LECTURE 3

Matrices

1. Matrices and Linear Systems

It is a familar task to solve simultaneous systems of linear equations. For example, to solve

$$x - 2y = 0 \tag{3.1}$$

$$x + y = 3 \tag{3.2}$$

We might add the first equation to 2 tiimes the second equation we obtain

$$3x + 0 = 6$$

from which we obtain $x = 2$; and then by substituting this number for x into either of the first pair of equations yields $y = 1$. One of the reasons for introducing the notion of matrices is to organize and simplify calculations of this sort.

Before introducing matrices, let me first cast the example above into a more general context.

DEFINITION 3.1. *Let $x_1, \dots, x_n$ be a set of n variables. An* **m by n *linear system*** *is a set of m linear equations in n unknowns; that is to say, a set of equations of the form*

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m \end{aligned} \tag{3.3}$$

The numbers a_{ij} are referred to as the coefficients; more explicitly the number a_{ij} is coefficient of the variable x_j in the i^{th} equation.

Specification of a linear system is thus made by stating a set of $m \times n$ coefficients $\{a_{ij}\}$ and a set of m numbers $\{b_i\}$. This we do as follows:

DEFINITION 3.2. *An* **m *by* n *matrix*** *is an arrangement of mn numbers a_{ij}, into an ordered rectangular array with m rows and n columns, such as*

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

A **m-*dimensional column vector*** *is a m by 1 matrix, such as*

$$\mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

A **n-*dimensional row vector*** *is a 1 by n matrix, such as*

$$\mathbf{c} = \begin{bmatrix} c_1 & c_2 & \cdots & c_n \end{bmatrix}$$

In terms of this new language, the specification of a m by n linear system is equivalent to the specification of an m by n matrix and an m by n matrix $\mathbf{A}$ and an m-dimensional column vector $\mathbf{b}$. The justification for this new terminology is not just the fact that it allows us to describe a set of linear equations in a generic fashion - but also because it leads to a formal point of view that distills from the presentation (3.3) of a linear system the fundamental components of a linear system:

- an n-dimensional column vector $\mathbf{x}$ of unknowns
$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$
- an m by n matrix of coefficients
$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ldots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$
- an m-dimensional column vector $\mathbf{b}$ of values
$$\mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

This notation allows us to write (formally at this point), the linear system (3.3) as

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ldots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} \tag{3.4}$$

or, more succinctly

$$\mathbf{A}\mathbf{x} = \mathbf{b} \tag{3.5}$$

EXAMPLE 3.3. Find the matrix $\mathbf{A}$ and the column vector $\mathbf{b}$ corresponding to the following linear system

$$\begin{aligned} x_1 + 2x_3 + x_4 &= 1 \\ 2x_1 - x_2 &= 3 \\ x_2 + x_3 - x_4 &= -2 \end{aligned} \tag{3.6}$$

- To help you see the appropriate matrix $\mathbf{A}$ let me write this system of equations in a fashion that all the variables appear in each equation
$$\begin{aligned} (1)x_1 + (0)x_2 + (2)x_3 + (1)x_4 &= 1 \\ (2)x_1 + (-1)x_2 + (0)x_3 + (0)x_4 &= 3 \\ (0)x_1 + (1)x_2) + (1)x_3 + (-1)x_4 &= -2 \end{aligned}$$
The matrix $\mathbf{A}$ is formed from the coefficients of the variables in the corresponding linear system; more precisely, the matrix entry in the i^{th} row of the j^{th} column is the coefficient of the the j^{th} variable in the i^{th} equation. Thus
$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 2 & 1 \\ 2 & -1 & 0 & 0 \\ 0 & 1 & 1 & -1 \end{bmatrix}$$

The components of the column vector $\mathbf{b}$ are the values on the right hand side of the corresponding linear system; more precisely, the i^{th} component of $\mathbf{b}$ is the value on the right hand side of the i^{th} equation. Thus,

$$\mathbf{b} = \begin{bmatrix} 1 \\ 3 \\ -2 \end{bmatrix}$$

2. Matrix Multiplication

In fact, the expression on the left hand side of (3.4) is more than just a formal way of representing the left hand side of (3.3); the two are actually identical. This identification comes about via the following definition.

