# Chapter 5: Matrices

Daniel Chan

UNSW

Term 1 2022

# In this chapter

- Matrices were first introduced in the Chinese "Nine Chapters on the Mathematical Art" to solve linear eqns.
- In the mid-1800s, senior wrangler Arthur Cayley studied matrices in their own right and showed how they have an interesting and useful algebra associated to them.
- We will look at Cayley's ideas and extend vector arithmetic to matrices and even show there is matrix multiplication akin to multiplying numbers.
- These ideas will not only shed light on solving linear eqns, they will also be useful later when you look at multivariable functions and mappings.

# Some new notation for matrices

Recall an $m \times n$-matrix is an array of (for us) scalars (real or complex).

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

## Notation

- We abbreviate the above to $A = (a_{ij})$ and call $a_{ij}$ the *ij-th entry of A*.
- Also write $[A]_{ij}$ for $a_{ij}$.
- We say the *size of A* is $m \times n$ because it has
- $M_{mn}(\mathbb{R})$ (resp $M_{mn}(\mathbb{C})$) denote the set of all $m \times n$-matrices with real entries (resp complex entries). Sometimes abbreviate to $M_{mn}$ if the scalars are understood or irrelevant.

**E.g.** A length $m$ *column vector* is an

# Revise matrix-vector product

Let $A = (a_{ij}) = (\boldsymbol{a}_1|\boldsymbol{a}_2| \ \dots|\boldsymbol{a}_n) \in M_{mn}$. Then

$$A\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \qquad\qquad\qquad\qquad = x_1\boldsymbol{a}_1 + x_2\boldsymbol{a}_2 + \dots + x_n\boldsymbol{a}_n.$$

Alternatively, the $i$-th entry of $A\boldsymbol{x}$ is

$$[A\boldsymbol{x}]_i = a_{i1}x_1 + \dots + a_{in}x_n = (a_{i1} \ \dots \ a_{in})\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.$$

Note similarity with dot products.

$A$ induces the linear function $T : \mathbb{R}^n \longrightarrow \mathbb{R}^m : \boldsymbol{x} \mapsto A\boldsymbol{x}$.
**Note** We will write all our results for matrices with real entries, but there are obvious analogues over the complexes.

# Arithmetic of matrices

Just as for vectors, we can define matrix addition and scalar multiplication to be entry-wise addition and scalar multiplication

**E.g.**

$$\begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & -1 \end{pmatrix} + \begin{pmatrix} 3 & 4 & 6 \\ 2 & -1 & 5 \end{pmatrix} =$$

$$7 \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & -1 \end{pmatrix} =$$

In formulas

## Matrix arithmetic

For $A, B \in M_{mn}(\mathbb{R}), \lambda \in \mathbb{R}$, the entries of $A + B, \lambda A \in M_{mn}(\mathbb{R})$ are

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

**N.B.** We don't define the sum of matrices of different sizes (just as is the case for vectors).

# Linear combinations and subtraction

**E.g.** We can also form linear combinations of matrices

$$\begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & -1 \end{pmatrix} + (-1)\begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & -1 \end{pmatrix} =$$

## Definition

- The *zero matrix* $\mathbf{0}$ has all entries 0. (There's one for each size $m \times n$.)
  $A + \mathbf{0} =$
- The *negative of* $A \in M_{mn}$ is $-A := (-1)A$. Hence $A + (-A) =$
- The *difference* $A - B = A + (-B)$ if $A, B$ have the same size.

# Another distributive & associative law

## Proposition

For $A, B \in M_{mn}(\mathbb{R}), \lambda \in \mathbb{R}, \boldsymbol{x} \in \mathbb{R}^n$

- $(A+B)\boldsymbol{x} = A\boldsymbol{x} + B\boldsymbol{x}$.
- $(\lambda A)\boldsymbol{x} = \lambda(A\boldsymbol{x})$.

**Proof.** Suppose $n = 2$ (else need more space) so $A = (\boldsymbol{a}_1|\boldsymbol{a}_2), B =$

**Upshot** Recall that in calculus, you define the sum and scalar multiple of functions pointwise, $(f+g)(x) = f(x) + g(x), (\lambda f)(x) = \lambda f(x)$.
The above formulas show that the linear function corresponding to $A + B$ which sends $\boldsymbol{x} \mapsto (A+B)\boldsymbol{x} = A\boldsymbol{x} + B\boldsymbol{x}$ is the pointwise sum of the functions corresponding to $A$ and $B$. The same goes for the scalar multiple.

