Review of first-year linear algebra

In MATH 2740, we rely on notions you acquired in MATH 1210/1220/1300

So let us (briefly) go over material in these courses

I also add (for some of you) a few things that will be handy and establish some terminology that we use throughout the course

OUTLINE

Sets and elements

Definition 1 (Set)

A **set** X is a collection of **elements**

We write $x \in X$ or $x \notin X$ to indicate that the element x belongs to the set X or does not belong to the set X, respectively

Definition 2 (Subset)

Let X be a set. The set S is a **subset** of X, which is denoted $S \subset X$, if all its elements belong to X

Not used here but worth noting: we say S is a **proper subset** of X and write $S \subsetneq X$, if it is a subset of X and not equal to X

Quantifiers

A shorthand notation for "for all elements x belonging to X" is $\forall x \in X$

For example, if $X = \mathbb{R}$, the *field* of real numbers, then $\forall x \in \mathbb{R}$ means "for all real numbers x"

A shorthand notation for "there exists an element x in the set X" is $\exists x \in X$

$\forall$ and $\exists$ are **quantifiers**

Intersection and union of sets

Let X and Y be two sets

Definition 3 (Intersection)

The intersection of X and Y, $X \cap Y$, is the set of elements that belong to X **and** to Y,

$$X \cap Y = \{x : x \in X \textbf{ and } x \in Y\}$$

Definition 4 (Union)

The union of X and Y, $X \cup Y$, is the set of elements that belong to X **or** to Y,

$$X \cup Y = \{x : x \in X \textbf{ or } x \in Y\}$$

In mathematics, or=and/or in common parlance. We also have an **exclusive or** (xor)

A teeny bit of logic

In a logical sense, a **proposition** is an assertion (or statement) whose truth value (true or false) can be asserted. For example, a theorem is a proposition that has been shown to be true. "The sky is blue" is also a proposition

Let A be a proposition. We generally write

$$A$$

to mean that A is true, and

$$\textbf{not } A$$

to mean that A is false. **not** A is the **contraposition** of A (or **not** A is the contraposite of A)

A teeny bit of logic (cont.)

Let A, B be propositions. Then

- $A \Rightarrow B$ (read A implies B) means that whenever A is true, then so is B
- $A \Leftrightarrow B$, also denoted A if and only if B (A iff B for short), means that $A \Rightarrow B$ **and** $B \Rightarrow A$
  We also say that A and B are **equivalent**

Let A and B be propositions. Then

$$(A \Rightarrow B) \Leftrightarrow (\textbf{not } B \Rightarrow \textbf{not } A)$$

Necessary or sufficient conditions

Suppose we want to establish whether a given statement P is true, depending on the truth value of a statement H. Then we say that

- H is a **necessary condition** if $P \Rightarrow H$
  (It is necessary that H be true for P to be true; so whenever P is true, so is H)

- H is a **sufficient condition** if $H \Rightarrow P$
  (It suffices for H to be true for P to also be true)

- H is a **necessary and sufficient condition** if $H \Leftrightarrow P$, i.e., H and P are equivalent

Playing with quantifiers

For the quantifiers $\forall$ (for all) and $\exists$ (there exists),

$$\exists \text{ is the contraposite of } \forall$$

Therefore, for example, the contraposite of

$$\forall x \in X, \exists y \in Y$$

is

$$\exists x \in X, \forall y \in Y$$

Complex numbers

Definition 5 (Complex numbers)

A **complex number** is an ordered pair (a, b), where $a, b \in \mathbb{R}$.
Usually written $a + ib$ or $a + bi$, where $i^2 = -1$ (i.e., $i = \sqrt{-1}$)
The set of all complex numbers is denoted $\mathbb{C}$,

$$\mathbb{C} = \{a + ib : a, b \in \mathbb{R}\}$$

Definition 6 (Addition and multiplication on $\mathbb{C}$)

Letting $a + ib$ and $c + id \in \mathbb{C}$, addition on $\mathbb{C}$ is defined by

$$(a + ib) + (c + id) = (a + c) + i(b + d)$$

and multiplication on $\mathbb{C}$ is defined by

$$(a + ib)(c + id) = (ac - bd) + i(ad + bc)$$

Latter is easy to obtain using regular multiplication and $i^2 = -1$

Properties

$\forall \alpha, \beta, \gamma \in \mathbb{C}$,

$\alpha + \beta = \beta + \alpha$ and $\alpha\beta = \beta\alpha$ **[commutativity]**

