Summary of Day 6

William Gunther

May 27, 2014

1 Objectives

- Talk about relationship between homogeneous equations and linear independence/dependence and prove some theorems.
- Discover another way to approach linear independence with row vectors in a matrix.
- Begin looking at matrices as objects, and operations on matrices.
- Connect matrix multiplication with systems of equations.

2 Summary

- As we've discussed, there are non-trivial solutions to a homogeneous system where the columns are the vectors from a set S if and only if the vectors in S are linearly dependent.

  Example Determine if $S = \{[3, 1, 4], [-2, 1, -1]\}$ are linearly independent or linearly dependent.

  Solution. First we set up a matrix where the columns are the vectors. Note: we are used to setting up an augmented matrix, but since the augmented part of the matrix with be $\mathbf{0}$ we will omit it because it really won't contribute anything to our analysis; any row operations will preserve that column.

  $$\begin{pmatrix} 3 & -2 \\ 1 & 1 \\ 4 & -1 \end{pmatrix} \to \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}$$

  This tells us that there is a unique solution; namely, $\mathbf{0}$. Therefore, the vectors are linearly independent.

  Example Consider the vectors $S = \{[1, 3], [1, 6], [-1, 2]\}$. Determine if they are linearly independent or linearly dependent.

  Solution. To determine if they are linearly dependent we put them in a matrix as columns.

  $$\begin{pmatrix} 1 & 1 & -1 \\ 3 & 6 & 2 \end{pmatrix}$$

  This matrix is consistent, and has rank ≤ 2. Therefore, there is at least on free variable, so there are infinitely many solutions. Therefore, there are nontrivial solutions, and so the vectors are linearly dependent.

- We can actually say quite a lot just from the number of vectors on whether they are linearly dependent.

  Theorem Any set of m vectors from $\mathbb{R}^n$ is linearly dependent.

  Proof. Let S be a set of m vectors. Consider a matrix A consisting of columns of column vectors from S. Recall that a homogeneous system has infinitely many solutions if there are more equations than unknowns; that is the case here. Therefore, there is infinitely many solutions, and therefore nontrivial solutions. So S is linearly dependent. □

  Remark This would have allowed us to easily decide linear dependence in the second example above since there were 3 vectors in $\mathbb{R}^2$.

  Remark The above is *not* an if and only if. The converse fails; can you find a set of $\leq n$ vectors which is linearly dependent? Bonus points for finding the smallest set.

- We have thus far always answered questions about vectors in terms of equations by putting the vectors as columns of a matrix which we viewed as the coefficient matrix of a system of equations. Now let's briefly discuss another view which is to put the vectors as the *rows* of a matrix. It's best to illustrate this with an example Here we do a little book-keeping with our elementary row operations as we go

  Example Consider the vectors $[1,2,0],[1,1,-1],[1,4,2]$. We will put them as the rows in a matrix and then row reduce.

$$\begin{array}{c} R_1 \\ R_2 \\ R_3 \end{array} \begin{pmatrix} 1 & 2 & 0 \\ 1 & 1 & -1 \\ 1 & 4 & 2 \end{pmatrix} \to \begin{array}{r} R_1 \\ R_2 - R_1 \\ R_3 - R_1 \end{array} \begin{pmatrix} 1 & 2 & 0 \\ 0 & -1 & -1 \\ 0 & 2 & 2 \end{pmatrix}$$

$$\to \begin{array}{r} R_1 \\ R_2 - R_1 \\ R_3 - R_1 + 2(R_2 - R_1) \end{array} \begin{pmatrix} 1 & 2 & 0 \\ 0 & -1 & -1 \\ 0 & 0 & 0 \end{pmatrix}$$

  Note, what we have is the third row is now $\mathbf{0}$ and is equal to $R_3 - R_1 + 2(R_2 - R_1) = -3R_1 + 2R_2 + R_3$ of the original matrix; which tells us:

$$\mathbf{0} = -3\begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} + 2\begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} + \begin{pmatrix} 1 \\ 4 \\ 2 \end{pmatrix}$$

  which tells us that these vectors are linearly dependent!

- The task above was to row reduce so that we could get a row of all 0's. Getting a row of all 0's depends on the rank of the matrix.

  Theorem If A is a matrix with m rows then $\text{rank}(A) < m$ if and only if A is row equivalent to a matrix with a row of 0's.

  Proof. ($\Rightarrow$) The rank is the number of leading entries in the matrix. If the number of leading entries is less than the number of rows, then there must be a row without a leading entry. A row lacking a leading entry must be all 0's.

  ($\Leftarrow$) Conversely, suppose A row reduced to a matrix with a row of all 0's. Then the number of leading entries must be less than the number of rows and this particular row does not have a leading entry. $\square$

  Remark We can see that we could have defined the rank to be the number of nonzero rows in the reduced row echelon form of a matrix (or row echelon form) as each nonzero row in ref has exactly one leading entry, and all zero rows have no leading entries, so they are the same.

