Summary of Day 7

William Gunther

May 28, 2014

1 Objectives

- Connect matrix multiplication with systems of equations.
- See the 'true nature' of matrices as functions.
- Define some more matrix operations (e.g. transpose) and explore algebraic properties of matrices.

2 Summary

- Matrix multiplication is *not* commutative. In fact, it often doesn't even make sense to switch the order because of the size constraints!
- We can rephrase linear systems into this language of matrix multiplication. This will allow us when we see the true nature of matrices to prove/use a lot of the theorems we said about systems of equations using a completely different way of thinking about systems than we were capable of.

  Example Consider the system

$$\begin{aligned} x - 2y + 3z &= 0 \\ 2x + y - z &= 4 \end{aligned}$$

  Recall we can phrase this as an augmented matrix:

$$\left(\begin{array}{ccc|c} 1 & -2 & 3 & 0 \\ 2 & 1 & -1 & 4 \end{array}\right)$$

  We could also phrase it in this way:

$$\begin{pmatrix} 1 & -2 & 3 \\ 2 & 1 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 4 \end{pmatrix}$$

  Remark Very important!! Now, notice, asking whether there is a solution to a system can be rephrased as 'is there anything that I can multiply this matrix by (on the right) to give me this vector as a solution'

- Note that the notation we have used for an augmented matrix $(A \mid \mathbf{b})$ will now just be thought to be shorthand for the matrix equation $A\mathbf{x} = \mathbf{b}$. It is still useful notationally because of Gaussian and Gauss-Jordan elimination.
- We like to think of matrices as a function; that is, a matrix is a way of transforming one set of information to another. We'll now look at an informal example of how this works, and connect it with matrix multiplication. This is example 3.7 of the book.

  Example Ann and Bert are planning to go shopping for fruit for the next week. They each want to buy some apples, oranges, and grapefruit, but in differing amounts. The following two tables give the quantity of each item that each other like to buy, and how much it costs are two of the nearby markets: Sam's and Theo's.

	Apples	**Grapefruit**	**Orange**
Ann	6	3	10
Bert	4	8	5

	Sam's	**Theo's**
Apple	\$0.10	\$0.15
Grapefruit	\$0.40	\$0.30
Orange	\$0.10	\$0.20

We first make a person $\times$ fruit matrix

$$\begin{pmatrix} 6 & 3 & 10 \\ 4 & 8 & 5 \end{pmatrix}$$

This is like a function; if you plug in a person it gives you all the fruit that the person wants to buy! Next we make a fruit $\times$ store matrix:

$$\begin{pmatrix} .1 & .15 \\ .4 & .3 \\ .1 & .2 \end{pmatrix}$$

This is similarly like a function; if you look at a particular fruit, it tells you the price at each of the stores. Now we multiple the type matrices, which gives us a person $\times$ store matrix; this tells us how much the person would spend at each of the store by composing the two function we had.

$$\begin{pmatrix} 6 & 3 & 10 \\ 4 & 8 & 5 \end{pmatrix} \begin{pmatrix} .1 & .15 \\ .4 & .3 \\ .1 & .2 \end{pmatrix} = \begin{pmatrix} 2.8 & 3.8 \\ 4.1 & 4 \end{pmatrix}$$

So it's in Ann's interest to go to Sam's by a significant amount, but in Bert's to go to Theo's by a relatively small amount.

Exercise: What could you multiply by what to get the total each would spend at each store?

Example Consider S the set of people on Facebook. Consider a $|S| \times |S|$ matrix, A, where the entries are either 0 or 1 based on whether two people are friends are not. What does AA signify?

- Here is another teaser: A is a $m \times n$ matrices. What can you multiply by what to get the ith column of the matrix as a column vector? What about the ith row as a row vector?

  Let e_j be the **jth standard basis vector** of R^n; that is,

  $$e_j = [0, 0, \dots, 0, 1, 0 \dots, 0, 0]$$

  where the 1 is in the jth row.

  Theorem Consider e_i as a row vector. If A is a $m \times n$ matrix.

  - Consider e_i as a row vector in $\mathbb{R}^m$. Then $e_i A = i$th row of A.
  - Consider e_j as a column vector in $\mathbb{R}^n$. Then $Ae_j = j$th column of A.