$(\alpha + \beta) + \gamma = \alpha + (\beta + \gamma)$ and $(\alpha\beta)\gamma = \alpha(\beta\gamma)$ **[associativity]**

$\gamma + 0 = \gamma$ and $\gamma 1 = \gamma$ **[identities]**

$\forall \alpha \in \mathbb{C}$, $\exists \beta \in \mathbb{C}$ unique s.t. $\alpha + \beta = 0$ **[additive inverse]**

$\forall \alpha \neq 0 \in \mathbb{C}$, $\exists \beta \in \mathbb{C}$ unique s.t. $\alpha\beta = 1$ **[multiplicative inverse]**

$\gamma(\alpha + \beta) = \gamma\alpha + \gamma\beta$ **[distributivity]**

Additive & multiplicative inverse, subtraction, division

Definition 7

Let $\alpha, \beta \in \mathbb{C}$

- $-\alpha$ is the **additive inverse** of α, i.e., the unique number in $\mathbb{C}$ s.t. $\alpha + (-\alpha) = 0$
- **Subtraction** on $\mathbb{C}$:

$$\beta - \alpha = \beta + (-\alpha)$$

- For $\alpha \neq 0$, $1/\alpha$ is the **multiplicative inverse** of α, i.e., the unique number in $\mathbb{C}$ s.t.

$$\alpha(1/\alpha) = 1$$

- **Division** on $\mathbb{C}$:

$$\beta/\alpha = \beta(1/\alpha)$$

Definition 8 (Real and imaginary parts)

Let $z = a + ib$. Then $\mathrm{Re}\, z = a$ is **real part** and $\mathrm{Im}\, z = b$ is **imaginary part** of z

If ambiguous, write $\mathrm{Re}\,(z)$ and $\mathrm{Im}\,(z)$

Definition 9 (Conjugate and Modulus)

Let $z = a + ib \in \mathbb{C}$. Then

- **Complex conjugate** of z is

$$\bar{z} = a - ib$$

- **Modulus** (or **absolute value**) of z is

$$|z| = \sqrt{a^2 + b^2} \geq 0$$

Properties of complex numbers

Let $w, z \in \mathbb{C}$, then

- $z + \bar{z} = 2\mathrm{Re}\, z$
- $z - \bar{z} = 2i\mathrm{Im}\, z$
- $z\bar{z} = |z|^2$
- $\overline{w + z} = \bar{w} + \bar{z}$ and $\overline{wz} = \bar{w}\bar{z}$
- $\bar{\bar{z}} = z$
- $|\mathrm{Re}\, z| \leq |z|$ and $|\mathrm{Im}\, z| \leq |z|$
- $|\bar{z}| = |z|$
- $|wz| = |w|\,|z|$
- $|w + z| \leq |w| + |z|$ **[triangle inequality]**

Solving quadratic equations

Consider the polynomial

$$P(x) = a_0 + a_1 x + a_2 x^2$$

where $x, a_0, a_1, a_2 \in \mathbb{R}$. Letting

$$\Delta = a_1^2 - 4a_0 a_2$$

you know that if $\Delta > 0$, then

$$P(x) = 0$$

has two distinct *real* solutions,

$$x_1 = \frac{-a_1 - \sqrt{\Delta}}{2a_2} \quad \text{and} \quad x_2 = \frac{-a_1 + \sqrt{\Delta}}{2a_2}$$

if $\Delta = 0$, then there is a (multiplicity 2) unique *real* solution

$$x_1 = \frac{-a_1}{2a_2}$$

while if $\Delta < 0$, there is no solution

Solving quadratic equations with complex numbers

Consider the polynomial

$$P(x) = a_0 + a_1 x + a_2 x^2$$

where $x, a_0, a_1, a_2 \in \mathbb{R}$. If instead of seeking $x \in \mathbb{R}$, we seek $x \in \mathbb{C}$, then the situation is the same, except when $\Delta < 0$

In the latter case, note that

$$\sqrt{\Delta} = \sqrt{(-1)(-\Delta)} = \sqrt{-1}\sqrt{-\Delta} = i\sqrt{-\Delta}$$