  Theorem Let $S = \{\mathbf{v_1}, \ldots, \mathbf{v_m}\} \subseteq R^n$. Let

$$A = \begin{pmatrix} v_1 \\ \vdots \\ v_m \end{pmatrix}$$

  i.e., it is a matrix with rows from S. Then S is linearly dependent if and only if $\text{rank}(A) < m$.

  Proof. ($\Leftarrow$) By the above theorem, we can row reduce the matrix to get a row with $\mathbf{0}$. As in the example above, this gives us a linear combination of vectors from S which equal $\mathbf{0}$.

  ($\Rightarrow$) We know there are $c_1\mathbf{v_1} + \cdots + c_m\mathbf{v_m} = \mathbf{0}$ where not all c_i are nonzero. Remove all vectors from this list with 0 coefficient. This tells us what row operations to do to get a $\mathbf{0}$ in the matrix; i.e. first scale R_1 by c_1, R_2 by c_2, etc, and then perform the row sums. $\square$

- We now see that properties of matrices can tell us a lot about properties of systems of equations and sets of vectors. For much of the remainder of the course we will study matrices by themselves; we still have not unlocked the secret of what a matrix is, and won't for a few days. For now though, we will explore the algebra of matrices.

- As before, a $m \times n$ **matrix** is a grid of real numbers with m rows and n columns. All definitions and algorithms carry over, but we not exclusively be using matrices as a way to represent systems. For a matrix A we write a_{ij} to denote the entry in the ith row and the jth column. We say a matrix is **square** if $m = n$.

- There are some matrices of particular interest; a $m \times 1$ matrix is called a **column matrix** or **column vector**; a $1 \times n$ matrix is called a **row matrix** or **row matrix**. The **diagonal entries** are a matrix are the a_{ii} entries. A **diagonal matrix** is a square one where all non-diagonal entries are zero. A **scalar matrix** is a diagonal matrix where all diagonal entries are the same. A **identity matrix** is a scalar matrix where the diagonal is 1.

- We denote *the* identity matrix which is $n \times n$ by I_n, or when the size of the matrix is clear from context we just say I.

- Matrices are equal only when they are equal in all entries and have the same size.

- We define two operations on entries which should mirror our operations on vectors: addition and scalar multiplication.

  We define **matrix addition** between two $m \times n$ matrices A and B by their entires: let C denote $A+B$. Then $c_{ij} = a_{ij} + b_{ij}$; that is, the addition is just entrywise.

  Example

$$\begin{pmatrix} 1 & 3 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 1 & -3 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 1 & 1 \end{pmatrix}$$

- Similarly, we define **scalar multiplication** of A by the entry: let C denote kA. Then $c_{ij} = ka_{ij}$.

  Example

$$3 \begin{pmatrix} 2 & 0 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 6 & 0 \\ 3 & 3 \end{pmatrix}$$

- The **zero matrix** is the matrix with all zeros (so it is a scalar matrix where the diagonal entries are 0, if that helps, but it probably doesn't; the first description was probably good enough). To denote this matrix we write O (which is the letter, not the number), and if we want to emphasize the size we write $O_{m \times n}$.

  Theorem For any $m \times n$ matrix A we have

$$A + O = O + A = A$$

  (Note: I didn't specify the size of O because it's clear from context). □

- We will view the set of matrices algebraically. $O_{m \times n}$ is then the **additive identity** of the set of $m \times n$ matrices. We have **additive closure** in the set of $m \times n$ matrices; this means that if you add two $m \times n$ matrices you get an $m \times n$ matrix. It's also true that **additive inverses** exist for matrices; namely, $A + (-A) = O$. Therefore, the set of $m \times n$ matrices forms a special algebraic structure called a **group**. We won't talk about groups in this course, but if you've seen groups, or will see groups, this will give you another example!

- Addition and scalar multiplication are two important operations on matrices, but the one that really reveals the 'true nature' of a matrix is the operation of **matrix multiplication**. This one is a bit difficult to explain, so we'll go about it careful.

  If A if a $m \times n$ matrix and B is a $n \times r$ matrix (note: the sizes here are very important; read them again) then we the product (under matrix multiplication) of A and B as $C = AB$. We define C by it's entries:

$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}$$

  Remark Does this make sense? No, it really shouldn't. But, knowing the dot product it can be slightly rephrased: what it says is that the ijth entry is the dot product of the ith row of the left matrix and the jth row of the right matrix.

  Example Calculate:

$$\begin{pmatrix} 1 & 3 & 0 \\ 0 & 1 & 1 \end{pmatrix}_{3} \begin{pmatrix} 1 & 0 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

Solution. I really don't know how to 'show' work for this, so I'll do it on the chalkboard. The answer is:

$$\begin{pmatrix} 4 & 6 & -1 \\ 1 & 2 & -1 \end{pmatrix}$$

<u>Remark</u> 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.

  <u>Example</u> 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:

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

<u>Remark</u> 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'