

Applied Linear Algebra in Data Analysis

Concepts in Vector Spaces

Sivakumar Balasubramanian

Department of Bioengineering
Christian Medical College, Bagayam
Vellore 632002

Vectors

- **Vectors** are ordered list of numbers (scalars). $\mathbf{v} = \begin{bmatrix} 1.2 \\ -0.1 \\ \vdots \\ -1.24 \end{bmatrix}$.

Note: Small bold letter will represent vectors. e.g. $\mathbf{a}, \mathbf{x}, \dots$

- Scalars can be any *field* $\mathbb{R}, \mathbb{C}, \mathbb{Z}, \mathbb{Q}$. Scalars will be represented using lower case normal font, e.g. $x, y, \alpha, \beta, \dots$
- Addition/multiplication operations performed on vectors will follow the rules of addition/multiplication of the corresponding scalar fields.
- We will typically encounter only \mathbb{R} and \mathbb{C} in this course.

Vectors

- ▶ Individual elements of a vector \mathbf{v} are indexed. The i^{th} element of \mathbf{v} is referred to as v_i .
- ▶ *Dimension* or *size* of a vector is number of elements in the vector.
- ▶ Set of n -real vectors is denoted by \mathbb{R}^n (similarly, \mathbb{C}^n)
- ▶ Vectors \mathbf{a} and \mathbf{b} are equal, if
 - ▶ both have the same size; and
 - ▶ $a_i = b_i, i \in \{1, 2, 3, \dots n\}$

Vectors

► Unit vector $\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$ Zero vector $\mathbf{0} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$ One vector $\mathbf{1} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}$

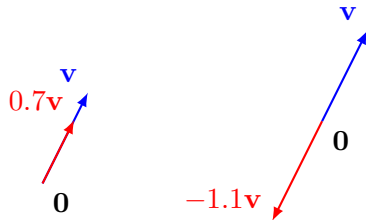
► Geometrically, real n -vectors can be thought of as points in \mathbb{R}^n space.



Vectors

- **Vector scaling:** Multiplication of a scalar and a vector.

$$\mathbf{w} = a\mathbf{v} = a \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ \vdots \\ v_n \end{bmatrix} = \begin{bmatrix} av_1 \\ av_2 \\ av_3 \\ \vdots \\ av_n \end{bmatrix} \quad a \in \mathbb{R}; \mathbf{w}, \mathbf{v} \in \mathbb{R}^n$$



Properties

- Scalar multiplication is *commutative*.

$$\alpha\mathbf{v} = \mathbf{v}\alpha$$

- Scalar multiplication is *associative*.

$$(\alpha\beta)\mathbf{v} = \alpha(\beta\mathbf{v})$$

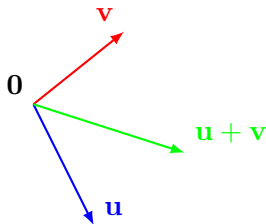
- Scalar multiplication is *distributive*.

$$(\alpha + \beta)\mathbf{v} = \alpha\mathbf{v} + \beta\mathbf{v}$$

Vectors

- **Vector addition:** Adding two vectors of the same dimension, element by element.

$$\mathbf{u} + \mathbf{v} = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ \vdots \\ u_n \end{bmatrix} + \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ \vdots \\ v_n \end{bmatrix} = \begin{bmatrix} u_1 + v_1 \\ u_2 + v_2 \\ u_3 + v_3 \\ \vdots \\ u_n + v_n \end{bmatrix} \quad \mathbf{u}, \mathbf{v} \in \mathbb{R}^n$$



Properties

- Vector addition is *commutative*.

$$\mathbf{a} + \mathbf{b} = \mathbf{b} + \mathbf{a}$$

- Vector addition is *associative*.

$$(\mathbf{a} + \mathbf{b}) + \mathbf{c} = \mathbf{a} + (\mathbf{b} + \mathbf{c})$$

- Zero vector has no effect.

$$\mathbf{a} + \mathbf{0} = \mathbf{a}$$

- Subtraction of vectors.

$$\mathbf{a} + (-1)\mathbf{a} = \mathbf{a} - \mathbf{a} = \mathbf{0}$$

Vector spaces

- ▶ A set of vectors V that is closed under **vector addition** and **vector scaling**.

