Lecture 8: Gaussian Elimination and Matrix Multiplication

February 7, 2024

Rows and Columns in a matrix product

For a $m \times p$ matrix A (m rows, p columns) and a $p \times n$ matrix B (p rows, n columns), the product AB is a $m \times n$ matrix and it is defined by any of the following equivalent (but separately useful) descriptions.

- For j from 1 to n, Column j of AB is the linear combination of the columns of A whose coefficients are the entries of Column j of B.
- For i from 1 to m, Row i of AB is the linear combination of the rows of B whose coefficients are the entries of Row i of A.
- For any position (i, j) in AB, the entry $(AB)_{ij}$ is the scalar product of the vectors given by Row i of A and Column j of B.

Columns of AB are linear combinations of columns of A.
Rows of AB are linear combinations of rows of B.

Note The description in terms of rows of B is one that we haven't seen until now.

The $n \times n$ identity matrix

For a positive integer n, the $n \times n$ identity matrix, denoted I_n, is the $n \times n$ matrix whose entries in the $(1,1), (2,2), \dots, (n,n)$ positions (the positions on the main diagonal) are all 1, and whose entries in all other positions (all off-diagonal positions) are 0. For example

$$I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad I_4 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

What is special about I_n? It is a identity element or neutral element for matrix multplication. Multiplying another matrix by it has no effect.

- If A is any matrix with n rows, then $I_n A = A$, and
- If B is any matrix with n columns, then $BI_n = B$.
- In particular, if C is a $n \times n$ matrix, then $CI_n = I_n C = C$.

Exercise Confirm these properties using the definitions on the last slide.

The Inverse of a Matrix

Let A be a square matrix of size $n \times n$. If there exists a $n \times n$ matrix B for which $AB = I_n$ and $BA = I_n$, then A and B are called inverses (or multiplicative inverses) of each other. If it does not already have another name, the inverse of A is denoted A^{-1}.

Example The matrices $\begin{pmatrix} 3 & 2 \\ -5 & -4 \end{pmatrix}$ and $\begin{pmatrix} 2 & 1 \\ -\frac{5}{2} & -\frac{3}{2} \end{pmatrix}$ are inverses of each other.

Not every square matrix has an inverse. For example the 2×2 matrix $\begin{pmatrix} 3 & 2 \\ -6 & -4 \end{pmatrix}$ does not.

Exercise Prove that a square matrix can only have one inverse.

Exercise If $AB = I_n$ for square matrices A and B, prove that $BA = I_n$.

Note For these exercises, assume (or show!) that $(PQ)R = P(RQ)$ for any matrices P, Q, R for which these products exist (we will prove this later - matrix multiplication is associative).

How to calculate the inverse of a matrix

Given a square matrix A, we want to find B so that $AB = I_n$. This means

- Column 1 of B is a solution to the linear system $Ax = e_1$, where x is a vector whose entries are variables $(x_1, \dots, x_n)$ and e_1 is the column vector with entries $1, 0, \dots, 0$.
- Column 2 of B is a solution to the linear system $Ax = e_2$ (where the column vector e_2 has 1 in its second position and zeros elsewhere).
- ... and so on for Column 3 to Column n

So A has an inverse if and only if each of $e_1, \dots, e_n$ is a linear combination of the columns[1] of A.

We can solve each of the n systems above using Gaussian elimination on its augmented matrix. Or we can solve all of them together by applying Gaussian elimination to the $n \times 2n$ matrix $[A|I_n]$.

[1] If this happens, then the "row versions" of $e_1, \dots, e_n$ are all linear combinations of the rows of A too. Why?

Gaussian elimination for inverse calculation

Example $A = \begin{bmatrix} 1 & -1 & 1 & 4 \\ 1 & 0 & 2 & 2 \\ 3 & -3 & 4 & 8 \\ 0 & -2 & -2 & 5 \end{bmatrix}$. Find A^{-1}.

$$\left[\begin{array}{cccc|cccc} 1 & -1 & 1 & 4 & 1 & 0 & 0 & 0 \\ 1 & 0 & 2 & 2 & 0 & 1 & 0 & 0 \\ 3 & -3 & 4 & 8 & 0 & 0 & 1 & 0 \\ 0 & -2 & -2 & 5 & 0 & 0 & 0 & 1 \end{array}\right] \rightarrow \left[\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 26 & -19 & -2 & -10 \\ 0 & 1 & 0 & 0 & 6 & -3 & -1 & -2 \\ 0 & 0 & 1 & 0 & -11 & 8 & 1 & 4 \\ 0 & 0 & 0 & 1 & -2 & 2 & 0 & 1 \end{array}\right]$$

Conclusion $A^{-1} = \begin{bmatrix} 26 & -19 & -2 & -10 \\ 6 & -3 & -1 & -2 \\ -11 & 8 & 1 & 4 \\ -2 & 2 & 0 & 1 \end{bmatrix}$. Check this!

If I_n is written in the first n columns of the RREF of $[A|I_n]$, the last n columns comprise A^{-1}.

Why did that work?

1. We were solving four linear systems simultaneously. All four had the same coefficient matrix A, their right hand sides were respectively ae_1, e_2, e_3, e_4.
2. The four leading 1's in the RREF mean that in each system (and any other with coefficient matrix A), there is a unique solution. Those unique solutions are respectively written in the last four columns. So Column 5 of the RREF is the unique column vector v with $Av = e_1$. This is the first column of A^{-1}.
3. If A didn't have an inverse, what would have happened?
   If A has no inverse, at least one of the four systems is inconsistent. In the row reduction, we encounter a row with 0 in the first four positions, but not in all the last four.