LINEAR ALGEBRA. LECTURE 2

Matrix Multiplication

We discuss four different ways of thinking about the product $AB = C$ of two matrices. If A is an $m\times n$ matrix and B is an $n\times p$ matrix, then C is an $m\times p$ matrix. We use c_{ij} to denote the entry in row i and column j of matrix C.

Standard (row times column)

The standard way of describing a matrix product is to say that c_{ij} equals the dot product of row i of matrix A and column j of matrix B. In other words,

$$c_{ij} = row_i\left(A\right)\cdot column_j\left(B\right) = a_{i1}b_{1j} + a_{i2}b_{2j} + a_{i3}b_{3j} \ldots + a_{in}b_{nj} + = \sum_{k=1}^{n} a_{ik}b_{kj}$$

In Figure 1, the 3, 2 entry of AB comes from row 3 and column 2:

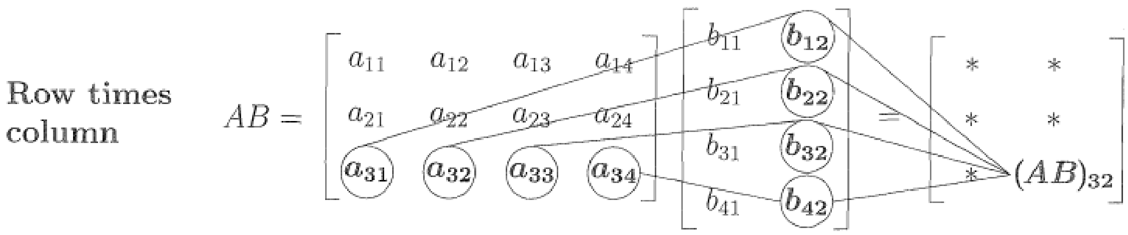

Figure 1: A 3x4 matrix A times a 4x2 matrix B is a 3x2 matrix $C=AB$.

Example 1.

$$C = AB = \begin{bmatrix} 2 & 3 \\ 4 & 0 \end{bmatrix}\begin{bmatrix} 1 & 2 & 0 \\ 5 & -1 & 0 \end{bmatrix} = \begin{bmatrix} 17 & 1 & 0 \\ 4 & 8 & 0 \end{bmatrix}$$

The entry 17 is (2)(1)+(3)(5), the inner product of the first row of A and first column of B. The entry 8 is (4)(2)+(0)(¡1), from the second row and second column. The third column is zero in B, so it is zero in AB. B consists of three columns side by side, and A multiplies each column separately. *Every column of C=AB is a combination of the columns of A.* Just as in a matrix-vector multiplication, the columns of A are multiplied by the entries in B.

Columns

The product of matrix A and column j of matrix B equals column j of matrix C. This tells us that the columns of C are combinations of columns of A.

$$column_j\left(C\right) = A\cdot column_j\left(B\right)$$

Rows

The product of row i of matrix A and matrix B equals row i of matrix C. So the rows of C are combinations of rows of B.

$$row_i\left(C\right) = B\cdot row_i\left(A\right)$$

Column times row

A column of A is an 3×1 vector and a row of B is a 1×2 vector. Their product is a matrix:

$$\begin{bmatrix}2\\3\\4\end{bmatrix}\begin{bmatrix}1 & 6\end{bmatrix}=\begin{bmatrix}2 & 12\\3 & 18\\4 & 24\end{bmatrix}$$

The columns of this matrix are multiples of the column of A and the rows are multiples of the row of B. If we think of the entries in these rows as the coordinates (2, 12) or (3, 18) or (4, 24), all these points lie on the same line; similarly for the two column vectors. Later we'll see that this is equivalent to saying that the *row space* of this matrix is a single line, as is the *column space*.

The product of A $m\times n$ and B $n\times 1$ is the sum of these "column times row" matrices:

$$AB=\sum_{k=1}^{n}\begin{bmatrix}a_{1k}\\ \vdots\\ \vdots\\ a_{mk}\end{bmatrix}\begin{bmatrix}b_{k1} & \cdots & b_{kn}\end{bmatrix}$$

Blocks

If we subdivide A and B into blocks that match properly, we can write the product $AB = C$ in terms of products of the blocks:

$$\begin{bmatrix}A_1 & A_2\\A_3 & A_4\end{bmatrix}\begin{bmatrix}B_1 & B_2\\B_3 & B_4\end{bmatrix}=\begin{bmatrix}C_1 & C_2\\C_3 & C_4\end{bmatrix}$$

Here $C_1=A_1B_1+A_2B_3$, $C_2=A_1B_2+A_2B_4$ and so on.