# Basic properties of matrix arithmetic

## Proposition

For $A, B \in M_{mn}$, and scalars $\lambda, \mu$

- $\lambda(\mu A) = (\lambda\mu)A$.
- $(\lambda + \mu)A =$
- $\lambda(A + B) =$

**Proof.** Just as for vectors e.g.

# Geometric example of linear combinations

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

$C := \frac{1}{2}(A + B)$

# Matrix multiplication

Let $A \in M_{mn}, B = (\boldsymbol{b}_1 | \dots | \boldsymbol{b}_p) \in M_{np}$. We define the *matrix product AB* to be the $m \times p$-matrix

$$AB = (A\boldsymbol{b}_1 | \dots | A\boldsymbol{b}_p) \in M_{mp}.$$

**E.g.** $\begin{pmatrix} 0 & 1 \\ 2 & 3 \\ 4 & 5 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 1 & 2 \end{pmatrix}$

Alternatively, the $ij$-th entry of $AB$ comes from "zipping up" the $i$th row of $A$ with the $j$th column of $B$: i.e. if $A = (a_{ij}), B = (b_{ij})$

$$[AB]_{ij} = (a_{i1} \ \dots \ a_{in}) \begin{pmatrix} b_{1j} \\ \vdots \\ b_{nj} \end{pmatrix} = a_{i1}b_{1j} + \dots + a_{in}b_{nj} = \sum_{l=1}^{n} a_{il}b_{lj}.$$

**Warning** The product $AB$ is only defined when no. columns $A$ = no. rows $B$.

# Associative law

## Associative law of matrix multiplication

Let $A \in M_{mn}, B = (\boldsymbol{b}_1 | \dots | \boldsymbol{b}_p) \in M_{np}, C = (\boldsymbol{c}_1 | \dots | \boldsymbol{c}_q) \in M_{pq}$. Then $(AB)C = A(BC)$.

**Proof.** It suffices show this for $C = \boldsymbol{c} = \begin{pmatrix} c_1 \\ \vdots \\ c_p \end{pmatrix}$ for assuming this case we see

$$\begin{aligned}(AB)C = ((AB)\boldsymbol{c}_1 | \dots | (AB)\boldsymbol{c}_q) &= (A(B\boldsymbol{c}_1) | \dots | A(B\boldsymbol{c}_q)) \\ &= A((B\boldsymbol{c}_1) | \dots | (B\boldsymbol{c}_q)) = A(BC).\end{aligned}$$

If $C = \boldsymbol{c}$ then

$$(AB)\boldsymbol{c} = (A\boldsymbol{b}_1 | \dots | A\boldsymbol{b}_p)\boldsymbol{c} = c_1 A\boldsymbol{b}_1 + \dots + c_p A\boldsymbol{b}_p = A(c_1\boldsymbol{b}_1 + \dots + c_p\boldsymbol{b}_p) = A(B\boldsymbol{c}).$$

# Functional interpretation of the associative law

The associative law says the function associated to $AB$ which maps $\boldsymbol{x} \mapsto (AB)\boldsymbol{x} = A(B\boldsymbol{x})$ is the composite $\boldsymbol{x} \mapsto B\boldsymbol{x} \mapsto A(B\boldsymbol{x})$ of the linear maps associated to $A$ and $B$.

**E.g.** Recall that $B = \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$ corresponds to reflection about the $x$-axis. The functional viewpoint shows that $B^2$ coresponds to the mapping

Let's check $\left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)\left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right) =$

**Remark** The definition of matrix multiplication was designed so that it reflects the composition of linear maps.

# Distributive laws & noncommutativity

## Distributive law

Let $A, B, C$ be matrices & $\lambda$ a scalar. The following formulas hold whenever the terms on one side are defined.

1. $A(B+C) = AB + AC$.
2. $(A+B)C = AC + BC$.
3. $(\lambda A)B = \lambda(AB) = A(\lambda B)$.

