THE UNIVERSITY OF CHICAGO

MSCA 37016
Advanced Linear Algebra for Machine Learning

Lecture 2

Danny Ng

3 Views of System of Linear Equations

- System of n linear equations in d unknown variables $x_1, \ldots, x_d$

$$\begin{aligned} x_1 + \; x_2 \qquad &= 5 \\ x_2 + x_3 &= 3 \\ 2x_1 + 2x_2 \qquad &= 10 \end{aligned}$$

- Matrix $\times$ vector $=$ another vector

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

with

$$A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 2 & 2 & 0 \end{pmatrix} \in \mathbb{R}^{n \times d}$$

$$\vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \mathbb{R}^d$$

$$\vec{b} = \begin{pmatrix} 5 \\ 3 \\ 10 \end{pmatrix} \in \mathbb{R}^n$$

- Linear combination of d vectors in $\mathbb{R}^n$ with coefficients $x_1, \ldots, x_d$

$$x_1 \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix} + x_2 \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}$$
$$+ x_3 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 5 \\ 3 \\ 10 \end{pmatrix}$$

3 Views of System of Linear Equations

$$\begin{aligned} x_1 + \; x_2 \qquad &= 5 \\ x_2 + x_3 &= 3 \\ 2x_1 + 2x_2 \qquad &= 10 \end{aligned}$$

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

with

$$A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 2 & 2 & 0 \end{pmatrix} \in \mathbb{R}^{n\times d}$$

$$\vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \quad \vec{b} = \begin{pmatrix} 5 \\ 3 \\ 10 \end{pmatrix}$$

$$x_1 \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix} + x_2 \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} + x_3 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 5 \\ 3 \\ 10 \end{pmatrix}$$

- Is the system **consistent** i.e. has at least one solution?
- Is vector $\vec{b}$ in the **range** / **column space** of the matrix A?
- Is the RHS vector **redundant** i.e. can be reproduced by the vectors on the LHS?

3 Views of System of Linear Equations

$$\begin{aligned} x_1 + \ \ x_2 \qquad\quad &= 0 \\ x_2 + x_3 &= 0 \\ 2x_1 + 2x_2 \qquad\quad &= 0 \end{aligned}$$

$$A\vec{x} = \vec{0}$$

with

$$A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 2 & 2 & 0 \end{pmatrix} \in \mathbb{R}^{n \times d}$$

$$\vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \quad \vec{0} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

$$x_1 \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix} + x_2 \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} + x_3 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

- Does the **homogeneous** system have more solutions besides the trivial $x_1 = \cdots = x_d = 0$?
- Is the matrix A invertible? More generally, does it have **full column rank**?
- Is the set of vectors on the LHS linearly independent or dependent?

Gaussian Elimination Overview

- **Gaussian elimination** is an efficient technique to solve a system of linear equations for all possible solutions (if exist)
- **Step 1:** Rewrite the system in a simpler form (without changing the solutions) using **elementary row operations**:
    - Swap any two rows
    - Multiply a row by a non-zero scalar
    - Add a multiple of one row to another row
- **Step 2:** It is simple enough when it is in **row echelon form** (looks somewhat like a triangular form):
    1. Any row of zeros is at the bottom
    2. First non-zero leading **pivot** entry of a non-zero row is in a column to the left of any leading entries below it
- **Step 3:** Use **backward substitution** to solve the simpler system and write the solution set in **parametric form**

Case 1: One Solution

$$\begin{aligned} 1x + 3y &= 11 \\ 3x + 2y &= 12 \end{aligned}$$

need 0

- **Step 1:** (2^{nd} row) $+(-3) \times$ (1^{st} row)

$$\begin{aligned} 1x + 3y &= 11 \\ 3x + 2y &= 12 \end{aligned} \quad \longmapsto \quad \begin{aligned} 1x + 3y &= 11 \\ -7y &= -21 \end{aligned}$$

pivots

- **Step 2:** It is now in row echelon form

- **Step 3:** 2^{nd} equation yields $y = 3$. Substitute back into 1^{st} equation yields $x = 2$. The only solution is $(x, y) = (2, 3)$.

Case 2: No Solution

$$\begin{aligned} 1x + 3y &= 11 \\ 2x + 6y &= -5 \end{aligned}$$

non-overlapping parallel lines

- **Step 1:** (2^{nd} row) $-2 \times$ (1^{st} row)

$$\begin{aligned} 1x + 3y &= 11 \\ 2x + 6y &= -5 \end{aligned} \quad \longmapsto \quad \begin{aligned} 1x + 3y &= 11 \\ 0x + 0y &= -27 \end{aligned}$$

pivot

- **Step 2:** It is now in row echelon form

- **Step 3:** But 2^{nd} equation $0 = -27$ is absurd. It can never be true no matter what x and y are. The system has no solution i.e. **inconsistent.**

Case 3: Infinite Solutions

eliminate $\rightarrow$ $$\left.\begin{aligned} 1x + 3y &= 11 \\ 5x + 15y &= 55 \end{aligned}\right\}\text{ same line}$$

