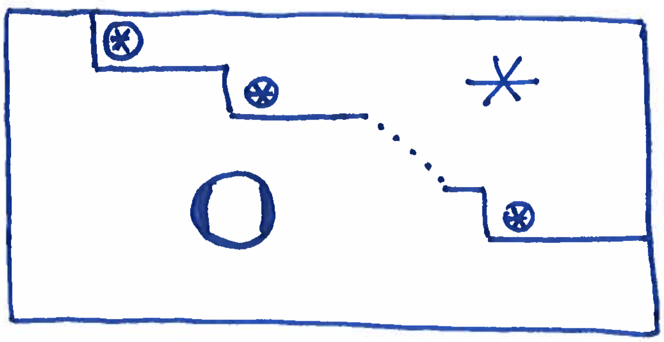
Last time we discussed how to solve systems with square nonsingular coefficient matrices using Gaussian elimination. These systems always have a unique solution.

Let us now turn to the general case: $Ax = b$, $A \in \mathbb{M}_{m \times n}$, $b \in \mathbb{R}^m$

The <u>key fact</u> is that any matrix A can be reduced to <u>row echelon form</u> by elementary row operations.

<u>Def</u>: An $m \times n$ matrix is said to be in row echelon form if it has the following "staircase" structure:

$U =$ 

$\circledast \neq 0$ are called pivots

$*$ is anything

$\}\ m-r$, $r = \operatorname{rank} A$

<u>Remark</u>: if A is square nonsingular, then

$U =$ 

$\bullet \neq 0$

$*$ is anything

this is a special case of row echelon form

<u>Remark</u> If $\circledast = 1$ and all entries above $\circledast$ are 0 $\Rightarrow$ reduced echelon form. It is unique.

The row echelon form U of matrix A is not unique, but the number of pivots is unique. This motivates the following definition.

<u>Def</u> The rank of a matrix A is the number of pivots in A's echelon form

Rank is one of the most important numerical quantities associated with a matrix.

Since there is at most one pivot per row ~~and~~ one pivot per column,

$$\boxed{0 \leq \operatorname{rank} A \leq \min\{m, n\}}$$

$\left[A \in \mathbb{M}_{n \times n} \text{ is nonsingular} \Leftrightarrow \operatorname{rank} A = n \right]$

In matrix language, the above key fact can be formulated as follows: there exists an $m \times m$ permutation matrix P and an $m \times m$ special low triangular L such that: $PA = LU$

To solve $Ax = b$, we therefore, proceed as follows:

1. $Ax = b \iff PAx = Pb \overset{\text{def}}{=} \tilde{b} \iff L\underbrace{Ux}_{y} = \tilde{b}$
2. Solve $Ly = \tilde{b}$ for y using forward substitution.
3. Now we need to solve $Ux = y$ for x.
   a) If the last $(m-r)$ entries of y are nonzeros $\Rightarrow$ no solution.
   b) Otherwise, we simply modify the back substitution procedure:
   - solve for the <u>basic variable</u> (associated with its pivot)
   - substitute the result into the previous equation

   $\Bigg\} \Rightarrow$ r basic variables will be expressed in terms of $(n-r)$ <u>free variables</u> (correspond to columns without pivots)

<u>Remark</u>

$[A|b]$ augmented matrix

$\Big\downarrow$ elementary row operations

$[U|y]$ row echelon form of $[A|b]$

Let $x = [\underbrace{x_1, \dots x_r}_{\text{basic}}, \underbrace{x_{r+1} \dots x_n}_{\text{free}}]^T$

(Free and basic variables are "mixed" in x, but we use this notation for symplicity)

Then $\begin{cases} x_1 = x_1(x_{r+1}, \dots, x_n) \\ \vdots \\ x_r = x_r(x_{r+1}, \dots, x_n) \end{cases}$

This is general solution.
Free variables can take any values, basic variables are are determined by free ones.
So, the general solution depends on $(n-r)$ parameters.

The following theorem summarizes the above discussion.

