Math 51
Lecture 15 – July 21, 2023

Topic(s): matrix algebra

For an $n \times n$ matrix A, its **diagonal** consists of the entries a_{ii} in position (i, i) for all $1 \le i \le n$. This is the diagonal going from upper left to lower right and is called the **main diagonal**.
More generally, for an $m \times n$ matrix B, its **diagonal** consists of its entries b_{ii}. An $m \times n$ matrix is called **diagonal** if all entries away from the diagonal vanish.

The $n \times n$ **identity matrix**, denoted I_n, is defined to be the diagonal $n \times n$ matrix

$$I_n = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}$$

whose diagonal entries are all equal to 1.

- The corresponding linear transformation $T_{I_n} : \mathbb{R}^n \to \mathbb{R}^n$ gives the same output as input. In other words, $T_{I_n}(\mathbf{x}) = \mathbf{x}$ for all $\mathbf{x} \in \mathbb{R}^n$.
- For any $m \times n$ matrix A, $I_m A = A = A I_n$.

One of the crucial facts from linear algebra (from Chapter 27) is that *every* $m \times n$ matrix A can be expressed as $A = RDR'$, where D is a diagonal $m \times n$ matrix with non-negative diagonal entries, and R and R' are $m \times m$ and $n \times n$ matrices which express "rigid motions" of $\mathbb{R}^m$ and $\mathbb{R}^n$, respectively. Qualitatively, this means that upon rotating the standard bases of $\mathbb{R}^m$ and $\mathbb{R}^n$ into some orthonormal bases, A looks diagonal in these new reference frames.

Example 1. Perform the following matrix multiplications. How does it affect the matrix with letters?

- $\begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} = \begin{bmatrix} a+2d & b+2e & c+2f \\ d & e & f \\ g & h & i \end{bmatrix} = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} + \begin{bmatrix} 2d & 2e & 2f \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$

  Adds twice the second row to the first row.

  in the (1, 2) entry

- $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} = \begin{bmatrix} a & b & c \\ d+2g & e+2h & f+2i \\ g & h & i \end{bmatrix} = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 2g & 2h & 2i \\ 0 & 0 & 0 \end{bmatrix}$

  Adds twice the third row to the second row

  in the (2,3) entry

- $\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} a & 2a+b & c \\ d & 2d+e & f \\ g & 2g+h & i \end{bmatrix} = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} + \begin{bmatrix} 0 & 2a & 0 \\ 0 & 2d & 0 \\ 0 & 2g & 0 \end{bmatrix}$

  Adds twice the first column to the second column.

  in the (1, 2) entry

An $m \times n$ matrix U is called **upper triangular** when its only possible nonzero entries are in or above the diagonal, or equivalently all entries "below" the diagonal vanish. In mathematical terms, $a_{ij} = 0$ whenever $i > j$.
If A and B are upper triangular matrices, AB is also upper triangular (in the case that AB makes sense).

Example 2. Perform the following matrix multiplications.

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

- $\begin{bmatrix} 1 & -4 & 3 \\ 0 & 1 & 6 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 4 & -27 \\ 0 & 1 & -6 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$

An $m \times n$ matrix L is called **lower triangular** when its only possible nonzero entries are in or below the diagonal, or equivalently all entries "above" the diagonal vanish. In mathematical terms, $a_{ij} = 0$ whenever $i < j$.
If A and B are lower triangular matrices, AB is also lower triangular (in the case that AB makes sense).

Example 3. Perform the following matrix multiplications.

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

- $\begin{bmatrix} 1 & 0 & 0 \\ -4 & 1 & 0 \\ 3 & 6 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 4 & 1 & 0 \\ -27 & -6 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$

Consider a scalar c and two $m \times n$ matrices

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

The **matrix sum** $A + B$ is defined to be the $m \times n$ matrix

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

and the **scalar multiple** cA is defined to be the $m \times n$ matrix

$$cA = \begin{bmatrix} ca_{11} & ca_{12} & \cdots & ca_{1n} \\ ca_{21} & ca_{22} & \cdots & ca_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ ca_{m1} & ca_{m2} & \cdots & ca_{mn} \end{bmatrix}.$$

Proposition 15.2.4. The linear transformations T_{A+B} and T_{cA} associated with the matrices $A + B$ and cA satisfy

$$\begin{aligned} T_{A+B}(\mathbf{x}) &= T_A(\mathbf{x}) + T_B(\mathbf{x}) \\ T_{cA}(\mathbf{x}) &= cT_A(\mathbf{x}) \end{aligned}$$

for all $\mathbf{x} \in \mathbb{R}^n$.

Properties of matrix algebra.