- **Step 1:** (2^{nd} row) $-5 \times$ (1^{st} row)

$$\begin{aligned} 1x + 3y &= 11 \\ 5x + 15y &= 55 \end{aligned} \quad \longmapsto \quad \begin{aligned} 1x + 3y &= 11 \\ 0x + 0y &= 0 \end{aligned}$$

$0 = 0$ is always true and redundant

- **Step 2:** It is now in row echelon form

- **Step 3:** The solution set is

$$\left\{\begin{pmatrix} x \\ y \end{pmatrix} : x = 11 - 3y\right\}$$

consists of infinitely many solutions

Solution Set in Parametric Form

$$1x + 3y = 11$$

leading / pivot variable x free variable y

- Any non-leading variable is a **free variable**

  e.g. You are free to choose $y = 2$ for a solution as long as you set $x = 11 - 3(2) = 5$ accordingly

- Rewrite the solution set in **parametric form**

$$\left\{\begin{pmatrix} 11 - 3y \\ y \end{pmatrix} : y \in \mathbb{R}\right\} = \left\{\begin{pmatrix} 11 \\ 0 \end{pmatrix} + y\begin{pmatrix} -3 \\ 1 \end{pmatrix} : y \in \mathbb{R}\right\}$$

scaling a vector

Gaussian Elimination Example

$$
\begin{aligned}
&& 4x_3 - 2x_4 &= 0 \\
x_1 - 2x_2 + && x_3 + 2x_4 &= 4 \\
-2x_1 + 4x_2 && - 5x_4 &= -8 \\
3x_1 - 6x_2 + && 5x_3 + 5x_4 &= 12
\end{aligned}
$$

- Use augmented shorthand notation

need non-zero pivot

$$
\left(\begin{array}{cccc|c} 0 & 0 & 4 & -2 & 0 \\ 1 & -2 & 1 & 2 & 4 \\ -2 & 4 & 0 & -5 & -8 \\ 3 & -6 & 5 & 5 & 12 \end{array}\right)
$$

swap 1st and 2nd rows

$$
\mapsto \left(\begin{array}{cccc|c} 1 & -2 & 1 & 2 & 4 \\ 0 & 0 & 4 & -2 & 0 \\ -2 & 4 & 0 & -5 & -8 \\ 3 & -6 & 5 & 5 & 12 \end{array}\right)
$$

need zeros below pivot

(3rd row)+2 ×(1st row)

(4th row)−3 ×(1st row)

another pivot

$$
\mapsto \left(\begin{array}{cccc|c} 1 & -2 & 1 & 2 & 4 \\ 0 & 0 & 4 & -2 & 0 \\ 0 & 0 & 2 & -1 & 0 \\ 0 & 0 & 2 & -1 & 0 \end{array}\right)
$$

(3rd row)−1/2 ×(2nd row)

(4th row)−1/2 ×(2nd row)

$$
\mapsto \underbrace{\left(\begin{array}{cccc|c} 1 & -2 & 1 & 2 & 4 \\ 0 & 0 & 4 & -2 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right)}_{\text{row echelon form}}
$$

Gaussian Elimination Example

- Row echelon form (is not unique e.g. depends on swap)

$$\left(\begin{array}{cccc|c} 1 & -2 & 1 & 2 & 4 \\ 0 & 0 & 4 & -2 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

- Redundant equations are eliminated

$$\begin{aligned} x_1 - 2x_2 + x_3 + 2x_4 &= 4 \\ 4x_3 - 2x_4 &= 0 \\ 0 &= 0 \\ 0 &= 0 \end{aligned}$$

2 pivot variables

- Free variables: x_2, x_4 (not always at the end)
- **Under-determined** system: $n < d$

- Write leading variables in terms of free variables by backward substitution

$$x_3 = 1/2\, x_4$$

$$x_1 = 4 + 2x_2 - x_3 - 2x_4 = 4 + 2x_2 - 5/2\, x_4$$

- Solution set in parametric form (not unique)

$$\left\{ \begin{pmatrix} 4 + 2x_2 - 5/2\, x_4 \\ x_2 \\ 1/2\, x_4 \\ x_4 \end{pmatrix} : x_2, x_4 \in \mathbb{R} \right\}$$

$$= \left\{ \begin{pmatrix} 4 \\ 0 \\ 0 \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} 2 \\ 1 \\ 0 \\ 0 \end{pmatrix} + x_4 \begin{pmatrix} -5/2 \\ 0 \\ 1/2 \\ 1 \end{pmatrix} : x_2, x_4 \in \mathbb{R} \right\}$$

Homogeneous vs Inhomogeneous Solution Set

- **Inhomogeneous** system: $\vec{b} \neq \vec{0}$

$$\begin{aligned} x_1 - x_2 + 2x_3 &= 5 \\ 2x_1 - 2x_2 + 4x_3 &= 10 \\ -x_1 + x_2 - 2x_3 &= -5 \end{aligned}$$

- Solution set in parametric form

$$\left\{ \underbrace{\begin{pmatrix} 5 \\ 0 \\ 0 \end{pmatrix}}_{\text{particular solution (when } x_2 = x_3 = 0)} + x_2 \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + x_3 \begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix} : x_2, x_3 \in \mathbb{R} \right\}$$