Since $\Delta < 0$, $-\Delta > 0$ and the square root is the usual one

Solving quadratic equations with complex numbers

To summarize, consider the polynomial

$$P(x) = a_0 + a_1 x + a_2 x^2$$

where $x, a_0, a_1, a_2 \in \mathbb{R}$. Letting

$$\Delta = a_1^2 - 4a_0 a_2$$

Then

$$P(x) = 0$$

has two solutions,

$$x_{1,2} = \frac{-a_1 \pm \sqrt{\Delta}}{2a_2}$$

where, if $\Delta < 0$, $x_1, x_2 \in \mathbb{C}$ and take the form

$$x_{1,2} = \frac{-a_1 \pm i\sqrt{-\Delta}}{2a_2}$$

Why this matters

Recall (we will come back to this later) that to find the *eigenvalues* of the matrix

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

we seek λ solutions to $\det(A - \lambda\mathbb{I}) = 0$, i.e., λ solutions to

$$|A - \lambda\mathbb{I}| = \begin{vmatrix} a_{11} - \lambda & a_{12} \\ a_{21} & a_{22} - \lambda \end{vmatrix} = (a_{11} - \lambda)(a_{22} - \lambda) - a_{12}a_{21} = 0$$

i.e., λ solutions to

$$\lambda^2 - (a_{11} + a_{22})\lambda + a_{11}a_{22} - a_{12}a_{21} = 0$$

Why this matters (cont.)

Let

$$P(\lambda) = \lambda^2 - (a_{11} + a_{22})\lambda + a_{11}a_{22} - a_{12}a_{21}$$

From previous discussion, letting

$$\begin{aligned}
\Delta &= (a_{11} + a_{22})^2 - 4(a_{11}a_{22} - a_{12}a_{21}) \\
&= a_{11}^2 + a_{22}^2 + 2a_{11}a_{22} - 4a_{11}a_{22} + 4a_{12}a_{21} \\
&= a_{11}^2 + a_{22}^2 - 2a_{11}a_{22} + 4a_{12}a_{21} \\
&= (a_{11} - a_{22})^2 + 4a_{12}a_{21}
\end{aligned}$$

we have two (potentially equal) solutions to $P(\lambda) = 0$

$$x_{1,2} = \frac{a_{11} + a_{22} \pm \sqrt{\Delta}}{2}$$

that are complex if $\Delta < 0$

Example: $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$

Vectors

A **vector** $\mathbf{v}$ is an ordered n-tuple of real or complex numbers

Denote $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$ (real or complex numbers). For $v_1, \ldots, v_n \in \mathbb{F}$,

$$\mathbf{v} = (v_1, \ldots, v_n) \in \mathbb{F}^n$$

is a vector. $v_1, \ldots, v_n$ are the **components** of $\mathbf{v}$

If unambiguous, we write v. Otherwise, $\mathbf{v}$ or $\vec{v}$

Vector space

Definition 10 (Vector space)

A **vector space** over $\mathbb{F}$ is a set V together with two binary operations, **vector addition**, denoted $+$, and **scalar multiplication**, that satisfy the relations:

1. $\forall \mathbf{u}, \mathbf{v}, \mathbf{w} \in V$, $\mathbf{u} + (\mathbf{v} + \mathbf{w}) = (\mathbf{u} + \mathbf{v}) + \mathbf{w}$
2. $\forall \mathbf{v}, \mathbf{w} \in V$, $\mathbf{v} + \mathbf{w} = \mathbf{w} + \mathbf{v}$
3. $\exists \mathbf{0} \in V$, the zero vector, such that $\mathbf{v} + \mathbf{0} = \mathbf{v}$ for all $\mathbf{v} \in V$
4. $\forall \mathbf{v} \in V$, there exists an element $\mathbf{w} \in V$, the additive inverse of $\mathbf{v}$, such that $\mathbf{v} + \mathbf{w} = \mathbf{0}$
5. $\forall \alpha \in \mathbb{R}$ and $\forall \mathbf{v}, \mathbf{w} \in V$, $\alpha(\mathbf{v} + \mathbf{w}) = \alpha\mathbf{v} + \alpha\mathbf{w}$
6. $\forall \alpha, \beta \in \mathbb{R}$ and $\forall \mathbf{v} \in V$, $(\alpha + \beta)\mathbf{v} = \alpha\mathbf{v} + \beta\mathbf{v}$
7. $\forall \alpha, \beta \in \mathbb{R}$ and $\forall \mathbf{v} \in V$, $\alpha(\beta\mathbf{v}) = (\alpha\beta)\mathbf{v}$
8. $\forall \mathbf{v} \in V$, $1\mathbf{v} = \mathbf{v}$