DEFINITION 3.4. *Let $\mathbf{a}$ be an n-dimensional row vector and let $\mathbf{b}$ be an n-dimensional column vector. Then the **matrix product** $\mathbf{ab}$ is the dot product of the vector $\mathbf{a}$ and the vector $\mathbf{b}$:*

$$\mathbf{ab} = \begin{bmatrix} a_1 & a_2 & \cdots & a_n \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} = a_1b_1 + a_2b_2 + \cdots + a_nb_n \tag{3.7}$$

DEFINITION 3.5. *Let $\mathbf{A}$ be an m by n matrix and let $\mathbf{x}$ be an n-dimensional column vector. Then the **matrix product** $\mathbf{Ax}$ is the m-dimensional column vector with components*

$$\mathbf{Ax} = \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \\ a_{mn}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{bmatrix} \tag{3.8}$$

In other words, the i^{th} entry in the column vector corresponding to the product $\mathbf{Ax}$ is the dot product of the i^{th} row of $\mathbf{A}$ (thought of as an n-dimensional vector) and the column vector $\mathbf{x}$ (thought of as an n-dimensional vector).

EXAMPLE 3.6. Compute the matrix product $\mathbf{Ax}$ where

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 2 & 1 \\ 2 & -1 & 0 & 0 \\ 0 & 1 & 1 & -1 \end{bmatrix} \quad , \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}$$

- We have

$$\mathbf{Ax} = \begin{bmatrix} 1 & 0 & 2 & 1 \\ 2 & -1 & 0 & 0 \\ 0 & 1 & 1 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} (1,0,2,1)\cdot(x_1,x_2,x_3,x_4) \\ (2,-1,0,0)\cdot(x_1,x_2,x_3,x_4) \\ (0,1,1,-1)\cdot(x_1,x_2,x_3,x_4) \end{bmatrix} = \begin{bmatrix} x_1 + 2x_3 + x_4 \\ 2x_1 - x_2 \\ x_2 + x_3 - x_4 \end{bmatrix}$$

Note that the matrix equation

$$\begin{bmatrix} 1 & 0 & 2 & 1 \\ 2 & -1 & 0 & 0 \\ 0 & 1 & 1 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \\ -2 \end{bmatrix}$$

is thus identical to the following linear system

$$\begin{aligned} x_1 + 2x_3 + x_4 &= 1 \\ 2x_2 - x_2 &= 3 \\ x_2 + x_3 - x_4 &= -2 \end{aligned}$$

Thus far we have defined matrix products $\mathbf{AB}$ only for two special cases:

(1) When the first matrix factor $\mathbf{A}$ is an n-dimensional row vector (i.e., a 1 by n matrix) and the second matrix factor is an n-dimensional column vector (i.e., an n by 1 matrix). The result of such a matrix product is always a real number corresponding to the dot product of the corresponding n-dimensional vectors.
(2) When the first matrix factor $\mathbf{A}$ is an m by n matrix and the second matrix factor is an n-dimensional column vector (i.e., an n by 1 matrix), the result of such a product is an m-dimensional column vector (i.e., an m by 1 matrix).

Note that these matrix multiplication always pair the components of the row of the first factor with the entries of columns of the second factor. I'll now give the general definition of the matrix product (which subsumes these two special cases).

Definition 3.7. *Let $\mathbf{A}$ be an m by n matrix and let $\mathbf{B}$ be an s by t matrix. If $n \neq s$ the matrix product $\mathbf{AB}$ is not defined (i.e. if the number of columns of $\mathbf{A}$ does not equal the numbe of rows of $\mathbf{B}$, the matrix product is not defined). If $n = s$, then the matrix product $\mathbf{AB}$ is defined and is the m by t matrix whose entries $(\mathbf{AB})_{ij}$ are prescribed by*

$$\begin{aligned} (\mathbf{AB})_{ij} &= a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj} \\ &= \sum_{k=1}^{n} a_{ik}b_{kj} \end{aligned}$$

In other words, the entry in j^{th} column of the i^{th} row of the product matrix $\mathbf{AB}$ is the dot product the vector correspondind to the i^{th} row of $\mathbf{A}$ and the vector corresponding to the j^{th} column of $\mathbf{B}$.

3. Examples of Matrix Multiplication

Recall from the preceding lecture our definition of matrix multiplication.