- Corresponding **homogeneous** system: $\vec{b} = \vec{0}$

$$\begin{aligned} x_1 - x_2 + 2x_3 &= 0 \\ 2x_1 - 2x_2 + 4x_3 &= 0 \\ -x_1 + x_2 - 2x_3 &= 0 \end{aligned}$$

- Solution set in parametric form

$$\left\{ \underbrace{x_2 \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + x_3 \begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix}}_{\text{all linear combinations but not shifted by a particular solution}} : x_2, x_3 \in \mathbb{R} \right\}$$

$$\boxed{\begin{pmatrix} \text{solution set of} \\ A\vec{x} = \vec{b} \end{pmatrix} = \begin{pmatrix} \text{any particular} \\ \text{solution } \vec{v} \end{pmatrix} + \begin{pmatrix} \text{solution set of} \\ A\vec{x} = \vec{0} \end{pmatrix}}$$

Maximal Linearly Independent Subset

- We know 4 vectors in $\mathbb{R}^3$ must be linearly dependent

$$\left\{\begin{pmatrix}1\\2\\-1\end{pmatrix},\begin{pmatrix}3\\6\\-3\end{pmatrix},\begin{pmatrix}1\\2\\2\end{pmatrix},\begin{pmatrix}2\\4\\-5\end{pmatrix}\right\}$$

- **Question:** How to identify one (of the possibly many) maximal linearly independent subset?

- Use row echelon form!

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

pivot columns

- Pivot columns form a maximal linearly independent subset

$$\left\{\begin{pmatrix}1\\2\\-1\end{pmatrix},\begin{pmatrix}1\\2\\2\end{pmatrix}\right\}$$

- Each free variable column must be expressible as a linear combination of this independent subset

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

$$\begin{pmatrix}2\\4\\-5\end{pmatrix} = 3\begin{pmatrix}1\\2\\-1\end{pmatrix} - 1\begin{pmatrix}1\\2\\2\end{pmatrix}$$

Gauss-Jordan Method: Invertible Case

- **Question:** How to compute the inverse A^{-1} of a matrix A?

$$A = \begin{pmatrix} 1 & 2 & -1 \\ 2 & 2 & 4 \\ 1 & 3 & -3 \end{pmatrix}$$

- **Step 1:** Construct the augmented matrix

$$(A|I_3) = \left(\begin{array}{ccc|ccc} 1 & 2 & -1 & 1 & 0 & 0 \\ 2 & 2 & 4 & 0 & 1 & 0 \\ 1 & 3 & -3 & 0 & 0 & 1 \end{array}\right)$$

- **Step 2:** Use Gaussian elimination to transform the first block into an identity matrix

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 9 & -3/2 & -5 \\ 0 & 1 & 0 & -5 & 1 & 3 \\ 0 & 0 & 1 & -2 & 1/2 & 1 \end{array}\right) = (I_3|A^{-1})$$

  - Use a top-down pass to transform the lower triangular entries to 0 i.e. row echelon form
  - Then use a bottom-up pass to transform the upper triangular entries to 0

Gauss-Jordan Method: Non-Invertible Case

- If a matrix A is not invertible

$$(A|I_3) = \left(\begin{array}{ccc|ccc} 1 & 2 & -1 & 1 & 0 & 0 \\ 2 & 2 & 4 & 0 & 1 & 0 \\ 1 & 0 & 5 & 0 & 0 & 1 \end{array}\right)$$

then the transformation will yield at least one row of zeros and get stuck

$$\left(\begin{array}{ccc|ccc} 1 & 2 & -1 & 1 & 0 & 0 \\ 0 & -2 & 6 & -2 & 1 & 0 \\ 0 & 0 & 0 & 1 & -1 & 1 \end{array}\right)$$

only 2 pivots

- Without full pivots, the 3 columns of A together is not linearly independent

Fact: A matrix $A \in \mathbb{R}^{d\times d}$ is invertible if and only if its d column vectors are linearly independent.

e.g. These matrices are obviously not invertible

$$A = \begin{pmatrix} 1 & 0 & 3 \\ 4 & 7 & 12 \\ 2 & 3 & 6 \end{pmatrix} \qquad B = \begin{pmatrix} 1 & 0 & 2 \\ 5 & 0 & -2 \\ -8 & 0 & 3 \end{pmatrix}$$

Span

- The **span** of a set of vectors $\vec{v}_1, \dots, \vec{v}_k \in \mathbb{R}^d$ is the collection of all possible linear combinations of these vectors

$$\operatorname{span}(\vec{v}_1, \dots, \vec{v}_k) = \{a_1\vec{v}_1 + \dots + a_k\vec{v}_k : a_1, \dots, a_k \in \mathbb{R}\}$$

e.g. Solution set of a previous homogeneous system is

$$\left\{\underbrace{x_2\begin{pmatrix}1\\1\\0\end{pmatrix} + x_3\begin{pmatrix}-2\\0\\1\end{pmatrix}}_{\text{all possible linear combinations}} : x_2, x_3 \in \mathbb{R}\right\} = \operatorname{span}\left(\begin{pmatrix}1\\1\\0\end{pmatrix}, \begin{pmatrix}-2\\0\\1\end{pmatrix}\right)$$