**Proof.** Easy ex similar to distributive law we proved for matrix-vector product.

**Noncommutativity** Note that if $AB$ is defined, $BA$ may not be, and even if it is, usually we have $AB \neq BA$.
**E.g.** Let $A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$, $B = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$

Hence $(A+B)^2 =$

# Miscellaneous matrix definitions

A matrix is said to be *square* if its no. rows = no. columns. The *diagonal* of a square matrix $A = (a_{ij})$ consists of the entries $a_{ii}$. The *$n \times n$-identity matrix $I$*, is the matrix with 1's on the diagonal and 0's elsewhere.
**E.g.**

## Formula

$IA = A, BI = B$ whenever the products are defined.

**Proof.** Just multiply matrices. We'll check here

**Upshot** In particular, we see the linear function associated to $I$ is the identity map $\boldsymbol{x} \mapsto \boldsymbol{x}$.

# Transpose

The *transpose* of an $m \times n$-matrix $A$, is the $n \times m$-matrix $A^T$ gotten by turning all the rows of $A$ into columns (or equivalently, flipping the matrix about the row $i =$ column $j$ diagonal).

**E.g.** $\begin{pmatrix} 2 & 5 & 1 \\ 1 & 3 & 2 \end{pmatrix}^T =$

More formally, the entries of $A^T$ are given by $[A^T]_{ij} = [A]_{ji}$.

## Formulas

Let $A, B$ be matrices & $\lambda$ a scalar. The following hold when one side is defined.

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

**Proof.** 1) & 2) are easy and say the function $A \mapsto A^T$ is linear. For 3)

$$[(B^T A^T)]_{ij} = \sum_l [B^T]_{il}[A^T]_{lj} = \sum_l [B]_{li}[A]_{jl} =$$

**Relation with dot product**

Let $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^n$. Then though $\boldsymbol{ab}$ is not defined we can define

$$\boldsymbol{a}^T \boldsymbol{b} = (a_1 \dots a_n) \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} = a_1 b_1 + \dots + a_n b_n = \boldsymbol{a} \cdot \boldsymbol{b}.$$

**Symmetric matrices**

A square matrix is *symmetric* if $A^T = A$ and *anti-symmetric* if $A^T = -A$ (Why is square in the defn?).

**E.g.** $\boldsymbol{bb}^T \in M_{nn}$ is symmetric since

In fact if $|\boldsymbol{b}| = 1$, then the linear function $\mathbb{R}^n \longrightarrow \mathbb{R}^n$ associated to $\boldsymbol{bb}^T$ is projection onto $\boldsymbol{b}$ for it sends

$$\boldsymbol{x} \mapsto \boldsymbol{bb}^T \boldsymbol{x} =$$

# Matrix Inverse

To attempt matrix "division" we need the matrix inverse.

## Inverse

A matrix $W$ is an *inverse* for the matrix $A$ if

$$AW = WA = I$$

If such a matrix exists we say that $A$ is *invertible*.
A matrix can have at most one inverse, denoted $A^{-1}$.

**Proof.** If $W'$ is another inverse then

**E.g.** $A = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}, \quad B = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}$

# Inverses and solving $A\boldsymbol{x} = \boldsymbol{b}$

## Proposition

Let $A \in M_{mn}$ be an invertible matrix. Then for any $\boldsymbol{b} \in \mathbb{R}^n$ the eqn $A\boldsymbol{x} = \boldsymbol{b}$ has $\boldsymbol{x} = A^{-1}\boldsymbol{b}$ as its unique soln.

**Proof.** $\boldsymbol{x} = A^{-1}\boldsymbol{b}$ is a soln since
It is the only possible soln since $A\boldsymbol{x} = \boldsymbol{b} \implies$

**Corollary** Any invertible matrix $A \in M_{mn}$ is square.
**Proof.** The inverse $W = A^{-1}$ must be $n \times m$ for $AW, WA$ to be defined.
$A\boldsymbol{x} = \boldsymbol{0}$ has a unique soln $\implies m$

# Formulas involving matrix inverses

## Formula

Let $A, B \in M_{nn}$ be invertible.

1. $A^{-1}$ is invertibe & $(A^{-1})^{-1} =$
2. $AB$ is invertible & $(AB)^{-1} = B^{-1}A^{-1}$.
3. $A^T$ is invertible & $(A^T)^{-1} = (A^{-1})^T$.

