MATH 1030 Chapter 6

The lecture is based on Beezer, A first course in Linear algebra. Ver 3.5 Downloadable at http://linear.ups.edu/download.html .

The print version can be downloaded at http://linear.ups.edu/download/fcla-3.50-print.pdf .

Reference.

- Beezer, Ver 3.5 Section Matrix Operations, Section Matrix Multiplication.

- Strang, Sect 1.4 and Sect 1.6.

Exercises with solutions can be downloaded at http://linear.ups.edu/download/fcla-3.50-solution-manual.pdf (Replace $\mathbb{C}$ by $\mathbb{R}$)Section MO (p52-56), all. Section MM (p57-60), all except T12 and T35.

6.1   Matrix Equality, Addition, Scalar Multiplication

Recall M_{mn} is the set of $m \times n$ matrices with real entries. Throughout the section, unless otherwise stated,

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

Definition 6.1 (Matrix Equality)**.** The $m \times n$ matrices A and B are **equal**, written $A = B$ provided: $[A]_{ij} = [B]_{ij}$ for all $1 \le i \le m$, $1 \le j \le n$, that is:

$$a_{ij} = b_{ij} \quad \text{for all } i, j.$$

Definition 6.2 (Matrix Addition)**.** Given $m \times n$ matrices A and B, define the **sum** of A and B as an $m \times n$ matrix, written $A + B$, according to

$$[A + B]_{ij} = [A]_{ij} + [B]_{ij}$$

i.e.,

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

Example 6.3. If

$$
A = \begin{bmatrix} 2 & -3 & 4 \\ 1 & 0 & -7 \end{bmatrix} \quad B = \begin{bmatrix} 6 & 2 & -4 \\ 3 & 5 & 2 \end{bmatrix}
$$

then:

$$
A + B = \begin{bmatrix} 2 & -3 & 4 \\ 1 & 0 & -7 \end{bmatrix} + \begin{bmatrix} 6 & 2 & -4 \\ 3 & 5 & 2 \end{bmatrix}
$$

$$
= \begin{bmatrix} 2+6 & -3+2 & 4+(-4) \\ 1+3 & 0+5 & -7+2 \end{bmatrix} = \begin{bmatrix} 8 & -1 & 0 \\ 4 & 5 & -5 \end{bmatrix}
$$

Definition 6.4 (Matrix Scalar Multiplication)**.** Given $m \times n$ matrix A and a scalar $\lambda \in \mathbb{R}$, the **scalar multiple** of A by λ is the $m \times n$ matrix, written λA, defined as follows:

$$
[\lambda A]_{ij} = \lambda [A]_{ij},
$$

i.e.,

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

Computationally, scalar matrix multiplication is very easy.

Example 6.5. If

$$
A = \begin{bmatrix} 2 & 8 \\ -3 & 5 \\ 0 & 1 \end{bmatrix}
$$

and $\lambda = 7$, then

$$
\lambda A = 7 \begin{bmatrix} 2 & 8 \\ -3 & 5 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 7(2) & 7(8) \\ 7(-3) & 7(5) \\ 7(0) & 7(1) \end{bmatrix} = \begin{bmatrix} 14 & 56 \\ -21 & 35 \\ 0 & 7 \end{bmatrix}.
$$

Definition 6.6 (Zero Matrix). The $m \times n$ **zero matrix** is written as $\mathcal{O} = \mathcal{O}_{m \times n}$ and defined by $[\mathcal{O}]_{ij} = 0$, for all $1 \leq i \leq m$, $1 \leq j \leq n$, i.e.

$$\mathcal{O} = \mathcal{O}_{m \times n} = \begin{bmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}.$$

Definition 6.7 (Additive Inverse). The additive inverse of a matrix $A \in M_{mn}$, denoted by $-A$ is defined by $-A = (-1)A$, i.e.

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

Below are some obvious properties satisfied by the addition and scalar multiplication of matrices:

1. **Associativity of Matrix Addition** For all A, B, $C \in M_{mn}$, we have:
$$A + (B + C) = (A + B) + C.$$

2. **Commutativity of Matrix Addition** For all A, $B \in M_{mn}$, we have:
$$A + B = B + A.$$

3. **Additive Identity of Matrix Addition** $A + \mathcal{O} = A$ for all $A \in M_{mn}$.

4. **Existence of Additive Inverse** For any $m \times n$ matrix A, we have:
$$A + (-A) = \mathcal{O}_{m \times n}.$$

5. **Associativity of Scalar Multiplication** For all α, $\beta \in \mathbb{R}$ and $A \in M_{mn}$, we have:
$$\alpha(\beta A) = (\alpha \beta)A.$$

6. **Distributivity across Matrix Addition** For all $\alpha \in \mathbb{R}$ and A, $B \in M_{mn}$, we have:
$$\alpha(A + B) = \alpha A + \alpha B.$$

7. **Distributivity across Scalar Addition** For all α, $\beta \in \mathbb{R}$ and $A \in M_{mn}$, we have:
$$(\alpha + \beta)A = \alpha A + \beta A.$$

3

8. **Scalar Multiplication by** $1 \in \mathbb{R}$ For all $A \in M_{mn}$, we have $1A = A$.

Example 6.8. As an example, we prove here property 7, $(\alpha + \beta)A = \alpha A + \beta A$. We need to establish the equality of two matrices.

