

Rulebook

Every definition, theorem and proposition of the course, in order and with the book's numbers. No proofs and no discussion: for those, open the same number in the book or watch the lecture named beside each section. This file is generated from the book (rulebook.py), so it always matches it.

Chapter 1. Vectors and the geometry of $\mathbb{R}^n$

1.1 What linear algebra is about

Lecture 1

Definition 1.1.1 (The space $\mathbb{R}^n$).

Let n be a positive integer. The set $\mathbb{R}^n$ consists of all ordered lists of n real numbers,

$$\mathbb{R}^n = \left\{ \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} : x_1, \dots, x_n \in \mathbb{R} \right\}.$$

Elements of $\mathbb{R}^n$ are called *vectors*, and the numbers $x_1, \dots, x_n$ are the *components* (or *entries*) of the vector. Real numbers, in contrast, are called *scalars*.

Definition 1.1.2 (Equality of vectors).

Two vectors $\mathbf{x} = (x_1, \dots, x_n)$ and $\mathbf{y} = (y_1, \dots, y_n)$ in $\mathbb{R}^n$ are *equal*, written $\mathbf{x} = \mathbf{y}$, when $x_i = y_i$ for every $i = 1, \dots, n$.

1.2 Adding and scaling vectors

Lecture 2

Definition 1.2.1 (Vector addition).

Let $\mathbf{v}, \mathbf{w} \in \mathbb{R}^n$. Their *sum* is the vector in $\mathbb{R}^n$ obtained by adding matching entries:

$$\mathbf{v} + \mathbf{w} = \begin{bmatrix} v_1 + w_1 \\ v_2 + w_2 \\ \vdots \\ v_n + w_n \end{bmatrix}, \quad \text{that is,} \quad (\mathbf{v} + \mathbf{w})_i = v_i + w_i \quad (i = 1, \dots, n).$$

Definition 1.2.3 (Scalar multiplication).

Let $c \in \mathbb{R}$ be a scalar and $\mathbf{v} \in \mathbb{R}^n$. The *scalar multiple* $c\mathbf{v}$ is the vector in $\mathbb{R}^n$ obtained by multiplying every entry of $\mathbf{v}$ by c :

$$c\mathbf{v} = \begin{bmatrix} c v_1 \\ c v_2 \\ \vdots \\ c v_n \end{bmatrix}, \quad \text{that is,} \quad (c\mathbf{v})_i = c v_i \quad (i = 1, \dots, n).$$

Definition 1.2.5. (a) The *zero vector* of $\mathbb{R}^n$ is $\mathbf{0} = (0, 0, \dots, 0)$, the vector whose entries are all 0.

(b) The *negative* of $\mathbf{v} \in \mathbb{R}^n$ is $-\mathbf{v} = (-v_1, \dots, -v_n)$, the vector whose entries are those of $\mathbf{v}$ with their signs changed. Equivalently, $-\mathbf{v} = (-1)\mathbf{v}$.

(c) The *difference* of $\mathbf{v}, \mathbf{w} \in \mathbb{R}^n$ is $\mathbf{v} - \mathbf{w} = \mathbf{v} + (-\mathbf{w})$. Its entries are $(\mathbf{v} - \mathbf{w})_i = v_i - w_i$.

Theorem 1.2.7 (Rules of vector algebra).

For all vectors $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{R}^n$ and all scalars $a, b \in \mathbb{R}$:

- (1) $\mathbf{v} + \mathbf{w} = \mathbf{w} + \mathbf{v}$ (addition is commutative)
- (2) $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$ (addition is associative)
- (3) $\mathbf{v} + \mathbf{0} = \mathbf{v}$ (the zero vector)
- (4) $\mathbf{v} + (-\mathbf{v}) = \mathbf{0}$ (negatives)
- (5) $1 \mathbf{v} = \mathbf{v}$ (scaling by one)
- (6) $a(b \mathbf{v}) = (ab) \mathbf{v}$ (scaling twice)
- (7) $a(\mathbf{v} + \mathbf{w}) = a \mathbf{v} + a \mathbf{w}$ (distributive law over vector addition)
- (8) $(a + b) \mathbf{v} = a \mathbf{v} + b \mathbf{v}$ (distributive law over scalar addition)

Proposition 1.2.8.

Let $\mathbf{u}, \mathbf{v}, \mathbf{w}, \mathbf{x} \in \mathbb{R}^n$ and $a \in \mathbb{R}$. Then:

- (a) (cancellation) if $\mathbf{v} + \mathbf{x} = \mathbf{w} + \mathbf{x}$, then $\mathbf{v} = \mathbf{w}$;
- (b) (the zero vector is unique) if $\mathbf{v} + \mathbf{z} = \mathbf{v}$ for some $\mathbf{z} \in \mathbb{R}^n$, then $\mathbf{z} = \mathbf{0}$;
- (c) (negatives are unique) if $\mathbf{v} + \mathbf{u} = \mathbf{0}$, then $\mathbf{u} = -\mathbf{v}$;
- (d) $0 \mathbf{v} = \mathbf{0}$;
- (e) $a \mathbf{0} = \mathbf{0}$;
- (f) $(-1) \mathbf{v} = -\mathbf{v}$;
- (g) if $a \mathbf{v} = \mathbf{0}$, then $a = 0$ or $\mathbf{v} = \mathbf{0}$.

Proposition 1.2.10 (Midpoint).

The midpoint M of the segment from P to Q has position vector

$$\mathbf{m} = \mathbf{p} + \frac{1}{2}(\mathbf{q} - \mathbf{p}) = \frac{1}{2} \mathbf{p} + \frac{1}{2} \mathbf{q}.$$

Theorem 1.2.11.

The diagonals of a parallelogram bisect each other: they meet at a point that is the midpoint of each.

1.3 Linear combinations and span

Lecture 3

Definition 1.3.1 (Linear combination).

Let $\mathbf{v}_1, \dots, \mathbf{v}_k$ be vectors in $\mathbb{R}^n$ and let $c_1, \dots, c_k$ be scalars. The vector

$$c_1 \mathbf{v}_1 + c_2 \mathbf{v}_2 + \dots + c_k \mathbf{v}_k$$

is called a *linear combination* of $\mathbf{v}_1, \dots, \mathbf{v}_k$ with *weights* (or *coefficients*) $c_1, \dots, c_k$.

Definition 1.3.5 (Span).

The *span* of vectors $\mathbf{v}_1, \dots, \mathbf{v}_k$ in $\mathbb{R}^n$ is the set of all their linear combinations,

$$\text{Span}\{\mathbf{v}_1, \dots, \mathbf{v}_k\} = \{c_1 \mathbf{v}_1 + \dots + c_k \mathbf{v}_k : c_1, \dots, c_k \in \mathbb{R}\}.$$

A vector $\mathbf{b}$ is *in the span* if $\mathbf{b} = c_1 \mathbf{v}_1 + \dots + c_k \mathbf{v}_k$ for some scalars $c_1, \dots, c_k$. By convention the span of the empty set is $\{\mathbf{0}\}$.

Theorem 1.3.8.

Let $\mathbf{v} = (p, q)$ and $\mathbf{w} = (r, s)$ be vectors in $\mathbb{R}^2$. The following are equivalent:

- (a) $\text{Span}\{\mathbf{v}, \mathbf{w}\} = \mathbb{R}^2$;
- (b) $\mathbf{v}$ and $\mathbf{w}$ are not parallel;

(c) $ps - qr \neq 0$.

If $\mathbf{v}$ and $\mathbf{w}$ are parallel, their span is a line through the origin, or $\{\mathbf{0}\}$ when $\mathbf{v} = \mathbf{w} = \mathbf{0}$.

Theorem 1.3.10.

Let $\mathbf{v}_1, \dots, \mathbf{v}_k \in \mathbb{R}^n$ and $S = \text{Span}\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$.

- (a) $\mathbf{0} \in S$ and $\mathbf{v}_i \in S$ for every i .
- (b) If $\mathbf{x}, \mathbf{y} \in S$ then $\mathbf{x} + \mathbf{y} \in S$ (S is closed under addition).
- (c) If $\mathbf{x} \in S$ and $t \in \mathbb{R}$ then $t\mathbf{x} \in S$ (S is closed under scalar multiplication).
- (d) S is the smallest such set: if $W \subseteq \mathbb{R}^n$ contains $\mathbf{v}_1, \dots, \mathbf{v}_k$ and is closed under addition and scalar multiplication, then $S \subseteq W$.

Theorem 1.3.11.

Let $\mathbf{v}_1, \dots, \mathbf{v}_k, \mathbf{u} \in \mathbb{R}^n$. Then

$$\text{Span}\{\mathbf{v}_1, \dots, \mathbf{v}_k, \mathbf{u}\} = \text{Span}\{\mathbf{v}_1, \dots, \mathbf{v}_k\} \quad \text{if and only if} \quad \mathbf{u} \in \text{Span}\{\mathbf{v}_1, \dots, \mathbf{v}_k\}.$$

1.4 Dot product, length and angle

Lecture 4

Definition 1.4.1 (Norm).