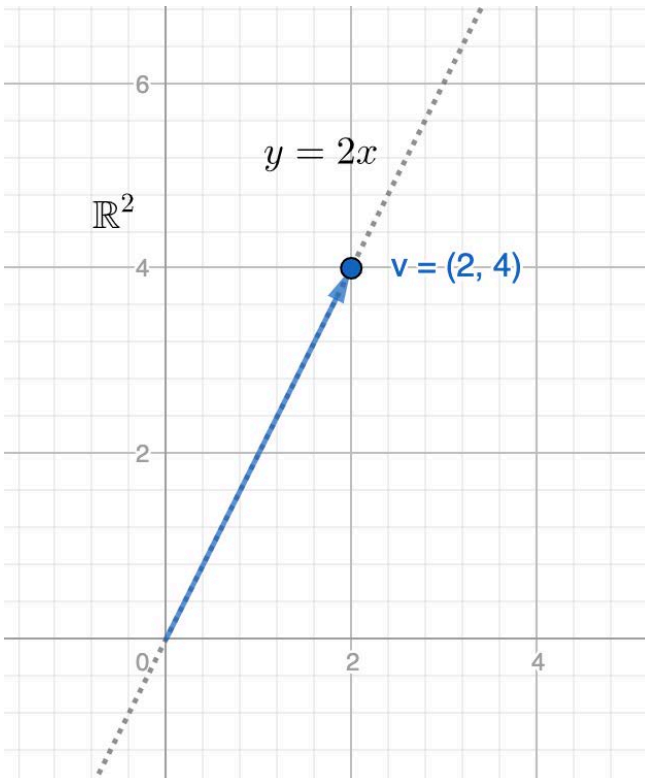
Geometry of Span

- **Line** through origin along the direction of $\vec{v}$

$$\operatorname{span}(\vec{v}) = \{\underbrace{a\vec{v}}_{\substack{\text{any scaling is a vector}\\ \text{on the same line}}} : a \in \mathbb{R}\}$$

- e.g.

$$\operatorname{span}\left(\begin{pmatrix}2\\4\end{pmatrix}\right) = \left\{a\begin{pmatrix}2\\4\end{pmatrix} : a \in \mathbb{R}\right\}$$
$$= \left\{\begin{pmatrix}x\\y\end{pmatrix} : y = 2x\right\}$$



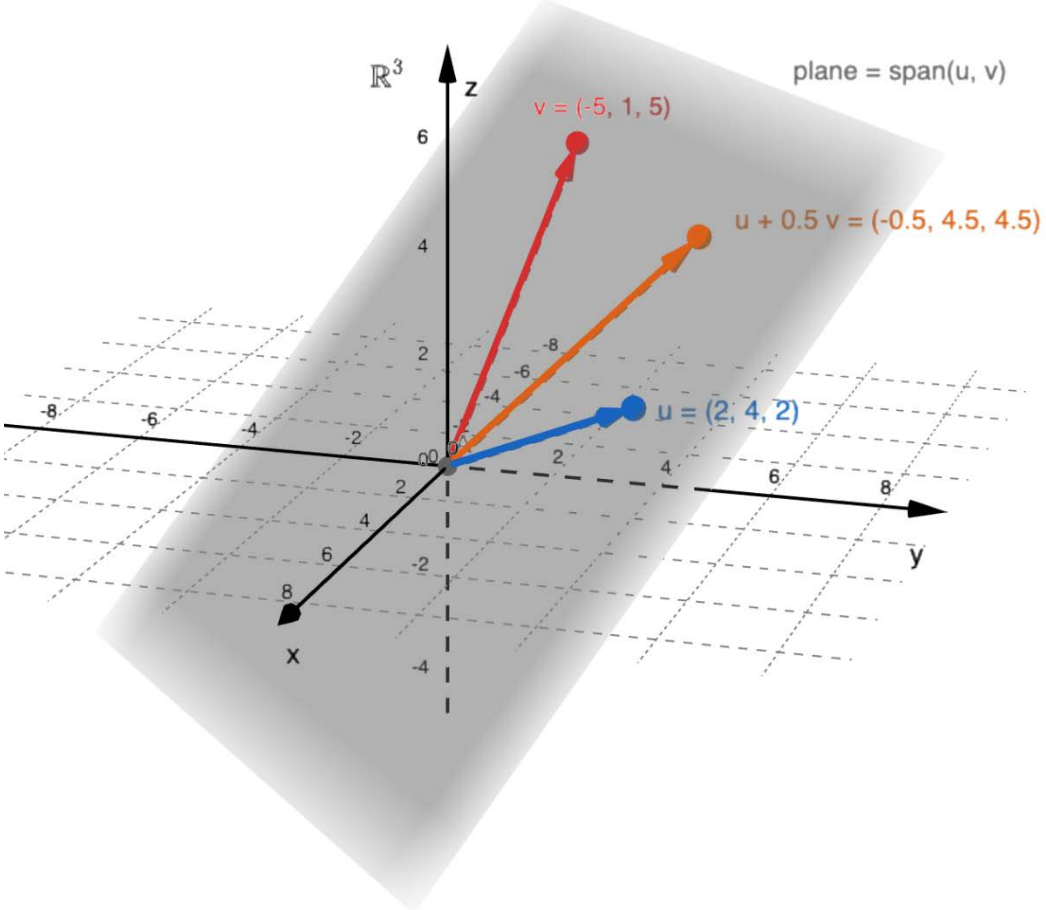
Geometry of Span

- **Plane** through origin containing the two (linearly independent) vectors $\vec{v}_1$ and $\vec{v}_2$