For any i and j, $1 \leq i \leq m$, $1 \leq j \leq n$,

$$
\begin{aligned}
[(\alpha + \beta)A]_{ij} &= (\alpha + \beta)\,[A]_{ij} \\
&= \alpha\,[A]_{ij} + \beta\,[A]_{ij} \\
&= [\alpha A]_{ij} + [\beta A]_{ij} \\
&= [\alpha A + \beta A]_{ij}\,.
\end{aligned}
$$

Hence by the definition of equality of matrices, $(\alpha + \beta)A = \alpha A + \beta A$.

6.2   Matrix-Vector Product

Definition 6.9 (Matrix-Vector Product). Suppose that A is an $m \times n$ matrix with columns $\mathbf{A}_1$, $\mathbf{A}_2$, $\mathbf{A}_3$, $\ldots$, $\mathbf{A}_n$ and $\mathbf{u}$ is a vector of size n. Then the **matrix-vector product** of A with $\mathbf{u}$ is the linear combination

$$
A\mathbf{u} = [\mathbf{u}]_1\,\mathbf{A}_1 + [\mathbf{u}]_2\,\mathbf{A}_2 + [\mathbf{u}]_3\,\mathbf{A}_3 + \cdots + [\mathbf{u}]_n\,\mathbf{A}_n
$$

Note that an $m \times n$ matrix A times a vector of size n will create a column vector of size m.

In particular, if A is (non-square) rectangular, then the size of the vector changes.

Example 6.10. Consider:

$$
A = \begin{bmatrix} 1 & 4 & 2 & 3 & 4 \\ -3 & 2 & 0 & 1 & -2 \\ 1 & 6 & -3 & -1 & 5 \end{bmatrix}, \quad \mathbf{u} = \begin{bmatrix} 2 \\ 1 \\ -2 \\ 3 \\ -1 \end{bmatrix}
$$

Then:

$$
\begin{aligned}
A\mathbf{u} &= 2\begin{bmatrix} 1 \\ -3 \\ 1 \end{bmatrix} + 1\begin{bmatrix} 4 \\ 2 \\ 6 \end{bmatrix} + (-2)\begin{bmatrix} 2 \\ 0 \\ -3 \end{bmatrix} + 3\begin{bmatrix} 3 \\ 1 \\ -1 \end{bmatrix} + (-1)\begin{bmatrix} 4 \\ -2 \\ 5 \end{bmatrix} \\
&= \begin{bmatrix} 7 \\ 1 \\ 6 \end{bmatrix}.
\end{aligned}
$$

6.2.1   Matrix Notation for Systems of Linear Equations

Theorem 6.11 (Systems of Linear Equations as Matrix Multiplication). *The solution set to the linear system $\mathcal{LS}(A, \mathbf{b})$ is equal to the set of solutions $\mathbf{x}$ to the vector equation $A\mathbf{x} = \mathbf{b}$.*

Proof of Systems of Linear Equations as Matrix Multiplication.

$$\mathbf{x} \text{ is a solution to } \mathcal{LS}(A, \mathbf{b})$$

$$\iff [\mathbf{x}]_1 \, \mathbf{A}_1 + [\mathbf{x}]_2 \, \mathbf{A}_2 + [\mathbf{x}]_3 \, \mathbf{A}_3 + \cdots + [\mathbf{x}]_n \, \mathbf{A}_n = \mathbf{b}$$

$$\iff \mathbf{x} \text{ is a solution to } A\mathbf{x} = \mathbf{b}$$

$\square$

Example 6.12. Consider the system of linear equations

$$2x_1 + 4x_2 - 3x_3 + 5x_4 + x_5 = 9$$

$$3x_1 + x_2 + x_4 - 3x_5 = 0$$

$$-2x_1 + 7x_2 - 5x_3 + 2x_4 + 2x_5 = -3$$

has coefficient matrix and vector of constants

$$A = \begin{bmatrix} 2 & 4 & -3 & 5 & 1 \\ 3 & 1 & 0 & 1 & -3 \\ -2 & 7 & -5 & 2 & 2 \end{bmatrix} \qquad \mathbf{b} = \begin{bmatrix} 9 \\ 0 \\ -3 \end{bmatrix}$$

and so will be described compactly by the vector equation $A\mathbf{x} = \mathbf{b}$.

Theorem 6.13 (Equality of Matrices and Matrix-Vector Products). *If A and B are $m \times n$ matrices such that $A\mathbf{x} = B\mathbf{x}$ for every $\mathbf{x} \in \mathbb{R}^n$, then $A = B$.*

Proof of Equality of Matrices and Matrix-Vector Products. Suppose $A\mathbf{x} = B\mathbf{x}$ for all $\mathbf{x} \in \mathbb{R}^n$. Then, in particular this equality holds for the standard unit vectors, defined as follows:

For $1 \leq j \leq n$, we define the **standard unit vector** $\mathbf{e}_j$ to be the column vector in $\mathbb{R}^n$ with the j-th entry equal 1 and all other entries equal to zero. For any $1 \leq i \leq m$ and $1 \leq j \leq n$, we have:

$$[A]_{ij} = [A\mathbf{e}_j]_i$$
$$= [B\mathbf{e}_j]_i$$
$$= [B]_{ij}$$

Hence $A = B$. $\square$

Remark. You might notice from studying the proof that the hypotheses of this theorem could be weakened i.e., made less restrictive). We need only suppose the equality of the matrix-vector products for the standard unit vectors or any other spanning set of $\mathbb{R}^n$. However, in practice, when we apply this theorem the stronger hypothesis will be in effect so this version of the theorem suffices for our purposes. (If we changed the statement of the theorem to have the less restrictive hypothesis, then we would call the theorem stronger.)