The *norm* (or *length*) of a vector $\mathbf{v} = (v_1, \dots, v_n) \in \mathbb{R}^n$ is the number

$$\|\mathbf{v}\| = \sqrt{v_1^2 + v_2^2 + \dots + v_n^2}.$$

Definition 1.4.3 (Dot product).

The *dot product* of $\mathbf{v} = (v_1, \dots, v_n)$ and $\mathbf{w} = (w_1, \dots, w_n)$ in $\mathbb{R}^n$ is the number

$$\mathbf{v} \cdot \mathbf{w} = v_1 w_1 + v_2 w_2 + \dots + v_n w_n = \sum_{i=1}^n v_i w_i.$$

Definition 1.4.5 (Distance).

The *distance* between $\mathbf{v}$ and $\mathbf{w}$ in $\mathbb{R}^n$ is $d(\mathbf{v}, \mathbf{w}) = \|\mathbf{w} - \mathbf{v}\|$.

Theorem 1.4.6 (Properties of the dot product).

For all $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{R}^n$ and every scalar c :

- (a) *symmetry*: $\mathbf{v} \cdot \mathbf{w} = \mathbf{w} \cdot \mathbf{v}$;
- (b) *additivity*: $\mathbf{u} \cdot (\mathbf{v} + \mathbf{w}) = \mathbf{u} \cdot \mathbf{v} + \mathbf{u} \cdot \mathbf{w}$;
- (c) *homogeneity*: $(c\mathbf{v}) \cdot \mathbf{w} = c(\mathbf{v} \cdot \mathbf{w})$;
- (d) *positivity*: $\mathbf{v} \cdot \mathbf{v} \geq 0$, with $\mathbf{v} \cdot \mathbf{v} = 0$ only when $\mathbf{v} = \mathbf{0}$.

Proposition 1.4.7.

For every $\mathbf{v} \in \mathbb{R}^n$ and scalar c , $\|c\mathbf{v}\| = |c| \|\mathbf{v}\|$.

Definition 1.4.8 (Unit vector).

A *unit vector* is a vector of norm 1. For $\mathbf{v} \neq \mathbf{0}$, the vector $\frac{1}{\|\mathbf{v}\|} \mathbf{v}$ is called the *normalisation* of $\mathbf{v}$.

Proposition 1.4.10 (Dot product and angle in the plane).

Let $\mathbf{v}, \mathbf{w}$ be nonzero vectors in $\mathbb{R}^2$ (or $\mathbb{R}^3$), and let θ be the angle between their arrows. Then

$$\mathbf{v} \cdot \mathbf{w} = \|\mathbf{v}\| \|\mathbf{w}\| \cos \theta.$$

Definition 1.4.13 (Orthogonal vectors).

Vectors $\mathbf{v}, \mathbf{w} \in \mathbb{R}^n$ are *orthogonal*, written $\mathbf{v} \perp \mathbf{w}$, when $\mathbf{v} \cdot \mathbf{w} = 0$.

Theorem 1.4.14 (Pythagoras in $\mathbb{R}^n$).

For $\mathbf{v}, \mathbf{w} \in \mathbb{R}^n$, $\|\mathbf{v} + \mathbf{w}\|^2 = \|\mathbf{v}\|^2 + \|\mathbf{w}\|^2$ if and only if $\mathbf{v} \perp \mathbf{w}$.

Theorem 1.4.15 (Cauchy–Schwarz inequality).

For all $\mathbf{v}, \mathbf{w} \in \mathbb{R}^n$,

$$|\mathbf{v} \cdot \mathbf{w}| \leq \|\mathbf{v}\| \|\mathbf{w}\|.$$

Equality holds if and only if one of the two vectors is a scalar multiple of the other.

Definition 1.4.16 (Angle in $\mathbb{R}^n$).

The *angle* between nonzero vectors $\mathbf{v}, \mathbf{w} \in \mathbb{R}^n$ is the unique $\theta \in [0, \pi]$ with

$$\cos \theta = \frac{\mathbf{v} \cdot \mathbf{w}}{\|\mathbf{v}\| \|\mathbf{w}\|}.$$

Theorem 1.4.18 (Triangle inequality).

For all $\mathbf{v}, \mathbf{w} \in \mathbb{R}^n$,

$$\|\mathbf{v} + \mathbf{w}\| \leq \|\mathbf{v}\| + \|\mathbf{w}\|.$$

Equality holds if and only if one of the vectors is a non-negative multiple of the other.

Proposition 1.4.19 (Triangle inequality for distances).

For all $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{R}^n$, $d(\mathbf{u}, \mathbf{w}) \leq d(\mathbf{u}, \mathbf{v}) + d(\mathbf{v}, \mathbf{w})$.

1.5 Lines, planes and the cross product

Lecture 5

Definition 1.5.1 (Line).

Let $\mathbf{p}, \mathbf{d} \in \mathbb{R}^n$ with $\mathbf{d} \neq \mathbf{0}$. The *line through $\mathbf{p}$ with direction vector $\mathbf{d}$* is the set

$$L = \{ \mathbf{p} + t\mathbf{d} : t \in \mathbb{R} \}.$$

The equation $\mathbf{x} = \mathbf{p} + t\mathbf{d}$ is called a *parametric equation* (or *parametric form*) of L , and t is the *parameter*.

Proposition 1.5.3 (When two lines are equal).

Let $L = \{ \mathbf{p} + t\mathbf{d} \}$ and $L' = \{ \mathbf{q} + s\mathbf{e} \}$ be lines in $\mathbb{R}^n$. Then $L = L'$ if and only if $\mathbf{q} \in L$ and $\mathbf{e}$ is a (nonzero) multiple of $\mathbf{d}$.

Proposition 1.5.4 (Lines in the plane as equations).

Let $\mathbf{n} = (a, b) \neq \mathbf{0}$ and $c \in \mathbb{R}$. Then the set $\{(x, y) \in \mathbb{R}^2 : ax + by = c\}$ is a line, with direction vector $(-b, a)$. Conversely, the line through $\mathbf{p}$ with direction $\mathbf{d} = (d_1, d_2)$ is the set of solutions of $\mathbf{n} \cdot \mathbf{x} = \mathbf{n} \cdot \mathbf{p}$, where $\mathbf{n} = (-d_2, d_1)$.

Definition 1.5.6 (Plane).

Let $\mathbf{p}, \mathbf{u}, \mathbf{v} \in \mathbb{R}^n$, where $\mathbf{u}$ and $\mathbf{v}$ are not parallel. The *plane through $\mathbf{p}$ with direction vectors $\mathbf{u}$ and $\mathbf{v}$* is the set

$$P = \{ \mathbf{p} + s\mathbf{u} + t\mathbf{v} : s, t \in \mathbb{R} \},$$

and $\mathbf{x} = \mathbf{p} + s\mathbf{u} + t\mathbf{v}$ is a *parametric equation* of P , with parameters s and t .

Proposition 1.5.8 (The solutions of one equation form a plane).

Let $\mathbf{n} = (a, b, c) \in \mathbb{R}^3$, $\mathbf{n} \neq \mathbf{0}$, and $d \in \mathbb{R}$. The set $\{\mathbf{x} \in \mathbb{R}^3 : \mathbf{n} \cdot \mathbf{x} = d\}$ is a plane, and both of its direction vectors can be chosen orthogonal to $\mathbf{n}$.

Definition 1.5.9 (Normal vector of a plane).

If a plane $P \subseteq \mathbb{R}^3$ is the solution set of $\mathbf{n} \cdot \mathbf{x} = d$ with $\mathbf{n} \neq \mathbf{0}$, then $\mathbf{n}$ is called a *normal vector* of P , and $\mathbf{n} \cdot \mathbf{x} = d$, or $ax + by + cz = d$, a *normal equation* of P . Equivalently, P is the set of $\mathbf{x}$ with $\mathbf{n} \cdot (\mathbf{x} - \mathbf{p}) = 0$, for any one point $\mathbf{p}$ of P .

Definition 1.5.12 (Cross product).

The *cross product* of $\mathbf{u} = (u_1, u_2, u_3)$ and $\mathbf{v} = (v_1, v_2, v_3)$ in $\mathbb{R}^3$ is the vector

$$\mathbf{u} \times \mathbf{v} = \begin{bmatrix} u_2 v_3 - u_3 v_2 \\ u_3 v_1 - u_1 v_3 \\ u_1 v_2 - u_2 v_1 \end{bmatrix}.$$

Theorem 1.5.14 (Properties of the cross product).

For all $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{R}^3$ and every scalar c :

- (a) *anticommutativity*: $\mathbf{v} \times \mathbf{u} = -(\mathbf{u} \times \mathbf{v})$; in particular $\mathbf{u} \times \mathbf{u} = \mathbf{0}$;
- (b) *additivity*: $\mathbf{u} \times (\mathbf{v} + \mathbf{w}) = \mathbf{u} \times \mathbf{v} + \mathbf{u} \times \mathbf{w}$;
- (c) *homogeneity*: $(c\mathbf{u}) \times \mathbf{v} = c(\mathbf{u} \times \mathbf{v}) = \mathbf{u} \times (c\mathbf{v})$;
- (d) *orthogonality*: $\mathbf{u} \cdot (\mathbf{u} \times \mathbf{v}) = 0$ and $\mathbf{v} \cdot (\mathbf{u} \times \mathbf{v}) = 0$.

