PODS Lab 2: Linear Algebra

Hamza Alshamy

Center for Data Science, NYU
ha2486@nyu.edu

01/31/25

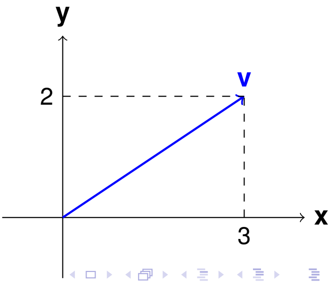
1. **Vectors**
2. **Vector Norms**
3. **Vector Projections**

Definition of Vectors

- **Vectors**: Ordered lists of numbers.
  - **Meaning:** Each entry provides a direction, and the magnitude of its value tells you its effect in that direction.
- **Dimensionality**: The number of components in a vector.
- **Orientation**: Column vs. Row vector.
  - **Transpose**: Converts a row vector to a column vector (and vice versa).

Example: A vector in $\mathbb{R}^2$:

$$\mathbf{v} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}, \quad \mathbf{v} \in \mathbb{R}^2$$



Examples of Vectors in Different Spaces

A vector in $\mathbb{R}^3$:

$$\mathbf{v} = \begin{bmatrix} 4 \\ -2 \\ 7 \end{bmatrix}, \quad \mathbf{v} \in \mathbb{R}^3$$

A vector in $\mathbb{R}^5$:

$$\mathbf{w} = \begin{bmatrix} 1 \\ -3 \\ 0 \\ 5 \\ 2 \end{bmatrix}, \quad \mathbf{w} \in \mathbb{R}^5$$

Dimensions:

- $\mathbb{R}^3 \rightarrow$ A three-dimensional vector.
- $\mathbb{R}^5 \rightarrow$ A five-dimensional vector.

Transpose of a Vector

Column Vector:

$$\mathbf{x} = \begin{bmatrix} 2 \\ -5 \\ 8 \end{bmatrix}, \quad \mathbf{x} \in \mathbb{R}^3$$

Row Vector (Transpose):

$$\mathbf{x}^T = \begin{bmatrix} 2 & -5 & 8 \end{bmatrix}$$

Key Idea: Transposing flips the dimensions of a vector.

Vector Operations

1. Vector Addition:

$$\mathbf{a}+\mathbf{b}=\begin{bmatrix}3\\1\end{bmatrix}+\begin{bmatrix}2\\-4\end{bmatrix}=\begin{bmatrix}5\\-3\end{bmatrix}$$

2. Scalar Multiplication:

$$3\mathbf{a}=3\begin{bmatrix}3\\1\end{bmatrix}=\begin{bmatrix}9\\3\end{bmatrix}$$

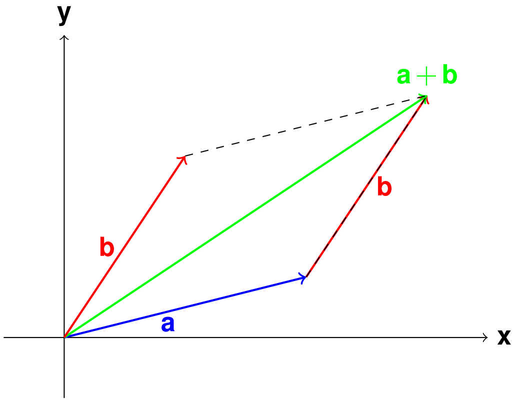
3. Dot Product:

$$\mathbf{a}\cdot\mathbf{b}^T=\sum_{i=1}^{n}a_ib_i=(3\times 2)+(1\times -4)=\ 6-4=2$$

Geometry of Vectors: Addition

Assume

$$\mathbf{a} = \begin{bmatrix} 4 \\ 1 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$$



- The *blue* vector **a** starts at the origin.
- The *red* vector **b** is placed at the tip of **a**.
- The *green* vector represents $\mathbf{a} + \mathbf{b}$.

Vector Norms

- **Norms:** Functions used to assign a magnitude/length to a given vector.
  - Different norms measure magnitude differently.
  - We choose the most appropriate norm depending on the context.
- **Definition:**

$$||\mathbf{x}||_p = \left(\sum_{i=1}^{n} |x_i|^p\right)^{\frac{1}{p}}$$
$$= \sqrt[p]{|x_1|^p + |x_2|^p + \ldots + |x_i|^p}$$

L1 Norm (AKA Manhattan Norm)

- The L1 norm measures distance as the sum of absolute values of vector components.
- Formula:

$$||\mathbf{x}||_1 = \sum_{i=1}^{n} |x_i|$$

- Example:

$$\mathbf{x} = \begin{bmatrix} 3 \\ -4 \\ 5 \end{bmatrix}, \quad ||\mathbf{x}||_1 = |3| + |-4| + |5| = 3 + 4 + 5 = 12$$

- **Interpretation:**
  - Think of it as the distance traveled on a grid (Manhattan distance).