- **(MM1)** If A is an $m \times n$ matrix and $\mathbf{x} \in \mathbb{R}^n$, the matrix-vector product $A\mathbf{x}$ is the same as the matrix-matrix product when thinking of $\mathbf{x}$ as an $n \times 1$ matrix.
- **(MM2)** $A(B + C) = AB + AC$ and $(A' + B')C' = A'C' + B'C'$.
- **(MM3)** $A(BC) = (AB)C$ and $A(cB) = c(AB) = (cA)B$ for any scalar c. Note that this implies $(AB)\mathbf{x} = A(B\mathbf{x})$.
- **(MM4)** If A is an $m \times n$ matrix, then $I_m A = A = AI_n$.

Example 4. In Exercise 15.5, we introduce the idea of plugging an $n \times n$ matrix A into a single-variable polynomial $f(t)$. One of the reasons this idea is so useful is that the "identities" of single-variable algebra carry over to this setting.
For example, one such identity is $(t-2)(t+4) = t^2 + 2t - 8$.

(a) Use the properties of matrix algebra to symbolically verify that this identity is valid for every square matrix A. That is, verify that

$$(A - 2I_n)(A + 4I_n) = A^2 + 2A - 8I_n$$

for every $n \times n$ matrix A.

$$\begin{aligned}(A-2I_n)(A+4I_n) &= AA + A(4I_n) + (-2I_n)A + (-2I_n)(4I_n) \\ &= A^2 + 4AI_n - 2I_nA - 8I_n \\ &= A^2 + 4A - 2A - 8I_n \\ &= A^2 + 2A - 8I_n\end{aligned}$$

(b) More generally, it can be shown that whenever there is a polynomial identity of the form $f(t) = h(t)g(t)+k(t)$ for polynomials f, g, h, k, then for every $n \times n$ matrix A, we have $f(A) = h(A)g(A)+k(A)$ as $n \times n$ matrices.

To apply this, let $p(t) = t^3 + 3t + 39$ and $A = \begin{bmatrix} 1 & 2 \\ -5 & 2 \end{bmatrix}$ as in Exercise 15.5. Check that $p(t) = (t+3)g(t)+3$, where $g(t) = t^2 - 3t + 12$. Use this to compute $p(A)$ as an explicit 2×2 matrix without performing any further matrix multiplications. You may assume the result of Exercise 15.5(b), which is $g(A) = 0$.

Ex. 15.5 (b): if $g(t) = t^2 - 3t + 12$, then $g(A) = 0$.

We check that

$$\begin{aligned}(t+3)g(t) + 3 &= (t+3)(t^2 - 3t + 12) + 3 \\ &= (t^3 + 0t^2 + 3t + 36) + 3 \\ &= t^3 + 3t + 39 = p(t)\end{aligned}$$

Hence, $p(A) = (A + 3I_2)g(A) + 3I_2 = 3I_2 = \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix}$.

We did not need to compute $A^3 + 3A + 39I_2$!

Example 5. Let D be a 5×5 diagonal matrix with diagonal entries $d_1, \ldots, d_5 \in \mathbb{R}$. Suppose A is a 5×5 matrix whose second and fourth columns, respectively, are $\begin{bmatrix} 2 \\ -1 \\ 0 \\ 8 \\ 9 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ 8 \\ -9 \\ 3 \\ 1 \end{bmatrix}$. For each of the following, compute it in terms of the d_i's or explain why there is not enough information given to do so:

(a) The fourth column of DA

$$\begin{bmatrix} d_1 & & & & \\ & d_2 & & & \\ & & d_3 & & \\ & & & d_4 & \\ & & & & d_5 \end{bmatrix} \begin{bmatrix} | & 2 & | & 1 & | \\ | & -1 & | & 8 & | \\ | & 0 & | & -9 & | \\ | & 8 & | & 3 & | \\ | & 9 & | & 1 & | \end{bmatrix} = \begin{bmatrix} | & | & | & d_1 & | \\ | & | & | & 8d_2 & | \\ | & | & | & -9d_3 & | \\ | & | & | & 3d_4 & | \\ | & | & | & d_5 & | \end{bmatrix}$$

(b) The second column of AD

$$\begin{bmatrix} | & 2 & | & 1 & | \\ | & -1 & | & 8 & | \\ | & 0 & | & -9 & | \\ | & 8 & | & 3 & | \\ | & 9 & | & 1 & | \end{bmatrix} \begin{bmatrix} d_1 & & & & \\ & d_2 & & & \\ & & d_3 & & \\ & & & d_4 & \\ & & & & d_5 \end{bmatrix} = \begin{bmatrix} | & 2d_2 & | & | & | \\ | & -d_2 & | & | & | \\ | & 0 & | & | & | \\ | & 8d_2 & | & | & | \\ | & 9d_2 & | & | & | \end{bmatrix}$$

Example 6.