Theorem 1.5.16 (Length of the cross product).

For all $\mathbf{u}, \mathbf{v} \in \mathbb{R}^3$,

$$\|\mathbf{u} \times \mathbf{v}\|^2 = \|\mathbf{u}\|^2 \|\mathbf{v}\|^2 - (\mathbf{u} \cdot \mathbf{v})^2. \quad (1)$$

If $\mathbf{u}$ and $\mathbf{v}$ are nonzero and θ is the angle between them, then $\|\mathbf{u} \times \mathbf{v}\| = \|\mathbf{u}\| \|\mathbf{v}\| \sin \theta$, which is the area of the parallelogram with sides $\mathbf{u}$ and $\mathbf{v}$.

Proposition 1.5.17 (Parallel vectors).

For $\mathbf{u}, \mathbf{v} \in \mathbb{R}^3$, $\mathbf{u} \times \mathbf{v} = \mathbf{0}$ if and only if $\mathbf{u}$ and $\mathbf{v}$ are parallel (one is a multiple of the other).

Proposition 1.5.19 (Triple cross product).

For all $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{R}^3$, $(\mathbf{u} \times \mathbf{v}) \times \mathbf{w} = (\mathbf{u} \cdot \mathbf{w}) \mathbf{v} - (\mathbf{v} \cdot \mathbf{w}) \mathbf{u}$.

Proposition 1.5.20 (Normal equation from a parametric plane).

Let P be the plane through $\mathbf{p}$ with directions $\mathbf{u}, \mathbf{v} \in \mathbb{R}^3$, and let $\mathbf{n} = \mathbf{u} \times \mathbf{v}$. Then $\mathbf{n} \neq \mathbf{0}$ and

$$P = \{\mathbf{x} \in \mathbb{R}^3 : \mathbf{n} \cdot (\mathbf{x} - \mathbf{p}) = 0\}.$$

In particular every plane in $\mathbb{R}^3$ has a normal equation.

Proposition 1.5.22 (Distance from a point to a plane).

Let P be the plane $\mathbf{n} \cdot \mathbf{x} = d$ in $\mathbb{R}^3$ and let $\mathbf{q} \in \mathbb{R}^3$. The point $\mathbf{f} = \mathbf{q} - \frac{\mathbf{n} \cdot \mathbf{q} - d}{\|\mathbf{n}\|^2} \mathbf{n}$ lies on P , and it is the point of P closest to $\mathbf{q}$: for every $\mathbf{x} \in P$ with $\mathbf{x} \neq \mathbf{f}$, $\|\mathbf{q} - \mathbf{x}\| > \|\mathbf{q} - \mathbf{f}\|$. The distance from $\mathbf{q}$ to P is therefore

$$\|\mathbf{q} - \mathbf{f}\| = \frac{|\mathbf{n} \cdot \mathbf{q} - d|}{\|\mathbf{n}\|}.$$

Chapter 2. Fields and linear systems

2.1 Fields

Lecture 6

Definition 2.1.1 (Field).

A *field* is a set $\mathbb{F}$ together with two operations, *addition* $(a, b) \mapsto a + b$ and *multiplication* $(a, b) \mapsto ab$ (also written $a \cdot b$), each taking two elements of $\mathbb{F}$ to an element of $\mathbb{F}$, such that for all $a, b, c \in \mathbb{F}$:

- (A1) $(a + b) + c = a + (b + c)$ (addition is associative)
 (A2) $a + b = b + a$ (addition is commutative)
 (A3) there is an element $0 \in \mathbb{F}$ with $a + 0 = a$ for every a (zero)
 (A4) for every a there is an element $-a \in \mathbb{F}$ with $a + (-a) = 0$ (negatives)
 (M1) $(ab)c = a(bc)$ (multiplication is associative)
 (M2) $ab = ba$ (multiplication is commutative)
 (M3) there is an element $1 \in \mathbb{F}$ with $1 \neq 0$ and $1a = a$ for every a (one)
 (M4) for every $a \neq 0$ there is an element $a^{-1} \in \mathbb{F}$ with $a a^{-1} = 1$ (inverses)
 (D) $a(b + c) = ab + ac$ (distributive law)

These nine conditions are the *field axioms*. The elements of a field are called *scalars*.

Proposition 2.1.4 (Uniqueness).

In a field $\mathbb{F}$:

- (a) there is only one zero and only one one;
- (b) each a has only one negative, and each $a \neq 0$ has only one inverse.

Proposition 2.1.5 (Arithmetic in a field).

Let a, b, c be elements of a field $\mathbb{F}$. Then:

- (a) $0a = 0$;
- (b) $(-1)a = -a$;
- (c) $-(-a) = a$, $(-a)b = -(ab)$ and $(-a)(-b) = ab$;
- (d) (no zero divisors) if $ab = 0$, then $a = 0$ or $b = 0$;
- (e) (cancellation) if $a + c = b + c$, then $a = b$; if $ac = bc$ and $c \neq 0$, then $a = b$.

Proposition 2.1.6 (One equation in one unknown).

Let $a, b \in \mathbb{F}$ with $a \neq 0$. The equation $ax = b$ has exactly one solution in $\mathbb{F}$, namely $x = a^{-1}b$.

Definition 2.1.7 (Complex numbers).

A *complex number* is an expression $z = a + bi$ with $a, b \in \mathbb{R}$; a is its *real part* and b its *imaginary part*. Two complex numbers are equal when their real parts and their imaginary parts agree. The set of complex numbers is $\mathbb{C}$, and its operations are

$$(a + bi) + (c + di) = (a + c) + (b + d)i, \quad (a + bi)(c + di) = (ac - bd) + (ad + bc)i.$$

The *conjugate* of $z = a + bi$ is $\bar{z} = a - bi$.

Theorem 2.1.9.

$\mathbb{C}$ is a field, with zero $0 = 0 + 0i$, one $1 = 1 + 0i$, negative $-(a + bi) = (-a) + (-b)i$, and, for $a + bi \neq 0$,

$$(a + bi)^{-1} = \frac{a - bi}{a^2 + b^2} = \frac{a}{a^2 + b^2} - \frac{b}{a^2 + b^2}i.$$

Definition 2.1.12 (Integers modulo n).

For an integer $n \geq 2$, the *integers modulo n* are the set $\mathbb{Z}_n = \{0, 1, \dots, n - 1\}$ with the operations

$$a \oplus b = (a + b) \bmod n, \quad a \odot b = (ab) \bmod n.$$

We usually write $a + b$ and ab for these operations, and say “in $\mathbb{Z}_n$ ” to make clear which ones are meant.

Proposition 2.1.13.

For every $n \geq 2$, $\mathbb{Z}_n$ satisfies all field axioms except possibly (M4), with zero 0, one 1, and negative $-a = n - a$ for $a \neq 0$ (and $-0 = 0$).

Proposition 2.1.15 (Euclid's lemma).

Let p be a prime number and a, b integers. If p divides ab , then p divides a or p divides b .

Theorem 2.1.16.

$\mathbb{Z}_n$ is a field if and only if n is prime.

Definition 2.1.18 (The space $\mathbb{F}^n$).

Let $\mathbb{F}$ be a field. $\mathbb{F}^n$ is the set of columns $\mathbf{v} = (v_1, \dots, v_n)$ with entries $v_i \in \mathbb{F}$. Addition and scalar multiplication are entry by entry, as in $\mathbb{R}^n$: $(\mathbf{v} + \mathbf{w})_i = v_i + w_i$ and $(c \mathbf{v})_i = c v_i$ for $c \in \mathbb{F}$.

2.2 Systems of linear equations

Lecture 7

Definition 2.2.1 (Linear equation).

A *linear equation* in the unknowns $x_1, \dots, x_n$ is an equation of the form

$$a_1x_1 + a_2x_2 + \dots + a_nx_n = b,$$

where the *coefficients* $a_1, \dots, a_n$ and the *constant term* b are elements of $\mathbb{F}$. A *solution* of the equation is a list $(s_1, \dots, s_n) \in \mathbb{F}^n$ such that $a_1s_1 + \dots + a_ns_n = b$.

Definition 2.2.3 (System of linear equations).

A *system of m linear equations in n unknowns* is a list of linear equations in the same unknowns $x_1, \dots, x_n$:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n &= b_1, \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n &= b_2, \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n &= b_m. \end{aligned}$$

The coefficient a_{ij} stands in equation i in front of unknown j : the first index names the equation (the row), the second the unknown (the column). A *solution* of the system is a list $(s_1, \dots, s_n) \in \mathbb{F}^n$ that is a solution of every equation at once. The set of all solutions is the *solution set*. The system is *consistent* if it has at least one solution and *inconsistent* if it has none.

Definition 2.2.5 (Equivalent systems).

Two systems of linear equations in the same unknowns are *equivalent* if they have the same solution set.

Theorem 2.2.7 (Zero, one or infinitely many).