L2 Norm (AKA Euclidean Norm)

- The L2 norm measures distance as the straight-line (Euclidean) distance from the origin.
- Formula:

$$||\mathbf{x}||_2 = \sqrt{\sum_{i=1}^{n} x_i^2}$$

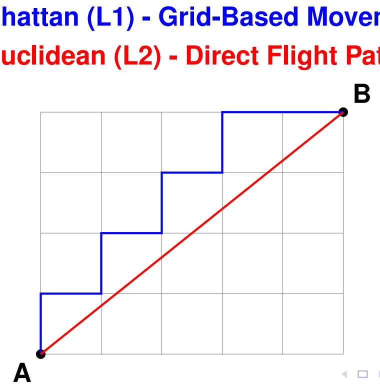
- Example:

$$\mathbf{x} = \begin{bmatrix} 3 \\ -4 \\ 5 \end{bmatrix}, \quad ||\mathbf{x}||_2 = \sqrt{3^2 + (-4)^2 + 5^2} = \sqrt{9 + 16 + 25} = \sqrt{50} \approx 7.1$$

- **Interpretation:**
  - Measures the direct shortest distance between two points.

Vector Norms - Geometry

- **Manhattan Norm** (L_1): Think of an Uber in Manhattan.
  - Movement is constrained to streets, so no diagonal travel is allowed.
- **Euclidean Norm** (L_2): Think of a drone taxi.
  - It can fly directly between two points.



L5 Norm?

- The L5 norm is a special case of the general Lp norm:

$$||\mathbf{x}||_p = \left(\sum_{i=1}^{n} |x_i|^p \right)^{\frac{1}{p}}$$

- Setting $p = 5$:

$$||\mathbf{x}||_5 = \left(\sum_{i=1}^{n} |x_i|^5 \right)^{\frac{1}{5}}$$

- Example:

$$\mathbf{x} = \begin{bmatrix} 3 \\ -4 \\ 5 \end{bmatrix}, \quad ||\mathbf{x}||_5 = \left(3^5 + (-4)^5 + 5^5\right)^{\frac{1}{5}} \approx 4.7$$

Norms Comparison

$$\textbf{Norms of } \mathbf{x} = \begin{bmatrix} 3 \\ -4 \\ 5 \end{bmatrix}$$

Norm	**Value**
L_1 (Manhattan)	12
L_2 (Euclidean)	7.1
L_5	4.7
L_{30}	5

- **Key point:** As norms increase, the magnitude of a vector converges to the largest value in that vector (e.g. $L_{30} \approx 5$).

Vector Projections: Intuition and Relevance

- **What is a vector projection?**
  - The projection of a vector **b** onto another vector **a** is the shadow of **b** when "dropped" onto **a**.
  - It represents the component of **b** that aligns with **a**.
- **Why is it relevant in Data Science?**
  - **Principal Component Analysis (PCA)**:
    - PCA projects data onto principal components to reduce dimensionality while preserving variance.
  - **Linear Regression**:
    - Regression finds the best fit line by projecting data points onto a regression line.
  - **Least Squares Optimization**:
    - The residuals (errors) in regression are minimized via perpendicular projections.

Derivation of the Vector Projection

Goal: Find the scalar β such that $\beta\mathbf{a}$ is as close as possible to $\mathbf{b}$.

Since the residual $\mathbf{b}-\beta\mathbf{a}$ is perpendicular to $\mathbf{a}$, we set:

$$(\mathbf{b}-\beta\mathbf{a})^T\mathbf{a}=0$$

Expanding the equation:

$$\mathbf{a}^T\mathbf{b}-\beta\mathbf{a}^T\mathbf{a}=0 \implies \mathbf{a}^T\mathbf{b}=\beta\mathbf{a}^T\mathbf{a}$$

Solving for β:

$$\beta=\frac{\mathbf{a}^T\mathbf{b}}{\mathbf{a}^T\mathbf{a}}$$

Thus, the projection of $\mathbf{b}$ onto $\mathbf{a}$ is:

$$\text{proj}_{\mathbf{a}}(\mathbf{b})=\frac{\mathbf{a}^T\mathbf{b}}{\mathbf{a}^T\mathbf{a}}\mathbf{a}$$