Norms

Definition 11 (Norm)

Let V be a vector space over $\mathbb{F}$, and $\mathbf{v} \in V$ be a vector. The **norm** of $\mathbf{v}$, denoted $\|\mathbf{v}\|$, is a function from V to $\mathbb{R}_+$ that has the following properties:

1. For all $\mathbf{v} \in V$, $\|\mathbf{v}\| \geq 0$ with $\|\mathbf{v}\| = 0$ iff $\mathbf{v} = \mathbf{0}$
2. For all $\alpha \in \mathbb{F}$ and all $\mathbf{v} \in V$, $\|\alpha \mathbf{v}\| = |\alpha| \, \|\mathbf{v}\|$
3. For all $\mathbf{u}, \mathbf{v} \in V$, $\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$

Let V be a vector space (for example, $\mathbb{R}^2$ or $\mathbb{R}^3$)

The **zero element** (or **zero vector**) is the vector $\mathbf{0} = (0, \ldots, 0)$

The **additive inverse** of $\mathbf{v} = (v_1, \ldots, v_n)$ is $-\mathbf{v} = (-v_1, \ldots, -v_n)$

For $\mathbf{v} = (v_1, \ldots, v_n) \in V$, the length (or Euclidean norm) of $\mathbf{v}$ is the **scalar**

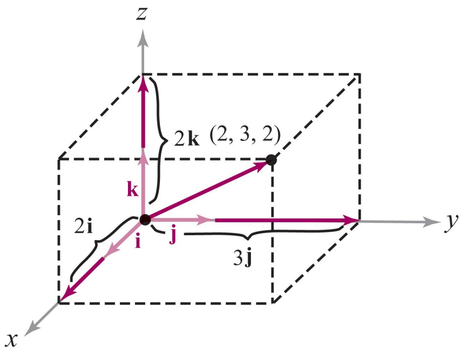
$$\|\mathbf{v}\| = \sqrt{v_1^2 + \cdots + v_n^2}$$

To **normalize** the vector $\mathbf{v}$ consists in considering $\tilde{\mathbf{v}} = \mathbf{v}/\|\mathbf{v}\|$, i.e., the vector in the same direction as $\mathbf{v}$ that has unit length

Standard basis vectors

Vectors $\mathbf{i} = (1,0,0)$, $\mathbf{j} = (0,1,0)$ and $\mathbf{k} = (0,0,1)$ are the **standard basis vectors** of $\mathbb{R}^3$. A vector $\mathbf{v} = (v_1, v_2, v_3)$ can then be written

$$\mathbf{v} = v_1\mathbf{i} + v_2\mathbf{j} + v_3\mathbf{k}$$



For V ($\mathbb{R}^n$), the standard basis vectors are usually denoted $\mathbf{e}_1, \ldots, \mathbf{e}_n$, with

$$\mathbf{e}_k = (\underbrace{0, \ldots, 0}_{k-1}, 1, \underbrace{0, \ldots, 0}_{n-k+1})$$

Dot product

Definition 12 (Dot product)

Let $\mathbf{a} = (a_1, \ldots, a_n) \in \mathbb{R}^n$, $\mathbf{b} = (b_1, \ldots, b_n) \in \mathbb{R}^n$. The **dot product** of **a** and **b** is the **scalar**

$$\mathbf{a} \bullet \mathbf{b} = \sum_{i=1}^{n} a_i b_i = a_1 b_1 + \cdots + a_n b_n$$

The dot product is a special case of **inner product**

Properties of the dot product

Theorem 13

For $\mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathbb{R}^n$ *and* $\alpha \in \mathbb{R}$,

- $\mathbf{a} \bullet \mathbf{a} = \|\mathbf{a}\|^2$ *(so* $\mathbf{a} \bullet \mathbf{a} \geq 0$, *with* $\mathbf{a} \bullet \mathbf{a} = 0$ *iff* $\mathbf{a} = \mathbf{0}$*)*
- $\mathbf{a} \bullet \mathbf{b} = \mathbf{b} \bullet \mathbf{a}$ *(*$\bullet$ *is commutative)*
- $\mathbf{a} \bullet (\mathbf{b} + \mathbf{c}) = \mathbf{a} \bullet \mathbf{b} + \mathbf{a} \bullet \mathbf{c}$ *(*$\bullet$ *distributive over* $+$*)*
- $(\alpha \mathbf{a}) \bullet \mathbf{b} = \alpha(\mathbf{a} \bullet \mathbf{b}) = \mathbf{a} \bullet (\alpha \mathbf{b})$
- $\mathbf{0} \bullet \mathbf{a} = 0$