Let $\mathbb{F}$ be $\mathbb{R}$, or more generally any field with infinitely many elements, such as $\mathbb{Q}$ or $\mathbb{C}$. A system of linear equations over $\mathbb{F}$ has either no solution, exactly one solution, or infinitely many solutions. More precisely, if $\mathbf{s} \neq \mathbf{t}$ are two solutions, then $\mathbf{s} + c(\mathbf{t} - \mathbf{s})$ is a solution for every $c \in \mathbb{F}$, and these solutions are distinct for distinct c .

Proposition 2.2.9.

A list $(s_1, \dots, s_n)$ solves the system exactly when $s_1\mathbf{a}_1 + \dots + s_n\mathbf{a}_n = \mathbf{b}$. Consequently the system is consistent if and only if

$$\mathbf{b} \in \text{Span}\{\mathbf{a}_1, \dots, \mathbf{a}_n\}.$$

Definition 2.2.11 (Matrix, coefficient matrix, augmented matrix).

An $m \times n$ matrix over $\mathbb{F}$ is a rectangular array of elements of $\mathbb{F}$ with m rows and n columns. The entry in row i and column j is written a_{ij} . The *coefficient matrix* of the system in Definition 2.2.3 is the $m \times n$ matrix $A = (a_{ij})$, with one row per equation and one column per unknown; its *augmented matrix* is the $m \times (n + 1)$ matrix obtained by adding the column $\mathbf{b}$ of constant terms,

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}, \quad [A \mid \mathbf{b}] = \left[\begin{array}{ccc|c} a_{11} & \cdots & a_{1n} & b_1 \\ \vdots & & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} & b_m \end{array} \right].$$

The vertical bar only marks where the constant terms begin.

2.3 Gaussian elimination

Lecture 8

Definition 2.3.1 (Elementary row operations).

The *elementary row operations* on a matrix are:

- (1) *swap*: exchange two rows, written $R_i \leftrightarrow R_j$;
- (2) *scaling*: multiply a row by a *nonzero* scalar c , written $R_i \rightarrow c R_i$;
- (3) *replacement*: add c times one row to a *different* row, written $R_i \rightarrow R_i + c R_j$ ($i \neq j$).

The arrow $\rightarrow$ is read “is replaced by”.

Definition 2.3.3 (Row equivalence).

Two matrices of the same size are *row equivalent* if one can be obtained from the other by a finite sequence of elementary row operations.

Theorem 2.3.4 (Row operations keep the solution set).

If two augmented matrices are row equivalent, then the corresponding linear systems have exactly the same solutions.

Definition 2.3.5 (Row echelon form).

In a nonzero row of a matrix, the first entry that is not zero is the *leading entry* of the row. A matrix is in *row echelon form* if

- (i) all rows consisting entirely of zeros are below all nonzero rows, and
- (ii) the leading entry of each nonzero row lies in a column strictly to the right of the leading entry of the row above it.

The leading entries of a matrix in row echelon form are called its *pivots*, and the columns that contain them are its *pivot columns*.

Theorem 2.3.7 (Existence of a row echelon form).

Every matrix is row equivalent to a matrix in row echelon form.

Proposition 2.3.11.

If a row echelon form of an augmented matrix $[A \mid \mathbf{b}]$ has a row $[0 \cdots 0 \mid c]$ with $c \neq 0$, then the system has no solution.

2.4 Reduced row echelon form

Lecture 9

Definition 2.4.1 (Reduced row echelon form).

A matrix is in *reduced row echelon form* if

- (i) all zero rows are at the bottom, and
- (ii) each leading entry is strictly to the right of the leading entry of the row above it,

that is, it is in row echelon form (Definition 2.3.5), and moreover

- (iii) every pivot equals 1, and
- (iv) each pivot is the only nonzero entry in its column.

Theorem 2.4.4 (Existence of the reduced row echelon form).

Every matrix is row equivalent to a matrix in reduced row echelon form.

Definition 2.4.6 (Basic and free variables).

Consider a linear system whose augmented matrix $[A \mid \mathbf{b}]$ is row equivalent to the matrix $[R \mid \mathbf{d}]$ in reduced row echelon form. The unknown x_j is a *basic variable* if column j of R is a pivot column, and a *free variable* otherwise. Basic variables are also called *pivot variables*.

Proposition 2.4.7 (Reading off the solutions).

Let $[R \mid \mathbf{d}]$ be in reduced row echelon form, with r nonzero rows and pivots in columns $j_1 < \cdots < j_r$, none of them the last column. Then for every choice of values of the free variables there is exactly one solution of the system with those values, namely

$$x_{j_i} = d_i - \sum_{k \text{ free}} r_{ik} x_k \quad (i = 1, \dots, r).$$

Theorem 2.4.9 (Uniqueness of the reduced row echelon form).

Every matrix is row equivalent to *exactly one* matrix in reduced row echelon form.

Proposition 2.4.11 (Row operations keep column relations).

If A and B are row equivalent $m \times n$ matrices with columns $\mathbf{a}_1, \dots, \mathbf{a}_n$ and $\mathbf{b}_1, \dots, \mathbf{b}_n$, then for all scalars $c_1, \dots, c_n$,

$$c_1 \mathbf{a}_1 + \cdots + c_n \mathbf{a}_n = \mathbf{0} \quad \Longleftrightarrow \quad c_1 \mathbf{b}_1 + \cdots + c_n \mathbf{b}_n = \mathbf{0}.$$

Proposition 2.4.12 (Columns of a reduced matrix).

Let R be an $m \times n$ matrix in reduced row echelon form with pivots in columns $j_1 < \cdots < j_r$, and columns $\mathbf{r}_1, \dots, \mathbf{r}_n$. Then:

- (a) the pivot columns are $\mathbf{r}_{j_i} = \mathbf{e}_i$ for $i = 1, \dots, r$;
- (b) column k is a pivot column if and only if $\mathbf{r}_k$ is *not* a linear combination of $\mathbf{r}_1, \dots, \mathbf{r}_{k-1}$ (for $k = 1$: if and only if $\mathbf{r}_1 \neq \mathbf{0}$);
- (c) if column k is not a pivot column, then $\mathbf{r}_k = \sum_{i: j_i < k} r_{ik} \mathbf{r}_{j_i}$, where r_{ik} is the entry of R in row i and column k .

Definition 2.4.13 (Pivot positions of a matrix).

The *pivot positions* of a matrix A are the positions of the pivots (the leading 1s) in its reduced row echelon form, and the *pivot columns* of A are the columns that contain them.

Proposition 2.4.14.

Every row echelon form of A has its pivots exactly in the pivot positions of A . In particular, all echelon forms of A have the same number of pivots.

Definition 2.5.2 (General solution, parametric vector form).

Let a consistent system have free variables $x_{j_1}, \dots, x_{j_k}$. Its *general solution* expresses every solution in terms of *parameters* $t_1, \dots, t_k \in \mathbb{F}$, the values given to the free variables. Written as

$$\mathbf{x} = \mathbf{p} + t_1 \mathbf{v}_1 + \dots + t_k \mathbf{v}_k,$$

with fixed vectors $\mathbf{p}, \mathbf{v}_1, \dots, \mathbf{v}_k \in \mathbb{F}^n$, it is in *parametric vector form*. The vector $\mathbf{p}$ (all parameters zero) is itself a solution, called a *particular solution*.

Theorem 2.5.3 (Existence and uniqueness).

Let E be a row echelon form of the augmented matrix $[A \mid \mathbf{b}]$ of a system of m linear equations in n unknowns over a field $\mathbb{F}$.

- (a) (*Existence*) The system is consistent if and only if the last column of E is not a pivot column, that is, if and only if E has no row $[0 \ \dots \ 0 \mid c]$ with $c \neq 0$.
- (b) (*Uniqueness*) Suppose the system is consistent, and let k be the number of free variables. Then for every choice of values of the free variables there is exactly one solution taking those values. In particular the solution is unique if and only if $k = 0$, that is, every column of A is a pivot column. If $k \geq 1$ and $\mathbb{F}$ is infinite (such as $\mathbb{R}$), there are infinitely many solutions.

Proposition 2.5.5 (More unknowns than equations).

If a consistent system of m equations in n unknowns has $n > m$, then it has at least $n - m$ free variables. Over $\mathbb{R}$ (or any infinite field) it therefore has infinitely many solutions.

Definition 2.5.7 (Homogeneous system).

A system of linear equations is *homogeneous* if all its right-hand sides are zero, that is, its augmented matrix is $[A \mid \mathbf{0}]$. The zero vector $\mathbf{x} = \mathbf{0}$ is always a solution, called the *trivial solution*; any other solution is *nontrivial*.

Proposition 2.5.8.

Row operations keep the last column of $[A \mid \mathbf{0}]$ zero. Consequently a homogeneous system has a nontrivial solution if and only if it has at least one free variable.

Theorem 2.5.9 (Nontrivial solutions).

A homogeneous system with more unknowns than equations ($n > m$) has a nontrivial solution. This holds over every field.

Theorem 2.5.11 (Structure of the solution set).

Let $\mathbf{p}$ be one solution of a system with augmented matrix $[A \mid \mathbf{b}]$, and let H be the solution set of the homogeneous system $[A \mid \mathbf{0}]$. Then the solution set of $[A \mid \mathbf{b}]$ is

$$\mathbf{p} + H = \{ \mathbf{p} + \mathbf{h} : \mathbf{h} \in H \}.$$

In words: every solution is the particular solution $\mathbf{p}$ plus a solution of the homogeneous system, and every such sum is a solution.

