Math 1600B Lecture 15, Section 2, 7 Feb 2014

Announcements:

Read Section 3.3 for next class. Work through recommended homework questions.

Office hour: next Monday, 12:30-1:30, MC103B.

Help Centers: Monday-Friday 2:30-6:30 in MC 106.

Partial review of Lectures 13 and 14:

Matrix multiplication

Definition: If A is $m \times n$ and B is $n \times r$, then the **product** $C = AB$ is the $m \times r$ matrix whose i, j entry is

$$c_{ij} = a_{i1} b_{1j} + a_{i2} b_{2j} + \cdots + a_{in} b_{nj} = \sum_{k=1}^{n} a_{ik} b_{kj}$$

$$= \mathrm{row}_i(A) \cdot \mathrm{col}_j(B).$$

To remember the shape of AB:

$$\begin{array}{ccccc} A & & B & = & AB \\ m \times n & & n \times r & & m \times r \end{array}$$

Note: In particular, if B is a column vector in $\mathbb{R}^n$, then AB is a column vector in $\mathbb{R}^m$. So one thing a matrix A can do is *transform* column vectors into column vectors. This point of view will be important later.

For the most part, matrix multiplication behaves like multiplication of real numbers, but there are several differences:

We can have $A \neq O$ but $A^k = O$ for some $k > 1$.
We can have $B \neq \pm I$, but $B^4 = I$.

We can have $AB \neq BA$.

But most expected properties **do** hold:

Theorem 3.3: Let A, B and C be matrices of the appropriate sizes, and let k be a scalar. Then:

(a) $A(BC) = (AB)C$	(associativity)
(b) $A(B+C) = AB + AC$	(left distributivity)
(c) $(A+B)C = AC + BC$	(right distributivity)
(d) $k(AB) = (kA)B = A(kB)$	(no cool name)
(e) $I_m A = A = AI_n$ if A is $m \times n$	(identity)

New material: Sections 3.1 and 3.2 continued.

Partitioned Matricies

Sometimes it is natural to view a matrix as **partitioned** into **blocks**. For example:

$$A = \begin{bmatrix} 1 & 0 & 0 & 2 & 1 \\ 0 & 1 & 0 & 1 & 3 \\ 0 & 0 & 1 & 4 & 0 \\ 0 & 0 & 0 & 1 & 7 \\ 0 & 0 & 0 & 7 & 2 \end{bmatrix} = \left[\begin{array}{ccc|cc} 1 & 0 & 0 & 2 & 1 \\ 0 & 1 & 0 & 1 & 3 \\ 0 & 0 & 1 & 4 & 0 \\ \hline 0 & 0 & 0 & 1 & 7 \\ 0 & 0 & 0 & 7 & 2 \end{array}\right] = \begin{bmatrix} I & D \\ O & C \end{bmatrix}$$

This can make matrix multiplication much easier when there are blocks that are zero or an identity matrix. For example, if

$$B = \left[\begin{array}{cc} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ \hline 1 & 0 \\ 0 & 1 \end{array}\right] = \begin{bmatrix} O \\ I \end{bmatrix}$$

then

$$AB = \begin{bmatrix} I & D \\ O & C \end{bmatrix} \begin{bmatrix} O \\ I \end{bmatrix} = \begin{bmatrix} IO + DI \\ O^2 + CI \end{bmatrix} = \begin{bmatrix} D \\ C \end{bmatrix} = \left[\begin{array}{cc} 2 & 1 \\ 1 & 3 \\ 4 & 0 \\ \hline 1 & 7 \\ 7 & 2 \end{array}\right]$$

You pretend that the submatrices are numbers and do matrix multiplication. As long as all of the sizes match up, this works. But keep the left/right order straight!

See **Example 3.12** for a larger, more complicated worked example.

The most common (and important) cases are when one or both of the matrices are partitioned into rows or columns. For example, if A is $m \times n$ and B is $n \times r$, and we partition B into its columns as $B = [\,\vec{b}_1 \mid \vec{b}_2 \mid \cdots \mid \vec{b}_r]$, then we have:

$$AB = A[\,\vec{b}_1 \mid \vec{b}_2 \mid \cdots \mid \vec{b}_r] = [\,A\vec{b}_1 \mid A\vec{b}_2 \mid \cdots \mid A\vec{b}_r],$$

where we think of A and the $\vec{b}_i$'s as scalars. The first column of AB consists of the dot products of the rows of A with the first column $\vec{b}_1$ of B.

Example on board: 2×3 times 3×2.

Note that each column of AB is a linear combination of the columns of A.

Similarly, if we partition A into rows, we can compute

$$AB = \left[\begin{array}{c} A_1 \\ \hline A_2 \\ \hline \vdots \\ \hline A_m \end{array}\right] B = \left[\begin{array}{c} A_1 B \\ \hline A_2 B \\ \hline \vdots \\ \hline A_m B \end{array}\right]$$

Same example on board.