6.3 Matrix Multiplication

Definition 6.14 (Matrix Multiplication). Suppose A is an $m \times n$ matrix and $\mathbf{B}_1$, $\mathbf{B}_2$, $\mathbf{B}_3$, ..., $\mathbf{B}_p$ are the columns of an $n \times p$ matrix B. Then the **matrix product** of A with B is the $m \times p$ matrix whose ith column is the matrix-vector product $A\mathbf{B}_i$. Symbolically,

$$AB = A\left[\mathbf{B}_1|\mathbf{B}_2|\mathbf{B}_3|\ldots|\mathbf{B}_p\right] = \left[A\mathbf{B}_1|A\mathbf{B}_2|A\mathbf{B}_3|\ldots|A\mathbf{B}_p\right].$$

Example 6.15. Let:

$$A = \begin{bmatrix} 1 & 2 & -1 & 4 & 6 \\ 0 & -4 & 1 & 2 & 3 \\ -5 & 1 & 2 & -3 & 4 \end{bmatrix}, \ B = \begin{bmatrix} 1 & 6 & 2 & 1 \\ -1 & 4 & 3 & 2 \\ 1 & 1 & 2 & 3 \\ 6 & 4 & -1 & 2 \\ 1 & -2 & 3 & 0 \end{bmatrix}.$$

Then:

$$AB = \left[A\begin{bmatrix} 1 \\ -1 \\ 1 \\ 6 \\ 1 \end{bmatrix} \middle| A\begin{bmatrix} 6 \\ 4 \\ 1 \\ 4 \\ -2 \end{bmatrix} \middle| A\begin{bmatrix} 2 \\ 3 \\ 2 \\ -1 \\ 3 \end{bmatrix} \middle| A\begin{bmatrix} 1 \\ 2 \\ 3 \\ 2 \\ 0 \end{bmatrix} \right]$$

$$= \begin{bmatrix} 28 & 17 & 20 & 10 \\ 20 & -13 & -3 & -1 \\ -18 & -44 & 12 & -3 \end{bmatrix}.$$

Remark. Is this the definition of matrix multiplication you expected? Perhaps our previous operations for matrices caused you to think that we might multiply two matrices of the same size, entry-by-entry? Notice that our current definition uses matrices of different sizes (though the number of columns in the first must

equal the number of rows in the second), and the result is of a third size. Notice too that in the previous example we cannot even consider the product BA, since the sizes of the two matrices in this order are not compatible. But it gets weirder than that. Many of your old ideas about multiplication will not apply to matrix multiplication, but some still will. So make no assumptions, and do not do anything until you have a theorem that says you can. Even if the sizes are right, matrix multiplication is not commutative – order matters.

Example 6.16. This example demonstrates that matrix multiplication is in general **not** commutative.

Let:
$$A = \begin{bmatrix} 1 & 3 \\ -1 & 2 \end{bmatrix}, \ B = \begin{bmatrix} 4 & 0 \\ 5 & 1 \end{bmatrix}.$$

Then:
$$AB = \begin{bmatrix} 19 & 3 \\ 6 & 2 \end{bmatrix}, \ BA = \begin{bmatrix} 4 & 12 \\ 4 & 17 \end{bmatrix}$$

So, $AB \neq BA$.

It should not be hard for you to construct other pairs of matrices that do not commute (try a couple of 3×3's). Can you find a pair of non-identical matrices that do commute?

Theorem 6.17 (Entries of Matrix Products). *Suppose A is an $m \times n$ matrix and B is an $n \times p$ matrix. Then for $1 \leq i \leq m$, $1 \leq j \leq p$, the individual entries of AB are given by:*

$$[AB]_{ij} = [A]_{i1} [B]_{1j} + [A]_{i2} [B]_{2j} + [A]_{i3} [B]_{3j} + \cdots + [A]_{in} [B]_{nj}$$
$$= \sum_{k=1}^{n} [A]_{ik} [B]_{kj}$$

Remark. In most books, this is used as the definition of AB.

Proof of Entries of Matrix Products. View the columns of A as column vectors and denote them from left to right by: $\mathbf{A}_1$, $\mathbf{A}_2$, $\mathbf{A}_3$, ..., $\mathbf{A}_n$. Similarly, denote the columns of B by: $\mathbf{B}_1$, $\mathbf{B}_2$, $\mathbf{B}_3$, ..., $\mathbf{B}_p$.