<u>Th</u> A system $Ax=b$ of m equations with n unknowns has either

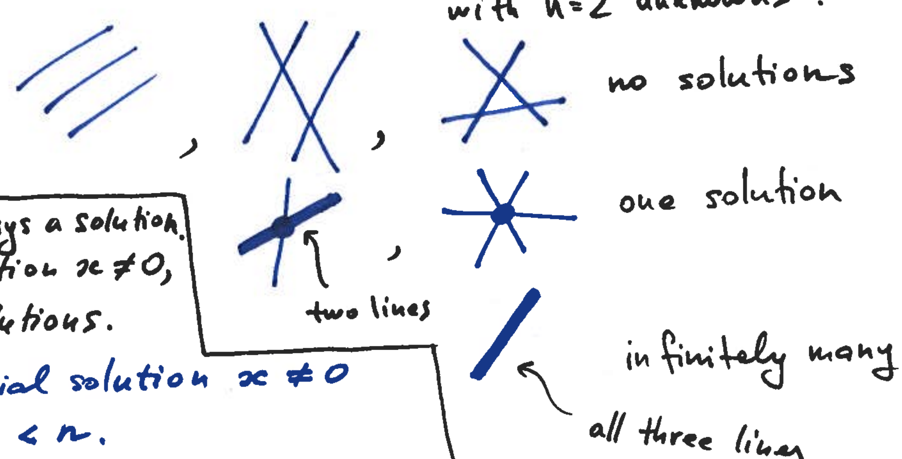
inconsistent 1) no solution $\iff$ rank $[A|b] >$ rank A

This characterization in terms of rank also follows from the fundamental Th. of Lin. Alg.

consistent $\begin{cases} \text{2) exactly one solution} \iff \text{rank}[A|b] = \text{rank } A \text{ and rank } A = n \\ \text{3) infinitely many solutions} \iff \text{rank}[A|b] = \text{rank } A \text{ and rank } A < n \end{cases}$

<u>Remark 1</u> Since rank $A \le m$, the system can have only one solution if $n \le m$
(only "tall" systems can have unique sol; "flat" systems cannot)

<u>Remark 2</u> Number of solutions is ~~either~~ $0, 1,$ or ∞. Here is a geometric interpret. for systems of $m=3$ equations with $n=2$ unknowns:



<u>Homogeneous Systems</u>

The system $Ax=0$ is called <u>homogeneous</u>

Homog. systems are always consistent: $x=0$ is always a solution.
If $Ax=0$ has a nontrivial solution $x \ne 0$, then it has infinitely many solutions.

<u>Corollary</u> 1) $Ax=0$ has a nontrivial solution $x \ne 0$ if and only if rank $A < n$.

2) If $m < n \implies Ax=0$ always has a nontr. sol.

3) If $m = n \implies Ax=0$ has nontrivial sol. if and only if A is singular

Determinants

for the goals of this class

This topic is a bit controversial. On one hand, determinants are of great theoretical importance in linear algebra, they appear (and very useful) in many other branches of mathematics, and have fascinating properties. On the other hand, like matrix inverses, they are almost completely irrelevant when it comes to large scale applications and practical computations (too expensive to compute). So, we will discuss them very briefly, mostly stating their properties which we will need ~~[illegible]~~ for later developments.

The determinant of a square matrix A is a scalar, denoted $\det A$.

sometimes $|A|$

There are many equivalent ways to define determinants. Let us start with "axiomatic" definition, which states several properties of the determinants, which uniquely define them, but do not tell us how to compute $\det A$.

Def A determinant is a function $\det : \mathbb{M}_{n\times n} \to \mathbb{R}$ such that

1) $\det I_n = 1$

2) $$\det \begin{bmatrix} - a^1 - \\ \vdots \\ - a^{i-1} - \\ - \alpha b + \beta c - \\ - a^{i+1} - \\ \vdots \\ - a^n - \end{bmatrix} = \alpha \det \begin{bmatrix} - a^1 - \\ \vdots \\ - a^{i-1} - \\ - b - \\ - a^{i+1} - \\ \vdots \\ - a^n - \end{bmatrix} + \beta \det \begin{bmatrix} - a^1 - \\ \vdots \\ - a^{i-1} - \\ - c - \\ - a^{i+1} - \\ \vdots \\ - a^n - \end{bmatrix}$$

linear in each row

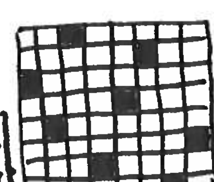
3) $$\det \begin{bmatrix} - a^1 - \\ \vdots \\ - a^i - \\ \vdots \\ - a^j - \\ \vdots \\ - a^n - \end{bmatrix} = - \det \begin{bmatrix} - a^1 - \\ \vdots \\ - a^j - \\ \vdots \\ - a^i - \\ \vdots \\ - a^n - \end{bmatrix}$$