If we partition A into rows and B into columns, we get

$$AB = \begin{bmatrix} A_1 \\ \hline A_2 \\ \hline \vdots \\ \hline A_m \end{bmatrix} [\,\vec{b}_1 \mid \vec{b}_2 \mid \cdots \mid \vec{b}_r\,] = \begin{bmatrix} A_1\vec{b}_1 & \cdots & A_1\vec{b}_r \\ \vdots & & \vdots \\ A_m\vec{b}_1 & \cdots & A_m\vec{b}_r \end{bmatrix}$$

which is just the usual description of AB, where the ij entry is the dot product of the ith row of A with the jth column of B!

(Outer products and Example 3.11 not covered.)

The Transpose and Symmetric Matrices

Here's another operation on matrices, which has no analog for real numbers:

Definition: The **transpose** of an $m \times n$ matrix A is the $n \times m$ matrix A^T whose ij entry is the ji entry of A.

Example 3.14: The transposes of

$$A = \begin{bmatrix} 1 & 3 & 2 \\ 5 & 0 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \quad \text{and} \quad C = [\,5 \quad -1 \quad 2\,]$$

are

$$A^T = \begin{bmatrix} 1 & 5 \\ 3 & 0 \\ 2 & 1 \end{bmatrix}, \quad B^T = \begin{bmatrix} a & c \\ b & d \end{bmatrix}, \quad \text{and} \quad C^T = \begin{bmatrix} 5 \\ -1 \\ 2 \end{bmatrix}.$$

Note that the columns and rows get interchanged.

One use of the transpose is to convert between row vectors and column vectors. In particular, we can use this to express the dot product in terms of matrix multiplication. If

$$\vec{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} \qquad \text{and} \qquad \vec{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$$

then

$$\vec{u}^T \vec{v} = [u_1 \; u_2 \; \cdots \; u_n] \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = u_1 v_1 + \cdots + u_n v_n = \vec{u} \cdot \vec{v}$$

Properties of the transpose

Theorem 3.4: Let A and B be matrices of the appropriate sizes, and let k be a scalar. Then:

(a) $(A^T)^T = A$ (b) $(A+B)^T = A^T + B^T$
(c) $(kA)^T = k(A^T)$ (d) $(AB)^T = B^T A^T$!
(e) $(A^r)^T = (A^T)^r$ for all nonnegative integers r

(a), (b) and (c) are easy to see. (d) is more of a surprise, so it is worth explaining:

Proof of (d): Suppose A is $m \times n$ and B is $n \times r$. Then both of $(AB)^T$ and $B^T A^T$ are $r \times m$. We have to check that the entries are equal:

$$\begin{aligned} [(AB)^T]_{ij} &= (AB)_{ji} = \text{row}_j(A) \cdot \text{col}_i(B) = \text{col}_j(A^T) \cdot \text{row}_i(B^T) \\ &= \text{row}_i(B^T) \cdot \text{col}_j(A^T) = [(B^T)(A^T)]_{ij}. \qquad \square \end{aligned}$$

Example on board

Note that (b) and (d) extend to several matrices. For example:

$$(A+B+C)^T = ((A+B)+C)^T = (A+B)^T + C^T = A^T + B^T + C^T$$

and

$$(ABC)^T = ((AB)C)^T = C^T(AB)^T = C^T B^T A^T$$

In particular, (e) follows: $(A^r)^T = (A^T)^r$.

Symmetric matrices

Definition: A square matrix A is **symmetric** if $A^T = A$. That is, $A_{ij} = A_{ji}$ for every i and j.

Example: These matrices are symmetric:

$$\begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix} \quad \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 3 & 5 & 6 \end{bmatrix} \quad \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

Example: These matrices are **not** symmetric:

$$\begin{bmatrix} 2 & 1 \\ 3 & 2 \end{bmatrix} \quad \begin{bmatrix} 1 & 2 & 1 \\ 5 & 4 & 2 \\ 1 & 5 & 1 \end{bmatrix} \quad \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

There are two ways to get a symmetric matrix from a non-symmetric matrix:

1. If A is square, then $A + A^T$ is symmetric. This is because

$$(A + A^T)^T = A^T + (A^T)^T = A^T + A = A + A^T.$$

Example on board.

2. And if B is any matrix, then $B^T B$ is symmetric. This is because

$$(B^T B)^T = B^T (B^T)^T = B^T B$$

The same kind of argument shows that BB^T is symmetric.

Example on board.

True/false: If A is symmetric, so is A^2. On board.

Challenge problems

Question: Find a 3×3 matrix A such that $A^2 \neq O$ but $A^3 = O$.

One solution is

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

Question: Find a 2×2 matrix A such that $A \neq I_2$ but $A^3 = I_2$.

One solution is

$$A = \begin{bmatrix} -1/2 & -\sqrt{3}/2 \\ \sqrt{3}/2 & -1/2 \end{bmatrix}$$

Where did I get this from? It is rotation by 120 degrees! Explain on board.

Similarly, for each n you can find a matrix such that $A^n = I$ but no lower power of A is the identity.