Summary of Day 12

William Gunther

June 5, 2014

1 Objectives

- Define dimension, and give geometric interpretation.
- Be able to write a vector in a different coordinate system relative to a different basis.
- Begin talking about linear transformations.

2 Summary

- So, matrices are (as I've said repeatedly) a special type of function. Let's unlock exactly what that means. First we need to review what a function is:

  A **function** f is a mapping from a **domain** A to a **codomain** B. A function has a few guarantees, namely:

  - Every element from A gets mapped to somewhere.
  - Every element gets mapped somewhere in B.
  - There is only one thing in B that each element of A gets mapped to.

  The last condition actually implies the rest, but it's nice to say them all separately.

  The **range** of the function is the stuff in the codomain that actually gets hit.

  Example $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = x^2$ has domain $\mathbb{R}$, codomain $\mathbb{R}$, and range $\{ x \in \mathbb{R} \mid x \geq 0 \}$

- $T : \mathbb{R}^n \to \mathbb{R}^m$ is called a **linear transformation** if:

  - $T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$ for all vectors $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$.
  - $T(c\mathbf{u}) = cT(\mathbf{u})$ for all scalars $\mathbf{u}$ and all vectors $\mathbf{u} \in \mathbb{R}^m$.

  As with subspaces, we can abbreviate this to say:

  - $T(0) = 0$
  - $T(\mathbf{u} + c\mathbf{v}) = T(\mathbf{u}) + cT(\mathbf{v})$ for all $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ and all scalars c.

  Or simply to

  - $T(d\mathbf{u} + c\mathbf{v}) = dT(\mathbf{u}) + cT(\mathbf{v})$ for all $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ and all scalars c, d.

  Example Consider the function $T : \mathbb{R}^2 \to \mathbb{R}^3$ defined by:

$$T \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ 2x - y \\ 3x + 4y \end{pmatrix}$$

This is a linear transformation. It is between two vectors spaces, and we can verify:

$$T\left(\begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} w \\ z \end{pmatrix}\right) = T\left(\begin{pmatrix} x+w \\ y+z \end{pmatrix}\right) = \begin{pmatrix} x+w \\ 2(x+w)-(y+z) \\ 3(x+w)+4(y+z) \end{pmatrix}$$
$$= \begin{pmatrix} x+w \\ 2x-y+2w-z \\ 3x+4y+3w+4z \end{pmatrix}$$
$$= \begin{pmatrix} x \\ 2x-y \\ 3x+4y \end{pmatrix} + \begin{pmatrix} w \\ 2w-z \\ 3w+4z \end{pmatrix} = T\begin{pmatrix} x \\ y \end{pmatrix} + T\begin{pmatrix} w \\ z \end{pmatrix}$$

- Matrices are important because...

  Theorem Let A be a $m \times n$ matrix. Defined $T_A : \mathbb{R}^n \to \mathbb{R}^m$ by:

  $$T_A(\mathbf{x}) = A\mathbf{x}$$

  this is a linear transformation.

  Proof. Because of the way distributivity and scalar multiplication work with matrices and matrix multiplication. □

- As vectors can represent the points of $\mathbb{R}^n$ (and give those points certain arithmetic properties) we can think of a linear transformation as a transformation of space. So it in the plane, it can distort, stretch, shrink, reflect, rotate, etc space (as long as it dones so linearly).

  For example, consider the linear transformation that arises from a 90 degree rotation around the origin. This transformation looks like:

  $$R\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -y \\ x \end{pmatrix}$$

  We can see this can actually be represented as a matrix, under matrix multiplication. To see this, let's look at where the standard basis goes:

  $$R\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \qquad R\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \end{pmatrix}$$

  As every element of space can be written as a linear transformation of these vectors in the following way:

  $$a\begin{pmatrix} 1 \\ 0 \end{pmatrix} + b\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} a \\ b \end{pmatrix}$$

  We can exploit the linearity and try to find a matrix with these two properties:

  $$A\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \qquad A\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \end{pmatrix}$$

  But, our hands are completely tied. Why? Well, what is the result of the matrix multiplication $A\begin{pmatrix} 1 \\ 0 \end{pmatrix}$? I know it has to be $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$, but at the same time it has to be the first column of the matrix. Same for $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ and $\begin{pmatrix} -1 \\ 0 \end{pmatrix}$. Therefore, the matrix has to look like:

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

  This behaves correctly on the standard basis, and therefore, it will work correctly on all vectors because of linearity!

  $$A\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} -b \\ a \end{pmatrix}$$

Theorem If $T : \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation then T can be represented by a matrix. Moreover, the matrix is constructed as follows: let $\mathbf{u_1} = T\mathbf{e_1}, \ldots, \mathbf{u_n} = T\mathbf{e_n}$, where $\mathbf{e_i}$ is the ith standard basis element. Then $A_T = (\mathbf{u_1} \ldots \mathbf{u_n})$, i.e. a matrix with the $\mathbf{u_i}$ as the columns of the matrix.

Proof. The above example illustrates the principle. □

Example The 90 degree rotation transformation is a specific case of a general rotation. Let's say we want to rotate the plane by 90 degrees. How should we do it?

- When we have two function $f : A \to B$ and $g : B \to C$ we can take the **composition** of f and g to get a function:

  $$g \circ f : A \to C$$

  Defined by: $(g \circ f)(x) = g(f(x))$ (first do f then do g to the result).

  This can be done with linear transformations, and with matrices, you just multiply the two matrices to compose the linear transformation they represent!

  Example What should you get when you multiply the 45 degree rotation matrix with itself? Or the 45 degree rotation matrix with the -45 degree rotation matrix?

- (This is not in the book for whatever reason as far as I can tell, so pay extra attention here) What does the column space and null space represent?

  The column space is the set of vectors you can write as a linear combination of the columns remember. When we perform the action of 'plugging' a vector into a linear transformation representing by a matrix then we are actually writing some matrix (the output) as a linear combination of the columns (think about the standard basis elements first to get a visual)

  Therefore, the column space corresponds to the **range** of the linear transformation. That is the stuff that you can write as linear combinations is exactly the stuff that you can get as output for the linear transformation.

  The null space is the stuff that when you 'plug in' to the matrix you get the $\mathbf{0}$. In his guise (when viewing matrices as linear transformations) the null space is usually called the **kernel**.

  Therefore the rank theorem tells us that the dimension of the range (which is the rank) plus the dimension of the kernel is equal to the dimension of the domain! In a way this is a like a 'preservation of matter'-type theorem. You have a n-dimensional space of vectors that your function is mapping out of. Some of it goes to 0, the rest of it doesn't, but both of those parts create subspaces. The sum of the dimension of these two parts is equal to the whole.

- We can also define an inverse linear transformation. If T is a linear transformation on $\mathbb{R}^n$ and there exists a R a linear transformation on $\mathbb{R}^n$ such that $T \circ R = R \circ T = \text{id}$ where id is the identity on $\mathbb{R}^n$ then we say T is invertible. Inverses are unique (you can prove it) and we write $T^{-1} = R$.

  This exactly parallels matrices because matrices and linear transformations are the same thing!!