Then, for $1 \leq i \leq m$, $1 \leq j \leq p$, we have:

$$
\begin{aligned}
[AB]_{ij} = [A\mathbf{B}_j]_i &= \left[[\mathbf{B}_j]_1 \, \mathbf{A}_1 + [\mathbf{B}_j]_2 \, \mathbf{A}_2 + \cdots + [\mathbf{B}_j]_n \, \mathbf{A}_n \right]_i \\
&= \left[[\mathbf{B}_j]_1 \, \mathbf{A}_1 \right]_i + \left[[\mathbf{B}_j]_2 \, \mathbf{A}_2 \right]_i + \cdots + \left[[\mathbf{B}_j]_n \, \mathbf{A}_n \right]_i \\
&= [\mathbf{B}_j]_1 \, [\mathbf{A}_1]_i + [\mathbf{B}_j]_2 \, [\mathbf{A}_2]_i + \cdots + [\mathbf{B}_j]_n \, [\mathbf{A}_n]_i \\
&= [B]_{1j} \, [A]_{i1} + [B]_{2j} \, [A]_{i2} + \cdots + [B]_{nj} \, [A]_{in} \\
&= [A]_{i1} \, [B]_{1j} + [A]_{i2} \, [B]_{2j} + \cdots + [A]_{in} \, [B]_{nj} \\
&= \sum_{k=1}^{n} [A]_{ik} \, [B]_{kj}
\end{aligned}
$$

$\square$

Example 6.18 (Product of Two Matrices Entry-by-Entry). Consider the matrices:

$$
A = \begin{bmatrix} 1 & 2 & -1 & 4 & 6 \\ 0 & -4 & 1 & 2 & 3 \\ -5 & 1 & 2 & -3 & 4 \end{bmatrix}, \quad
B = \begin{bmatrix} 1 & 6 & 2 & 1 \\ -1 & 4 & 3 & 2 \\ 1 & 1 & 2 & 3 \\ 6 & 4 & -1 & 2 \\ 1 & -2 & 3 & 0 \end{bmatrix}
$$

Suppose we just wanted the entry of AB in the second row, third column:

$$
\begin{aligned}
[AB]_{23} &= [A]_{21} \, [B]_{13} + [A]_{22} \, [B]_{23} + [A]_{23} \, [B]_{33} + [A]_{24} \, [B]_{43} + [A]_{25} \, [B]_{53} \\
&= (0)(2) + (-4)(3) + (1)(2) + (2)(-1) + (3)(3) = -3
\end{aligned}
$$

Notice how there are 5 terms in the sum, since 5 is the common dimension of the two matrices (column count for A, row count for B). In the conclusion of the above theorem, it would be the index k that would run from 1 to 5 in this computation. Here is a bit more practice. The entry of third row, first column:

$$
\begin{aligned}
[AB]_{31} &= [A]_{31} \, [B]_{11} + [A]_{32} \, [B]_{21} + [A]_{33} \, [B]_{31} + [A]_{34} \, [B]_{41} + [A]_{35} \, [B]_{51} \\
&= (-5)(1) + (1)(-1) + (2)(1) + (-3)(6) + (4)(1) = -18
\end{aligned}
$$

Try to compute all the other entries.

How to memorize the formula : To find the (i, j)-th entry of AB. (1) Find the i-th row of A (simply called the row below)(2) Find the j-th column of B (simply called the column below)(3) sum up the product the corresponding entries of the row and the column, i.e. (entry 1 of the row $\times$ entry 1 of the column) + (entry 2 of the row $\times$ entry 2 of the column) + $\cdots$

Example 6.19. Find the $(3, 2)$ entry of AB in the previous example. The 3-rd row of A is $\begin{bmatrix} -5 & 1 & 2 & -3 & 4 \end{bmatrix}$.

The 2-nd column of B is $\begin{bmatrix} 6 \\ 4 \\ 1 \\ 4 \\ -2 \end{bmatrix}$.

Let's do the multiplication:

Row	-5	1	2	-3	4
Column	6	4	1	4	-2
Product	-30	4	2	-12	-8

The sum is:
$$-30 + 4 + 2 - 12 - 8 = -44.$$

6.3.1 Properties of Matrix Multiplication

In this subsection, we collect properties of matrix multiplication and its interaction with the zero matrix, the identity matrix, matrix addition, scalar matrix multiplication and the transpose.

Theorem 6.20 (Matrix Multiplication and the Zero Matrix)**.** *Suppose A is an $m \times n$ matrix. Then*

1. $A\mathcal{O}_{n \times p} = \mathcal{O}_{m \times p}$

2. $\mathcal{O}_{p \times m} A = \mathcal{O}_{p \times n}$

Proof of Matrix Multiplication and the Zero Matrix. We will prove (1) and leave (2) to you. Using the entry-by-entry definition of matrix multiplication, for $1 \leq i \leq m$, $1 \leq j \leq p$, we compute

$$[A\mathcal{O}_{n \times p}]_{ij} = \sum_{k=1}^{n} [A]_{ik} [\mathcal{O}_{n \times p}]_{kj}$$

$$= \sum_{k=1}^{n} [A]_{ik} \, 0$$

$$= \sum_{k=1}^{n} 0$$

$$= 0$$

$$= [\mathcal{O}_{m \times p}]_{ij}$$

So the matrices $A\mathcal{O}_{n \times p}$ and $\mathcal{O}_{m \times p}$ are equal. $\qquad\square$

Definition 6.21. The **identity matrix** I_m is the $m \times m$ square matrix whose diagonal entries are all equal to 1, and all off-diagonal entries are equal to zero. That is:

$$[I_m]_{ij} = \begin{cases} 1, & \text{if } i = j; \\ 0, & \text{if } i \neq j. \end{cases}$$

For example:

$$I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad I_5 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

Theorem 6.22 (Matrix Multiplication and Identity Matrix). *Suppose that A is an $m \times n$ matrix. Then*

