Lecture 5 : Matrices and Matrix Operations

Math 3013
Oklahoma State University

January 21, 2022

Lecture 5 : Matrices and Matrix Operations

Agenda:

1. Matrix Multiplication
2. Examples of Matrix Multiplication
3. Other Matrix Operations

Matrices

Recall

Definition

An $n \times m$ **matrix** is a rectangular arrangement of nm real numbers with n rows and m columns.

The usual way we will indicate a generic $n \times m$ matrix:

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \ddots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{bmatrix}$$

Here a_{ij} is the entry in the i^{th} row (reading left to right) and j^{th} column (reading top to bottom).

Matrix Multiplication

• Suppose **A** is a matrix with m columns and **B** is a matrix with m rows. Then the matrix product **AB** is defined.
If **A** is an $n \times m$ matrix and **B** is an $m \times p$ matrix, the matrix product **AB** is the $n \times p$ matrix with entries

$$(\mathbf{AB})_{ij} = Row_i(\mathbf{A}) \cdot Col_j(\mathbf{B})$$

• If the number of columns of the first factor matrix **A** is not the same as the number of rows of the second factor matrix **B**, then the product **AB** is **undefined**.)
• the matrix dimensions follow the rule

$$(n \times m) \cdot (m \times p) \longrightarrow (n \times p)$$

Examples of Matrix Multiplication

Example

$$\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 2 & -1 \\ 1 & -2 \end{bmatrix} \qquad \text{is undefined}$$

since

$$\underset{\text{mismatch}}{(3 \times \underset{\uparrow}{1})(\underset{\uparrow}{3} \times 2)} \quad \longrightarrow \quad \text{undefined matrix product}$$

(matrix multiplication requires the number of columns in the first factor to equal the number of rows in the second factor)

Example: Multiplying a Vector by Matrix

If $\mathbf{a} \equiv (a_1, \ldots, a_n) \in \mathbb{R}^n$ is an ordered list of n numbers, there are two different ways of interpreting $\mathbf{A}$ as a matrix

$$\begin{array}{ccc} & & \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} \quad \text{(an } n \times 1 \text{ matrix)} \\ & \nearrow & \\ (a_1, \ldots, a_n) & & \\ & \searrow & \\ & & [a_1, \ldots, a_n] \quad \text{(an } 1 \times n \text{ matrix)} \end{array}$$

Example: Multiplying a Vector by Matrix, Example

$$\begin{bmatrix} 1 & 2 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} \qquad \text{is defined}$$

since for this product we have

$$(2 \times 2)(2 \times 1) \quad \longrightarrow \quad (2 \times 1)$$

$$\uparrow \quad \uparrow$$

match

Explicitly computing the product:

$$\begin{bmatrix} 1 & 2 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} (1,2) \cdot (1,-1) \\ (-1,2) \cdot (1,-1) \end{bmatrix}$$

$$= \begin{bmatrix} (1)(1) + (2)(-1) \\ (-1)(1) + (2)(-1) \end{bmatrix}$$

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

On the other hand,

$$\begin{aligned}
\begin{bmatrix} 1 & -1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ -1 & 2 \end{bmatrix} &= [(1,-1)\cdot(1,-1) \quad (1,-1)\cdot(2,2)] \\
&= \begin{bmatrix} (1)(1)+(-1)(-1) & (1)(2)+(-1)(2) \end{bmatrix} \\
&= \begin{bmatrix} 2 & 0 \end{bmatrix}
\end{aligned}$$

Comparing these two results

$$\begin{aligned}
\begin{bmatrix} 1 & 2 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} &= \begin{bmatrix} -1 \\ -3 \end{bmatrix} \\
\begin{bmatrix} 1 & -1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ -1 & 2 \end{bmatrix} &= [2 \quad 0]
\end{aligned}$$

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

Example

Let $\mathbf{A} = \begin{bmatrix} 2 & 1 \\ -1 & 1 \end{bmatrix}$ and $\mathbf{B} = \begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix}$.

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

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

So

$$\mathbf{AB} \neq \mathbf{BA}$$

Thus, if the order of factors changes the value of a product of matrices can also change.

Indeed, it can happen that a product **AB** exists but **BA** is not even defined:

If **A** is an $n \times m$ matrix and **B** is a $m \times p$ matrix, then

$$\mathbf{AB} \quad \sim \quad (n \times m)(m \times p) \quad \longrightarrow \quad (n \times p)$$

so the product **AB** is defined and it will be a $n \times p$ matrix.
On the other hand,

$$\mathbf{BA} \qquad \sim \qquad (m \times p)(n \times m) \longrightarrow \text{undefined}$$

$$\uparrow \quad \uparrow$$

match only when $p = n$

Moral: One must always maintain the order of factors when multiplying matrices.