We summarize these three different ways to look at matrix multiplication.

(I) Each entry of AB is the product of a *row* and a *column*: $(AB)_{ij}$ = (row i of A) times (column j of B).

(II) Each column of AB is the product of a matrix and a column: column j of AB = A times (column j of B).

(III) Each row of AB is the product of a row and a matrix: row i of AB = (row i of A) times B.

This leads back to the key properties of matrix multiplication. Suppose the shapes of three matrices A, B, C (possibly rectangular) permit them to be multiplied. The rows in A and B multiply the columns in B and C. Then the key property is this:

Matrix multiplication is associative: $(AB)C = A(BC)$. Just write ABC

Matrix operations are distributive: $A(B+C) = AB+AC$ and $(B+C)D = BD+CD$

Matrix multiplication is not commutative: Usually $AB \neq BA$.

Inverses

Square matrices

If A is a square matrix, the most important question you can ask about it is whether it has an inverse A^{-1}. If it does, then $A^{-1}A = I = AA^{-1}$ and we say that A is *invertible* or *nonsingular*. If A is *singular* – i.e. A does not have an inverse – its determinant is zero and we can find some non-zero vector $\mathbf{x}$ for which $A\mathbf{x} = 0$. For example:

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

In this example, three times the first column minus one times the second column equals the zero vector; the two column vectors lie on the same line.

Finding the inverse of a matrix is closely related to solving systems of linear equations:

$$\underset{A}{\begin{bmatrix} 1 & 3 \\ 2 & 7 \end{bmatrix}} \underset{A^{-1}}{\begin{bmatrix} a & b \\ c & d \end{bmatrix}} = \underset{I}{\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}}$$

can be read as saying "A times column j of A^{-1} equals column j of the identity matrix". This is just a special form of the equation $Ax = b$.

Gauss-Jordan Elimination

We can use the method of elimination to solve two or more linear equations at the same time. Just augment the matrix with the whole identity matrix I:

$$\left[\begin{array}{cc|cc} 1 & 3 & 1 & 0 \\ 2 & 7 & 0 & 1 \end{array}\right] \rightarrow \left[\begin{array}{cc|cc} 1 & 3 & 1 & 0 \\ 0 & 1 & -2 & 1 \end{array}\right] \rightarrow \left[\begin{array}{cc|cc} 1 & 0 & 7 & -3 \\ 0 & 1 & -2 & 1 \end{array}\right]$$

(Once we have used Gauss' elimination method to convert the original matrix to upper triangular form, we go on to use Jordan's idea of eliminating entries in the upper right portion of the matrix.)

$$A^{-1} = \begin{bmatrix} 7 & -3 \\ -2 & 1 \end{bmatrix}$$

As in the last lecture, we can write the results of the elimination method as the product of a number of elimination matrices E_{ij} with the matrix A. Letting E be the product of all the E_{ij}, we write the result of this Gauss-Jordan elimination using block matrices: $E[\ A \mid I\ ] = [\ I \mid E\ ]$. But if $EA = I$, then $E = A^{-1}$

Factorization into *A* = *LU*

One goal of today's lecture is to understand Gaussian elimination in terms of matrices; to find a matrix L such that $A = LU$. We start with some useful facts about matrix multiplication.

Inverse of a product

The inverse of a matrix product AB is $B^{-1}A^{-1}$:

$$AB = C \quad \Rightarrow \quad (AB)^{-1} = C^{-1}$$

$$A^{-1}AB = A^{-1}C \quad \Rightarrow \quad B^{-1}A^{-1}AB = B^{-1}A^{-1}C = I \quad \Rightarrow \quad C^{-1} = (AB)^{-1} = B^{-1}A^{-1}$$

Transpose of a product

We obtain the *transpose* of a matrix by exchanging its rows and columns. In other words, the entry in row i column j of A is the entry in row j column i of A^T.

The transpose of a matrix product AB is B^TA^T. For any invertible matrix A, the inverse of A^T is $(A^{-1})^T$.