$$\forall \mathbf{x}, \mathbf{y} \in V, \quad \mathbf{x} + \mathbf{y} \in V$$

$$\forall \mathbf{x} \in V, \text{ and } \alpha \in F, \quad \alpha \mathbf{x} \in V$$

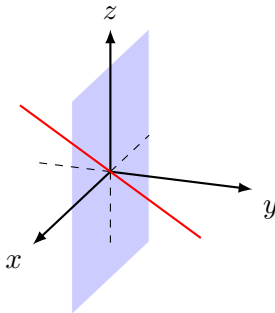
- ▶ For a set to be a vector space, it must satisfy the following properties: $\mathbf{x}, \mathbf{y}, \mathbf{z} \in V$
 - ▶ *Commutativity*: $\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}$
 - ▶ *Associativity of vector addition*: $(\mathbf{x} + \mathbf{y}) + \mathbf{z} = \mathbf{x} + (\mathbf{y} + \mathbf{z})$
 - ▶ *Additive identity*: $\mathbf{x} + \mathbf{0} = \mathbf{0} + \mathbf{x} = \mathbf{x}$ ($\mathbf{0} \in V$)
 - ▶ *Additive inverse*: $\exists -\mathbf{x} \in V, \mathbf{x} + (-\mathbf{x}) = \mathbf{0}$
 - ▶ *Associativity of scalar multiplication*: $\alpha(\beta \mathbf{x}) = (\alpha\beta)\mathbf{x}$
 - ▶ *Distributivity of scalar sums*: $(\alpha + \beta)\mathbf{x} = \alpha\mathbf{x} + \beta\mathbf{x}$
 - ▶ *Distributivity of vector sums*: $\alpha(\mathbf{x} + \mathbf{y}) = \alpha\mathbf{x} + \alpha\mathbf{y}$
 - ▶ *Scalar multiplication identity*: $1\mathbf{x} = \mathbf{x}$
- ▶ We will mostly deal with \mathbb{R}^n and \mathbb{C}^n vector spaces in this course.

Subspaces

- ▶ A **subspace** S of a vector space V is a subset of V and is itself a vector space.

$$S \subset V, \quad \forall \mathbf{x}, \mathbf{y} \in S, \alpha \mathbf{x} + \beta \mathbf{y} \in S, \quad \alpha, \beta \in F$$

- ▶ The zero vector is called the **trivial subspace** of a vector space V .
- ▶ For example, in \mathbb{R}^3 all planes and lines passing through the origin are subspaces of \mathbb{R}^3 .



Linear independence

- A collection of vectors $\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \dots, \mathbf{x}_n\}$, $\mathbf{x}_i \in \mathbb{R}^m$ $i \in \{1, 2, 3, \dots, n\}$ is called *linearly dependent* if,

$$\sum_{i=1}^n \alpha_i \mathbf{x}_i = 0, \text{ hold for some } \alpha_1, \alpha_2, \dots, \alpha_n \in \mathbb{R}, \text{ such that } \exists \alpha_i \neq 0$$

- Another way to state this: A collection of vectors is *linearly dependent* if at least one of the vectors in the collection can be expressed as a linear combination of the other vectors in the collection, i.e.

$$\mathbf{x}_i = - \sum_{j=1, j \neq i}^n \left(\frac{\alpha_j}{\alpha_i} \right) \mathbf{x}_j$$

Linear independence

- A collection of vectors is *linearly independent* if it is **not** *linearly dependent*.

$$\sum_{i=1}^n \alpha_i \mathbf{x}_i = 0 \implies \alpha_1 = \alpha_2 = \alpha_3 \dots = \alpha_n = 0$$

Span of a set of vectors

- ▶ Consider a set of vectors $S = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3 \dots \mathbf{v}_r\}$ where $\mathbf{v}_i \in \mathbb{R}^n, 1 \leq i \leq r$.
- ▶ The **span** of the set S is defined as the set of all linear combinations of the vectors \mathbf{v}_i ,
$$\text{span}(S) = \{\alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 + \dots + \alpha_r \mathbf{v}_r\}, \alpha_i \in \mathbb{R}$$
- ▶ Is $\text{span}(S)$ a subspace of \mathbb{R}^n ?