1. $AI_n = A$

2. $I_m A = A$

Proof of Matrix Multiplication and Identity Matrix. Again, we will prove (1) and leave (2) to you. Using the entry-by-entry definition of matrix multiplication, for $1 \leq i \leq m$, $1 \leq j \leq n$, we compute

$$[AI_n]_{ij} = \sum_{k=1}^{n} [A]_{ik} [I_n]_{kj}$$

$$= [A]_{ij} [I_n]_{jj} + \sum_{\substack{k=1 \\ k \neq j}}^{n} [A]_{ik} [I_n]_{kj}$$

$$= [A]_{ij} (1) + \sum_{k=1, k \neq j}^{n} [A]_{ik} (0)$$

$$= [A]_{ij} + \sum_{k=1, k \neq j}^{n} 0$$

$$= [A]_{ij}.$$

So the matrices A and AI_n are equal entrywise. By the definition of matrix equality, they are equal matrices. □

Remark. It is the previous theorem that gives the identity matrix its name. It is a matrix that behaves with matrix multiplication like the scalar 1 does with scalar multiplication. To multiply by the identity matrix is to have no effect on the other matrix.

Theorem 6.23 (Matrix Multiplication Distributes Across Addition). *Suppose that A is an $m \times n$ matrix and B and C are $n \times p$ matrices and D is a $p \times s$ matrix. Then:*

1. $A(B + C) = AB + AC$

2. $(B + C)D = BD + CD$

Proof of Matrix Multiplication Distributes Across Addition. We will do (1), you do (2). Entry-by-entry, for $1 \le i \le m$, $1 \le j \le p$,

$$
\begin{aligned}
[A(B + C)]_{ij} &= \sum_{k=1}^{n} [A]_{ik} [B + C]_{kj} \\
&= \sum_{k=1}^{n} [A]_{ik} ([B]_{kj} + [C]_{kj}) \\
&= \sum_{k=1}^{n} [A]_{ik} [B]_{kj} + [A]_{ik} [C]_{kj} \\
&= \sum_{k=1}^{n} [A]_{ik} [B]_{kj} + \sum_{k=1}^{n} [A]_{ik} [C]_{kj} \\
&= [AB]_{ij} + [AC]_{ij} \\
&= [AB + AC]_{ij}
\end{aligned}
$$

So the matrices $A(B + C)$ and $AB + AC$ are equal, entry-by-entry. Hence by the definition of matrix equality, we can say they are equal matrices. $\square$

Theorem 6.24 (Matrix Multiplication and Scalar Matrix Multiplication). *Suppose A is an $m \times n$ matrix and B is an $n \times p$ matrix. Let α be a scalar. Then $\alpha(AB) = (\alpha A)B = A(\alpha B)$.*

Proof of Matrix Multiplication and Scalar Matrix Multiplication. These are equalities of matrices. We will do the first one, the second is similar and will be good

practice for you. For $1 \le i \le m$, $1 \le j \le p$,

$$[\alpha(AB)]_{ij} = \alpha\,[AB]_{ij}$$

$$= \alpha\sum_{k=1}^{n} [A]_{ik}\,[B]_{kj}$$

$$= \sum_{k=1}^{n} \alpha\,[A]_{ik}\,[B]_{kj}$$

$$= \sum_{k=1}^{n} [\alpha A]_{ik}\,[B]_{kj}$$

$$= [(\alpha A)B]_{ij}$$

So the matrices $\alpha(AB)$ and $(\alpha A)B$ are equal, entry-by-entry, and by the definition of matrix equality we can say they are equal matrices. $\qquad\square$

Theorem 6.25 (Matrix Multiplication is Associative). *Suppose A is an $m \times n$ matrix, B is an $n \times p$ matrix and D is a $p \times s$ matrix. Then $A(BD) = (AB)D$.*

Proof of Matrix Multiplication is Associative. A matrix equality, so we will go entry-by-entry, no surprise there. For $1 \le i \le m$, $1 \le j \le s$,

$$[A(BD)]_{ij} = \sum_{k=1}^{n} [A]_{ik}\,[BD]_{kj}$$

$$= \sum_{k=1}^{n} [A]_{ik}\left(\sum_{\ell=1}^{p} [B]_{k\ell}\,[D]_{\ell j}\right)$$

$$= \sum_{k=1}^{n}\sum_{\ell=1}^{p} [A]_{ik}\,[B]_{k\ell}\,[D]_{\ell j}$$

We can switch the order of the summation since these are finite sums.

$$= \sum_{\ell=1}^{p}\sum_{k=1}^{n} [A]_{ik}\,[B]_{k\ell}\,[D]_{\ell j}$$

As $[D]_{\ell j}$ does not depend on the index k, we can use distributivity to move it outside of the inner sum.

$$= \sum_{\ell=1}^{p} [D]_{\ell j} \left(\sum_{k=1}^{n} [A]_{ik} [B]_{k\ell} \right)$$

$$= \sum_{\ell=1}^{p} [D]_{\ell j} [AB]_{i\ell}$$

$$= \sum_{\ell=1}^{p} [AB]_{i\ell} [D]_{\ell j}$$

$$= [(AB)D]_{ij}$$

Hence, $(AB)D = A(BD)$. □

Alternatively,

Proof of Matrix Multiplication is Associative. Write:

$$A = (\mathbf{A}_1 | \mathbf{A}_2 | \cdots | \mathbf{A}_n)$$

$$B = (\mathbf{B}_1 | \mathbf{B}_2 | \cdots | \mathbf{B}_p)$$

$$D = (\mathbf{D}_1 | \mathbf{D}_2 | \cdots | \mathbf{D}_s)$$

Then, for any $1 \le j \le s$, the j-th column of $A(BD)$ is the j-th column of:

$$A (B\mathbf{D}_1 | B\mathbf{D}_2 | \cdots | B\mathbf{D}_s),$$

which is equal to:

$$A(B\mathbf{D}_j).$$

The j-th column of $(AB)D$ is:

$$(AB)\mathbf{D}_j.$$

Hence, it suffices to show that:

$$A(B\mathbf{D}_j) = (AB)\mathbf{D}_j$$

for any $1 \le j \le s$.