$$\operatorname{span}(\vec{v}_1, \vec{v}_2) \\ = \{\underbrace{a_1\vec{v}_1 + a_a\vec{v}_2}_{\text{any linear combination is a vector on the same plane}} : a_1, a_2 \in \mathbb{R}\}$$

- e.g.

$$\operatorname{span}\left(\begin{pmatrix}2\\4\\2\end{pmatrix}, \begin{pmatrix}-5\\1\\5\end{pmatrix}\right)$$



Vector Space

- A **vector space** V is a collection of vectors that is:
  1. Closed under vector addition:
     If $\vec{u}$ and $\vec{v}$ are in V, then $\vec{u} + \vec{v}$ is in V
  2. Closed under scalar multiplication:
     If $\vec{v}$ is in V and $a \in \mathbb{R}$, then $a\vec{v}$ is in V

Vector Space Properties

- If V is a vector space, then:
  - Must contain the zero vector $\vec{0} \in V$
  - Closed under any linear combination
    If $\vec{v}_1, \dots, \vec{v}_k \in V$ and $a_1, \dots, a_k \in \mathbb{R}$, then

$$a_1\vec{v}_1 + \cdots + a_k\vec{v}_k \in V$$

Interpretation: A vector space is such a "rich" collection that no matter how we linearly combine the vectors in the space, the resulting vector is also in the space

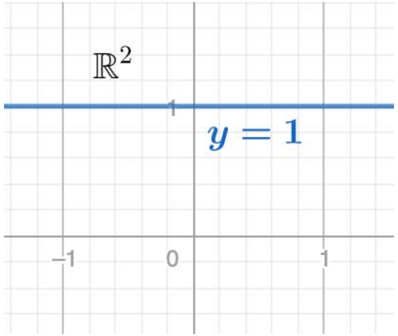
Examples of Vector Space in $\mathbb{R}^2$

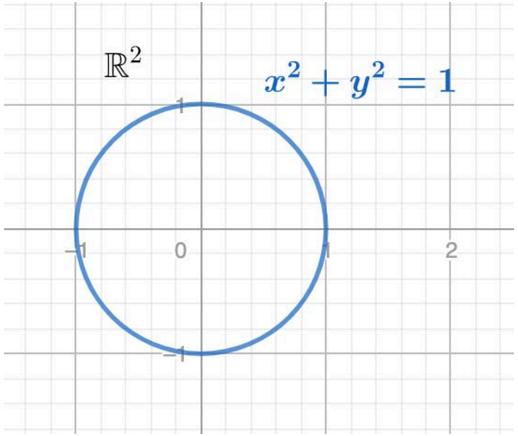
- Examples of vector space: origin itself, a line through origin, entire Euclidean space

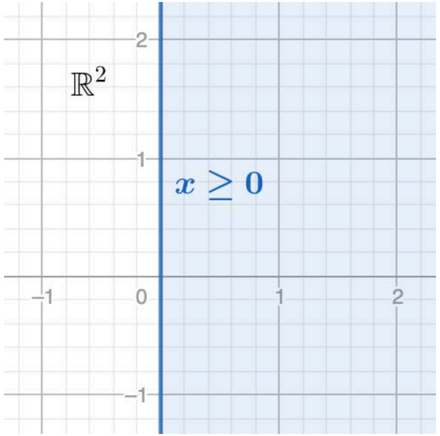
$$V = \left\{\begin{pmatrix}0\\0\end{pmatrix}\right\} \qquad V = \left\{\begin{pmatrix}x\\y\end{pmatrix} : x + 2y = 0\right\} \qquad V = \mathbb{R}^2$$

- Subsets in $\mathbb{R}^2$ but not vector space

$$S = \left\{\begin{pmatrix}x\\1\end{pmatrix} : x \in \mathbb{R}\right\} \qquad S = \left\{\begin{pmatrix}x\\y\end{pmatrix} : x^2 + y^2 = 1\right\} \qquad S = \left\{\begin{pmatrix}x\\y\end{pmatrix} : x \geq 0\right\}$$







Span is a Vector Space

- A span is always a vector space (basically by design)
- Proof?