**Proof.** All easy e.g.

**E.g.** Simplify $(ABA^{-1})^{-2}A$

# Computing inverses of matrices: example

**Q** Find the inverse of $A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 3 & 2 \\ -1 & 1 & -1 \end{pmatrix}$

**A**

# The algorithm for inverting matrices

Let $A \in M_{nn}$. To determine invertibility and invert $A$ if invertible we:

1. Form the augmented $n \times 2n$-matrix $(A|I)$.
2. Apply ERO's until we get row echelon form $(U|B)$.
3. If $U$ has non-leading columns, we stop as solns to $A\boldsymbol{x} = \mathbf{0}$ are not unique so $A$ is not invertible.
4. If $U$ has no non-leading columns, then $A$ is invertible & we can apply EROS to transform $(U|B)$ to the form $(I|C)$ with $C = A^{-1}$.

**Q** Is $A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 3 & 2 \\ 2 & 4 & 3 \end{pmatrix}$ invertible? If so, find its inverse.

# Why our algorithm works

Suppose EROs send $(A|I) \xrightarrow{EROs} (I|C)$ where $C = (\mathbf{c}_1|\dots|\mathbf{c}_n)$.
Restrict Gaussian elimination

$$(A|\mathbf{e}_i) \xrightarrow{EROs} (I|\mathbf{c}_i),$$

i.e. $A\mathbf{c}_i = \mathbf{e}_i$.

$$AC = (A\mathbf{c}_1|\dots|A\mathbf{c}_n) = (\mathbf{e}_1|\dots|\mathbf{e}_n) =??$$

**Challenge Q** Show that we also have $CA = I$. Hint: Observe $ACA = A = AI$ and explain why you can cancel the $A$'s.

# Invertibility for square matrices

## Theorem

Suppose that $A \in M_{nn}$ has row-echelon form $U$. Then the following are equivalent:

1. $A$ is invertible.
2. $U$ has no zero rows.
3. $U$ has no nonleading columns.
4. $A\boldsymbol{x} = \boldsymbol{0}$ has a unique solution $\boldsymbol{x} = \boldsymbol{0}$.
5. For each $\boldsymbol{b} \in \mathbb{R}^n$, $A\boldsymbol{x} = \boldsymbol{b}$ has a unique solution.

**Proof** We check 5) $\implies$ 4) $\implies$ 3) $\implies$ 2) $\implies$ 1) $\implies$ 5)

**Corollary** Suppose that $A \in M_{nn}$. If $XA = I$ then $A$ is invertible with inverse $X$.
**Challenge Q.** Prove this by checking (4) above.

# Inverse of $2 \times 2$-matrices

Recall the $2 \times 2$-determinant

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = \det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = a_{11}a_{22} - a_{12}a_{21}.$$

## Inverse formula

Let $A = (a_{ij}) \in M_{22}$. Then $A$ is invertible if $\det A \neq 0$ in which case

$$A^{-1} = \frac{1}{\det(A)} \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}.$$

**Proof.** Just multiply

**Remark** We'll generalise the relationship between determinants and invertibility later. The formula in this case is best appreciated if you understand the linear mapping $\mathbb{R}^2 \longrightarrow \mathbb{R}^2$ associated to $A$.

# Minors of a square matrix

We define the *determinant* of $A \in M_{nn}$, using induction on $n$. We've defined it for $n = 2$ (and 3), and suppose it's defined for $(n-1) \times (n-1)$-matrices. First

## Definition

For $1 \leq i, j \leq n$, the *(row $i$, column $j$) minor* of $A$, denoted $|A_{ij}|$, is the determinant of the $(n-1) \times (n-1)$ matrix obtained by deleting row $i$ and column $j$ from $A$.

**E.g.**Let $A = \begin{pmatrix} 1 & 2 & 3 \\ -1 & 0 & 1 \\ 4 & -2 & 5 \end{pmatrix}$.

$$|A_{21}| =$$

**Defn** For a $1 \times 1$-matrix $(a)$ we define its *determinant* to be $\det(a) = a$.

There are many definitions of the determinant. To keep the exposition elementary, we'll use a generalisation of the one for $3 \times 3$-determinants.

# Determinant

This definition is via *expanding along the first row*.

## Definition

Let $A \in M_{nn}$ with $n \geq 2$. Then the *determinant of $A$* is

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

**E.g.** Let $A = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 2 & 3 & 0 & 0 \\ -1 & 1 & 5 & 0 \\ 7 & -2 & 3 & 1 \end{pmatrix}$. Then $\det(A) =$

# Determinant of lower triangular matrices

We say $A = (a_{ij})$ is *lower triangular*, if $a_{ij} = 0$ whenever $i < j$.