Given any $1 \le j \le s$, to simplify the notation, let $\vec{v} = \mathbf{D}_j$. We have:

$$A(B\vec{v}) = A(v_1 \mathbf{B}_1 + \cdots + v_p \mathbf{B}_p) = v_1(A\mathbf{B}_1) + \cdots + v_p(A\mathbf{B}_p)$$
$$= (A\mathbf{B}_1 | \cdots | A\mathbf{B}_p)\vec{v}$$
$$= (AB)\vec{v}.$$

This completes the proof. □

Remark. The above result says matrix multiplication is associative; it means we do not have to be careful about how we parenthesize an expression with just several matrices multiplied together. So this is where we draw the line on explaining every last detail in a proof. We will frequently add, remove, or rearrange parentheses with no comment.

6.4   Row Operations and Matrix Multiplication

In this section, we will discuss the relation between elementary row operations and matrix multiplication.

Recall the definition of Definition 4.14 (Row Operations) on matrices.

Theorem 6.26. *Let $A \in M_{mn}$. Let B be a matrix obtained by applying one of the above row operations on A. Let J be a matrix obtained by applying the same row operation on I_m. Then*
$$JA = B.$$

Proof of Theorem 6.26. Exercise. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

Example 6.27. Let:
$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} \\ a_{31} & a_{32} & a_{33} & a_{34} & a_{35} \\ a_{41} & a_{42} & a_{43} & a_{44} & a_{45} \end{bmatrix}$$

Consider the row operation $3R_2 + R_3$.

$$A \xrightarrow{3R_2+R_3} B = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} \\ 3a_{21}+a_{31} & 3a_{22}+a_{32} & 3a_{23}+a_{33} & 3a_{24}+a_{34} & 3a_{25}+a_{35} \\ a_{41} & a_{42} & a_{43} & a_{44} & a_{45} \end{bmatrix}$$

$$I_4 \xrightarrow{3R_2+R_3} J = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 3 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Verify that:
$$JA = B.$$

Definition 6.28. A matrix J which corresponds to an elementary row operation as in Theorem 6.26 is called an **elementary matrix**.

6.5 Invertible Matrices

Definition 6.29. An $n \times n$ square matrix A is said to be **invertible** if there exists an $n \times n$ matrix B such that:

$$AB = BA = I_n.$$

We call B an **inverse** of A.

Theorem 6.30. *If A, B, B' are $n \times n$ matrices such that:*

$$AB = BA = I_n \quad and \quad AB' = B'A = I_n,$$

then $B = B'$. Hence, the inverse of an invertible matrix A is unique. We denote it by A^{-1}.

Fact 6.31. *Let A be an $n \times n$ matrix.*

- *If A is invertible, so is A^{-1}, with:*

$$\left(A^{-1}\right)^{-1} = A.$$

- *The matrix A is invertible if and only if it is row-equivalent to I_n.*

- *The matrix A is invertible if and only if it is **nonsingular**, that is:*

$$A\vec{x} = \vec{0}$$

  if and only if $\vec{x} = \vec{0}$.

- *If the matrix A is invertible, then so is its transpose A^t, with:*

$$\left(A^t\right)^{-1} = \left(A^{-1}\right)^t.$$

- *If A, B are invertible $n \times n$ matrices, then AB is also invertible, with:*

$$(AB)^{-1} = B^{-1}A^{-1}.$$

Theorem 6.32. *Elementary matrices are invertible.*

We may now prove Theorem Theorem 4.17 (Row-Equivalent Matrices represent Equivalent Systems).

Proof of Row-Equivalent Matrices represent Equivalent Systems. Suppose $A = [A'|\vec{a}]$ and $B = [B'|\vec{b}]$. If A and B are row-equivalent, then there exists a sequence of elementary matrices $J_1, J_2, \ldots J_l$ such that:

$$[B'|\vec{b}] = B = J_l \cdots J_2 J_1 A = J_l \cdots J_2 J_1 [A'|\vec{a}]$$

Let $J = J_l \cdots J_2 J_1$. Then, we have:

$$B' = JA' \quad \text{and} \quad \vec{b} = J\vec{a}.$$

For any $\vec{v}$ which is a solution to $\mathcal{LS}(A', \vec{a})$, by definition we have:

$$A'\vec{v} = \vec{a}.$$

So,

$$B'\vec{v} = JA'\vec{v} = J\vec{a} = \vec{b},$$

which implies that $\vec{v}$ is also a solution to $\mathcal{LS}(B', \vec{b})$.