Let $V = \text{span}(\vec{v}_1, \dots, \vec{v}_k)$. If $\vec{u}$ and $\vec{v} \in V$ so

$$\vec{u} = a_1\vec{v}_1 + \cdots + a_k\vec{v}_k$$
$$\vec{v} = b_1\vec{v}_1 + \cdots + b_k\vec{v}_k$$

then

$$\vec{u} + \vec{v} = (a_1 + b_1)\vec{v}_1 + \cdots + (a_k + b_k)\vec{v}_k \in V$$

$$s\vec{u} = (sa_1)\vec{v}_1 + \cdots + (sa_k)\vec{v}_k \in V$$

Vector Space is a Span

- A (finite-dimensional) vector space can always be expressed as a span (of some vectors)

e.g. The xy-plane in $\mathbb{R}^3$

$$\left\{\begin{pmatrix} x \\ y \\ z \end{pmatrix} : z = 0\right\} = \text{span}\left(\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}\right)$$

- For our purpose, it is enough to think of "vector space" as a synonym for "span" (of a set of vectors)

Vector Space in $\mathbb{R}^d$

- The only vector spaces / spans in $\mathbb{R}^d$ are either:
    1. The singleton set of the origin $\{\vec{0}\}$
    2. A line through the origin
    3. A plane through the origin
    4. … etc
- It must contain the origin!

Subspace

- A **subspace** is a vector space that is also a subset of another vector space

e.g. xy-plane is a subspace of $\mathbb{R}^3$

$$\left\{\begin{pmatrix} x \\ y \\ z \end{pmatrix} : z = 0\right\} \subsetneq \mathbb{R}^3$$

e.g. x-axis is a subspace of both the xy-plane and $\mathbb{R}^3$

$$\operatorname{span}\left(\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}\right) \subsetneq xy\text{–plane} \subsetneq \mathbb{R}^3$$

Null Space of Matrix

- The **null space** / **kernel** of a matrix A is

$$\mathrm{null}(A) = \{\vec{x} \in \mathbb{R}^d : A\vec{x} = \vec{0}\} = \begin{pmatrix} \text{solution set of} \\ A\vec{x} = \vec{0} \end{pmatrix}$$

- It is always a vector space in $\mathbb{R}^d$

Proof?

If $\vec{v}$ and $\vec{u} \in \mathrm{null}(A)$, then

$$A(\vec{v} + \vec{u}) = A\vec{v} + A\vec{u} = \vec{0}$$
$$A(s\vec{v}) = s(A\vec{v}) = s\vec{0} = \vec{0}$$

so $\vec{v} + \vec{u}$ and $s\vec{v} \in \mathrm{null}(A)$ too

Or parametric form of solution set of homogeneous system is a span

Column Space of Matrix

- The **column space** / **range** of a matrix A is

$$\mathrm{range}(A) = \mathrm{span}\begin{pmatrix} \text{column} \\ \text{vectors of } A \end{pmatrix}$$

- It is always a vector space in $\mathbb{R}^n$

e.g.

$$A = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix} \quad \mathrm{range}(A) = \left\{ \begin{pmatrix} x \\ 0 \end{pmatrix} : x \in \mathbb{R} \right\}$$

$$B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \mathrm{range}(B) = \mathbb{R}^3$$

- **Fact:** An inhomogeneous system $A\vec{x} = \vec{b}$ is consistent iff $\vec{b} \in \mathrm{range}(A)$

Spanning Set

- **Previous fact:** A vector space V can always be expressed as a span of some vectors

$$V = \mathrm{span}(\vec{v}_1, \ldots, \vec{v}_k)$$

Such a set of vectors $\vec{v}_1, \ldots, \vec{v}_k$ is called a **spanning set** of V

- In this case, every $\vec{v} \in V$ can be expressed as a linear combination

$$\vec{v} = a_1\vec{v}_1 + \cdots + a_k\vec{v}_k$$

with some $a_1, \ldots, a_k \in \mathbb{R}$

Non-Unique and Redundancy

- Spanning set is not unique and can have redundancy

e.g.

or any two linearly independent vectors

$$\begin{aligned}\mathbb{R}^2 &= \mathrm{span}\left(\begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}\right)\\ &= \mathrm{span}\left(\begin{pmatrix}1\\2\end{pmatrix}, \begin{pmatrix}3\\4\end{pmatrix}\right)\\ &= \mathrm{span}\left(\underbrace{\begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}, \begin{pmatrix}1\\1\end{pmatrix}}_{\text{not linearly independent}}\right)\end{aligned}$$

Basis

- A **basis** $\{\vec{v}_1, \dots, \vec{v}_k\}$ of a vector space V is a linearly independent spanning set
  1. **Spanning**: It has enough vectors to span / reach the entire vector space V
  2. **Linear independence**: It has just enough vectors to avoid redundancy i.e. not unnecessarily large

$$V = \text{span}\underbrace{(\vec{v}_1, \dots, \vec{v}_k)}_{\text{also linearly independent}}$$

Examples of Basis of $\mathbb{R}^2$

e.g. Not a spanning set so is not a basis

$$\left\{\begin{pmatrix}1\\2\end{pmatrix}\right\} \qquad \left\{\begin{pmatrix}1\\2\end{pmatrix}, \begin{pmatrix}-3\\-6\end{pmatrix}\right\}$$

e.g. Spanning set but not linearly independent so still not a basis

$$\left\{\begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}, \begin{pmatrix}1\\1\end{pmatrix}\right\}$$

e.g. Basis in $\mathbb{R}^2$ (there are many)

$$\left\{\begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}\right\} \qquad \left\{\begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}1\\1\end{pmatrix}\right\} \qquad \left\{\begin{pmatrix}1\\2\end{pmatrix}, \begin{pmatrix}3\\4\end{pmatrix}\right\}$$

- Basis is not unique. A vector space has many bases.