Definition 3.8. *Let $\mathbf{A}$ be an m by n matrix and let $\mathbf{B}$ be an s by t matrix. If $n \neq s$ the matrix product $\mathbf{AB}$ is not defined (i.e. if the number of columns of $\mathbf{A}$ does not equal the numbe of rows of $\mathbf{B}$, the matrix product is not defined). If $n = s$, then the matrix product $\mathbf{AB}$ is defined and is the m by t matrix whose entries $(\mathbf{AB})_{ij}$ are prescribed by*

$$\begin{aligned} (\mathbf{AB})_{ij} &= a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj} \\ &= \sum_{k=1}^{n} a_{ik}b_{kj} \end{aligned}$$

In other words, the entry in j^{th} column of the i^{th} row of the product matrix $\mathbf{AB}$ is the dot product of the vector corresponding to the i^{th} row of $\mathbf{A}$ and the vector corresponding to the j^{th} column of $\mathbf{B}$.

Let's now compute some illustrative examples

Example 3.9.

$$\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 2 & -1 \\ 1 & -2 \end{bmatrix} \qquad \text{does not exist}$$

Because we need the same number of columns in the first factor as there are rows in the second factor.

Example 3.10.

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

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

So even though the 2 by 1 matrix $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$ and the 1 by 2 matrix $\begin{bmatrix} 1 & -1 \end{bmatrix}$ correspond to the same 2-dimensional vector $(1, -1)$, their products with the 2 by 2 matrix $\begin{bmatrix} 1 & 2 \\ -1 & 2 \end{bmatrix}$ are not the same.

Example 3.11.

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

$$\begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 0 \\ -4 & 1 \end{bmatrix}$$

So the product $\mathbf{AB}$ of two matrices $\mathbf{A}$ and $\mathbf{B}$ is not necessarily the same as the product $\mathbf{BA}$. In other words, matrix multiplication is not commutative in general. Indeed, it can happen that $\mathbf{AB}$ exists but $\mathbf{BA}$ is not even defined.

Note that this circumstance partially explains the paradox of the first example. Let $\mathbf{A}$ denote the 2 by 2 matrix $\begin{bmatrix} 1 & 2 \\ -1 & 2 \end{bmatrix}$. If we interpret the vector $(1, -1)$ as a 2 by 1 matrix $\mathbf{v}$, then only the product $\mathbf{Av}$ is defined; and if we interpret the vector $(1, -1)$ as a 1 by 2 matrix then only the product $\mathbf{vA}$ is defined

Example 3.12.

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

Recall that for real numbers $x^2 = 0$ implies $x = 0$. This is evidently not the case for matrices: it can happen that $\mathbf{A}^2 = \mathbf{0}$ but $\mathbf{A}$ is not equal to the zero matrix $\mathbf{0}$.

Example 3.13.

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

Recall that for real numbers $xy = 0$ implies either $x = 0$ or $y = 0$. This is evidently not the case for matrices: it can happen that $\mathbf{AB} = \mathbf{0}$ but neither $\mathbf{A}$ or $\mathbf{B}$ is equal to the zero matrix $\mathbf{0}$.

Example 3.14.

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

And so muliplying any 3 by 3 matrix $\mathbf{A}$ by the matrix

$$\mathbf{I} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

just replicates the matrix $\mathbf{A}$:

$$\mathbf{AI} = \mathbf{IA} = \mathbf{A}$$

3.1. The Identity Matrix. The preceding example generalizes to arbitrary n by n matrices (i.e. "square matrices"). This motivates the following definition.

DEFINITION 3.15. *Let* $\mathbf{I}$ *be the* n *by* n *matrix whose entries are given by*

$$I_{ij} = \begin{cases} 1 & \textit{if} \;\; i = j \\ 0 & \textit{if} \;\; i \neq j \end{cases}$$

In other words, $\mathbf{I}$ *is an* n *by* n *matrix with* 1*'s along the diagonal (running from the upper left to the lower right) and* 0*'s everywhere else. We call such a matrix the* n *by* n ***identity matrix****. It has the property that* $\mathbf{IA} = \mathbf{AI} = \mathbf{A}$ *for all* n *by* n *matrices* $\mathbf{A}$ *except the* $\mathbf{0}$ *matrix.*