2.6 Linear independence

Definition 2.6.1 (Linear independence).

Vectors $\mathbf{v}_1, \dots, \mathbf{v}_k \in \mathbb{F}^n$ are *linearly independent* if the only scalars $c_1, \dots, c_k \in \mathbb{F}$ with

$$c_1 \mathbf{v}_1 + c_2 \mathbf{v}_2 + \dots + c_k \mathbf{v}_k = \mathbf{0}$$

are $c_1 = c_2 = \cdots = c_k = 0$. They are *linearly dependent* if there are scalars $c_1, \dots, c_k$, *not all zero*, with $c_1 \mathbf{v}_1 + \cdots + c_k \mathbf{v}_k = \mathbf{0}$; such an equation is called a *linear dependence relation*.

Theorem 2.6.3.

Let $k \geq 2$. Vectors $\mathbf{v}_1, \dots, \mathbf{v}_k$ are linearly dependent if and only if at least one of them is a linear combination of the others.

Proposition 2.6.5.

Let $\mathbf{v}, \mathbf{w}, \mathbf{v}_1, \dots, \mathbf{v}_k, \mathbf{u}_1, \dots, \mathbf{u}_m \in \mathbb{F}^n$.

- (a) A single vector $\mathbf{v}$ is independent if and only if $\mathbf{v} \neq \mathbf{0}$.
- (b) Every list that contains $\mathbf{0}$ is dependent.
- (c) Two vectors $\mathbf{v}, \mathbf{w}$ are dependent if and only if one of them is a scalar multiple of the other.
- (d) If $\mathbf{v}_1, \dots, \mathbf{v}_k$ are dependent, then so are $\mathbf{v}_1, \dots, \mathbf{v}_k, \mathbf{u}_1, \dots, \mathbf{u}_m$. Equivalently, if a list is independent, so is every list obtained from it by deleting vectors.

Theorem 2.6.7.

Vectors $\mathbf{v}_1, \dots, \mathbf{v}_k$ are linearly independent if and only if $\mathbf{v}_1 \neq \mathbf{0}$ and, for each $j = 2, \dots, k$,

$$\mathbf{v}_j \notin \text{Span}\{\mathbf{v}_1, \dots, \mathbf{v}_{j-1}\}.$$

Theorem 2.6.8 (Elimination test).

Let A be the matrix with columns $\mathbf{v}_1, \dots, \mathbf{v}_k$, and let E be a row echelon form of A . The following are equivalent:

- (a) $\mathbf{v}_1, \dots, \mathbf{v}_k$ are linearly independent;
- (b) the homogeneous system $[A \mid \mathbf{0}]$ has only the trivial solution;
- (c) every column of E is a pivot column.

If column j of E has no pivot, setting $c_j = 1$, every other non-pivot unknown to 0, and solving for the pivot unknowns by back substitution gives a linear dependence relation.

Theorem 2.6.10.

If $k > n$, then any k vectors in $\mathbb{F}^n$ are linearly dependent.

Chapter 3. Matrices and linear transformations

3.1 The matrix-vector product

Lecture 12

Definition 3.1.1 (Matrix-vector product).

Let A be an $m \times n$ matrix over $\mathbb{F}$ with columns $\mathbf{a}_1, \dots, \mathbf{a}_n$, and let $\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{F}^n$. The *product* of A and $\mathbf{x}$ is the vector

$$A\mathbf{x} = x_1 \mathbf{a}_1 + x_2 \mathbf{a}_2 + \cdots + x_n \mathbf{a}_n \in \mathbb{F}^m,$$

the linear combination of the columns of A with the entries of $\mathbf{x}$ as weights. The product is defined only when $\mathbf{x}$ has as many entries as A has columns.

Proposition 3.1.4 (Row rule).

For $i = 1, \dots, m$, the i -th entry of $A\mathbf{x}$ is

$$(A\mathbf{x})_i = a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n.$$

Equivalently, $(A\mathbf{x})_i = \mathbf{r}_i \cdot \mathbf{x}$, where $\mathbf{r}_i = (a_{i1}, \dots, a_{in})$ is the i -th row of A and $\cdot$ is the dot product on $\mathbb{F}^n$ (Sections 1.4 and 2.2).

Definition 3.1.5 (Identity matrix).

The *identity matrix* I_n is the $n \times n$ matrix with 1 in every diagonal position (i, i) and 0 in every other position. For example,

$$I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

We write I when the size is clear.

Theorem 3.1.6 (Linearity of the product).

Let A be an $m \times n$ matrix over $\mathbb{F}$. For all $\mathbf{u}, \mathbf{v} \in \mathbb{F}^n$ and all $c, d \in \mathbb{F}$:

- (a) $A(\mathbf{u} + \mathbf{v}) = A\mathbf{u} + A\mathbf{v}$;
- (b) $A(c\mathbf{u}) = c(A\mathbf{u})$;
- (c) $A(c\mathbf{u} + d\mathbf{v}) = cA\mathbf{u} + dA\mathbf{v}$.

Theorem 3.1.7 (Three forms of a linear system).

For $\mathbf{s} \in \mathbb{F}^n$ the following are equivalent:

- (1) $\mathbf{s}$ is a solution of the system above;
- (2) $s_1\mathbf{a}_1 + \cdots + s_n\mathbf{a}_n = \mathbf{b}$;
- (3) $A\mathbf{s} = \mathbf{b}$.

In particular, the system, the vector equation $x_1\mathbf{a}_1 + \cdots + x_n\mathbf{a}_n = \mathbf{b}$ and the matrix equation $A\mathbf{x} = \mathbf{b}$ have the same solution set.

Theorem 3.1.9.

For an $m \times n$ matrix A over $\mathbb{F}$ the following are equivalent:

- (a) $A\mathbf{x} = \mathbf{b}$ has a solution for every $\mathbf{b} \in \mathbb{F}^m$;
- (b) every $\mathbf{b} \in \mathbb{F}^m$ is a linear combination of the columns of A ;
- (c) the columns of A span $\mathbb{F}^m$, that is, $\text{Span}\{\mathbf{a}_1, \dots, \mathbf{a}_n\} = \mathbb{F}^m$;
- (d) an echelon form of A has a pivot in every row.

Theorem 3.1.11 (Structure of the solution set, matrix form).

Let $\mathbf{p} \in \mathbb{F}^n$ satisfy $A\mathbf{p} = \mathbf{b}$. Then the solutions of $A\mathbf{x} = \mathbf{b}$ are exactly the vectors

$$\mathbf{x} = \mathbf{p} + \mathbf{h}, \quad \text{where } A\mathbf{h} = \mathbf{0}.$$

Proposition 3.1.13 (Uniqueness).

For an $m \times n$ matrix A with columns $\mathbf{a}_1, \dots, \mathbf{a}_n$, the following are equivalent:

- (a) whenever $A\mathbf{x} = \mathbf{b}$ has a solution, it has only one;
- (b) the homogeneous system $A\mathbf{x} = \mathbf{0}$ has only the trivial solution;
- (c) the columns $\mathbf{a}_1, \dots, \mathbf{a}_n$ are linearly independent;
- (d) an echelon form of A has a pivot in every column.

3.2 Linear transformations of $\mathbb{R}^n$

Lecture 13

Definition 3.2.1 (Linear transformation).

A transformation $T: \mathbb{F}^n \rightarrow \mathbb{F}^m$ is *linear* if

- (i) $T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$ for all $\mathbf{u}, \mathbf{v} \in \mathbb{F}^n$, and
- (ii) $T(c\mathbf{u}) = cT(\mathbf{u})$ for all $\mathbf{u} \in \mathbb{F}^n$ and all $c \in \mathbb{F}$.

Proposition 3.2.2 (Matrix transformations are linear).

For every $m \times n$ matrix A over $\mathbb{F}$, the matrix transformation T_A is linear.

Proposition 3.2.3 (First consequences).

Let $T: \mathbb{F}^n \rightarrow \mathbb{F}^m$ be linear. Then

- (a) $T(\mathbf{0}) = \mathbf{0}$;
- (b) $T(c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k) = c_1T(\mathbf{v}_1) + \cdots + c_kT(\mathbf{v}_k)$ for all $\mathbf{v}_1, \dots, \mathbf{v}_k \in \mathbb{F}^n$ and $c_1, \dots, c_k \in \mathbb{F}$.

Conversely, a transformation T with $T(c\mathbf{u} + d\mathbf{v}) = cT(\mathbf{u}) + dT(\mathbf{v})$ for all $\mathbf{u}, \mathbf{v} \in \mathbb{F}^n$ and $c, d \in \mathbb{F}$ is linear.

Definition 3.2.6 (Standard basis vectors).

For $j = 1, \dots, n$, the *standard basis vector* $\mathbf{e}_j \in \mathbb{F}^n$ has 1 in position j and 0 in every other position:

$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \quad \dots, \quad \mathbf{e}_n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}.$$

Theorem 3.2.7 (The standard matrix).