Dimension

- **Fact**: All bases of a vector space V must have the same number of vectors. This common number is the **dimension**

$$\dim(V)$$

of the vector space

e.g.

$$\dim(\mathbb{R}^2) = 2 \qquad \dim(\mathbb{R}^3) = 3$$

- Dimension describes the "size" of a vector space (which is a collection of infinitely many vectors)

Dimension Properties

- For a d-dimensional vector space V:

1. A set $\{\vec{v}_1, \dots, \vec{v}_k\}$ of $k < d$ vectors cannot span V
2. A set $\{\vec{v}_1, \dots, \vec{v}_k\}$ of $k > d$ vectors cannot be linearly independent
3. If a set $\{\vec{v}_1, \dots, \vec{v}_d\}$ is linearly independent, then it must be a basis
4. If a set $\{\vec{v}_1, \dots, \vec{v}_d\}$ is a spanning set, then it must be a basis

Standard Basis of $\mathbb{R}^d$

- The **standard basis** of $\mathbb{R}^d$ is

$$\{\vec{e}_1, \dots, \vec{e}_d\}$$

with

$$\vec{e}_i = (0, \dots, 0,1,0, \dots, 0)$$

i.e. all zeros except at the i^{th} entry

e.g.

$$\left\{\begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}\right\} \text{ for } \mathbb{R}^2$$

$$\left\{\begin{pmatrix}1\\0\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\end{pmatrix}, \begin{pmatrix}0\\0\\1\end{pmatrix}\right\} \text{ for } \mathbb{R}^3$$

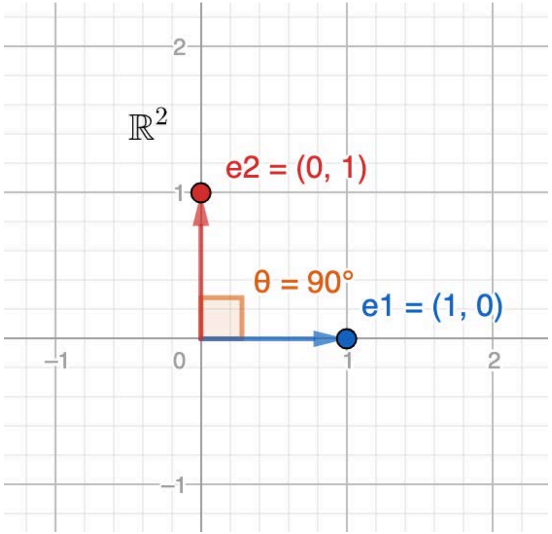
Standard Basis Properties

- The standard basis is:

$\left.\begin{array}{l}\text{1. Unit vectors}\\ \text{2. Mutually orthogonal}\end{array}\right\}$ "orthonormal"

(implies linearly independent)

3. Axis-aligned



Orthonormal Basis of $\mathbb{R}^d$

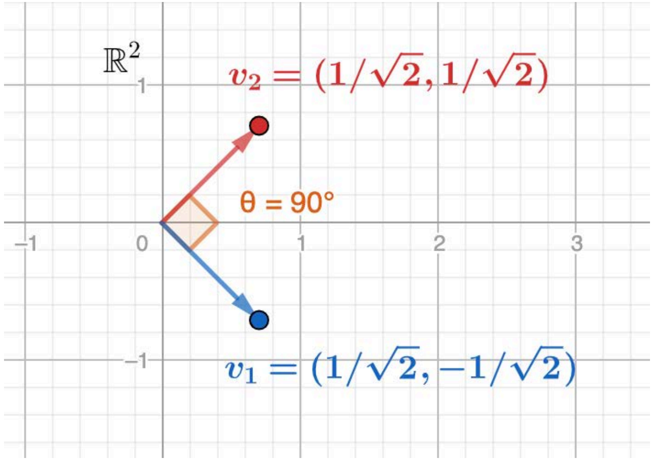
- An **orthonormal basis** of $\mathbb{R}^d$ is:
  1. Unit vectors
  2. Mutually orthogonal
     (implies linearly independent)
  3. ~~Axis-aligned~~

e.g.

$$\left\{\begin{pmatrix}1/\sqrt{2}\\-1/\sqrt{2}\end{pmatrix},\begin{pmatrix}1/\sqrt{2}\\1/\sqrt{2}\end{pmatrix}\right\} \text{ for } \mathbb{R}^2$$

Orthonormal Basis Properties

- A “rotated” version of the standard basis



- Columns of an orthogonal matrix $Q \in \mathbb{R}^{d\times d}$ form an orthonormal basis

$$Q = \begin{pmatrix}1/\sqrt{2} & 1/\sqrt{2}\\-1/\sqrt{2} & 1/\sqrt{2}\end{pmatrix}$$

