Lecture 11

Matrices

Although the mathematical functions describing natural phenomena can be very complicated, one can say that locally, everything behaves linearly. Think of the tangent of a curve at a point: it is a *line* which approximates the function locally around a point. This is one reason why *linear algebra* is so useful.

Some of the main players in linear algebra are matrices. Matrices help to streamline the solution to system of linear equations:

Example 11.1. Solve the system

$$\begin{cases} x + 2y & = & 1 & \quad (1) \\ -2x - 3y & = & 2 & \quad (2). \end{cases}$$

To solve this we transform the equations: First, equation (2) is transformed into "equation (1) added twice to equation (2)", that is

$$(2) \longrightarrow 2 \cdot (1) + (2).$$

This gives the new system

$$\begin{cases} x + 2y & = & 1 & \quad (1) \\ y & = & 4 & \quad (2). \end{cases}$$

The point here was to cancel all xs in (2). Now cancel y in (1) via

$$(1) \longrightarrow (1) - 2 \cdot (2)$$

to obtain

$$\begin{cases} x & = & -7 & \quad (1) \\ y & = & 4 & \quad (2). \end{cases}$$

We have now obtained the solution.

As the above example demonstrates, what matters is not the variables x and y themselves, but only the coefficients of the equations, that is the tables of numbers

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

which encode the system of equations. These tables of numbers are called *matrices*. In general, a matrix is a rectangular array of numbers, called its *entries*. A matrix is said to be an $m \times n$ matrix if it has m rows and n columns. For example,

$$\begin{pmatrix} 3 & \frac{1}{2} \\ -1 & 0 \\ 57 & \pi \end{pmatrix}$$

is a 3×2 matrix.

Here are some important special cases of matrices:

- $n = 1$: A $m \times 1$ matrix $\begin{pmatrix} \vdots \\ \vdots \end{pmatrix}$ is a *column vector* (or just vector).
- $m = 1$: A $1 \times m$ matrix $\begin{pmatrix} \cdots\cdots \end{pmatrix}$ is a *row vector*.
- $n = m$: An $n \times n$ matrix is a *square matrix*.

We can write a matrix

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix}$$

where the a_{ij} are the entries, for $i = 1, 2, 3$ and $j = 1, 2$. This is a 3×2 matrix, but we can extend this notation to any $m \times n$ matrix:

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ & \vdots & & \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}.$$

We can write this compactly as (a_{ij}), remembering that $i = 1, \ldots, m$ and $j = 1, \ldots, n$.

The entries a_{ij} can be real of complex numbers. Assume first that $a_{ij} \in \mathbb{R}$. We then let

$$\mathrm{Mat}_{m \times n}(\mathbb{R})$$

be the set of all $m \times n$ matrices with real entries, and

$$\mathrm{Mat}_n(\mathbb{R})$$

be the set of $n \times n$ square matrices.

Matrix multiplication

There is a way to multiply two matrices, which is a bit unusual at first, but turns out to be very useful.

Example 11.2. Let A and B be the matrices

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}, \qquad B = \begin{pmatrix} 5 \\ 6 \end{pmatrix}.$$

Then the product is

$$AB = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} 5 \\ 6 \end{pmatrix} = \begin{pmatrix} 1 \cdot 5 + 2 \cdot 6 \\ 3 \cdot 5 + 4 \cdot 6 \end{pmatrix} = \begin{pmatrix} 17 \\ 39 \end{pmatrix}.$$

Note that the product of the square matrix A with the vector matrix B is another vector matrix.

In general, a product of $m \times n$ matrix A and an $n \times l$ matrix B is $m \times l$ matrix C, and

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

The reason why it is useful to define matrix multiplication like this is that we can write the system of equations in (11.1) as

$$\begin{pmatrix} 1 & 2 \\ -2 & -3 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x + 2y \\ -2x - 3y \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}.$$

We have therefore replaced a system with *two* equations by a *single* matrix equation. This is helpful if we had a system of 1000 equations, especially if we want to solve it using a computer (which, surely, we want).