Let $T: \mathbb{F}^n \rightarrow \mathbb{F}^m$ be linear. There is exactly one $m \times n$ matrix A over $\mathbb{F}$ with

$$T(\mathbf{x}) = A\mathbf{x} \quad \text{for every } \mathbf{x} \in \mathbb{F}^n,$$

namely the matrix whose columns are the images of the standard basis vectors:

$$A = \begin{bmatrix} T(\mathbf{e}_1) & T(\mathbf{e}_2) & \cdots & T(\mathbf{e}_n) \end{bmatrix}.$$

It is called the *standard matrix* of T . In particular, the linear transformations $\mathbb{F}^n \rightarrow \mathbb{F}^m$ are exactly the matrix transformations T_A .

Theorem 3.2.9 (Rotations).

The rotation R_θ is linear, with standard matrix

$$\begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}.$$

Proposition 3.2.12 (Projection onto a line).

For every $\mathbf{x} \in \mathbb{R}^n$ there is exactly one point $\mathbf{p} \in L$ with $\mathbf{x} - \mathbf{p}$ orthogonal to $\mathbf{u}$, namely

$$\mathbf{p} = P(\mathbf{x}) = (\mathbf{x} \cdot \mathbf{u}) \mathbf{u},$$

and it is the point of L closest to $\mathbf{x}$. The map $P: \mathbb{R}^n \rightarrow \mathbb{R}^n$, the *orthogonal projection onto L* , is linear, and its standard matrix has entry $u_i u_j$ in row i , column j . In $\mathbb{R}^2$, with $\mathbf{u} = (u_1, u_2)$,

$$P = \begin{bmatrix} u_1^2 & u_1 u_2 \\ u_1 u_2 & u_2^2 \end{bmatrix}.$$

Proposition 3.2.14 (Reflection across a line).

The reflection $Q: \mathbb{R}^n \rightarrow \mathbb{R}^n$ across $L = \text{Span}\{\mathbf{u}\}$ is linear, and its standard matrix has entry $2u_i^2 - 1$ in each diagonal position (i, i) and $2u_i u_j$ in position (i, j) for $i \neq j$. In $\mathbb{R}^2$,

$$Q = \begin{bmatrix} 2u_1^2 - 1 & 2u_1 u_2 \\ 2u_1 u_2 & 2u_2^2 - 1 \end{bmatrix}.$$

3.3 Matrix multiplication

Lecture 14

Definition 3.3.1 (Matrix product).

Let A be an $m \times n$ matrix and B an $n \times p$ matrix with columns $\mathbf{b}_1, \dots, \mathbf{b}_p$. The *product* AB is the $m \times p$ matrix

$$AB = [\mathbf{A}\mathbf{b}_1 \quad \mathbf{A}\mathbf{b}_2 \quad \cdots \quad \mathbf{A}\mathbf{b}_p],$$

whose j -th column is A times the j -th column of B . The product is defined only when A has as many columns as B has rows.

Theorem 3.3.2 (The product is composition).

Let A be $m \times n$ and B be $n \times p$. Then for every $\mathbf{x} \in \mathbb{F}^p$,

$$(AB)\mathbf{x} = A(B\mathbf{x}).$$

Proposition 3.3.3 (Composition of linear transformations).

Let $S: \mathbb{F}^p \rightarrow \mathbb{F}^n$ and $T: \mathbb{F}^n \rightarrow \mathbb{F}^m$ be linear transformations with standard matrices B and A (Theorem 3.2.7). Then the composite $T \circ S: \mathbb{F}^p \rightarrow \mathbb{F}^m$, $\mathbf{x} \mapsto T(S(\mathbf{x}))$, is a linear transformation, and its standard matrix is AB .

Theorem 3.3.5 (Row-column rule).

Let $A = (a_{ik})$ be $m \times n$ and $B = (b_{kj})$ be $n \times p$. For $i = 1, \dots, m$ and $j = 1, \dots, p$,

$$(AB)_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj} = \sum_{k=1}^n a_{ik}b_{kj}.$$

In words: entry (i, j) of AB is row i of A times column j of B , entry by entry and added up; it is the dot product $\mathbf{r}_i \cdot \mathbf{b}_j$ of the i -th row of A with the j -th column of B .

Definition 3.3.7 (Sum and scalar multiple of matrices).

Let $A = (a_{ij})$ and $B = (b_{ij})$ be $m \times n$ matrices and $c \in \mathbb{F}$. The *sum* $A + B$ and the *scalar multiple* cA are the $m \times n$ matrices with entries

$$(A + B)_{ij} = a_{ij} + b_{ij}, \quad (cA)_{ij} = c a_{ij}.$$

The *zero matrix* O (or $O_{m \times n}$) has every entry 0, and $-A = (-1)A$. The sum is defined only for matrices of the same size.

Proposition 3.3.8.

Let A and B be $m \times n$ matrices, $c \in \mathbb{F}$ and $\mathbf{x} \in \mathbb{F}^n$. Then

$$(A + B)\mathbf{x} = A\mathbf{x} + B\mathbf{x}, \quad (cA)\mathbf{x} = c(A\mathbf{x}).$$

Theorem 3.3.9 (Rules of matrix multiplication).

Let $c \in \mathbb{F}$, and let A, B, C be matrices of sizes for which each side below is defined. Then:

- (a) $(AB)C = A(BC)$ (associativity)
- (b) $A(B + C) = AB + AC$ (distributive law, left)
- (c) $(A + B)C = AC + BC$ (distributive law, right)
- (d) $c(AB) = (cA)B = A(cB)$ (scalars)
- (e) $I_m A = A = A I_n$ for every $m \times n$ matrix A (identity)

Definition 3.3.10 (Powers).

Let A be an $n \times n$ matrix. Its powers are $A^0 = I_n$, $A^1 = A$, $A^2 = AA$, and in general $A^{k+1} = A^k A$ for $k \geq 0$.

Definition 3.4.1 (Transpose).

The *transpose* of an $m \times n$ matrix A is the $n \times m$ matrix A^T whose rows are the columns of A :

$$(A^T)_{ij} = a_{ji} \quad (1 \leq i \leq n, 1 \leq j \leq m).$$

Theorem 3.4.3 (Rules for the transpose).

Let $c \in \mathbb{F}$. Whenever the sizes allow the operations on the left:

- (a) $(A^T)^T = A$;
- (b) $(A + B)^T = A^T + B^T$;
- (c) $(cA)^T = cA^T$;
- (d) $(AB)^T = B^T A^T$.

Proposition 3.4.5.

Let A be an $m \times n$ matrix, $\mathbf{x} \in \mathbb{F}^n$ and $\mathbf{y} \in \mathbb{F}^m$. Then $(A\mathbf{x}) \cdot \mathbf{y} = \mathbf{x} \cdot (A^T \mathbf{y})$.

Definition 3.4.6 (Symmetric and skew-symmetric matrices).

A square matrix A is *symmetric* if $A^T = A$, that is, $a_{ij} = a_{ji}$ for all i, j . It is *skew-symmetric* if $A^T = -A$, that is, $a_{ji} = -a_{ij}$ for all i, j .

Proposition 3.4.8.

Symmetric matrices can be built from any matrix:

- (a) For every $m \times n$ matrix A , the $n \times n$ matrix $A^T A$ and the $m \times m$ matrix AA^T are symmetric. Moreover $(A^T A)_{ij} = \mathbf{a}_i \cdot \mathbf{a}_j$, where $\mathbf{a}_1, \dots, \mathbf{a}_n$ are the columns of A .
- (b) For every square matrix A , the matrix $A + A^T$ is symmetric and $A - A^T$ is skew-symmetric.
- (c) If $1 + 1 \neq 0$ in $\mathbb{F}$ (for instance over $\mathbb{R}$ or $\mathbb{C}$), every square matrix is, in exactly one way, the sum of a symmetric and a skew-symmetric matrix:

$$A = \frac{1}{2} (A + A^T) + \frac{1}{2} (A - A^T).$$

Definition 3.4.10 (Triangular and diagonal matrices).

A square matrix $A = (a_{ij})$ is

- *upper triangular* if $a_{ij} = 0$ whenever $i > j$ (all entries below the main diagonal are zero);
- *lower triangular* if $a_{ij} = 0$ whenever $i < j$ (all entries above the main diagonal are zero);
- *diagonal* if $a_{ij} = 0$ whenever $i \neq j$, that is, if it is both upper and lower triangular.

The diagonal matrix with diagonal entries $d_1, \dots, d_n$ is written $\text{diag}(d_1, \dots, d_n)$.

Theorem 3.4.11 (Products of triangular matrices).

Let A and B be $n \times n$ matrices.

- (a) If A and B are upper triangular, then AB is upper triangular and $(AB)_{ii} = a_{ii}b_{ii}$ for every i .
- (b) If A and B are lower triangular, then AB is lower triangular and $(AB)_{ii} = a_{ii}b_{ii}$ for every i .
- (c) $\text{diag}(d_1, \dots, d_n) \text{diag}(e_1, \dots, e_n) = \text{diag}(d_1e_1, \dots, d_ne_n)$. In particular, diagonal matrices of the same size commute.

Theorem 3.4.13 (Block multiplication).