Span of a set of vectors

- ▶ We say that the subspace $\text{span}(S)$ is spanned by the *spanning set* S . $\longrightarrow S$ spans $\text{span}(S)$.

- ▶ **Sum of subspaces** X, Y is defined as the sum of all possible vectors from X and Y .

$$X + Y = \{\mathbf{x} + \mathbf{y} \mid \mathbf{x} \in X, \mathbf{y} \in Y\}$$

- ▶ Sum of two subspace is also a subspace.

Inner Product

- ▶ **Standard inner product** is defined as the following,

$$\mathbf{x}^\top \mathbf{y} = \sum_{i=1}^n x_i y_i, \quad \mathbf{x}, \mathbf{y} \in \mathbb{R}^n$$

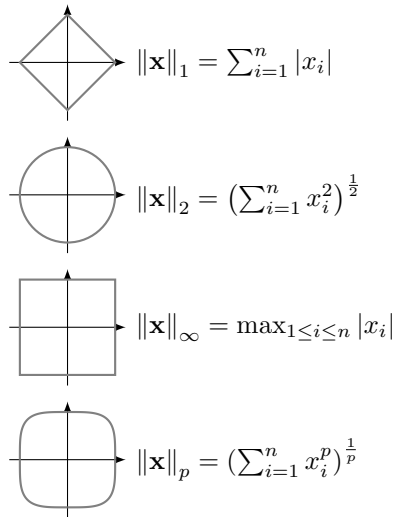
For complex vectors: $\mathbf{x}^* \mathbf{y} = \sum_{i=1}^n \bar{x}_i y_i, \quad \mathbf{x}, \mathbf{y} \in \mathbb{C}^n$

- ▶ **Properties**

- ▶ $\mathbf{x}^\top \mathbf{x} > 0, \quad \forall \mathbf{x} \neq 0$ and $\mathbf{x}^\top \mathbf{x} = 0 \Leftrightarrow \mathbf{x} = 0$
- ▶ *Commutative*: $\mathbf{x}^\top \mathbf{y} = \mathbf{y}^\top \mathbf{x}$
- ▶ *Associativity with scalar multiplication*: $(\alpha \mathbf{x})^\top \mathbf{y} = \alpha (\mathbf{x}^\top \mathbf{y})$
- ▶ *Distributivity with vector addition*: $(\mathbf{x} + \mathbf{y})^\top \mathbf{z} = \mathbf{x}^\top \mathbf{z} + \mathbf{y}^\top \mathbf{z}$

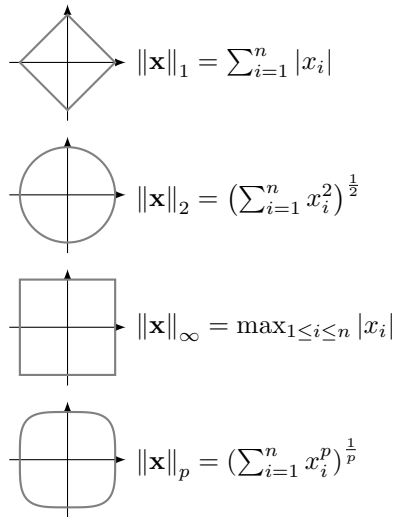
Norm

- ▶ Norm is a measure of the size of a vector.
- ▶ *Euclidean norm* of a n -vector $\mathbf{x} \in \mathbb{R}^n$ is defined as,
 $\|\mathbf{x}\|_2 = \sqrt{\mathbf{x}^\top \mathbf{x}} = \sqrt{\sum_{i=1}^n x_i^2}$.
- ▶ $\|\mathbf{x}\|_2$ is a measure of the length of the vector \mathbf{x} .
- ▶ Any function of the form $\|\bullet\| : \mathbb{R}^n \rightarrow \mathbb{R}_{\geq 0}$ is a valid norm, provided it satisfies the following properties.