Some results stemming from the dot product

Theorem 14

If θ is the angle between the vectors $\mathbf{a}$ and $\mathbf{b}$, then

$$\mathbf{a} \bullet \mathbf{b} = \|\mathbf{a}\| \, \|\mathbf{b}\| \, \cos\theta$$

Corollary 15 (Cauchy-Schwarz inequality)

For any two vectors $\mathbf{a}$ and $\mathbf{b}$, we have

$$|\mathbf{a} \bullet \mathbf{b}| \leq \|\mathbf{a}\| \, \|\mathbf{b}\|$$

with equality if and only if $\mathbf{a}$ is a scalar multiple of $\mathbf{b}$, or one of them is $\mathbf{0}$.

Theorem 16

$\mathbf{a}$ and $\mathbf{b}$ are orthogonal if and only if $\mathbf{a} \bullet \mathbf{b} = 0$.

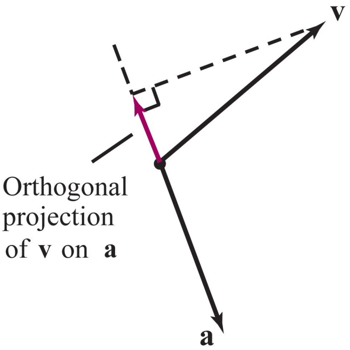
Scalar and vector projections

Scalar projection of **v** onto **a** (or component of **v** along **a**):

$$\mathrm{comp}_{\mathbf{a}}\mathbf{v} = \frac{\mathbf{a} \bullet \mathbf{v}}{\|\mathbf{a}\|}$$

Vector (or orthogonal) projection of **v** onto **a**:

$$\mathrm{proj}_{\mathbf{a}}\mathbf{v} = \left(\frac{\mathbf{a} \bullet \mathbf{v}}{\|\mathbf{a}\|}\right)\frac{\mathbf{a}}{\|\mathbf{a}\|} = \frac{\mathbf{a} \bullet \mathbf{v}}{\|\mathbf{a}\|^2}\mathbf{a}$$



Linear systems

Definition 17 (Linear system)

A **linear system** of m equations in n unknowns takes the form

$$\begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \cdots & + & a_{2n}x_n & = & b_2 \\ \vdots & & \vdots & & \vdots & & \vdots & & \vdots \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \cdots & + & a_{mn}x_n & = & b_n \end{array} \tag{1}$$

The a_{ij}, x_j and b_j could be in $\mathbb{R}$ or $\mathbb{C}$, although here we typically assume they are in $\mathbb{R}$

The aim is to find $x_1, x_2, \ldots, x_n$ that satisfy all equations simultaneously

Theorem 18 (Nature of solutions to a linear system)

A linear system can have

- *no solution*
- *a unique solution*
- *infinitely many solutions*

Operations on linear systems

You learned to manipulate linear systems using

- Gaussian elimination
- Gauss-Jordan elimination

with the aim to put the system in **row echelon form** (REF) or **reduced row echelon form** (RREF)

Matrices and linear systems

Writing

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}, \quad \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \quad \text{and} \quad \mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$$

where A is an $m \times n$ **matrix**, $\mathbf{x}$ and $\mathbf{b}$ are n (column) **vectors** (or $n \times 1$ matrices), then the linear system in the previous slide takes the form

$$A\mathbf{x} = \mathbf{b}$$

Notation for vectors

We usually assume vectors are column vectors and thus write, e.g.,

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = (x_1, x_2, \dots, x_n)^T$$

Here, ^{T} is the **transpose operator** (more on this soon)

Consider the system

$$A\mathbf{x} = \mathbf{b}$$

If $\mathbf{b} = \mathbf{0}$, the system is **homogeneous** and always has the solution $\mathbf{x} = 0$ and so the "no solution" option in Theorem **??** goes away

Definition 19 (Matrix)

An m-by-n or $m \times n$ matrix is a rectangular array of elements of $\mathbb{R}$ or $\mathbb{C}$ with m rows and n columns,

$$A = [a_{ij}] = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}$$

We always list indices as "row,column"

