Intro to Matrices

Chris Tisdell

youtube.com/DrChrisTisdell

Where are we going?

- We will learn about an new area of mathematics, called “matrices”.
- Matrices give us a simplifying framework with which we can write linear systems in very simple terms.

As we know, linear systems arise frequently, for example, in:

- engineering (mechanical vibrations and control)
- economics (supply / demand dynamics)
- life–sciences (predator–prey models)
- technology (graphics in screens, printing).

What is a matrix?

A *matrix* is rectangular block of numbers, formed into rows and columns

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

The number a_{ij} denotes the entry of A that occupies the ith row and jth column.
A matrix with m rows and n columns is said to have size $m \times n$ (read "m by n").

1. An $m \times 1$ matrix is called a column vector, while a $1 \times n$ matrix is called a row vector.
2. An $n \times n$ matrix is called a square matrix.
3. We use M_{mn} to stand for the set of all $m \times n$ matrices. When all the entries of the matrices are real, we denote the set by $M_{mn}(\mathbb{R})$.
4. When we say "let $A = (a_{ij})$", we mean we know the size of A and we let a_{ij} be the entry in the ith row and jth column for all i, j.
5. On the other hand, for a given matrix A, we denote the entry in the ith row and jth column by $[A]_{ij}$.
6. The *zero* matrix (written $\mathbf{0}$) is a matrix in which every entry is zero.

Addition of matrices

If A and B are $m \times n$ matrices, then the *sum* $C = A + B$ is the $m \times n$ matrix whose entries are

$$[C]_{ij} = [A]_{ij} + [B]_{ij} \qquad \text{for all} \quad i, j.$$

that is

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} + \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m1} & b_{m2} & \cdots & b_{mn} \end{pmatrix}$$

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

Note that the sum of two matrices with different sizes is not defined.

Let A, B, C be matrices, $\mathbf{0}$ be a zero matrix. Let λ, μ be scalars. Assume all the expressions are defined. The following rules are true:

1. $A + B = B + A$.
2. $(A + B) + C = A + (B + C)$.
3. $A + \mathbf{0} = \mathbf{0} + A = A$.

Example: If

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

and

$$C = \begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix}$$

then compute $A + C$.

Scalar multiplication with a matrix

If λ is a scalar, then the *scalar multiple* $B = \lambda A$ is the $m \times n$ matrix whose entries are

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

In other words

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

For any matrix A, the *negative* of A is the matrix $-A$ of the same size with entries $[-A]_{ij} = -[A]_{ij}$

Let $A,\ B,\ C$ be matrices, $\mathbf{0}$ be a zero matrix. Let λ, μ be scalars. Assume all the expressions are defined. The following rules are true:

1. $A + (-A) = (-A) + A = \mathbf{0}$
2. $\lambda(\mu A) = (\lambda\mu)A$
3. $\lambda(A + B) = \lambda A + \lambda B$
4. $(\lambda + \mu)A = \lambda A + \mu A$.

Example: If $A = \begin{pmatrix} 2 & 1 \\ 0 & -2 \end{pmatrix}$ and $C = \begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix}$ then compute $3A - C$.

Matrix multiplication is a row–times–column process: we get the (row i, column j) entry of AB by going across the ith row of A and down the jth column of B, multiplying and adding as we go.

If A is an $m \times n$ matrix and B is an $n \times p$ matrix, then the *product* AB is the $m \times p$ matrix whose entries are given by the formula

$$\begin{aligned} [AB]_{ij} &= [A]_{i1}[B]_{1j} + \cdots + [A]_{in}[B]_{nj} \\ &= \sum_{k=1}^{n} [A]_{ik}[B]_{kj} \quad \text{for } 1 \leqslant i \leqslant m,\ 1 \leqslant j \leqslant p. \end{aligned}$$

Let A be an $m \times n$ matrix and B be an $r \times s$ matrix. The product AB is defined **only** when the number of columns of A is the same as the number of rows of B, i.e. $n = r$. If $n = r$, the size of AB will be $m \times s$.

Let A, B, C be matrices. Let λ, μ be scalars. Assume all the expressions are defined. The following rules are true:

1. $(AB)C = A(BC)$
2. $(A + B)C = AC + BC$
3. $A(B + C) = AB + AC$
4. $AB \neq BA$ in general.

Example: If $A = \begin{pmatrix} 2 & 1 \\ 0 & -2 \end{pmatrix}$ and $C = \begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix}$ then compute AC.

The identity matrix

An *identity* matrix (written I) is a square matrix with 1s on the diagonal and 0s off the diagonal, with the diagonal running top to bottom, left to right. The identity matrix satisfies

$$AI = IA = A$$

for all square matrices A.
Example: For all 2×2 matrices A, verify that $AI = IA = A$ for the identity matrix $I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.

Transpose of a matrix

The *transpose* of an $m \times n$ matrix A is the $n \times m$ matrix A^T (read "A transpose") with entries given by

$$[A^T]_{ij} = [A]_{ji}.$$

An easy way to remember the transpose is the switch the first row to become the first column, the 2nd row to become the 2nd column and so on.

