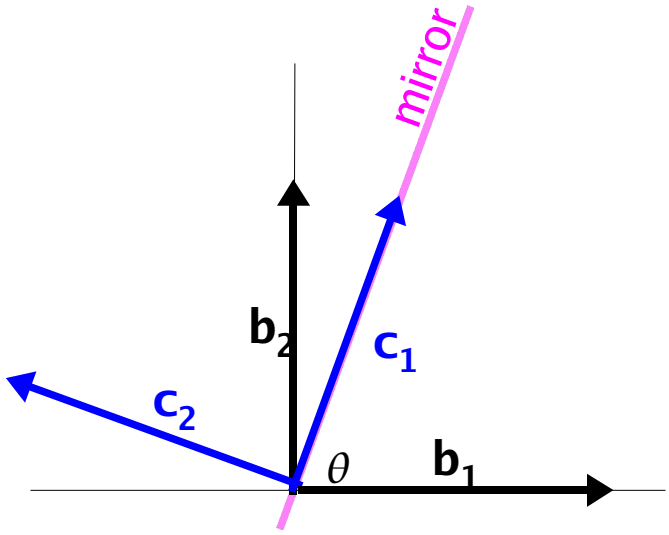

Figure 6: Reflection in the mirror at angle θ to the $\mathbf{b}_1$ axis. Change of basis from $B = \mathbf{b}_1, \mathbf{b}_2$ to $C = \mathbf{c}_1, \mathbf{c}_2$ in order to diagonalise this linear transformation. The basis C is obtained from B by counterclockwise rotation through θ.

for the reflection matrix in (9)? We see that $\mathrm{Tr}A = 0$, $\det A = -1$ and so $\triangle A = 4$, confirming that they are two distinct real roots. It is easy to check that these eigenvalues, as given by (19), are ± 1.

In this case we should be able diagonalise the matrix. Let us go through the details to see how they work. We will take advantage of knowing the directions in which the eigenvectors must lie. Let us assume that the new eigenvectors, $\mathbf{c}_1$ and $\mathbf{c}_2$, are chosen as shown in Figure 6, and that they have the same length as the original basis vectors, $\mathbf{b}_1$ and $\mathbf{b}_2$, respectively. We need to calculate the change of basis matrix I_{BC} but you should be able to see from Figure 6 that we already know a matrix which changes the basis from B to C. It is the rotation through the same angle θ as the mirror subtends against the $\mathbf{b}_1$ axis. Hence, I_{BC} is given by (8). Evidently, $I_{CB} = I_{BC}^{-1}$ must be given by (8) with $-\theta$ in place of θ. Accordingly, we expect that

$$\begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta \end{pmatrix} \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} .$$

I will leave it to you to check this. You will need the addition formulae for trigonometric functions. There is a simple way to work these out starting from $\exp(ia) = \cos a + i \sin a$ and using the addition formula for the exponential function: $\exp(a+b) = \exp(a)\exp(b)$.

There is one final case to discuss when $\triangle A = 0$. This is the awkward case. The matrix has two equal real roots. There is only one eigenvalue, repeated twice. Sometimes this is harmless and we can find two eigenvectors having the same eigenvalue. The identity matrix is an example. The following matrix, however,

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} ,$$

has the eigenvalue 1 repeated twice but only one eigenvector (what is it?). It cannot be diagonalised.

If we slightly reorganise what we have learned, we see that a 2×2 matrix A can fall into one of 3 mutually disjoint classes.

Class (1) $\triangle A < 0$ so that $p_A(\lambda)$ has two complex conjugate roots. In this case A cannot be diagonalised but it is similar to a rotation matrix and through a change of basis it can be brought into the form

$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix} .$$

Notice that this is the same form as the complex number matrices that we discussed in §4 of Part I.

Class (2) $\triangle A \geq 0$, $p_A(\lambda)$ has two real roots and two distinct eigenvectors. A can be diagonalised and through a change of basis it can be brought into the form

$$\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}.$$

Class (3) $\triangle A = 0$, $p_A(\lambda)$ has one repeated real root but only one eigenvector. A cannot be diagonalised but through a change of basis it can be brought into the form

$$\begin{pmatrix} a & b \\ 0 & a \end{pmatrix}.$$

$$\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}$$

$$\mathbf{x} + (\mathbf{y} + \mathbf{z}) = (\mathbf{x} + \mathbf{y}) + \mathbf{z}$$

$$\mathbf{0} + \mathbf{x} = \mathbf{x} + \mathbf{0} = \mathbf{x}$$

$$\mathbf{x} + -\mathbf{x} = -\mathbf{x} + \mathbf{x} = \mathbf{0}$$

$$(\lambda\mu)\mathbf{x} = \lambda(\mu\mathbf{x})$$

$$\lambda(\mathbf{x} + \mathbf{y}) = \lambda\mathbf{x} + \lambda\mathbf{y}$$

$$(\lambda + \mu)\mathbf{x} = \lambda\mathbf{x} + \mu\mathbf{x}$$

$$0\mathbf{x} = \mathbf{0} \quad 1\mathbf{x} = \mathbf{x} \quad (-1)\mathbf{x} = -\mathbf{x}$$

Figure 7: The rules satisfied by vector addition and scalar multiplication. $\mathbf{0}$ denotes the zero vector and $-\mathbf{x}$ denotes the inverse vector, as defined in the text.

Index