We denote $\mathcal{M}_{mn}(\mathbb{F})$ or $\mathbb{F}^{mn}$ the set of $m \times n$ matrices with entries in $\mathbb{F} = \{\mathbb{R}, \mathbb{C}\}$. Often, we omit $\mathbb{F}$ in $\mathcal{M}_{mn}$ if the nature of $\mathbb{F}$ is not important

When $m = n$, we usually write $\mathcal{M}_n$

Basic matrix arithmetic

Let $A \in \mathcal{M}_{mn}, B \in \mathcal{M}_{mn}$ be matrices (of the same size) and $c \in \mathbb{F} = \{\mathbb{R}, \mathbb{C}\}$ be a scalar

- **Scalar multiplication**

$$cA = [ca_{ij}]$$

- **Addition**

$$A + B = [a_{ij} + b_{ij}]$$

- **Subtraction** (addition of $-B = (-1)B$ to A)

$$A - B = A + (-1)B = [a_{ij} + (-1)b_{ij}] = [a_{ij} - b_{ij}]$$

- **Transposition** of A gives a matrix $A^T = \mathcal{M}_{nm}$ with

$$A^T = [a_{ji}], \quad j = 1, \ldots, n, \quad i = 1, \ldots, m$$

Matrix multiplication

The (matrix) **product** of A and B, AB, requires the "inner dimensions" to match, i.e., the number of columns in A must equal the number of rows in B

Suppose that is the case, i.e., let $A \in \mathcal{M}_{mn}$, $B \in \mathcal{M}_{np}$. Then the i, j entry in $C := AB$ takes the form

$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}$$

Recall that the matrix product is not commutative, i.e., in general, $AB \neq BA$ (when both those products are defined, i.e., when $A, B \in \mathcal{M}_n$)

Special matrices

Definition 20 (Zero and identity matrices)

The **zero** matrix is the matrix 0_{mn} whose entries are all zero. The **identity** matrix is a square $n \times n$ matrix $\mathbb{I}_n$ with all entries on the main diagonal equal to one and all off diagonal entries equal to zero

Definition 21 (Symmetric matrix)

A square matrix $A \in \mathcal{M}_n$ is **symmetric** if $\forall i, j = 1, \ldots, n$, $a_{ij} = a_{ji}$. In other words, $A \in \mathcal{M}_n$ is symmetric if $A = A^T$

Properties of symmetric matrices

Theorem 22

1. *If $A \in \mathcal{M}_n$, then $A + A^T$ is symmetric*
2. *If $A \in \mathcal{M}_{mn}$, then $AA^T \in \mathcal{M}_m$ and $A^TA \in \mathcal{M}_n$ are symmetric*

X symmetric $\iff$ $X = X^T$, so use $X =$ the matrix whose symmetric property you want to check

1. True if $A + A^T = (A + A^T)^T$. We have

$$(A + A^T)^T = A^T + (A^T)^T = A^T + A = A + A^T$$

2. AA^T symmetric if $AA^T = (AA^T)^T$. We have

$$(AA^T)^T = (A^T)^TA^T = AA^T$$

A^TA works similarly

Determinants

Definition 23 (Determinant)

Let $A \in \mathcal{M}_n$ with $n \geq 2$. The **determinant** of A is the *scalar*

$$\det(A) = |A| = \sum_{j=1}^{n} a_{ij} C_{ij}$$

where $C_{ij} = (-1)^{i+j} \det(A_{ij})$ is the (i,j)-**cofactor** of A and A_{ij} is the submatrix of A from which the ith row and jth column have been removed

This is a cofactor expansion along the ith row

This is a recursive formula: it gives result in terms of n $\mathcal{M}_{n-1}$ matrices, to which it must in turn be applied, all the way down to

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

Two special matrices and their determinants

Definition 24

$A \in \mathcal{M}_n$ is **upper triangular** if $a_{ij} = 0$ when $i > j$, **lower triangular** if $a_{ij} = 0$ when $j > i$, **triangular** if it is *either* upper or lower triangular and **diagonal** if it is *both* upper and lower triangular

When A diagonal, we often write $A = \operatorname{diag}(a_{11}, a_{22}, \ldots, a_{nn})$

Theorem 25

Let $A \in \mathcal{M}_n$ be triangular or diagonal. Then

$$\det(A) = \prod_{i=1}^{n} a_{ii} = a_{11}a_{22}\cdots a_{nn}$$

Inversion/Singularity

Definition 26 (Matrix inverse)

