Lecture 3a

Matrix Addition and Scalar Multiplication

(pages 115-118)

Mathematics is a strange world. You never know what's going to turn out to be useful. But it happens that this funny thing called a "matrix" has a great deal of value beyond just being an abbreviation for a system of linear equations. So much so, in fact, that we will spend almost all our remaining time in this course studying matrices! But the journey of a thousand miles (or, as it happens, 7 weeks), begins with a bunch of definitions:

Definition: A **matrix** is a rectangular array of numbers. (For this class, "numbers" always refers to "real numbers".) We say that A is an $m \times n$ matrix when A has m rows and n columns, such as

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1j} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2j} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{ij} & \cdots & a_{in} \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mj} & \cdots & a_{mn} \end{bmatrix}$$

Definition: Two matrices A and B are **equal** if and only if they have the same size (that is, the same number of rows and the same number of columns) and their corresponding entries are equal. That is, if $a_{ij} = b_{ij}$ for all $1 \leq i \leq m$ and $1 \leq j \leq n$.

Notation: We sometimes denote the ij-th entry of a matrix A by $(A)_{ij}$. This is taken to be the same thing as a_{ij}.

Example: $\begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & -3 \\ 0 & 0 & 0 \end{bmatrix} \neq \begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & -3 \end{bmatrix}$, as these matrices are not the same size.

We see that matrix equality is defined "entrywise", quite similar to the way that we defined vector equality. This is not a coincidence (although, lets face it, how else would you define "equal") and just as we did with vectors, we will now define matrix addition and scalar multiplication entrywise:

Definition: Let A and B be $m \times n$ matrices. We define **addition** of matrices by

$$(A + B)_{ij} = (A)_{ij} + (B)_{ij}$$

That is, the ij-th entry of $A + B$ is the sum of the ij-th entry of A with the ij-th entry of B.

Example: $\begin{bmatrix} 1 & -1 \\ 2 & 8 \\ -10 & -5 \end{bmatrix} + \begin{bmatrix} 2 & 12 \\ 3 & -9 \\ 7 & -6 \end{bmatrix} = \begin{bmatrix} (1+2) & (-1+12) \\ (2+3) & (8-9) \\ (-10+7) & (-5-6) \end{bmatrix} = \begin{bmatrix} 3 & 11 \\ 5 & -1 \\ -3 & -11 \end{bmatrix}$

Definition: Let A be an $m \times n$ matrix, and $t \in \mathbb{R}$ a scalar. We define the **scalar multiplication** of matrices by

$$(tA)_{ij} = t(A)_{ij}$$

That is, the ij-th entry of tA is t times the ij-th entry of A.

Example: $2\begin{bmatrix} 1 & -1 \\ 2 & 8 \\ -10 & -5 \end{bmatrix} = \begin{bmatrix} 2(1) & 2(-1) \\ 2(2) & 2(8) \\ 2(-10) & 2(-5) \end{bmatrix} = \begin{bmatrix} 2 & -2 \\ 4 & 16 \\ -20 & -10 \end{bmatrix}$

Continuing to retrace the steps we took in Chapter 1, we will write $-A$ for the scalar product $(-1)A$ and $A - B$ for $A + (-1)B$. We also define a zero matrix as follows:

Notation: We write $O_{m,n}$ for the $m \times n$ matrix whose entries are all zero. Unlike the zero vector, it can be helpful to keep track of the size of the zero matrix we are using.

And just as with vectors, we get the EXACT SAME theorem of useful properties.