A* = *LU

We've seen how to use elimination to convert a suitable matrix A into an upper triangular matrix U. This leads to the factorization $A = LU$, which is very helpful in understanding the matrix A. Recall that (when there are no row exchanges) we can describe the elimination of the entries of matrix A in terms of multiplication by a succession of elimination matrices E_{ij}, so that $A \to E_{21}A \to E_{31}E_{21}A \to E_{32}E_{31}E_{21}A \to \ldots \to U$. In the two by two case this looks like:

$$\begin{array}{ccccc} E_{21} & A & & U \\ \begin{bmatrix} 1 & 0 \\ -4 & 1 \end{bmatrix} & \begin{bmatrix} 2 & 1 \\ 8 & 7 \end{bmatrix} & = & \begin{bmatrix} 2 & 1 \\ 0 & 3 \end{bmatrix} \end{array}$$

We can convert this to a factorization $A = LU$ by "canceling" the matrix E_{21}; multiply by its inverse to get $E_{21}^{-1}E_{21}A = E_{21}^{-1}U$.

$$\begin{array}{cccc} A & & L & U \\ \begin{bmatrix} 2 & 1 \\ 8 & 7 \end{bmatrix} & = & \begin{bmatrix} 1 & 0 \\ 4 & 1 \end{bmatrix} & \begin{bmatrix} 2 & 1 \\ 0 & 3 \end{bmatrix} \end{array}$$

The matrix U is upper triangular with pivots on the diagonal. The matrix L is *lower triangular* and has ones on the diagonal. Sometimes we will also want to factor out a diagonal matrix whose entries are the factors:

$$\begin{array}{ccccc} A & & L & D & U \\ \begin{bmatrix} 2 & 1 \\ 8 & 7 \end{bmatrix} & = & \begin{bmatrix} 1 & 0 \\ 4 & 1 \end{bmatrix} & \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix} & \begin{bmatrix} 1 & 1/2 \\ 0 & 1 \end{bmatrix} \end{array}$$

In the three dimensional case, if $E_{32}E_{31}E_{21}A = U$ then $A = E_{21}^{-1}E_{31}^{-1}E_{32}^{-1}U = LU$. For example, suppose E_{31} is the identity matrix and E_{32} and E_{21} are as shown below:

$$\begin{array}{ccccc} E_{32} & & E_{21} & & E \\ \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -5 & 1 \end{bmatrix} & & \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} & = & \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 10 & -5 & 1 \end{bmatrix} \end{array}$$

The 10 in the lower left corner arises because we subtracted twice the first row from the second row, then subtracted five times the new second row from the third.

The factorization $A = LU$ is preferable to the statement $EA = U$ because the combination of row subtractions does not have the effect on L that it did on E. Here $L = E^{-1} = E_{21}^{-1}E_{32}^{-1}$:

$$\begin{array}{ccccc} E_{21}^{-1} & & E_{32}^{-1} & & L \\ \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} & & \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 5 & 1 \end{bmatrix} & = & \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 5 & 1 \end{bmatrix} \end{array}$$

Notice the *0* in row three column one of $L = E^{-1}$, where E had a 10. If there are no row exchanges, the multipliers from the elimination matrices are copied directly into L.

How expensive is elimination?

Some applications require inverting very large matrices. This is done using a computer, of course. How hard will the computer have to work? How long will it take?

When using elimination to find the factorization $A = LU$ we just saw that we can build L as we go by keeping track of row subtractions. We have to remember L and (the matrix which will become) U; we don't have to store A or E_{ij} in the computer's memory.

How many operations does the computer perform during the elimination process for an nxn matrix? A typical operation is to multiply one row and then subtract it from another, which requires on the order of n operations. There are n rows, so the total number of operations used in eliminating entries in the first column is about n^2. The second row and column are shorter; that product costs about $(n\text{-}1)^2$ operations, and so on. The total number of operations needed to factor A into LU is on the order of n^3:

$$1^2 + 2^2 + ... + (n-1)^2 + n^2 \approx \int_0^n x^2 dx = \frac{1}{3} n^3$$

While we're factoring A we're also operating on b. Going forward, we subtract multiples of b_l from the lower components b_2 , . .. , b_n. This is n - 1 steps. The second stage takes only n - 2 steps, because b_i is not involved. The last stage of forward elimination takes one step. Now start back substitution. Computing x_n uses one step (divide by the last pivot). The next unknown uses two steps. When we reach x_1 it will require n steps *(n*-1 substitutions of the other unknowns, then division by the first pivot). The total count on the right side, from b to c to x-*forward to the bottom and back to the top-is* exactly n^2 *:*

$$\left[(n-1)+(n-2)+...+1\right]+\left[1+2+...+(n-1)+n\right]=n^2$$

To see that sum, pair off *(n* - 1) with 1 and *(n* - 2) with 2. The pairings leave n terms, each equal to n. That makes n^2 .That costs about n^2 operations, which is hardly worth counting compared to $n^3/3$.