$A \in \mathcal{M}_n$ is **invertible** (or **nonsingular**) if $\exists A^{-1} \in \mathcal{M}_n$ s.t.

$$AA^{-1} = A^{-1}A = \mathbb{I}$$

A^{-1} is the **inverse** of A. If A^{-1} does not exist, A is **singular**

Theorem 27

Let $A \in \mathcal{M}_n$, $\mathbf{x}, \mathbf{b} \in \mathbb{F}^n$. Then

- *A invertible* $\iff \det(A) \neq 0$
- *If A invertible, A^{-1} is unique*
- *If A invertible, then $A\mathbf{x} = \mathbf{b}$ has the unique solution $\mathbf{x} = A^{-1}\mathbf{b}$*

Revisiting matrix arithmetic

With addition, subtraction, scalar multiplication, multiplication, transposition and inversion, you can perform arithmetic on matrices essentially as on scalar, if you bear in mind a few rules

- The sizes have to be compatible
- The order is important since matrix multiplication is not commutative
- Transposition and inversion change the order of products:

$$(AB)^T = B^T A^T \text{ and } (AB)^{-1} = B^{-1} A^{-1}$$

Eigenvalues / Eigenvectors / Eigenpairs

Definition 28

Let $A \in \mathcal{M}_n$. A vector $\mathbf{x} \in \mathbb{F}^n$ such that $\mathbf{x} \neq \mathbf{0}$ is an **eigenvector** of A if $\exists \lambda \in \mathbb{F}$ called an **eigenvalue**, s.t.

$$A\mathbf{x} = \lambda \mathbf{x}$$

A couple $(\lambda, \mathbf{x})$ with $\mathbf{x} \neq \mathbf{0}$ s.t. $A\mathbf{x} = \lambda\mathbf{x}$ is an **eigenpair**

If $(\lambda, \mathbf{x})$ eigenpair, then for $c \neq 0$, $(\lambda, c\mathbf{x})$ also eigenpair since $A(c\mathbf{x}) = cA\mathbf{x} = c\lambda\mathbf{x}$ and dividing both sides by c..

Similarity

Definition 29 (Similarity)

$A, B \in \mathcal{M}_n$ are **similar** ($A \sim B$) if $\exists P \in \mathcal{M}_n$ invertible s.t.

$$P^{-1}AP = B$$

Theorem 30 ($\sim$ is an equivalence relation)

$A, B, C \in \mathcal{M}_n$, then

- $A \sim A$ *($\sim$ **reflexive**)*
- $A \sim B \implies B \sim A$ *($\sim$ **symmetric**)*
- $A \sim B$ *and* $B \sim C \implies A \sim C$ *($\sim$ **transitive**)*

Similarity (cont.)

Theorem 31

$A, B \in \mathcal{M}_n$ with $A \sim B$. Then

- $\det A = \det B$
- *A invertible $\iff$ B invertible*
- *A and B have the same eigenvalues*

Diagonalisation

Definition 32 (Diagonalisability)

$A \in \mathcal{M}_n$ is **diagonalisable** if $\exists D \in \mathcal{M}_n$ diagonal s.t. $A \sim D$

In other words, $A \in \mathcal{M}_n$ is diagonalisable if there exists a diagonal matrix $D \in \mathcal{M}_n$ and a nonsingular matrix $P \in \mathcal{M}_n$ s.t. $P^{-1}AP = D$

Could of course write $PAP^{-1} = D$ since P invertible, but $P^{-1}AP$ makes more sense for computations

Theorem 33

$A \in \mathcal{M}_n$ diagonalisable $\iff$ A has n linearly independent eigenvectors

Corollary 34 (Sufficient condition for diagonalisability)

$A \in \mathcal{M}_n$ has all its eigenvalues distinct $\implies$ A diagonalisable

For $P^{-1}AP = D$: in P, put the linearly independent eigenvectors as columns and in D, the corresponding eigenvalues

Linear combination and span

Definition 35 (Linear combination)

Let V be a vector space. A **linear combination** of a set $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ of vectors in V is a *vector*

$$c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k$$

where $c_1, \ldots, c_k \in \mathbb{F}$

Definition 36 (Span)

The set of all linear combinations of a set of vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k$ is the **span** of $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$,

$$\operatorname{span}(\mathbf{v}_1, \ldots, \mathbf{v}_k) = \{c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k : c_1, \ldots, c_k \in \mathbb{F}\}$$