Theorem 3.1.1 Let A, B, and C be $m \times n$ matrices and let s and t be real scalars. Then

(1) $A + B$ is an $m \times n$ matrix (closed under addition)
(2) $A + B = B + A$ (addition is commutative)
(3) $(A + B) + C = A + (B + C)$ (addition is associative)
(4) There exists a matrix, denoted by $O_{m,n}$, such that $A + O_{m,n} = A$ (zero matrix)
(5) For each matrix A, there exists an $m \times n$ matrix $(-A)$, with the property that $A + (-A) = O_{m,n}$ (additive inverses)
(6) sA is an $m \times n$ matrix (closed under scalar multiplication)
(7) $s(tA) = (st)A$ (scalar multiplication is associative)
(8) $(s + t)A = sA + tA$ (distributive law)
(9) $s(A + B) = sA + sB$ (distributive law)
(10) $1A = A$ (scalar multiplicative identity)

As in Chapter 1, the textbook leaves the proofs of these properties "to the reader", and as in chapter 1 I think that's robbing you of a wonderful learning experience. This time, I'll give a proof of property (3). Note that in order to prove two matrices are equal, we need to show that corresponding entries are equal. As such, my proof will focus on an arbitrary entry.

Proof of Theorem 3.1.1, (3): Let A, B, C, be $m \times n$ matrices. Then for any $1 \leq i \leq m$ and $1 \leq j \leq n$ we have the following:

$$\begin{array}{rcll} ((A+B)+C)_{ij} & = & (A+B)_{ij} + (C)_{ij} & \text{definition of addition} \\ & = & ((A)_{ij} + (B)_{ij}) + (C)_{ij} & \text{definition of addition} \\ & = & (A)_{ij} + ((B)_{ij} + (C)_{ij}) & \text{associativity of addition of real numbers} \\ & = & (A)_{ij} + (B+C)_{ij} & \text{definition of addition} \\ & = & (A+(B+C))_{ij} & \text{definition of addition} \end{array}$$

And since $((A+B)+C)_{ij} = (A+(B+C))_{ij}$ for all applicable i and j, the definition of equality tells us that $(A+B)+C = A+(B+C)$.

To finish off this introduction to matrices, there are a couple of terms that apply to matrices that we'll want to make use of.

Definition: We say that a matrix that has the same number of columns and rows (that is, an $n \times n$ matrix for some n) is a **square matrix**.

Example: $\begin{bmatrix} 1 & -1 & 5 & -3 \\ 2 & 7 & 6 & 8 \\ 0 & -3 & -8 & 13 \\ -7 & 7 & 4 & -3 \end{bmatrix}$ is a square matrix because is has 4 rows and 4 columns. $\begin{bmatrix} 5 & 2 & -3 & 5 \\ 9 & 0 & 2 & -1 \\ 3 & -8 & -2 & -9 \end{bmatrix}$ is not a square matrix, because is it has 3 rows but 4 columns.

Definition: A square matrix U is said to be **upper triangular** if the entries beneath the main diagonal are all zero–that is, when $u_{ij} = 0$ whenever $i > j$. This means that the only non-zero entries are in the "upper" part of the matrix.

Example: $U = \begin{bmatrix} 5 & 0 & 7 & -4 \\ 0 & 3 & -6 & 7 \\ 0 & 0 & 13 & -9 \\ 0 & 0 & 0 & 5 \end{bmatrix}$ is an upper triangular matrix.

Definition: A square matrix L is said to be **lower triangular** if the entries above the main diagonal are all zero–that is, when $l_{ij} = 0$ whenever $i < j$. This means that the only non-zero entries are in the "lower" part of the matrix.

Example: $L = \begin{bmatrix} -3 & 0 & 0 & 0 \\ 6 & 7 & 0 & 0 \\ -2 & 10 & 3 & 0 \\ -2 & -5 & 9 & 3 \end{bmatrix}$ is a lower triangular matrix.

Definition: A matrix D that is both upper and lower triangular is called a **diagonal matrix**–that is, $d_{ij} = 0$ for all $i \neq j$. In this case, the non-zero entries are only on the “diagonal” (also known as the “main diagonal”) part of the matrix.

<u>Notation</u>: We denote an $n \times n$ diagonal matrix by

$$D = \text{diag}(d_{11}, d_{22}, \cdots, d_{nn})$$

Example: $D = \begin{bmatrix} 3 & 0 & 0 & 0 \\ 0 & -2 & 0 & 0 \\ 0 & 0 & 7 & 0 \\ 0 & 0 & 0 & -3 \end{bmatrix}$ is a diagonal matrix. We can write $D = \text{diag}(3, -2, 7, -3)$.