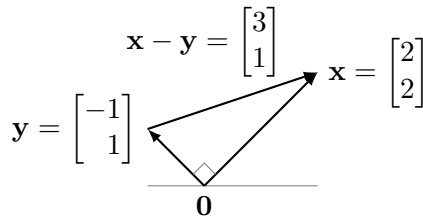
► Properties

- *Definiteness.* $\|\mathbf{x}\| = 0 \iff x = 0$
 - *Non-negativity.* $\|\mathbf{x}\| \geq 0$
 - *Non-negative homogeneity.* $\|\beta\mathbf{x}\| = |\beta| \|\mathbf{x}\|, \beta \in \mathbb{R}$
 - *Triangle inequality.* $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$
- p -norm: $\|\mathbf{x}\|_p = (\sum_{i=1}^n |x_i|^p)^{\frac{1}{p}}$
- Norm of difference between two vectors is a measure of the distance between the vectors. $d = \|\mathbf{x} - \mathbf{y}\|_2$.



Orthogonality

- Orthogonality is the idea of two vectors being perpendicular, $\mathbf{x} \perp \mathbf{y}$.



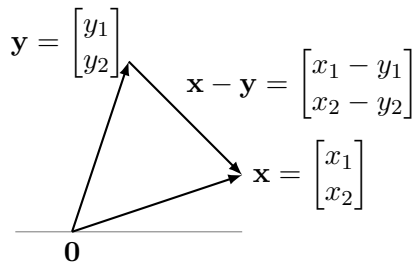
Using the Pythagorean theorem, $\|\mathbf{x} - \mathbf{y}\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2$

$$\|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 - 2\mathbf{x}^\top \mathbf{y} = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 \implies \mathbf{x}^\top \mathbf{y} = 0$$

- We extend this to the n -dimensional case and define two vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ being orthogonal, if

$$\mathbf{x}^\top \mathbf{y} = \sum_{i=1}^n x_i y_i = 0$$

Angle between vectors



- ▶ Inner products are used for projecting a vector onto another vector or a subspace.
- ▶ It is also a measure of similarity between two vectors, $\cos(\theta) = \frac{\mathbf{x}^\top \mathbf{y}}{\|\mathbf{x}\| \|\mathbf{y}\|}$
- ▶ **Cauchy-Bunyakovski-Schwartz Inequality:**

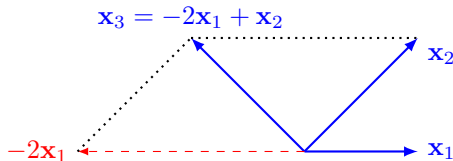
$$|\mathbf{x}^\top \mathbf{y}| \leq \|\mathbf{x}\| \|\mathbf{y}\|, \quad \mathbf{x}, \mathbf{y} \in \mathbb{R}^n$$

Basis

Consider a vector $\mathbf{y} = \sum_{i=1}^n \alpha_i \mathbf{x}_i$. What can we say about the coefficients α_i s when the collection $\{\mathbf{x}_i\}_{i=1}^n$ is,

- ▶ linearly independent $\implies \alpha_i$ s are *unique*.
- ▶ linearly dependent $\implies \alpha_i$ s are not *unique*.

Consider \mathbb{R}^2 vector space. $\mathbf{x}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $\mathbf{x}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, $\mathbf{x}_3 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$.

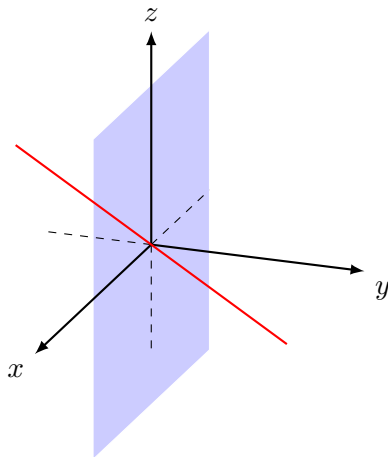


Independence-Dimension inequality: What is the maximum possible size of a linearly independent collection?

A linearly independent collection of n -vectors can at most have n vectors.

Basis

How many vectors can we choose from the following vectors before the set becomes linearly dependent?



Basis

- ▶ A linearly independent set of n -vectors from \mathbb{R}^n , of size n , is called a *basis* for \mathbb{R}^n .
- ▶ Any n -vector from \mathbb{R}^n can be represented as a *unique* linear combination of the elements of the basis.
- ▶ Consider the basis $\{\mathbf{x}_i\}_{i=1}^n$, $\mathbf{x}_i \in \mathbb{R}^n$. Any vector $\mathbf{y} \in \mathbb{R}^n$ can be represented as a linear combination of \mathbf{x}_i s, $\mathbf{y} = \sum_{i=1}^n \alpha_i \mathbf{x}_i$. This is called the *expansion of \mathbf{y}* in the $\{\mathbf{x}_i\}_{i=1}^n$ basis.