Basis for Null Space

- The homogeneous system of linear equations

$$A\vec{x} = \vec{0} \quad \text{with} \quad A = \begin{pmatrix} 1 & -1 & 2 \\ 2 & -2 & 4 \\ -1 & 1 & -2 \end{pmatrix}$$

has solution set

$$\left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = a \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + b \begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix} : a, b \in \mathbb{R} \right\}$$

a basis for $\text{null}(A)$

- **Fact**: Gaussian elimination always yields linearly independent vectors

Matrix Nullity

- The **nullity** of a matrix $A \in \mathbb{R}^{r \times c}$ is

$$\text{nullity}(A) = \dim\big(\text{null}(A)\big)$$

e.g.

$$\text{nullity} \begin{pmatrix} 1 & -1 & 2 \\ 2 & -2 & 4 \\ -1 & 1 & -2 \end{pmatrix} = 2$$

Basis for Column Space

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

$$\mathrm{range}(A) = \mathrm{span}(\text{all columns of } A)$$

- **Question**: How to obtain a basis for $\mathrm{range}(A)$?

- **Answer:** Find a linearly independent subset of the columns of A with the same span

Reduce Spanning Set to Basis

- Use row echelon form to identify a maximal linearly independent subset

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

- 1st and 3rd columns of A form a basis for its column space

$$\mathrm{range}(A) = \mathrm{span}\left(\begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}\right)$$

Matrix Rank

- The **rank** of a matrix $A \in \mathbb{R}^{r \times c}$ is

$$\operatorname{rank}(A) = \dim\big(\operatorname{range}(A)\big)$$

- Rank is the number of linearly independent column vectors in A

e.g.

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

- Rank is the number of pivots in row echelon form of A

Rank 1 Matrix

- If a matrix A has rank 1, then each column is a scalar multiple of the 1st column

$$A = \begin{pmatrix} | & | & & | \\ s_1 A_1 & s_2 A_1 & \cdots & s_c A_1 \\ | & | & & | \end{pmatrix}$$

- It can be rewritten as an **outer product** of two vectors

e.g.

$$\begin{pmatrix} 1 & 10 & 100 \\ 2 & 20 & 200 \\ 3 & 30 & 300 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \times \begin{pmatrix} 1 & 10 & 100 \end{pmatrix}$$

Matrix Rank Properties

- Let $A \in \mathbb{R}^{r \times c}$ be a matrix
  1. $\operatorname{rank}(A) = \operatorname{rank}(A^T)$
  2. $\operatorname{rank}(A)$ is also the number of linearly independent rows vectors in A
  3. $\operatorname{rank}(A) \leq \min(r, c)$

e.g.

$$\operatorname{rank}\begin{pmatrix} 1 & 2 \\ 2 & 4 \\ 3 & 6 \end{pmatrix} = 1$$

- If B is another matrix, then

$$\operatorname{rank}(AB) \leq \min\big(\operatorname{rank}(A), \operatorname{rank}(B)\big)$$

Full Rank

- A matrix $A \in \mathbb{R}^{r \times c}$ has **full rank** if

$$\operatorname{rank}(A) = \min(r, c)$$

- It has as many linearly independent rows (or columns) as possible given the shape of A

e.g.

full rank

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}$$

not full rank

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

- **Fact**: A square matrix $A \in \mathbb{R}^{d \times d}$ is invertible iff it has full rank d
- For non-square matrix, having full rank is as close to the concept of invertible as possible

Rank-Nullity Dimension Theorem

- Let $A \in \mathbb{R}^{n\times d}$ be a matrix with $\text{rank}(A) = k$. The homogeneous system $A\vec{x} = \vec{0}$ of linear equations has

$$\underbrace{\begin{pmatrix}\text{number of}\\ \text{variables}\end{pmatrix}}_{d} = \underbrace{\begin{pmatrix}\text{number of pivots /}\\ \text{leading variables}\end{pmatrix}}_{\text{rank}(A) = k} + \underbrace{\begin{pmatrix}\text{number of}\\ \text{free variables}\end{pmatrix}}_{\text{nullity}(A) = d-k}$$

$$= \begin{pmatrix}\text{number of}\\ \text{non} - \text{redundant}\\ \text{equations}\end{pmatrix} + \dim(\text{solution set})$$

- This is the **fundamental theorem of linear algebra**!
- **Fact**: If a system is under-determined i.e. $d > n$, then $\dim(\text{solution set}) = d - k \geq d - n > 0$ so it has infinitely many solutions.

Matrix Invertibility Criteria

- The following statements are equivalent:

1. The matrix $A \in \mathbb{R}^{d \times d}$ is invertible
2. The matrix A^T is invertible
3. The determinant $\det(A) \neq 0$
4. The d row (or column) vectors of A are linearly independent / span $\mathbb{R}^d$ / is a basis for $\mathbb{R}^d$
5. The matrix A has full rank d
6. The system $A\vec{x} = \vec{0}$ only has trivial solution i.e. $\vec{x} = \vec{0}$
7. $\text{null}(A) = \{\vec{0}\}$
8. $\text{nullity}(A) = 0$
9. The system $A\vec{x} = \vec{b}$ has unique solution i.e. $\vec{x} = A^{-1}\vec{b}$