Conversely, by Theorem 6.32 the matrix J is invertible, hence,

$$A' = J^{-1}B' \quad \text{and} \quad \vec{a} = J^{-1}\vec{b}.$$

It then follows from the same arguments used before that any solution $\vec{v}$ to $\mathcal{LS}(B', \vec{b})$ is also a solution to $\mathcal{LS}(A', \vec{a})$.

We conclude that $\mathcal{LS}(A', \vec{a})$ and $\mathcal{LS}(B', \vec{b})$ have the same solution set, hence they are equivalent linear systems. $\qquad\square$

6.6   Transposes and Symmetric Matrices

We now describe one more common operation which can be performed on matrices. Informally, to transpose a matrix is to build a new matrix by swapping its rows and columns.

Definition 6.33 (Transpose of a Matrix)**.** Given an $m \times n$ matrix A, its **transpose** is the $n \times m$ matrix A^t given by

$$\left[A^t\right]_{ij} = [A]_{ji}, \quad 1 \le i \le n, 1 \le j \le m,$$

i.e.

$$A^t = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \ldots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \ldots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \ldots & a_{mn} \end{bmatrix}^t$$

$$= \begin{bmatrix} a_{11} & a_{21} & a_{31} & \ldots & a_{m1} \\ a_{12} & a_{22} & a_{32} & \ldots & a_{m2} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{1n} & a_{2n} & a_{3n} & \ldots & a_{mn} \end{bmatrix}.$$

Example 6.34. Suppose:

$$D = \begin{bmatrix} 3 & 7 & 2 & -3 \\ -1 & 4 & 2 & 8 \\ 0 & 3 & -2 & 5 \end{bmatrix}.$$

Then,

$$D^t = \begin{bmatrix} 3 & -1 & 0 \\ 7 & 4 & 3 \\ 2 & 2 & -2 \\ -3 & 8 & 5 \end{bmatrix}.$$

Definition 6.35 (Symmetric Matrix)**.** A matrix A is said to be **symmetric** if $A = A^t$, i.e.

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$$

with

$$a_{ij} = a_{ji} \text{ for all } i, j.$$

Example 6.36. The matrix:

$$E = \begin{bmatrix} 2 & 3 & -9 & 5 & 7 \\ 3 & 1 & 6 & -2 & -3 \\ -9 & 6 & 0 & -1 & 9 \\ 5 & -2 & -1 & 4 & -8 \\ 7 & -3 & 9 & -8 & -3 \end{bmatrix}$$

is symmetric.

Theorem 6.37 (Symmetric Matrices are Square)**.** *Suppose that A is a symmetric matrix. Then A is square.*

Proof of Symmetric Matrices are Square. Suppose A is a $n \times m$ matrix. Then A^t is a $m \times n$ matrix. In order for A and A^t to be equal, they must have the same dimension. Hence $n = m$. □

17

Definition 6.38. A matrix A is said to be **skew-symmetric** (or **antisymmetric**) if:
$$A^t = -A$$

Example 6.39.
$$A = \begin{bmatrix} 0 & -5 & 7 & 11 \\ 5 & 0 & 8 & -2 \\ -7 & -8 & 0 & 15 \\ -11 & 2 & -15 & 0 \end{bmatrix}$$

Remark. Notice that if A is skew-symmetric, then:

- It must be square.

- It's diagonal entries must all be equal to zero.

Theorem 6.40 (Transpose and Matrix Addition). *Suppose that A and B are $m \times n$ matrices. Then $(A + B)^t = A^t + B^t$.*

Proof of Transpose and Matrix Addition. For $1 \leq i \leq n$, $1 \leq j \leq m$,
$$\begin{aligned} \left[(A + B)^t \right]_{ij} &= [A + B]_{ji} \\ &= [A]_{ji} + [B]_{ji} \\ &= \left[A^t \right]_{ij} + \left[B^t \right]_{ij} \\ &= \left[A^t + B^t \right]_{ij} \end{aligned}$$

Since the matrices $(A+B)^t$ and $A^t + B^t$ agree at each entry, they are equal. $\square$

Theorem 6.41 (Transpose and Matrix Scalar Multiplication). *Suppose that $\alpha \in \mathbb{R}$ and A is an $m \times n$ matrix. Then $(\alpha A)^t = \alpha A^t$.*

Proof of Transpose and Matrix Scalar Multiplication. For $1 \leq i \leq m$, $1 \leq j \leq n$,
$$\begin{aligned} \left[(\alpha A)^t \right]_{ji} &= [\alpha A]_{ij} \\ &= \alpha [A]_{ij} \\ &= \alpha \left[A^t \right]_{ji} \\ &= \left[\alpha A^t \right]_{ji}. \end{aligned}$$

Since the matrices $(\alpha A)^t$ and αA^t agree at each entry, they are equal. $\square$

Theorem 6.42 (Transpose of a Transpose). *Suppose that A is an $m \times n$ matrix. Then $(A^t)^t = A$.*

Proof of Transpose of a Transpose. For $1 \leq i \leq m$, $1 \leq j \leq n$,

$$\left[\left(A^t \right)^t \right]_{ij} = \left[A^t \right]_{ji}$$

$$= [A]_{ij} \, .$$

Since the matrices $\left(A^t \right)^t$ and A agree at each entry, they are equal. $\qquad\square$

Theorem 6.43 (Matrix Multiplication and Transposes)**.** *Suppose A is an $m \times n$ matrix and B is an $n \times p$ matrix. Then $(AB)^t = B^t A^t$.*

