Mathematics 307—September 7, 1995

Two dimensions

I will start the course with a discussion of linear algebra in two and then three dimensions. This part of the course will be to some extent a review of things you have already seen, but with a few new twists to familiar things and a few new things as well.

Suppose P to be a flat surface in space. Pick a point O on it, which will serve as an origin. If x is any point on P, we can then determine a vector by requiring that its tail be at O and its head at x. Since P is flat, all of the vector will in fact lie in P. Since we can scale vectors by constants, we can speak of scaling points on P also; and similarly we can add points by applying vector addition to the vectors they determine. The choice of origin and the operations of scalar multiplication and vector addition make P into a **vector space**.

Linear coordinates

A pair of vectors u, v in P are said to be **linearly dependent** if they lie along the same line, or in other words if one of them is a scalar multiple of the other. Otherwise they are said to be **linearly independent**. If e_1 and e_2 are linearly independent vectors on P then the set of all linear combinations

$$x_1 e_1 + x_2 e_2$$

make up all of P. Any vector in P has a unique representation of this form, and the coefficients corresponding to a point (or vector) in P are called its **coordinates** determined by the choice of the pair e_1, e_2. Choosing coordinates on P allows you to refer to all points explicitly (or as explicitly as you can refer to real numbers), and the operations of scalar multiplication and vector addition can be expressed in simple terms by means of coordinates. *If x has coordinates $x_1, x_2)$ then the coordinates of cx are (cx_1, cx_2). If x has coordinates (x_1, x_2) and y has coordinates (y_1, y_2) then the coordinates of $x+y$ are (x_1+y_1, x_2+y_2).*

The point of thinking of P as an arbitrary plane in space is to reinforce the idea that no one coordinate system on P is necessarily special unless more data are given. If we are given coordinates in space itself, then generally we can choose at least three coordinate systems on P by projecting onto one of the three coordinate planes. Of course many other choices are possible, as we shall understand later on, and as far as the **linear structure** of P is concerned they are all equally good.

The original pair of linearly independent vectors e_1, e_2 determining the coordinate system are said to be its **basis**. The coordinates of a point depend on the choice of basis, and we can ask how coordinates with respect to two different coordinate systems, or two different bases, are related.

So suppose that e_1, e_2 are one basis and f_1, f_2 another. The f's can be expressed as linear combinations of the e's, say

$$\begin{aligned} f_1 &= ae_1 + be_2 \\ f_2 &= ce_1 + de_2 \end{aligned}$$

and we can write this, at least formally, in terms of matrix multiplication:

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

This is not quite what you have seen as matrix multiplication before, where matrices had only numbers as coefficients, but it makes perfect sense according to the same rules. I will use this sort of expression quite often; I hope to show that it leads to very simple explanations of formulas that at first sight look quite complicated. I can write the above formula as what I call the **basic equation relating two bases**:

$$F = EA$$

where

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

and the columns of A are the coordinates of the f's in terms of the e's. We shall see later that in this equation the 2×2 matrix A must be **non-singular**, which is to say that its determinant is not zero, or that it has a matrix inverse.

This formula can be used to answer the question about different coordinates. Suppose x is a vector in P, and that E, F are two bases. Let x_E and x_F be the two column vectors we make up from the coordinates of x with respect to E and F. If

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

then we can write the equation

$$x = x_1 e_1 + x_2 e_2 = [\, e_1 \quad e_2 \,] \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = E x_E$$

and if we want to find the column vector x_F we can determine it according to the specification

$$x = F x_F \, .$$

Well, we have

$$x = E x_E$$

and also

$$F = EA \, .$$

We can get from this to

$$F = EA^{-1}, \quad E x_E = F A^{-1} x_E$$

and deduce immediately by comparison that

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

Example. Say that

$$x_E = \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \quad f_1 = e_1 + e_2, \quad f_2 = e_1 - e_2 \, .$$

Then

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

since if

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

then

$$A^{-1} = \frac{\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}}{\det A} \, .$$

Linear transformations

A linear transformation T of P is a map from P to itself which preserves its linear structure. Geometrically, this means that T maps parallelograms into paralleograms, and grids into (skewed) grids. Algebraically it means that

$$T\,cx = cTx$$
$$T(x+y) = Tx + Ty$$

and in terms of coordinates T must be expressed in linear terms: if x has coordinates (x_1, x_2) then Tx has coordinates $(ax_1 + bx_2, cx_1 + dx_2)$ for some constants a etc. If the coordinates of Tx are (y_1, y_2) we have then a matrix equation

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

or

$$(Tx)_E = Ax_E$$

for any given basis E. The matrix is determined the formula for coordinates, but it can also be determined by this fundamental rule: *if we are given a basis E and a linear transformation T then the matrix assocaited to T by the choice of E is the 2×2 matrix whose columns are made up of the coordinates of Te_1 and Te_2.* This is because we can see that the coordinates of Te_1 are

$$\begin{bmatrix} a \\ c \end{bmatrix}$$

and those of Te_2 are

$$\begin{bmatrix} b \\ d \end{bmatrix} .$$

Examples.

The identity map

It takes any point to itself. It takes e_1 to e_1 and e_2 to e_2. Its matrix is

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

Reflection in the y-axis

It takes e_1 to $-e_1$, e_2 to itself. The matrix is

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

Uniform scaling by c

It takes any x to cx. Its matrix is

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

Non-uniform scaling

Suppose we scale along the x-axis by a and along the y-axis by b. The matrix is

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

Projection

A special case of scaling is where we do no scaling in the x-direction, but collapse completely vertically. This amounts to **perpendicular** or **orthogonal projection** onto the x-axis. The matrix is

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

Rotation

Rotation in the positive direction (counter-clockwise) by θ has matrix

$$\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} .$$

Shear

Sliding parallel to the x-axis is called a **horizontal shear**. The matrix is of the form

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