  Proof. We'll just prove the first as the second is similar; the size of the resultant matrix is clearly $1 \times n$. The $1j$th entry is the vector e_i dotted with the jth row of A; but, since e_i is 1 only in the ith coordinate and 0 elsewhere, only the ith row of the jth column contributes to the sum. Therefore, the $1j$th entry of the resultant matrix is just the ith row of the jth column of A. Therefore, the resultant matrix is just the ith row of A. □

- The example of Facebook above reveals often it's nice to multiply a matrix by itself; you can only do this with square matrices. This is called a **matrix power** and we write it as A^k where k is a nonnegative integer.

  Theorem The following familiar rules occur:

  - $A^r A^s = A^{r+s}$
  - $(A^r)^s = A^r s$

- The next operation on matrices seems like an unnatural one, but consider the example above with the person $\times$ fruit matrix it should be natural; the idea that the people were the rows was arbitrary, and in order to multiply the matrices one often needs to switch the order of association. For this we introduce the **transpose** of a matrix; The transpose of a matrix is as follow:

If A is $(c_1\, c_2\, \ldots\, c_n)$ where c_i are column vectors then A^T (pronounced A transpose) is defined to be

$$A^T = \begin{pmatrix} r_1 \\ r_2 \\ \vdots \\ r_n \end{pmatrix}$$

where r_i is the row vector corresponding to c_i

We could also defined A^T entrywise: that is, the ijth entry of A^T is the jith entry of A.

<u>**Example**</u>

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

- A matrix that will prove to be important later is a **symmetric matrix**; A is symmetric if $A = A^T$.
- Now we will explore the algebraic properties of some of these operations

  <u>**Theorem**</u> If A, B, C are all $m \times n$ matrices and c, d are real numbers then:

  1. $A + B = B + A$
  2. $(A + B) + C = A + (B + C)$
  3. $A + O = A$
  4. $A + (-A) = O$
  5. $c(A + B) = cA + cB$
  6. $(c + d)A = cA + dA$
  7. $c(dA) = (cd)A$
  8. $1A = A$

  These properties should look a lot like the properties for vectors; in fact, we can carry over a lot of what we did for vectors for matrices.

- A **linear combination** of matrices $A_1, \ldots, A_k$ is $c_1A_1 + \ldots c_kA_k$ where c_i are scalars and the sizes of the A_i make sense (i.e. they are all the same). The definitions of span, linear dependence, and linear independence also carry over.

  Really we're doing nothing new here as the next example should make clear:

  <u>**Example**</u> Determine if the following matrices are linearly dependent:

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

  Solution. We are looking for non-zero solutions to this equation:

  $$c_1 \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} + c_2 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + c_3 \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} = O$$

  But you can see this is really expressing the following simultaneous system:

  - $c_2 + c_3 = 0$
  - $c_1 + c_3 = 0$
  - $-c_1 + c_3 = 0$
  - $c_2 = 0$

  which can be expressed as this matrix:

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

  You'll see this has a unique solution of $\mathbf{0}$ if you row reduce.

- The algebra of matrix multiplication is a little more unorderly. It does have some nice properties though.

  <u>Theorem</u> If A, B, C are matrices of appropriate size and c a scalar then:

  1. $A(BC) = (AB)C$
  2. $A(B + C) = (AB) + (AC)$
  3. $(A + B)C = AC + BC$
  4. $k(AB) = (kA)B = A(kB)$
  5. $I_m A = AI_n = A$ if A is $m \times n$.

- Many properties that we're used to with real numbers fail in general for matrix multiplication. Particularly, these two very important properties of real number multiplication fail spectacularly:

  – Commutativity: It is *not* true that $AB = BA$ (even when the sizes make sense).

  – Invertibility: It is not true that every matrix is invertible (we will talk about this next).

- Transposes also work more or less how you'd expect with perhaps a few surprises.

  <u>Theorem</u> If A, B matrices of appropriate size and c a scalar then:

  1. $(A^T)^T = A$
  2. $(kA)^T = kA^T$
  3. $(A^r)^T = \left(A^T\right)^r$
  4. $(A + B)^T = A^T + B^T$
  5. $(AB)^T = B^T A^T$

  Proof. We will just prove the last property which may be the most surprising. We will do this by calculating the ijth entry of each matrix, and show they're the same.

  The ijth entry of the lhs is the jith entry of AB (by the definition of transpose) which is the jth row of A dotted with the ith column of B.

  The ijth entry of the rhs is the ith row of B^T dotted with the jth column of A^T which in turn is the ith column of B dotted with the jth row of A by the definition of transpose. Dotting is commutative, so these two quantities are the same. $\square$