Let A be $m \times n$ and B be $n \times p$. Cut the rows of A into groups of sizes $m_1, \dots, m_R$, the columns of A into groups of sizes $n_1, \dots, n_S$, the rows of B into the *same* groups $n_1, \dots, n_S$, and the columns of B into groups of sizes $p_1, \dots, p_T$. Then AB , cut into rows $m_1, \dots, m_R$ and columns $p_1, \dots, p_T$, has blocks

$$(AB)_{rt} = \sum_{s=1}^S A_{rs} B_{st} = A_{r1} B_{1t} + A_{r2} B_{2t} + \dots + A_{rS} B_{St}.$$

In particular, for 2×2 block matrices with matching cuts,

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} = \begin{bmatrix} A_{11}B_{11} + A_{12}B_{21} & A_{11}B_{12} + A_{12}B_{22} \\ A_{21}B_{11} + A_{22}B_{21} & A_{21}B_{12} + A_{22}B_{22} \end{bmatrix}.$$

Definition 3.4.15 (Trace).

The *trace* of an $n \times n$ matrix A is the sum of its diagonal entries:

$$\operatorname{tr} A = a_{11} + a_{22} + \dots + a_{nn} = \sum_{i=1}^n a_{ii}.$$

Theorem 3.4.16 (Properties of the trace).

Let A and B be $n \times n$ matrices and $c \in \mathbb{F}$.

- (a) $\operatorname{tr}(A + B) = \operatorname{tr} A + \operatorname{tr} B$ and $\operatorname{tr}(cA) = c \operatorname{tr} A$;
- (b) $\operatorname{tr}(A^T) = \operatorname{tr} A$;
- (c) more generally, if A is $m \times n$ and B is $n \times m$, then $\operatorname{tr}(AB) = \operatorname{tr}(BA)$.

Proposition 3.4.18.

There are no $n \times n$ real matrices A and B with $AB - BA = I_n$.

3.5 Inverse matrices

Lecture 16

Definition 3.5.1 (Invertible matrix, inverse).

Let A be an $n \times n$ matrix over $\mathbb{F}$. We say that A is *invertible* if there is an $n \times n$ matrix B such that

$$AB = I_n \quad \text{and} \quad BA = I_n.$$

Such a matrix B is called an *inverse* of A .

Proposition 3.5.4 (Uniqueness of the inverse).

An $n \times n$ matrix has at most one inverse: if B and C are both inverses of A , then $B = C$.

Theorem 3.5.5 (Solving with the inverse).

Let A be an invertible $n \times n$ matrix. Then for every $\mathbf{b} \in \mathbb{F}^n$ the system $A\mathbf{x} = \mathbf{b}$ has exactly one solution, namely

$$\mathbf{x} = A^{-1}\mathbf{b}.$$

Theorem 3.5.6 (Inverse of a 2×2 matrix).

Let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ over $\mathbb{F}$.

- (a) If $ad - bc \neq 0$, then A is invertible and

$$A^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

(b) If $ad - bc = 0$, then A is not invertible.

Theorem 3.5.8 (Rules for inverses).

Let A and B be invertible $n \times n$ matrices.

- (a) A^{-1} is invertible, and $(A^{-1})^{-1} = A$.
- (b) AB is invertible, and $(AB)^{-1} = B^{-1}A^{-1}$.
- (c) A^T is invertible, and $(A^T)^{-1} = (A^{-1})^T$.

Proposition 3.5.9 (One product is enough).

Let A and X be $n \times n$ matrices with $AX = I$. Then $XA = I$ as well, so A is invertible and $X = A^{-1}$.

Theorem 3.5.10 (Gauss–Jordan inversion).

Let A be an $n \times n$ matrix over $\mathbb{F}$.

- (a) If $[A \mid I_n]$ row reduces to $[I_n \mid X]$, then A is invertible and $X = A^{-1}$.
- (b) If the reduced row echelon form of A is not I_n , then A is not invertible.
- (c) Hence A is invertible if and only if its reduced row echelon form is I_n , and in that case $[A \mid I_n]$ row reduces to $[I_n \mid A^{-1}]$.

3.6 The invertible matrix theorem

Lecture 17

Definition 3.6.1 (Elementary matrix).

An $m \times m$ matrix E is an *elementary matrix* if it is obtained from the identity I_m by a single elementary row operation (Definition 2.3.1). We call E the elementary matrix of that operation.

Theorem 3.6.3 (Row operations are multiplications).

Let E be the $m \times m$ elementary matrix of a row operation. Then for every $m \times n$ matrix A over $\mathbb{F}$, the product EA is the matrix obtained by performing that row operation on A .

Proposition 3.6.5 (Sequences of row operations).

Suppose row operations with elementary matrices $E_1, E_2, \dots, E_k$, used in this order, take the $m \times n$ matrix A to U . Then

$$U = E_k \cdots E_2 E_1 A.$$

In particular, if A and U are row equivalent, then $U = PA$ for a product P of elementary matrices.

Theorem 3.6.6 (Elementary matrices are invertible).

Every elementary matrix E is invertible, and E^{-1} is the elementary matrix of the reverse row operation.

Theorem 3.6.8 (Products of elementary matrices).

Let A be an $n \times n$ matrix over $\mathbb{F}$.

- (a) If $E_k \cdots E_1 A = I_n$ for elementary matrices $E_1, \dots, E_k$, then A is invertible and $A^{-1} = E_k \cdots E_1$.
- (b) A is invertible if and only if A is a product of elementary matrices.

Theorem 3.6.10 (The invertible matrix theorem).

Let A be an $n \times n$ matrix over $\mathbb{F}$. The following statements are equivalent; that is, either they are all true, or they are all false.

- (1) A is invertible.
- (2) There is an $n \times n$ matrix C with $CA = I_n$.
- (3) The equation $A\mathbf{x} = \mathbf{0}$ has only the trivial solution $\mathbf{x} = \mathbf{0}$.

- (4) A has n pivot positions.
- (5) The reduced row echelon form of A is I_n .
- (6) The columns of A are linearly independent.
- (7) For every $\mathbf{b} \in \mathbb{F}^n$, the equation $A\mathbf{x} = \mathbf{b}$ has at least one solution.
- (8) The columns of A span $\mathbb{F}^n$.
- (9) For every $\mathbf{b} \in \mathbb{F}^n$, the equation $A\mathbf{x} = \mathbf{b}$ has exactly one solution.
- (10) There is an $n \times n$ matrix D with $AD = I_n$.
- (11) A^T is invertible.
- (12) A is a product of elementary matrices.

Proposition 3.6.14.

Let A and B be $n \times n$ matrices. If AB is invertible, then A and B are both invertible.

3.7 LU factorisation

Lecture 18

Definition 3.7.3 (Unit lower triangular, LU factorisation).

A square matrix is *unit lower triangular* if it is lower triangular (Definition 3.4.10) and every entry on its main diagonal is 1. An *LU factorisation* of an $m \times n$ matrix A over $\mathbb{F}$ is a product

$$A = LU,$$

where L is an $m \times m$ unit lower triangular matrix and U is an $m \times n$ matrix in row echelon form.

Theorem 3.7.4 (Elimination without swaps is a factorisation).

Let A be an $m \times n$ matrix over $\mathbb{F}$. Suppose that forward elimination brings A to a row echelon form U using only replacements, with no row swaps: in stage p ($p = 1, \dots, m-1$), each row $i > p$ is replaced by

$$R_i \rightarrow R_i - l_{ip}R_p \quad (l_{ip} \in \mathbb{F}),$$

with $l_{ip} = 0$ wherever no step is needed. Let L be the $m \times m$ unit lower triangular matrix whose entry in row i , column p is l_{ip} for $i > p$. Then

$$A = LU.$$

Theorem 3.7.5 (Solving with an LU factorisation).

Let $A = LU$, where A is $m \times n$, L is $m \times m$ unit lower triangular and U is $m \times n$. Let $\mathbf{b} \in \mathbb{F}^m$.

- (a) The system $L\mathbf{y} = \mathbf{b}$ has exactly one solution $\mathbf{y}$, found from the top down by *forward substitution*:

$$y_1 = b_1, \quad y_i = b_i - l_{i1}y_1 - \dots - l_{i,i-1}y_{i-1} \quad (i = 2, \dots, m).$$

- (b) For this $\mathbf{y}$, a vector $\mathbf{x} \in \mathbb{F}^n$ satisfies $A\mathbf{x} = \mathbf{b}$ if and only if $U\mathbf{x} = \mathbf{y}$.

Proposition 3.7.7 (Operation counts).

Let A be an $n \times n$ matrix whose forward elimination needs no row swaps, as in Theorem 3.7.4.

- (a) Computing U and the multipliers takes at most $\frac{n^3 - n}{3}$ multiplications and divisions.
- (b) If the diagonal entries of U are all nonzero, then for each $\mathbf{b} \in \mathbb{F}^n$, forward substitution takes $\frac{n(n-1)}{2}$ multiplications, and back substitution takes $\frac{n(n-1)}{2}$ multiplications and n divisions: n^2 operations in all.

Definition 3.7.9 (Permutation matrix).

An $m \times m$ matrix P is a *permutation matrix* if its rows are the rows of I_m in some order, each used once: $\text{row}_i(P) = \mathbf{e}_{\sigma(i)}^T$ for $i = 1, \dots, m$, where $\sigma(1), \dots, \sigma(m)$ are the numbers $1, \dots, m$ in some order.