Properties of the Transpose

Let A, B be matrices of the same size and λ, μ be scalars. We have:

1. $(A^T)^T = A$
2. $(\lambda A + \mu B)^T = \lambda A^T + \mu B^T$
3. $(AB)^T = B^T A^T$.

A matrix X is said to be symmetric if $X^T = X$.

a) Explain why X must be a square matrix.

b) Let A be an $n \times n$ matrix. Prove that $A + A^T$ is symmetric.

Inverse of a matrix

We solve the equation $ax = b$ by dividing both sides by a. How about solving a matrix equation $AX = B$? We cannot divide by a matrix, but we do have inverses for certain matrices.

A matrix X is said to be an *inverse* of a matrix A if and only if both

$$AX = I \qquad \text{and} \qquad XA = I,$$

where I is an identity (or unit) matrix of the appropriate size. When a matrix has an inverse, the matrix is said to be *invertible*.

Properties of Inverse.

1. The inverse of an invertible matrix is unique. We shall denote the inverse of A by A^{-1}.
2. All invertible matrices are square. However, not all square matrices are invertible.
3. When A is a square matrix, if $AX = I$ or $XA = I$ then $X = A^{-1}$.
4. If A is invertible, then $(A^{-1})^{-1} = A$.
5. If A and B are invertible, then $(AB)^{-1} = B^{-1}A^{-1}$.
6. If A is invertible, then $(A^T)^{-1} = (A^{-1})^T$.

How to calculate the inverse of a matrix.

Finding inverse of a square matrix A.

1. Reduce the augmented matrix $(A|I)$ to row–echelon form $(U|C)$. If U has a zero row, then A is not invertible.
2. Otherwise, reduce $(U|C)$ to reduced row–echelon form $(I|B)$. The inverse of A is B.

If A is invertible then $A\mathbf{x} = \mathbf{b}$ has a unique solution for all $\mathbf{b}$.

Find the inverse of

$$\begin{pmatrix} 1 & 3 \\ 2 & 0 \end{pmatrix}.$$

Determinant of a matrix

The *determinant* of a square matrix A, denoted by $\det(A)$ or $|A|$, is defined recursively as follows.

- If $A = (a)$, define $\det(A) = a$.
- If $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, then $\det(A) = ad - bc$.
- If A is an $n \times n$ matrix, then

$$\begin{aligned} |A| &= a_{11}|A_{11}| - a_{12}|A_{12}| + a_{13}|A_{13}| \\ &\qquad + \cdots + (-1)^{n-1} a_{1n}|A_{1n}| \\ &= \sum_{k=1}^{n} (-1)^{1+k} a_{1k}|A_{1k}|, \end{aligned}$$

where $|A_{ij}|$ is the determinant of the matrix obtained from A by deleting row i and column j from A. The determinant $|A_{ij}|$ is called the (row i, column j) *minor* of A.

More on determinants

1. $\det\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}$ can be written as $\begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix}$ when $n > 1$.
2. The definition above for the $n \times n$ determinant is also called expanding along the first row.
3. $\det(I) = 1$.

Properties of determinants

- $\det(A^T) = \det(A)$.
- If B is the matrix formed by interchanging two rows (or columns) of A, then $\det(B) = -\det(A)$. Hence, we can evaluate a determinant by expanding along any row or any column by

$$|A| = \sum_{k=1}^{n} (-1)^{i+k} a_{ik} |A_{ik}| = \sum_{k=1}^{n} (-1)^{k+j} a_{kj} |A_{kj}|.$$

Practically, we should choose a row or a column with most of the entries 0 and the signs chosen from the following array.

$$\begin{pmatrix} + & - & + & - & \cdots \\ - & + & - & + & \cdots \\ + & - & + & - & \cdots \\ - & + & - & + & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

- If A and B are identical except that the ith row (or the jth column) of B is λ times the ith row (or the jth column) of A, then $\det(B) = \lambda \det(A)$, that is

$$\begin{vmatrix} a & b & c \\ d & e & f \\ \lambda h & \lambda i & \lambda j \end{vmatrix} = \lambda \begin{vmatrix} a & b & c \\ d & e & f \\ h & i & j \end{vmatrix}.$$

- If B is obtained from A by adding a multiple of a row (column) to another row (column), then $\det(A) = \det(B)$.
- $\det(AB) = \det(A)\det(B)$.

The above properties provide us a mean to perform the row operations and keep track of how the determinant changes.

The determinant of a square matrix in row-echelon form is easy to evaluate. As illustrated by the following example, we can evaluate a determinant efficiently by row–reduction.

Find $|A| = \begin{vmatrix} a & 2a & -3a \\ 1 & a+3 & 2a-1 \\ 4 & 4-a & 4a+4 \end{vmatrix}$. Determine the values of a that ensure the linear system $A\mathbf{x} = \mathbf{b}$ has a solution for each vector $\mathbf{b}$

Prove that $\det(A^{-1}) = \dfrac{1}{\det(A)}$, if A is invertible.