Basis

- ▶ The numbers α_i are called the *coefficients* of the expansion of \mathbf{y} in the $\{\mathbf{x}_i\}_{i=1}^n$ basis.
- ▶ **Orthogonal vectors:** A set of vectors $\{\mathbf{x}_i\}_{i=1}^n$ is (*mutually*) *orthogonal* if $\mathbf{x}_i \perp \mathbf{x}_j$ for all $i, j \in \{1, 2, 3, \dots, n\}$ and $i \neq j$.
- ▶ This set is called **orthonormal** if its elements are all of unit length $\|\mathbf{x}_i\|_2 = 1$ for all $i \in \{1, 2, 3, \dots, n\}$.

$$\mathbf{x}_i^\top \mathbf{x}_j = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases}$$

Representing a Vector in an Orthonormal Basis

- ▶ An orthonormal collection of vectors is linearly independent.
- ▶ Consider an orthonormal basis $\{\mathbf{x}_i\}_{i=1}^n$. The expansion of a vector \mathbf{y} is given by,

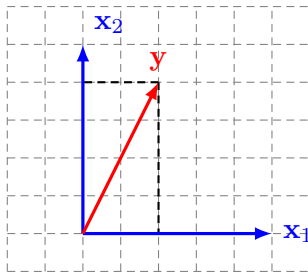
$$\mathbf{y} = \alpha_1 \mathbf{x}_1 + \alpha_2 \mathbf{x}_2 + \alpha_3 \mathbf{x}_3 + \dots + \alpha_n \mathbf{x}_n$$

$$\mathbf{x}_i^\top \mathbf{y} = \alpha_1 \mathbf{x}_i^\top \mathbf{x}_1 + \alpha_2 \mathbf{x}_i^\top \mathbf{x}_2 + \alpha_3 \mathbf{x}_i^\top \mathbf{x}_3 + \dots + \alpha_n \mathbf{x}_i^\top \mathbf{x}_n = \alpha_i$$

Representing a Vector in an Orthonormal Basis

► Thus, we can rewrite this as,

$$\mathbf{y} = (\mathbf{y}^\top \mathbf{x}_1) \mathbf{x}_1 + (\mathbf{y}^\top \mathbf{x}_2) \mathbf{x}_2 + (\mathbf{y}^\top \mathbf{x}_3) \mathbf{x}_3 + \dots + (\mathbf{y}^\top \mathbf{x}_n) \mathbf{x}_n$$



Dimension of a Vector Space

- ▶ There are an infinite number of bases for a vector space.
- ▶ There is one thing that is common among all these bases – the number of basis vectors.
- ▶ This number is a property of the vector space, and represents the “degrees of freedom” of the space. This is called the **dimension** of the vector space.

Dimension of a Vector Space

- ▶ A subspace of dimension m can have at most m independent vectors.
- ▶ Notice that the word “dimension” of a vector space is different from the “dimension” of a vector.
- ▶ E.g. Vectors from \mathbb{R}^3 are three dimensional vectors. But the yz -plane in \mathbb{R}^3 is a 2 dimensional subspace of \mathbb{R}^3 .

Linear Functions

- Let f be a function which maps vectors from \mathbb{R}^n to scalar real numbers. It can be represented as the following,

$$f : \mathbb{R}^n \longrightarrow \mathbb{R}; \quad y = f(\mathbf{x}) = f(x_1, x_2, x_3, \dots, x_n)$$

- Criteria for f to be a linear function:

$$\textbf{Superposition} : f(\alpha \mathbf{x} + \beta \mathbf{y}) = \alpha f(\mathbf{x}) + \beta f(\mathbf{y})$$

where $\alpha, \beta \in \mathbb{R}$ and $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$.

Linear Functions

- **Inner product** is a linear function in one of the arguments.

$$f(x) = \mathbf{w}^\top \mathbf{x} = w_1x_1 + w_2x_2 + w_3x_3 + \dots + w_nx_n$$

- Any linear function can be represented in the form $f(\mathbf{x}) = \mathbf{w}^\top \mathbf{x}$ with an appropriately chosen \mathbf{w} .