(a) Let $A = \begin{bmatrix} 4 \\ 2 \\ -2 \end{bmatrix}$ and $B = \begin{bmatrix} 3 & -5 & 2 \end{bmatrix}$. Compute BA and AB. (Note that the 1×1 matrix BA has as its single entry, the dot product, also sometimes called the **inner product**, of the two vectors; the 3×3 matrix product AB is then sometimes called the **outer product**.)

$$BA = \begin{bmatrix} 3 & -5 & 2 \end{bmatrix} \begin{bmatrix} 4 \\ 2 \\ -2 \end{bmatrix} = \begin{bmatrix} -2 \end{bmatrix} \quad \text{and} \quad AB = \begin{bmatrix} 4 \\ 2 \\ -2 \end{bmatrix} \begin{bmatrix} 3 & -5 & 2 \end{bmatrix} = \begin{bmatrix} 12 & -20 & 8 \\ 6 & -10 & 4 \\ -6 & 10 & -4 \end{bmatrix}.$$

(b) What do you notice about how the rows of AB are related to each other, and likewise for the columns?

The row vectors are scalar multiples of each other, and the column vectors are scalar multiples of each other as well.

(c) Let $C = \begin{bmatrix} 5 & -10 & 30 \\ 2 & -4 & 12 \\ -3 & 6 & -18 \end{bmatrix}$. Find a 3×1 matrix A and 1×3 matrix B for which $C = AB$. (Matrices arising in this way, as a nonzero column multiplied on the left against a nonzero row, are called "rank 1" matrices and are very important throughout data analysis.)

One such decomposition is $C = \begin{bmatrix} 5 \\ 2 \\ -3 \end{bmatrix} \begin{bmatrix} 1 & -2 & 6 \end{bmatrix}$.

There are infinitely many possibilities.

Example 7. Let $A = \begin{bmatrix} 1 & -1 \\ 2 & 1 \end{bmatrix}$ and $B = \begin{bmatrix} 2 & 0 \\ -1 & -2 \end{bmatrix}$. Denote the columns of A as $\mathbf{c}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ and $\mathbf{c}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$; denote the rows of B as $\mathbf{r}_1 = \begin{bmatrix} 2 & 0 \end{bmatrix}$ and $\mathbf{r}_2 = \begin{bmatrix} -1 & -2 \end{bmatrix}$.

(a) Compute AB directly.

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

(b) Symbolically writing $A = \begin{bmatrix} \mathbf{c}_1 & \mathbf{c}_2 \end{bmatrix}$ and $B = \begin{bmatrix} \mathbf{r}_1 \\ \mathbf{r}_2 \end{bmatrix}$, if we were to treat the $\mathbf{c}_i$'s and $\mathbf{r}_j$'s as if they were scalars (which they aren't!) then we might wonder if AB is equal to

$$\begin{bmatrix} \mathbf{c}_1 & \mathbf{c}_2 \end{bmatrix}\begin{bmatrix} \mathbf{r}_1 \\ \mathbf{r}_2 \end{bmatrix} = \mathbf{c}_1\mathbf{r}_1 + \mathbf{c}_2\mathbf{r}_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}\begin{bmatrix} 2 & 0 \end{bmatrix} + \begin{bmatrix} -1 \\ 1 \end{bmatrix}\begin{bmatrix} -1 & -2 \end{bmatrix}.$$

Check that this really gives the right answer!

We compute

$$\begin{bmatrix} 1 \\ 2 \end{bmatrix}\begin{bmatrix} 2 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 4 & 0 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} -1 \\ 1 \end{bmatrix}\begin{bmatrix} -1 & -2 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ -1 & -2 \end{bmatrix}.$$

Indeed, $\begin{bmatrix} 2 & 0 \\ 4 & 0 \end{bmatrix} + \begin{bmatrix} 1 & 2 \\ -1 & -2 \end{bmatrix} = \begin{bmatrix} 3 & 2 \\ 3 & -2 \end{bmatrix}$.

(c) In general, if A is an $m \times n$ matrix with columns $\mathbf{c}_1, \ldots, \mathbf{c}_n$ and B is an $n \times p$ matrix with rows $\mathbf{r}_1, \ldots, \mathbf{r}_n$, then the $m \times p$ product matrix AB is always equal to

$$AB = \mathbf{c}_1\mathbf{r}_1 + \mathbf{c}_2\mathbf{r}_2 + \cdots + \mathbf{c}_n\mathbf{r}_n.$$

Ultimately, this is seen via a close inspection of the mechanics of the process of matrix multiplication.

Use this to compute $\begin{bmatrix} 1 & 7 & 0 \\ 4 & -1 & -1 \\ -3 & 0 & 2 \end{bmatrix}\begin{bmatrix} 0 & 5 & -1 \\ -2 & 3 & 0 \\ 2 & 0 & -3 \end{bmatrix}$.

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

$$= \begin{bmatrix} 0 & 5 & -1 \\ 0 & 20 & -4 \\ 0 & -15 & 3 \end{bmatrix} + \begin{bmatrix} -14 & 21 & 0 \\ 2 & -3 & 0 \\ 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ -2 & 0 & 3 \\ 4 & 0 & -6 \end{bmatrix} = \boxed{\begin{bmatrix} -14 & 26 & -1 \\ 0 & 17 & -1 \\ 4 & -15 & -3 \end{bmatrix}}.$$