"alternating" function

It turns out that there exists unique function that satisfies these properties:

$$\det A = \sum_{\sigma \in S_n} (\operatorname{sign} \sigma)\, \underbrace{a_{\sigma(1),1} \cdots a_{\sigma(n),n}}_{\text{example}}$$

← Leibniz formula



Here $\sigma = \{3,5,1,\cancel{7},4,2,8,6\}$

Here σ is a permutation of set $\{1,2,\dots,n\}$, which reoders this set of integers. S_n is the set of all such permutations. For example, if $n=3$, then

$$S_3 = \begin{Bmatrix} \sigma_1 = \{1,2,3\} & \sigma_4 = \{2,3,1\} \\ \sigma_2 = \{2,1,3\} & \sigma_5 = \{3,1,2\} \\ \sigma_3 = \{1,3,2\} & \sigma_6 = \{3,2,1\} \end{Bmatrix}$$

$\text{sign}(\sigma)$ is the sign (also called signature) of permutation σ.

$$\text{sign}(\sigma) = \begin{cases} +1 & \text{if the reodering given by } \sigma \text{ can be achieved by successively interchanging two entries an } \underline{\text{even}} \text{ number of times} \\ -1 & \text{if } \underline{\text{odd}} \text{ number of interchages is required} \end{cases}$$

For example,

σ	sign
$\{1,2,3\}$	1
$\{1,3,2\}$	-1
$\{2,1,3\}$	-1
$\{2,3,1\}$	1
$\{3,1,2\}$	1
$\{3,2,1\}$	-1

So, the $\det A$ is a sum of all possible terms of the form $a_{\sigma(1),1} \cdot \ldots \cdot a_{\sigma(n),n}$. Each term is a product of matrix entries, such that each row and each column contributes only one entry to each term.

If $n=2$, then Leibniz formula reduces to

$$\det \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \overbrace{\text{sign}(\{1,2\})}^{+} a_{11} a_{22} + \overbrace{\text{sign}(\{2,1\})}^{-} a_{21} a_{12}$$

$$= \underline{a_{11}a_{22} - a_{12}a_{21}} \quad \text{(well known expression)}$$

Check that for $n=3$, it also gives a well known result.

Leibniz formula, however, is not used in practice, since it contains $n!$ terms, which, as soon as n is of moderate size, is a huge number. For example,

if $n=10$, then $n! \sim 3 \cdot 10^6$

if $n=100$ $n! \sim 9 \cdot 10^{157}$.

On the other hand, determinants have very nice properties, which are useful for theoretical developments.

Properties

1. $\det(AB) = \det A \cdot \det B$
2. $\det A^{-1} = \dfrac{1}{\det A}$ $\quad \left(\begin{array}{c} A \text{ is nonsingular} \\ \Leftrightarrow \det A \neq 0 \end{array} \right)$
3. $\det A^T = \det A$
4. $\det(\alpha A) = \alpha^n \det A$
5. If A is a triangular matrix

$$\begin{bmatrix} * & & 0 \\ & \ddots & \\ * & & * \end{bmatrix} \quad \text{or} \quad \begin{bmatrix} * & & * \\ & \ddots & \\ 0 & & * \end{bmatrix}$$

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

6. $|\det A| = \text{Vol}(a_1, \ldots, a_n)$

Geom. interpretation $\quad = \text{Vol}(a^1 \ldots a^n)$

Remark:

Properties 1) and 5) suggest how to compute $\det A$ of a nonsingular matrix:

If A is nonsingular $\Rightarrow$

$PA = LU \Rightarrow \det P \cdot \det A = \det L \cdot \det U$

- $\det P = (-1)^k$, $k = \#$ row interchanges
- $\det L = 1$
- $\det U = \prod_{i=1}^{n} u_{ii}$

$$\Rightarrow \boxed{\det A = (-1)^k \prod_{i=1}^{n} u_{ii}}$$

Gaussian elimination for computing dets.

This is, generally, the most efficient way to compute dets

The notion of a vector space unifies many seemingly unrelated sets with operations, such as ordinary vectors, sets of functions, matrices, operators, etc. As a result, vector spaces and associated constructions provide the common language of linear algebra.

A vector space is the abstract reformulation of the properties of $\mathbb{R}^n$.