Finite/infinite-dimensional vector spaces

Theorem 37

The span of a set of vectors in V is the smallest subspace of V containing all the vectors in the set

Definition 38 (Set of vectors spanning a space)

If $\text{span}(\mathbf{v}_1, \ldots, \mathbf{v}_k) = V$, we say $\mathbf{v}_1, \ldots, \mathbf{v}_k$ **spans** V

Definition 39 (Dimension of a vector space)

A vector space V is **finite-dimensional** if some set of vectors in it spans V. A vector space V is **infinite-dimensional** if it is not finite-dimensional

Linear (in)dependence

Definition 40 (Linear independence/Linear dependence)

A set $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ of vectors in a vector space V is **linearly independent** if

$$(c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k = 0) \Leftrightarrow (c_1 = \cdots = c_k = 0),$$

where $c_1, \ldots, c_k \in \mathbb{F}$. A set of vectors is **linearly dependent** if it is not linearly independent.

If linearly dependent, assume w.l.o.g. that $c_1 \neq 0$, then

$$\mathbf{v}_1 = -\frac{c_2}{c_1}\mathbf{v}_2 - \cdots - \frac{c_k}{c_1}\mathbf{v}_k$$

i.e., $\mathbf{v}_1$ is a linear combination of the other vectors in the set

Theorem 41

Let V be a finite-dimensional vector space. Then the **cardinal** *(number of elements) of every linearly independent set of vectors is less than or equal to the number of elements in every spanning set of vectors*

E.g., in $\mathbb{R}^3$, a set with 4 or more vectors is automatically linearly dependent

Basis

Definition 42 (Basis)

Let V be a vector space. A **basis** of V is a set of vectors in V that is both linearly independent and spanning

Theorem 43 (Criterion for a basis)

A set $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ *of vectors in a vector space* V *is a basis of* V $\iff$ $\forall \mathbf{v} \in V$, $\mathbf{v}$ *can be written uniquely in the form*

$$\mathbf{v} = c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k,$$

where $c_1, \ldots, c_k \in \mathbb{F}$

Plus/Minus Theorem

Theorem 44 (Plus/Minus Theorem)

S a nonempty set of vectors in vector space V

- *If S is linearly independent and $V \ni \mathbf{v} \notin \mathrm{span}(S)$, then $S \cup \{\mathbf{v}\}$ is linearly independent*
- *If $\mathbf{v} \in S$ is linear combination of other vectors in S, then $\mathrm{span}(S) = \mathrm{span}(S - \{\mathbf{v}\})$*

More on bases

Theorem 45 (Basis of finite-dimensional vector space)

Every finite-dimensional vector space has a basis

Theorem 46

Any two bases of a finite-dimensional vector space have the same number of vectors

Definition 47 (Dimension)

The **dimension** $\dim V$ of a finite-dimensional vector space V is the number of vectors in any basis of the vector space

Theorem 48 (Dimension of a subspace)

Let V be a finite-dimensional vector space and $U \subset V$ be a subspace of V. Then $\dim U \leq \dim V$

Constructing bases

Theorem 49

Let V be a finite-dimensional vector space. Then every linearly independent set of vectors in V with $\dim V$ elements is a basis of V

Theorem 50

Let V be a finite-dimensional vector space. Then every spanning set of vectors in V with $\dim V$ elements is a basis of V

To finish: the “famous” “growing result”

Theorem 51

Let $A \in \mathcal{M}_n$. The following statements are equivalent (TFAE)

1. *The matrix A is invertible*
2. $\forall \mathbf{b} \in \mathbb{F}^n$, $A\mathbf{x} = \mathbf{b}$ *has a unique solution* ($\mathbf{x} = A^{-1}\mathbf{b}$)
3. *The only solution to* $A\mathbf{x} = \mathbf{0}$ *is the trivial solution* $\mathbf{x} = \mathbf{0}$
4. $RREF(A) = \mathbb{I}_n$
5. *The matrix A is equal to a product of elementary matrices*
6. $\forall \mathbf{b} \in \mathbb{F}^n$, $A\mathbf{x} = \mathbf{b}$ *has a solution*
7. *There is a matrix $B \in \mathcal{M}_n$ such that $AB = \mathbb{I}_n$*
8. *There is an invertible matrix $B \in \mathcal{M}_n$ such that $AB = \mathbb{I}_n$*
9. $\det(A) \neq 0$
10. 0 *is not an eigenvalue of A*