Mathematics 307—October 11, 1995

Linear transformations and coordinate changes in two dimensions

We can summarize what we have already seen in two results. (1) Linear coordinate systems are those with respect to which scalar multiplication and vector addition have linear expressions in the coordinates. Linear coordinate systems correspond to bases for the two-dimensional plane or, equivalently, pairs of linearly independent vectors in the plane.

Proposition. *Suppose that $E = [e_1\ e_2]$ is one basis and $F = [f_1\ f_2]$ another with the matrix relationship*

$$\begin{aligned} f_1 &= ae_1 + ce_2 \\ f_2 &= be_1 + de_2 \end{aligned}$$

or

$$[\, f_1 \quad f_2 \,] = [\, e_1 \quad e_2 \,] \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \qquad F = EA$$

If a vector x has coordinate column vector

$$x_E = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

with respect to E, so that

$$x = Ex_E$$

and x_F with respect to F, then

$$x_F = A^{-1} x_E \ .$$

One point here is the role that a matrix plays: it acts by a simple multiplication on the right on pairs of vectors:

$$[\, e_1 \quad e_2 \,] \begin{bmatrix} a & b \\ c & d \end{bmatrix} = [\, ae_1 + ce_2 \quad be_1 + de_2 \,]$$

(2) Suppose a basis E chosen. If T is a linear transformation then it corresponds to a matrix

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

whose columns are the E-coordinates of Te_1, Te_2.

Proposition. *If x has E coordinate vector x_E then Tx has E coordinates $M_E x_E$.*

In other words

$$T\colon \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \mapsto \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

or

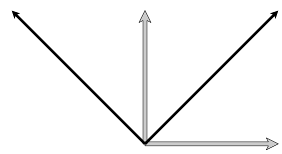
$$(Tx)_E = M_E x_E \ .$$

A linear transformation is essentially a geometric object, and a matrix amounts to an algebraic representation of it. The entries of this matrix are in some sense the coordinates of the linear transformation in terms of the coordinate system at hand.

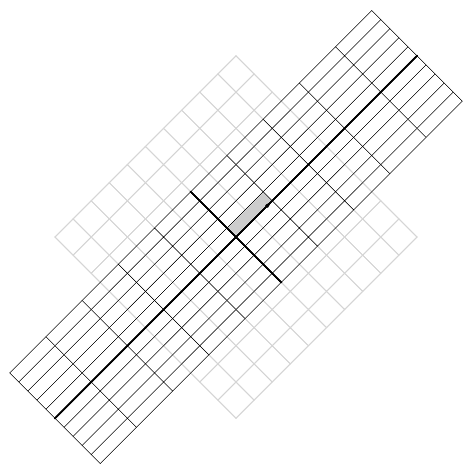
The first new point in these notes is how the matrix of a linear transformation changes with the coordinate system. I will begin with an example.

Example. Suppose

$$f_1 = e_1 + e_2$$
$$f_2 = -e_1 + e_2$$



Define the linear transformation to be a scale change, but instead of one along the axes, it will scale by a factor of 2 along the line $y = x$ and by a factor of $1/2$ along the line $y = -x$.



These directions are the lines through f_1 and f_2, so that

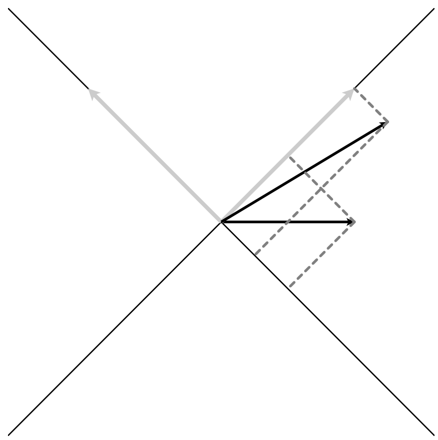
$$Tf_1 = 2f_1$$
$$Tf_2 = f_2/2$$

The matrix of the transformation T in the F-coordinate system has as its columns the coordinate vectors of Tf_1 and Tf_2 in the F-coordinate system. The F-coordinates of $2f_1$ are $(2, 0)$ and those of $f_2/2$ are $(0, 1/2)$, so the matrix of T in F-coordinates is

$$M_F = \begin{bmatrix} 2 & 0 \\ 0 & 1/2 \end{bmatrix} .$$

What is the matrix of T in the E-coordinate system? We will use two methods to find the answer.

The first is more complicated but more direct. The columns of the matrix are the vectors Te_1, Te_2 expressed as linear combinations of e_1 and e_2. So we must figure out what the effect of T is on e_1 and e_2.



Try e_1 first. The basic trick here is to resolve e_1 into its components along the lines $y = x$ and $y = -x$. As you can see from the picture and verify, this resolution is

$$e_1 = (1/2)f_1 - (1/2)f_2$$

and *since T is linear*

$$\begin{aligned} Te_1 &= (1/2)Tf_1 - (1/2)Tf_2 \\ &= 2(1/2)f_1 - (1/2)(1/2)f_2 \\ &= f_1 - (1/4)f_2 \\ &= (e_1 + e_2) - (1/4)(-e_1 + e_2) \\ &= (5/4)e_1 + (3/4)e_2 \end{aligned}$$

so the first column of M_E is

$$\begin{bmatrix} 5/4 \\ 3/4 \end{bmatrix}$$

For the second we must find Te_2. The picture shows that

$$e_2 = (1/2)f_1 + (1/2)f_2$$

so that

$$\begin{aligned} Te_2 &= (1/2)Tf_1 + (1/2)Tf_2 \\ &= 2(1/2)f_1 + (1/2)(1/2)f_2 \\ &= f_1 + (1/4)f_2 \\ &= (e_1 + e_2) + (1/4)(-e_1 + e_2) \\ &= (3/4)e_1 + (5/4)e_2 \end{aligned}$$