<u>Def</u> A <u>vector space</u> is a set V equipped with two operations:

1. Addition: $(u,v) \mapsto w = u+v$, $u,v,w \in V$
2. Multiplication by a scalar: $(v,\alpha) \mapsto w = \alpha \cdot v$, $v,w \in V$, $\alpha \in \mathbb{R}$ could be $\mathbb{C}$ or other field $\mathbb{F}$

that satisfy the following properties:

(a) $u+v = v+u$

(b) $u+(v+w) = (u+v)+w$

(c) $\exists\, 0 \in V : \forall v \in V \quad 0+v = v+0 = v$

(d) $\forall v \in V \ \exists (-v) : v + (-v) = 0$

(e) $(\alpha+\beta)v = (\alpha v) + (\beta v)$
$\alpha(v+u) = (\alpha v) + (\alpha u)$

(f) $\alpha(\beta v) = (\alpha\beta)v$

(g) $1 \cdot v = v$

<u>Example 1</u> The prototypical example of a vector space is $\mathbb{R}^n = \left\{ v = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix},\ v_i \in \mathbb{R} \right\}$

$$v+u = \begin{pmatrix} v_1+u_1 \\ \vdots \\ v_n+u_n \end{pmatrix} \qquad \alpha \cdot v = \begin{pmatrix} \alpha v_1 \\ \vdots \\ \alpha v_n \end{pmatrix} \qquad \mathbb{0} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} \qquad -v = \begin{pmatrix} -v_1 \\ \vdots \\ -v_n \end{pmatrix}$$

<u>Example 2</u> Let $\mathbb{M}_{m\times n}$ denote the space of all real matrices of size $m \times n$.

$\Rightarrow$ $\mathbb{M}_{m\times n}$ is a vector space under the usual laws of matrix addition and scalar multiplication.

<u>Remark</u> $\mathbb{R}^n \equiv \mathbb{M}_{n\times 1}$

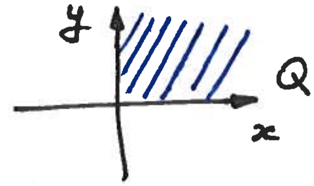
<u>Example 3</u> Let $\mathbb{P}^{(n)}$ be the set of all polynomials of deg $\le n$

$$\mathbb{P}^{(n)} = \{ p(x) = a_n x^n + \dots + a_1 x + a_0\ ,\ a_i \in \mathbb{R} \}$$

$\Rightarrow$ $\mathbb{P}^{(n)}$ is a vector space under the usual $p(x)+q(x)$ and $\alpha \cdot p(x)$.

<u>Remark</u> There is nothing special in polynomials or, more generally, functions on $\mathbb{R}$. Let S be any set, let $\mathcal{F}(S)$ be the set of all functions $f : S \to \mathbb{R}$. Define $(f+g)(x) = f(x)+g(x)$ and $(\alpha f)(x) = \alpha \cdot f(x)$ $\Rightarrow$ $\mathcal{F}(S)$ is a vector space.

All these examples are more or less trivial. Here is a less trivial one.



Example 4 Let $Q = \{(x,y) : x, y > 0\} \subset \mathbb{R}^2$

Let us define $(x_1 y_1) \oplus (x_2 y_2) = (x_1 x_2, y_1 y_2)$

$\alpha \cdot (x,y) = (x^\alpha, y^\alpha)$ $\Rightarrow$ $(Q, \oplus, \cdot)$ is a vector space.

The zero element is $\mathbb{0} = (1,1)$, the inverse $\ominus(x,y) = (\frac{1}{x}, \frac{1}{y})$.

Def Let V be a vector space. A subset $W \subset V$ is called a subspace if W is itself a vector space (under the same $u+v$ and $\alpha \cdot v$)

Question: Given a subset $W \subset V$, to check whether it is a subspace, do we need to check all axioms (a)–(g)?

Answer: No, we only need to check that W is closed under "+" and "·" (that is W respects the vector space operations)

Theorem 1: A nonempty subset $W \subset V$ of a vector space is a subspace

$\Longleftrightarrow \begin{cases} 1.\ \forall u, v \in W,\ u+v \in W \\ 2.\ \forall v \in W, \forall \alpha \in \mathbb{R},\ \alpha v \in W \end{cases}$

$\Longleftrightarrow$ 3. $\forall \alpha, \beta \in \mathbb{R}, \forall u, v \in W,\ \alpha v + \beta u \in W$

V

Not sure what I want to say with this picture.

Example 1 Every vector space V has two trivial subspaces: $W = V$ and $W = \{\mathbb{0}\}$.

Example 2 Let $V = \mathbb{R}^3 \Rightarrow W = \{(x,y,0)\} \subset V$ is a subspace.