Another deviation from the ordinary multiplication of numbers

Recall that for real numbers $xy = 0$ implies either $x = 0$ or $y = 0$. Now consider

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

This example shows that for matrices, it can happen that $\mathbf{AB} = \mathbf{0}$ with neither $\mathbf{A}$ or $\mathbf{B}$ equal to the zero matrix $\mathbf{0}$.

The Identity Matrix

Consider the following matrix products

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

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

And so multiplying any 3×3 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 $\mathbf{A}$:

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

The Identity Matrix, Cont'd

The preceding example generalizes to arbitrary $n \times n$ matrices (usually called "square matrices") and motivates the following definition.

Definition

The $n \times n$ **Identity matrix** is the the $n \times n$ matrix $\mathbf{I}$ whose entries are given by

$$I_{ij} = \begin{cases} 1 & \text{if} \quad i = j \\ 0 & \text{if} \quad i \neq j \end{cases} \qquad \Longleftrightarrow \qquad \mathbf{I} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \vdots \\ \vdots & & \ddots & \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$

In other words, $\mathbf{I}$ is the $n \times n$ matrix with 1's along the diagonal (running from the upper left to the lower right) and 0's everywhere else.

The Identity Matrix, Cont'd

The identity matrix $\mathbf{I}$ has the property that

$$\begin{aligned} \mathbf{AI} &= \mathbf{A} \quad \text{whenever } \mathbf{AI} \text{ is defined} \\ \mathbf{IA} &= \mathbf{A} \quad \text{whenever } \mathbf{IA} \text{ is defined} \end{aligned}$$

Addition and Scalar Multiplication of Matrices

Definition

Suppose $\mathbf{A}$ and $\mathbf{B}$ are both $n \times m$ matrices. Then the **matrix sum** $\mathbf{A} + \mathbf{B}$ is defined as the $n \times m$ matrix with entries

$$(\mathbf{A} + \mathbf{B})_{ij} = A_{ij} + B_{ij}$$

Definition

Let $\lambda \in \mathbb{R}$, and let $\mathbf{A}$ be an $n \times m$ matrix. Then the **scalar product** of $\mathbf{A}$ by λ is the $n \times m$ matrix $\lambda\mathbf{A}$ with entries

$$(\lambda\mathbf{A})_{ij} = \lambda A_{ij}$$

Example

Let

$$\mathbf{A} = \begin{bmatrix} 1 & -1 & 2 \\ 1 & 0 & 1 \end{bmatrix} \quad , \quad \mathbf{B} = \begin{bmatrix} 0 & 1 & -1 \\ 2 & 1 & 0 \end{bmatrix}$$

Then

$$\mathbf{A} + \mathbf{B} = \begin{bmatrix} 1+0 & -1+1 & 2-1 \\ 1+2 & 0+1 & 1+0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 \\ 3 & 1 & 1 \end{bmatrix}$$

$$3\mathbf{A} = \begin{bmatrix} (3)(1) & (3)(-1) & (3)(2) \\ (3)(1) & (3)(0) & (3)(1) \end{bmatrix} = \begin{bmatrix} 3 & -3 & 6 \\ 3 & 0 & 3 \end{bmatrix}$$

The Transpose of a Matrix

Definition

Suppose $\mathbf{A}$ is an $n \times m$ matrix. The **transpose** of $\mathbf{A}$ is the $m \times n$ matrix $\mathbf{A}^t$ with entries

$$\left(\mathbf{A}^t\right)_{ij} = A_{ji} \quad ; \quad \text{for } i = 1, \dots, m \quad j = 1, \dots, n$$

(note how the row and column indices have been reversed). Equivalently, $\mathbf{A}^t$ is the matrix obtained from $\mathbf{A}$ by converting, in order, the rows of $\mathbf{A}$ into columns.

Example

If

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

then

$$\mathbf{A}^t = \begin{bmatrix} 1 & -2 & 3 \\ 3 & 1 & -1 \end{bmatrix}$$

Remarks

- If $\mathbf{A}$ is an $n \times m$ matrix then $\mathbf{A}^2 = \mathbf{AA}$ is defined only if $m = n$. However, the products $\mathbf{AA}^t$ and $\mathbf{A}^t\mathbf{A}$ are always well-defined.

$$\begin{aligned} \mathbf{AA}^t &\sim (n \times m \text{ matrix})\,(m \times n \text{ matrix}) = (n \times n \text{ matrix}) \\ \mathbf{A}^t\mathbf{A} &\sim (m \times n \text{ matrix})\,(n \times m \text{ matrix}) = (m \times m \text{ matrix}) \end{aligned}$$

- If $\mathbf{v}$ is a column vector (i.e., an $n \times 1$ matrix), then $\mathbf{v}^t$ is a row vector (i.e, a $1 \times n$ matrix) and

$$\begin{aligned} \mathbf{v}^t\mathbf{v} &= [v_1 \;\; v_2 \;\; \cdots \;\; v_n] \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} \\ &= \left[(v_1)^2 + (v_2)^2 + \cdots + (v_n)^2\right] = \left[\|\mathbf{v}\|^2\right] \end{aligned}$$

Remarks, Cont'd

- A **square matrix** is a matrix with the same number of rows as it has columns.
- A square matrix $\mathbf{A}$ is called **symmetric** if $\mathbf{A}^t = \mathbf{A}$.