We can also multiply two square matrices:

Example 11.3. Let

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}, \qquad B = \begin{pmatrix} 0 & 2 \\ 0 & 0 \end{pmatrix}.$$

Then

$$AB = \begin{pmatrix} 1 \cdot 0 + 2 \cdot 0 & 1 \cdot 2 + 2 \cdot 0 \\ 3 \cdot 0 + 4 \cdot 0 & 3 \cdot 2 + 4 \cdot 0 \end{pmatrix} = \begin{pmatrix} 0 & 2 \\ 0 & 6 \end{pmatrix}$$

and

$$BA = \begin{pmatrix} 0 & 2 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 0 \cdot 1 + 2 \cdot 3 & 0 \cdot 2 + 2 \cdot 4 \\ 0 \cdot 1 + 0 \cdot 3 & 0 \cdot 2 + 0 \cdot 4 \end{pmatrix} = \begin{pmatrix} 6 & 8 \\ 0 & 0 \end{pmatrix}.$$

So we see that AB is *not* equal to BA! Moreover,

$$BB = B^2 = \begin{pmatrix} 0 & 2 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 2 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix},$$

which is called the zero matrix $0_{2\times 2}$ (or simply 0). We see that it can happen that the square of a non-zero matrix is zero!

We can multiply a 2×2 matrix with a 2×1 one, but not with a 3×1 or bigger vector matrix. In general, we can multiply an $m \times n$ matrix by an $n \times k$ one. For example, a 2×3 one by a 3×2 one:

$$\begin{pmatrix} 1 & -4 & 3 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 3 & 2 \\ -1 & 0 \\ 0 & 5 \end{pmatrix} = \begin{pmatrix} 1 \cdot 3 + (-4)(-1) + 3 \cdot 0 & 1 \cdot 2 + (-4) \cdot 0 + 3 \cdot 5 \\ 0 \cdot 3 + 0 \cdot (-1) + 1 \cdot 0 & 0 \cdot 2 + 0 \cdot 0 + 1 \cdot 5 \end{pmatrix} = \begin{pmatrix} 7 & 17 \\ 0 & 5 \end{pmatrix}.$$

Similarly, we have

$$\begin{pmatrix} 1 & 2 \end{pmatrix} \begin{pmatrix} 3 \\ 4 \end{pmatrix} = (1 \cdot 3 + 2 \cdot 4) = (11),$$

and

$$\begin{pmatrix} 3 \\ 4 \end{pmatrix} \begin{pmatrix} 1 & 2 \end{pmatrix} = \begin{pmatrix} 3 \cdot 1 & 3 \cdot 2 \\ 4 \cdot 1 & 4 \cdot 2 \end{pmatrix} = \begin{pmatrix} 3 & 6 \\ 4 & 8 \end{pmatrix}.$$

Summary

- If $A \in \text{Mat}_{m\times n}(\mathbb{R})$ and $B \in \text{Mat}_{n\times k}(\mathbb{R})$, we can multiply A and B and $AB \in \text{Mat}_{m\times k}(\mathbb{R})$.
- We can only multiply two matrices A and B if A has the same number of columns as B has rows.
- We can have $AB \neq BA$. If $AB = AB$ (which happens sometimes) the matrices A and B are said to *commute*.
- We can have $AB = 0$, even though $A \neq 0$ and $B \neq 0$.

Other operations on matrices

- Addition of matrices is easy. Just add element-wise: $(A+B)_{ij} = a_{ij} + b_{ij}$. For example,

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} + \begin{pmatrix} 0 & 2 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1+0 & 2+2 \\ 3+0 & 4+0 \end{pmatrix} = \begin{pmatrix} 1 & 4 \\ 3 & 4 \end{pmatrix}.$$

  Since $a+b = b+a$ for any real numbers a, b, it is clear that $A+B = B+A$, for two matrices A, B.
  Note that we can only add two matrices if they are of the *same* size.

- If $\lambda \in \mathbb{R}$ is a scalar and $A = (a_{ij}) \in \text{Mat}_{m\times n}(\mathbb{R})$, then $\lambda A = (\lambda a_{ij})$. For example, if $\lambda = 2$ and $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$, we have

$$2 \cdot \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 2 & 4 \\ 6 & 8 \end{pmatrix}.$$

- If A is a matrix, we can turn its rows into columns (and columns into rows; same thing). The result is called the *transpose*: A^T of A, for example:

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}, \qquad A^T = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}.$$

- If A is an $m \times n$ matrix, then $-A$ is the matrix $0_{m\times n} - A$, where $0_{m\times n}$ is the zero matrix of that size. In other words, to get $-A$ just change sign on each of the entries of A.