Remark In $\mathbb{R}^3$, there 4 different types of subspaces:

1. $W = \{\mathbb{0}\}$
2. a line passing through $\mathbb{0}$
3. a plane passing through $\mathbb{0}$
4. $W = \mathbb{R}^3$.

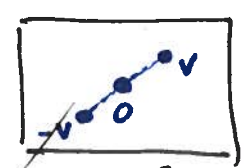
Example 3 Let $V = \mathbb{R}^n$ and W is the set of solutions

$$W = \{x \in \mathbb{R}^n : a_1 x_1 + \ldots + a_n x_n = 0\}$$

Example 4 Let $V = \mathcal{F}(I)$, where $I = [a,b]$. Let $W \subset \mathcal{F}(I)$ be the set of all solutions to the linear differential equation $u'' + au' + bu = 0$.

Question: In Theorem 1, are both (1) and (2) important?

Is there an example, where $W \subset V$, $W \neq \emptyset$, and $\forall u, v \in W$, $u+v \in W$, but W is not a subspace? Yes $\longrightarrow$

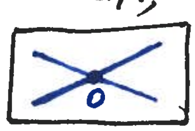


$V = \mathbb{R}^2$

$W = \{-v, 0, v\}$

What about $W \subset V$, $W \neq \emptyset$, $\alpha v \in W$ $\forall \alpha, v$, but W is not a subspace? Yes

$W = \{\text{union of two lines}\}$



<u>Span</u> is a universal method for constructing subspaces of vector spaces.

Def Let V be a vector space and $v_1, \dots, v_k \in V$.

A sum $\sum_{i=1}^{k} \alpha_i v_i = \alpha_1 v_1 + \dots + \alpha_k v_k$, where $\alpha_i \in \mathbb{R}$, is called a <u>linear combination</u> of the elements $v_1 \dots v_k$. The subset

$W = \left\{ \sum_{i=1}^{k} \alpha_i v_i ,\ \alpha_i \in \mathbb{R} \right\}$ consisting of all linear combinations is called the <u>span</u> of $v_1 \dots v_k$ and denoted $\boxed{W = \operatorname{span}(v_1, \dots, v_k)}$

Theorem 2 The span $W = \operatorname{span}(v_1 \dots v_k)$ is a subspace of V.

Proof: We need to show that W is closed under the vector space operations.

1. Let $u = \sum \alpha_i v_i \in W \Rightarrow \beta \cdot u = \sum \underbrace{(\alpha_i \beta)}_{\hat{\alpha}_i} v_i = \sum \hat{\alpha}_i v_i \in W$

2. Let $u_1 = \sum \alpha_i^{(1)} v_i$, $u_2 = \sum \alpha_i^{(2)} v_i$ $\Rightarrow u_1 + u_2 = \sum \underbrace{(\alpha_i^{(1)} + \alpha_i^{(2)})}_{\hat{\alpha}_i} v_i = \sum \hat{\alpha}_i v_i \in W$ $\square$

Example 1 Let $V = \mathcal{F}(\mathbb{R})$ be the vector space of all functions $f: \mathbb{R} \to \mathbb{R}$.

Let $f_1 = 1,\ f_2 = x,\ \dots,\ f_{n+1} = x^n \Rightarrow \operatorname{span}(1, x, \dots, x^n) = \mathbb{P}^{(n)}$

Example 2 The following example often appears in many applications.

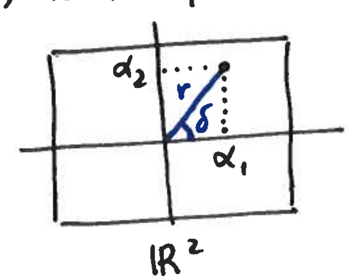
Let us again $V = \mathcal{F}(\mathbb{R})$ and let $f_1(x) = \cos wx$, $f_2(x) = \sin wx$ where $w \in \mathbb{R}$ (fixed)

The span is then

$$\operatorname{span}(f_1, f_2) = \{\alpha_1 \cos wx + \alpha_2 \sin wx\} = \{r \cos(wx - \delta)\} \qquad r \geq 0,\ \delta \in [0, 2\pi)$$

r: amplitude, δ: phase shift

Since $r\cos(wx - \delta) = \underbrace{r\cos\delta}_{\alpha_1} \cos wx + \underbrace{r \sin\delta}_{\alpha_2} \sin wx$, we can think of (r, δ) as the polar coordinates of (α_1, α_2)



Test question: Suppose that z is a linear combination of u and v. Is it true that v is always a linear combination of u and z?

Answer: No. For example, $z = u \neq v$