Proposition 3.7.10 (Properties of permutation matrices).

Let P be an $m \times m$ permutation matrix with $\text{row}_i(P) = \mathbf{e}_{\sigma(i)}^T$.

- (a) For every $m \times n$ matrix A , $\text{row}_i(PA) = \text{row}_{\sigma(i)}(A)$: multiplying by P on the left puts the rows of A in the same order as the rows of I_m in P .
- (b) The product of two $m \times m$ permutation matrices is a permutation matrix.
- (c) P is invertible, and $P^{-1} = P^T$.

Theorem 3.7.11 ($PA = LU$).

Every $m \times n$ matrix A over $\mathbb{F}$ has a factorisation

$$PA = LU,$$

where P is an $m \times m$ permutation matrix, L is an $m \times m$ unit lower triangular matrix and U is an $m \times n$ matrix in row echelon form.

Proposition 3.7.14.

Every unit lower triangular $n \times n$ matrix L is invertible, and L^{-1} is unit lower triangular.

Theorem 3.7.15 (Uniqueness of the LU factorisation).

Let A be an invertible $n \times n$ matrix. If $A = L_1U_1 = L_2U_2$, where L_1, L_2 are unit lower triangular and U_1, U_2 are in row echelon form, then $L_1 = L_2$ and $U_1 = U_2$.

Chapter 4. Vector spaces

4.1 Subspaces of $\mathbb{R}^n$

Lecture 19

Definition 4.1.1 (Subspace).

A subset H of $\mathbb{F}^n$ is a *subspace* of $\mathbb{F}^n$ if

- (a) $\mathbf{0} \in H$;
- (b) H is closed under addition: if $\mathbf{u}, \mathbf{v} \in H$, then $\mathbf{u} + \mathbf{v} \in H$;
- (c) H is closed under scalar multiplication: if $\mathbf{u} \in H$ and $c \in \mathbb{F}$, then $c\mathbf{u} \in H$.

Proposition 4.1.4 (Closed under linear combinations).

Let H be a subspace of $\mathbb{F}^n$. If $\mathbf{v}_1, \dots, \mathbf{v}_k \in H$ and $c_1, \dots, c_k \in \mathbb{F}$, then $c_1\mathbf{v}_1 + \dots + c_k\mathbf{v}_k \in H$.

Theorem 4.1.6 (Spans are subspaces).

Let $\mathbf{v}_1, \dots, \mathbf{v}_k \in \mathbb{F}^n$ and $S = \text{Span}\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$.

- (a) S is a subspace of $\mathbb{F}^n$ that contains $\mathbf{v}_1, \dots, \mathbf{v}_k$.
- (b) S is the smallest such subspace: every subspace H of $\mathbb{F}^n$ that contains $\mathbf{v}_1, \dots, \mathbf{v}_k$ contains S .

Proposition 4.1.7 (The subspaces of $\mathbb{F}^2$).

Every subspace H of $\mathbb{F}^2$ is one of: the zero subspace $\{\mathbf{0}\}$; a line $\text{Span}\{\mathbf{v}\}$ with $\mathbf{v} \neq \mathbf{0}$; or $\mathbb{F}^2$ itself.

Definition 4.1.8 (Column space).

Let A be an $m \times n$ matrix over $\mathbb{F}$ with columns $\mathbf{a}_1, \dots, \mathbf{a}_n$. The *column space* of A is

$$\text{Col } A = \text{Span}\{\mathbf{a}_1, \dots, \mathbf{a}_n\} \subseteq \mathbb{F}^m.$$

Proposition 4.1.9.

Let A be an $m \times n$ matrix over $\mathbb{F}$.

- (a) $\text{Col } A = \{ Ax : x \in \mathbb{F}^n \}$, and $\text{Col } A$ is a subspace of $\mathbb{F}^m$.
- (b) For $b \in \mathbb{F}^m$, the system $Ax = b$ has a solution if and only if $b \in \text{Col } A$.

Definition 4.1.11 (Null space).

Let A be an $m \times n$ matrix over $\mathbb{F}$. The *null space* of A is the solution set of the homogeneous system $Ax = 0$:

$$\text{Nul } A = \{ x \in \mathbb{F}^n : Ax = 0 \} \subseteq \mathbb{F}^n.$$

Theorem 4.1.12.

For every $m \times n$ matrix A over $\mathbb{F}$, $\text{Nul } A$ is a subspace of $\mathbb{F}^n$.

Proposition 4.1.15 (A spanning list for $\text{Nul } A$).

Let A be an $m \times n$ matrix over $\mathbb{F}$. If $Ax = 0$ has no free variables, then $\text{Nul } A = \{0\}$. Otherwise, write its general solution in parametric vector form, $x = t_1 v_1 + \cdots + t_k v_k$, one vector for each free variable. Then $\text{Nul } A = \text{Span}\{v_1, \dots, v_k\}$.

Theorem 4.1.16.

Let A be an $m \times n$ matrix over $\mathbb{F}$.

- (a) $Ax = b$ has a solution for every $b \in \mathbb{F}^m$ if and only if $\text{Col } A = \mathbb{F}^m$, if and only if an echelon form of A has a pivot in every row.
- (b) $Ax = b$ has at most one solution for every $b \in \mathbb{F}^m$ if and only if $\text{Nul } A = \{0\}$, if and only if an echelon form of A has a pivot in every column.
- (c) If $m = n$, then A is invertible if and only if $\text{Nul } A = \{0\}$, if and only if $\text{Col } A = \mathbb{F}^n$.

4.2 Basis and dimension in $\mathbb{R}^n$

Lecture 20

Definition 4.2.1 (Basis).

Let H be a subspace of $\mathbb{F}^n$. A list $\mathcal{B} = (b_1, \dots, b_k)$ of vectors in H is a *basis* of H if

- (i) $b_1, \dots, b_k$ are linearly independent, and
- (ii) $b_1, \dots, b_k$ span H , that is, $\text{Span}\{b_1, \dots, b_k\} = H$.

Theorem 4.2.4 (Unique representation).

Let $\mathcal{B} = (b_1, \dots, b_k)$ be a basis of a subspace H of $\mathbb{F}^n$. Then every $x \in H$ can be written as

$$x = c_1 b_1 + c_2 b_2 + \cdots + c_k b_k$$

with scalars $c_1, \dots, c_k \in \mathbb{F}$ in exactly one way.

Theorem 4.2.6 (Shrinking a spanning list).

Let $H = \text{Span}\{v_1, \dots, v_k\}$ be a subspace of $\mathbb{F}^n$ with $H \neq \{0\}$. Then some of the vectors $v_1, \dots, v_k$ form a basis of H .

Theorem 4.2.8 (Growing an independent list).

Let H be a subspace of $\mathbb{F}^n$.

- (a) Every linearly independent list $u_1, \dots, u_m$ of vectors in H can be extended to a basis $(u_1, \dots, u_m, w_1, \dots, w_l)$ of H , with $l \geq 0$. Here $m \geq 0$: the empty list ($m = 0$) is independent, and the theorem extends it too.

(b) H has a basis, and every basis of H has at most n vectors.

Theorem 4.2.9.

Let H be a subspace of $\mathbb{F}^n$ spanned by k vectors $\mathbf{w}_1, \dots, \mathbf{w}_k$. Then every linearly independent list of vectors in H has at most k vectors. Equivalently, any $m > k$ vectors in H are linearly dependent.

Theorem 4.2.11 (Invariance of dimension).

Let H be a subspace of $\mathbb{F}^n$. Any two bases of H have the same number of vectors.

Definition 4.2.12 (Dimension).

Let H be a subspace of $\mathbb{F}^n$. The *dimension* of H , written $\dim H$, is the number of vectors in any basis of H . In particular $\dim\{\mathbf{0}\} = 0$.

Theorem 4.2.14.

Let A be an $m \times n$ matrix over $\mathbb{F}$ with columns $\mathbf{a}_1, \dots, \mathbf{a}_n$, and let R be its reduced row echelon form, with r pivots.

- (a) The pivot columns of A form a basis of $\text{Col } A$. Hence $\dim \text{Col } A = r$.
- (b) Write the general solution of $A\mathbf{x} = \mathbf{0}$ in parametric vector form, $\mathbf{x} = t_1\mathbf{s}_1 + \dots + t_f\mathbf{s}_f$, with one parameter for each free variable (Definition 2.5.2). Then $(\mathbf{s}_1, \dots, \mathbf{s}_f)$ is a basis of $\text{Nul } A$. Hence $\dim \text{Nul } A = f$, the number of free variables.
- (c) $\dim \text{Col } A + \dim \text{Nul } A = n$.

Theorem 4.2.16.

Let H be a subspace of $\mathbb{F}^n$ with $\dim H = k$.

- (a) Any k linearly independent vectors in H form a basis of H .
- (b) If $k \geq 1$, any k vectors that span H form a basis of H .
- (c) If U is a subspace of $\mathbb{F}^n$ with $U \subseteq H$, then $\dim U \leq \dim H$, and $\dim U = \dim H$ only if $U = H$.