Example:

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

is a symmetric matrix.

Properties of the Transpose Operation

Theorem

Let $\mathbf{A}$ *and* $\mathbf{B}$ *be matrices and let* $\lambda \in \mathbb{R}$*.*

- $\left(\mathbf{A}^t\right)^t = \mathbf{A}$
- $(\lambda \mathbf{A})^t = \lambda \mathbf{A}^t$
- $(\mathbf{A} + \mathbf{B})^t = \mathbf{A}^t + \mathbf{B}^t$
- $(\mathbf{AB})^t = \mathbf{B}^t \mathbf{A}^t$ *(Note how the order of factors changed.)*

Back to Linear Systems

Recall matrices were introduced as a means of writing large systems as linear equations as a single matrix equation.

$$\left.\begin{array}{ccc} a_{11}x_1 + & \cdots & + a_{1m}x_m = b_1 \\ \vdots & & \vdots \\ a_{n1}x_1 + & \cdots & + a_{nm}x_m = b_n \end{array}\right\} \iff \mathbf{A}\mathbf{x} = \mathbf{b}$$

where

$$\mathbf{A} = \begin{bmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nm} \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_m \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}$$

Case 0: 0 equations in m unknowns

In this case, there are no conditions placed on the variables $x_1, \ldots, x_m$; and so each variable is allowed to range over the entire real line. The "solution vectors" are just

$$\left\{ \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix} \mid x_1, x_2, \ldots, x_m \in \mathbb{R} \right\} = \mathbb{R}^m$$

Case 1: 1 equation in m unknowns

We now impose a single linear equation

$$a_1 x_1 + a_2 x_2 + \cdots + a_m x_m = b$$

on m variables $x_1, \ldots, x_m$. So long as $a_m \neq 0$ (so that the term $a_m x_m$ actually contributes to the equation), this equation is readily solved for x_m :

$$x_m = \frac{b}{a_m} - \frac{a_1}{a_m} x_1 - \cdots - \frac{a_{m-1}}{a_m} x_{m-1}$$

Thus, a solution vector will be a vector of the form

$$
\begin{aligned}
\mathbf{x} &= \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{m-1} \\ \frac{b}{a_m} - \frac{a_1}{a_m}x_1 - \cdots - \frac{a_{m-1}}{a_m}x_{m-1} \end{bmatrix} \\
&= \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ \frac{b}{a_m} \end{bmatrix} + \begin{bmatrix} x_1 \\ 0 \\ \vdots \\ 0 \\ -\frac{a_1}{a_m}x_1 \end{bmatrix} + \cdots + \begin{bmatrix} 0 \\ 0 \\ \vdots \\ x_{m-1} \\ -\frac{a_{m-1}}{a_m}x_{m-1} \end{bmatrix}
\end{aligned}
$$

or

$$\mathbf{x} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ \frac{b}{a_m} \end{bmatrix} + x_1 \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ -\frac{a_1}{a_m} \end{bmatrix} + \cdots + x_{m-1} \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ -\frac{a_{m-1}}{a_m} \end{bmatrix}$$

Note how in this last form, the solution vector is expressed as a point on a $(m-1)$-dimensional hyperplane of vectors; i.e., a set of vectors of the form

$$\{\mathbf{p}_0 + t_1\mathbf{d}_1 + \cdots + t_{m-1}\mathbf{d}_{m-1} \mid t_1, \ldots, t_{m-1} \in \mathbb{R}\}$$

Geometric Construction of Solutions

We have seen that the solution space of a linear equation in m variables is a $(m-1)$-dimensional hyperplane in the vector space of variable values.

Thus, to satisfy a two linear equations in m variables, a point $\mathbf{x}$ must lie on both of the corresponding hyperplanes.

Satisfying multiple equations means that solution points live on multiple hyperplanes.

Geometric Construction of Solutions, Cont'd

Let $Eq_1, \ldots, Eq_n$ be a system of n linear equations in m unknowns. and let $\mathcal{H}_1, \ldots, \mathcal{H}_n$ be the corresponding hyperplanes

$$\mathcal{H}_i = \text{solution set of } Eq_i \quad , \quad i = 1, \ldots, n$$

Then a point in the solution set of a linear system must simultaneously live on each hyperplane $\mathcal{H}_i$
Thus, geometrically,

$$\text{solution set} = \bigcap_{i=1}^{n} \mathcal{H}_i$$

i.e., a point in the solution set must lie in the intersection of all the equation-hyperplanes.

Solving Systems of Linear Equations

$\longrightarrow$ **Monday's Lecture**