Proof of Matrix Multiplication and Transposes. Here we go again, entry-by-entry. For $1 \leq i \leq m$, $1 \leq j \leq p$,

$$\left[(AB)^t \right]_{ji} = [AB]_{ij}$$

$$= \sum_{k=1}^{n} [A]_{ik} [B]_{kj}$$

$$= \sum_{k=1}^{n} [B]_{kj} [A]_{ik}$$

$$= \sum_{k=1}^{n} \left[B^t \right]_{jk} \left[A^t \right]_{ki}$$

$$= \left[B^t A^t \right]_{ji}$$

So, $(AB)^t = B^t A^t$. $\qquad\square$

6.7   Block Matrix Multiplication

Recall from Example 4.5 that a matrix can be formed from other smaller matrices:

$$A = \left[\begin{array}{c|c|c|c} A_{11} & A_{12} & \cdots & A_{1t} \\ \hline A_{21} & A_{22} & \cdots & A_{2t} \\ \hline \vdots & \vdots & \cdots & \vdots \\ \hline A_{s1} & A_{s2} & \cdots & A_{st} \end{array} \right],$$

where instead of being scalars each A_{ij} is itself a matrix.

In certain situations, there is a convenient formula facilitating the multiplication of such "**block matrices**".

Theorem 6.44. *Let:*

$$
A = \left[\begin{array}{c|c|c|c}
A_{11} & A_{12} & \cdots & A_{1t} \\ \hline
A_{21} & A_{22} & \cdots & A_{2t} \\ \hline
\vdots & \vdots & \cdots & \vdots \\ \hline
A_{s1} & A_{s2} & \cdots & A_{st}
\end{array}\right], \quad
B = \left[\begin{array}{c|c|c|c}
B_{11} & B_{12} & \cdots & B_{1v} \\ \hline
B_{21} & B_{22} & \cdots & B_{2v} \\ \hline
\vdots & \vdots & \cdots & \vdots \\ \hline
B_{u1} & B_{u2} & \cdots & B_{uv}
\end{array}\right],
$$

such that $u = t$, *and, for* $1 \le i \le s, 1 \le j \le v, 1 \le k \le t$, *the matrix product:*

$$A_{ik} B_{kj}$$

is well-defined, i.e. the number of columns of A_{ik} is equal to the number of rows of B_{jk}, then:

$$
AB = \left[\begin{array}{c|c|c|c}
C_{11} & C_{12} & \cdots & C_{1v} \\ \hline
C_{21} & C_{22} & \cdots & C_{2t} \\ \hline
\vdots & \vdots & \cdots & \vdots \\ \hline
C_{s1} & C_{t2} & \cdots & C_{sv}
\end{array}\right],
$$

where, for $1 \le i \le s, 1 \le j \le v$,

$$C_{ij} = \sum_{k=1}^{t} A_{ik} B_{kj}$$

Example 6.45. Let:

$$
A = \left[\begin{array}{cccc}
-4 & 5 & 0 & 0 \\
1 & 2 & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1
\end{array}\right], \quad
B = \left[\begin{array}{cccc}
0 & 0 & 3 & 1 \\
0 & 0 & 2 & -1 \\
5 & 0 & 0 & 0 \\
-7 & 14 & 0 & 0
\end{array}\right],
$$

Viewing A and B as block matrices:

$$
A = \left[\begin{array}{cc|cc}
-4 & 5 & 0 & 0 \\
1 & 2 & 0 & 0 \\ \hline
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1
\end{array}\right], \quad
B = \left[\begin{array}{cc|cc}
0 & 0 & 3 & 1 \\
0 & 0 & 2 & -1 \\ \hline
5 & 0 & 0 & 0 \\
-7 & 14 & 0 & 0
\end{array}\right],
$$

we can very quickly compute:

$$
AB = \left[
\begin{array}{c|c}
\begin{bmatrix} -4 & 5 \\ 1 & 2 \end{bmatrix} \mathcal{O}_{2\times 2} + \mathcal{O}_{2\times 2} \begin{bmatrix} 5 & 0 \\ -7 & 14 \end{bmatrix} & \begin{bmatrix} -4 & 5 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 3 & 1 \\ 2 & -1 \end{bmatrix} + \mathcal{O}_{2\times 2}\mathcal{O}_{2\times 2} \\
\hline
\mathcal{O}_{2\times 2}\mathcal{O}_{2\times 2} + I_2 \begin{bmatrix} 5 & 0 \\ -7 & 14 \end{bmatrix} & \mathcal{O}_{2\times 2} \begin{bmatrix} 3 & 1 \\ 2 & -1 \end{bmatrix} + I_2\mathcal{O}_{2\times 2}
\end{array}
\right]
$$

$$
= \left[
\begin{array}{c|c}
\mathcal{O}_{2\times 2} & \begin{matrix} -2 & -9 \\ 7 & -1 \end{matrix} \\
\hline
\begin{matrix} 5 & 0 \\ -7 & 14 \end{matrix} & \mathcal{O}_{2\times 2}
\end{array}
\right]
$$

$$
= \begin{bmatrix}
0 & 0 & -2 & -9 \\
0 & 0 & 7 & -1 \\
5 & 0 & 0 & 0 \\
-7 & 14 & 0 & 0
\end{bmatrix}
$$