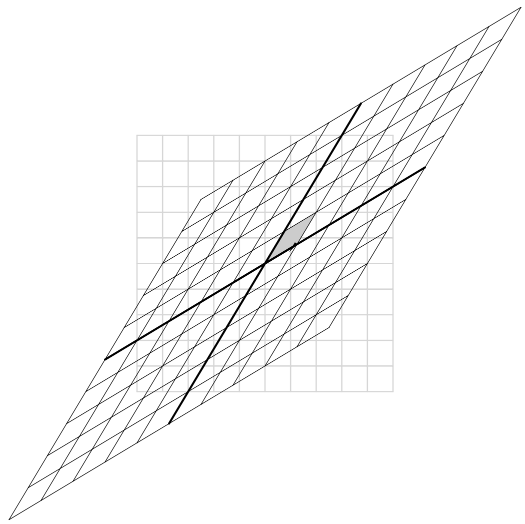
so that the second column is

$$\begin{bmatrix} 3/4 \\ 5/4 \end{bmatrix}$$

making the matrix M_E equal to

$$\begin{bmatrix} 5/4 & 3/4 \\ 3/4 & 5/4 \end{bmatrix} .$$

Its effect on a grid in the E coordinate system is this:



Keep in mind: *this is the same transformation as before, but its effect is portrayed on a different grid.*

In a second method: the meaning of the matrix M_E is that

$$(Tx)_E = M_E x_E$$

and similarly for F-coordinates

$$(Tx)_F = M_F x_F \ .$$

Recall also that

$$x_F = A^{-1} x_E, \quad x_E = A x_F \ .$$

If we apply this to Tx instead of x we have

$$(Tx)_F = A^{-1}(Tx)_E$$

But then if we substitute for $(Tx)_E$ from the first equation in this we get

$$(Tx)_F = A^{-1} M_E x_E = A^{-1} M_E A \, x_F$$

and if we compare this we the second equation we see that

$$M_F = A^{-1} M_E A \ .$$

This is the third main result of these notes.

Proposition. *Suppose E and F are are two bases with $F = EA$. If T is a linear transformation, M_E is its matrix with respect to E and M_F its matrix with respect to F then*

$$M_F = A^{-1} M_E A \ .$$

This calculation is essentially the same as the direct calculation we did first, although it might not not seem to be. (It is somewhat different from what I did in class.)

You have probably seen an equation like this in some earlier class, where it was said to express the fact that M_E and M_F are **similar**. This Proposition means that *similar matrices are the matrices of the same linear transformation, but for different choices of coordinate systems.*

Eigenvectors and coordinate change

If T is a linear transformation, an **eigenvector** of T is a vector (other than the zero vector) on which T acts by scaling. In other words, T takes it into a multiple of itself. The scale factor involved is called the **eigenvalue** of T associated to that eigenvector. In other words, $v \neq (0,0)$ is an eigenvector of T if and only if there exists a constant c such that

$$Tv = cv \ .$$

This is a geometric notion: it means that in the direction of v the transformation T acts by shrinking or stretching or whatever, but preserves the direction itself. In particular, this is a notion independent of any choice of coordinate system. But in order to find the eigenvectors and eigenvalues, unless you are lucky, you should expect to work with a coordinate system and a matrix. I recall that *if M is the matrix of T in some coordinate system then its eigenvalues are the roots of the polynomial equation*

$$\det(M - cI) = 0$$

where I is the identity matrix

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

This will be a quadratic equation. If c is an eigenvalue then the system of equations

$$M\, x_E = 0$$

will be singular, and will have a one-dimensional set of solutions. Any of these solutions other than $(0,0)$ will be an eigenvector.

If c_1 and c_2 are two real eigenvalues with eigenvectors x_1, x_2 and the eigenvectors are linearly independent (which will always be the case if $c_1 \neq c_2$) then we can choose this pair as a basis aand the matrix of T in this coordiante system will be

$$\begin{bmatrix} c_1 & 0 \\ 0 & c_2 \end{bmatrix} .$$

This is, in a sense, the point of eigenvectors: *the matrix of a linear transformation has a simple form if we choose a basis made up of eigenvectors.*

When the two-dimensional linear transformation has two linearly independent eigenvectors, it is said to be **diagonalizable**. Even it is diagonalizable, there may be some problems in using eigenvectors to visualize the linear transformation.

Example. Let T be rotation through $90°$. Its matrix in a suitable coordinate system is

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

It should be clear geometrically that there are no vectors in the plane that are scaled by T, since every vector other than the origin changes direction under the rotation. Nonetheless, T is diagonalizable: its eigenvalues are $\pm\sqrt{-1}$, and its eigenvectors are complex.

Not every linear transformation is diagonalizable, even if we allow complex eigenvalues. If T is a horizontal shear, then the only vectors whose direction does not change under T are those lying along the x-axis. In other words, for shears there exist real eigenvectors but there do not exist two linearly independent ones.

The next main result in this course is that, roughly speaking, *every linear transformation is either (1) diagonalizable in real directions; (2) the combination of a rotation and a uniform scale change; or (3) the combination of a shear with a uniform scale change.* The classification depends on whether T has (1) real eigenvalues and two linearly dependent eigenvectors; (2) conjugate complex eigenvalues; (3) two equal eigenvalues and only one line of eigenvectors.