The computation above shows more generally the following

## Proposition

If $A = (a_{ij})$ is lower triangular, then $\det(A)$ is just the product of the diagonal elements.

$$\det(A) = \prod_{i=1}^{n} a_{ii}$$

In particular,

$$\det(I) = 1.$$

# Determinant of the Tranpose

Let $A = (a_{ij}) \in M_{nn}$. Then

## Proposition

$\det(A) = \det(A^T)$.

**Proof.** Hard (& omitted) with our defn except for $2 \times 2$.

Since the transpose swaps columns with rows, we can compute $\det(A)$ by *expanding along the first column* as in the formula below

## Formula

$$\det(A) = a_{11}|A_{11}| - a_{21}|A_{21}| + \cdots + (-1)^{n+1}a_{n1}|A_{n1}|.$$

In particular, if $A$ is *upper triangular* in the sense that $a_{ij} = 0$ whenever $i > j$, then $\det(A)$ is the product of the diagonal entries.

# Determinant of products

## Theorem

$$\det(AB) = \det(A)\det(B)$$

Note that $A, B$ square and $AB$ defined $\implies$ $A, B$ have the same size.

**Proof.** Hard, see my MATH2601 notes.

**E.g.** Find the determinant of $A^{42}$ where $A = \begin{pmatrix} 1 & 0 & 0 \\ 4e^{\pi-\sqrt{2}} & \frac{1}{2} & 0 \\ 0 & \int_0^1 \frac{dx}{\sqrt{(1-x^2)(2-x^2)}} & 4 \end{pmatrix}$.

## Corollary

If $A$ is invertible, then $\det(A) \neq 0$ & $\det(A^{-1}) = \det(A)^{-1}$.

**Proof.**

# EROs via matrix multiplication

Consider *elementary matrices*

$E_i(c) =$ $\quad$ $E_{ij} =$ $\quad$ $E_{ij}(c) =$

## Proposition

1. $A \xrightarrow{R_i \leftrightarrow R_j} E_{ij}A.$
2. $A \xrightarrow{R_i = cR_i} E_i(c)A.$
3. $A \xrightarrow{R_i = R_i + cR_j} E_{ij}(c)A.$

**Why?** Just calculate eg

# EROs and determinants

Performing an ERO $A \xrightarrow{ERO} B$ changes the determinant by:

## Proposition

1. If $A \xrightarrow{R_i \leftrightarrow R_j} B$, then $\det(B) = -\det(A)$. i.e. swapping two rows of a matrix negates the determinant.
2. If $A \xrightarrow{R_i = cR_i} B$, then $\det(B) = c\det(A)$. i.e, multiplying a row by $c$ multiplies the determinant by $c$.
3. If $A \xrightarrow{R_i = R_i + cR_j} B$, then $\det(B) = \det(A)$. i.e,

In particular, if two rows of $A$ are the same, or $A$ has a row of 0s, then $\det(A) =$

**Why?** Prop last slide $\implies$ $B = EA$ for some elementary $E$. Hence $\det(B) = \det(E)\det(A)$ &

**E.g.** Suppose $A \in M_{33}$ has determinant 2. Find $\det(3A)$.

# Computing determinants in practice

**Warning** We don't compute determinants using the defn except for small or special matrices!

In practice, a) we apply EROs to reduce $A$ to row-echelon form $U$, b) record how we've changed the determinant (see previous slide), & c) note $U$ square & row-echelon $\implies$ upper triangular so has readily computable determinant.

**E.g.** Find the determinant of $A = \begin{pmatrix} 1 & -1 & 2 & 3 \\ -1 & 1 & 1 & 2 \\ -1 & 2 & 0 & 1 \\ 2 & 0 & 3 & -1 \end{pmatrix}$

# Conclusions from the algorithm for computing determinants

- We can transform $A$ into upper echelon form $U$ by using ERO's of form $R_i \leftrightarrow R_j$ & $R_i = R_i + cR_j$ only. Then $\det(A) = \pm \det(U)$.
- For this $U$, it has all leading columns iff all diagonal entries are non-zero iff $\det(U) \neq 0$.

This gives

**Theorem**

A square matrix $A$ is invertible iff $\det(A) \neq 0$.

**Remark** The formula for inverting $2 \times 2$-matrices generalises (see Cramer's rule in my MATH2601 notes). It has $\det(A)$ in the denominator for $A^{-1}$ and an otherwise well-defined numerator.