Numerical Linear Algebra
Vector Spaces

Franck K. Mutombo
franckm@aims.ac.za

Professor of Mathematics
University of Lubumbashi

use this email to contact me:
franck.mutombo@aims.ac.rw

October 2024

Linear Algebra I

Outline

Vector spaces

Vector spaces

Note:

We focus on the **field of real** numbers ($\mathbb{R}$) but most of the results can be generalized to the **field of complex** numbers ($\mathbb{C}$) in a straightforward fashion.

Vector spaces

A vector space or *linear space* (over the field $\mathbb{R}$) consists of

(a) a set of vectors $\mathcal{V}$

(b) an addition operation: $\mathcal{V} \times \mathcal{V} \to \mathcal{V}$

(c) a scalar multiplication operation: $\mathbb{R} \times \mathcal{V} \to \mathcal{V}$

(d) a distinguished element $\mathbf{0} \in \mathcal{V}$

Vector spaces

A vector space or *linear space* (over the field $\mathbb{R}$) consists of

(a) a set of vectors $\mathcal{V}$

(b) an addition operation: $\mathcal{V} \times \mathcal{V} \to \mathcal{V}$

(c) a scalar multiplication operation: $\mathbb{R} \times \mathcal{V} \to \mathcal{V}$

(d) a distinguished element $\mathbf{0} \in \mathcal{V}$

and satisfies the following properties:

1. $\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}, \quad \forall \mathbf{x}, \mathbf{y} \in \mathcal{V}$ (commutative under addition)
2. $(\mathbf{x} + \mathbf{y}) + \mathbf{z} = \mathbf{x} + (\mathbf{y} + \mathbf{z}), \quad \forall \mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathcal{V}$ (associative under addition)
3. $\mathbf{0} + \mathbf{x} = \mathbf{x}, \quad \forall \mathbf{x} \in \mathcal{V}$ ($\mathbf{0}$ being additive identity)
4. $\forall \mathbf{x} \in \mathcal{V} \ \exists \, (-\mathbf{x})$ such that $\mathbf{x} + (-\mathbf{x}) = \mathbf{0}$ ($-\mathbf{x}$ being additive inverse)
5. $(\alpha\beta)\mathbf{x} = \alpha(\beta\mathbf{x}), \quad \forall \alpha, \beta \in \mathbb{R} \quad \forall \mathbf{x} \in \mathcal{V}$ (associative under scalar mult.)
6. $\alpha(\mathbf{x} + \mathbf{y}) = \alpha\mathbf{x} + \alpha\mathbf{y}, \quad \forall \alpha \in \mathbb{R} \quad \forall \mathbf{x}, \mathbf{y} \in \mathcal{V}$ (distributive)
7. $1\mathbf{x} = \mathbf{x}, \quad \forall \mathbf{x} \in \mathcal{V}$ ($\mathbf{1}$ being multiplicative identity)

Vector spaces contd.

Example (Vector space)

- $\mathcal{V}_1 = \{\mathbf{0}\}$ for $\mathbf{0} \in \mathbb{R}^n$
- $\mathcal{V}_2 = \mathbb{R}^n$
- $\mathcal{V}_3 = \sum_{i=1}^{k} \alpha_i \mathbf{x}_i$ for $\alpha_i \in \mathbb{R}$, $k < n$, and $\mathbf{x}_i \in \mathbb{R}^n$

Vector spaces contd.

Example (Vector space)

- $\mathcal{V}_1 = \{\mathbf{0}\}$ for $\mathbf{0} \in \mathbb{R}^n$
- $\mathcal{V}_2 = \mathbb{R}^n$
- $\mathcal{V}_3 = \sum_{i=1}^{k} \alpha_i \mathbf{x}_i$ for $\alpha_i \in \mathbb{R}$, $k < n$, and $\mathbf{x}_i \in \mathbb{R}^n$

It is straight forward to show that $\mathcal{V}_1$, $\mathcal{V}_2$, and $\mathcal{V}_3$ satisfy properties 1–7 above. However, using a subspace argument is much simpler.

Vector spaces contd.

Example (Vector space)

- $\mathcal{V}_1 = \{\mathbf{0}\}$ for $\mathbf{0} \in \mathbb{R}^n$
- $\mathcal{V}_2 = \mathbb{R}^n$
- $\mathcal{V}_3 = \sum_{i=1}^{k} \alpha_i \mathbf{x}_i$ for $\alpha_i \in \mathbb{R}$, $k < n$, and $\mathbf{x}_i \in \mathbb{R}^n$

It is straight forward to show that $\mathcal{V}_1$, $\mathcal{V}_2$, and $\mathcal{V}_3$ satisfy properties 1–7 above. However, using a subspace argument is much simpler.

Definition (Subspace)

A **subspace** is a vector space that is a *subset* of another vector space.

Vector spaces contd.

Example (Vector space)

- $\mathcal{V}_1 = \{\mathbf{0}\}$ for $\mathbf{0} \in \mathbb{R}^n$
- $\mathcal{V}_2 = \mathbb{R}^n$
- $\mathcal{V}_3 = \sum_{i=1}^{k} \alpha_i \mathbf{x}_i$ for $\alpha_i \in \mathbb{R}$, $k < n$, and $\mathbf{x}_i \in \mathbb{R}^n$

It is straight forward to show that $\mathcal{V}_1$, $\mathcal{V}_2$, and $\mathcal{V}_3$ satisfy properties 1–7 above. However, using a subspace argument is much simpler.

Definition (Subspace)

A **subspace** is a vector space that is a *subset* of another vector space.

Example (Subspace)

$\mathcal{V}_3$ (and actually $\mathcal{V}_1$ as well as $\mathcal{V}_2$) in the example above is subspace of $\mathbb{R}^n$.

Vector spaces contd.

Definition (Linear independence)

A set of vectors, $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\}$, is **linearly independent** if

$$\alpha_1\mathbf{x}_1 + \alpha_2\mathbf{x}_2 + \cdots + \alpha_n\mathbf{x}_n = \mathbf{0} \;\Rightarrow\; \alpha_1 = \alpha_2 = \cdots = \alpha_n = 0.$$

Vector spaces contd.

Definition (Linear independence)

A set of vectors, $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$, is **linearly independent** if

$$\alpha_1 \mathbf{x}_1 + \alpha_2 \mathbf{x}_2 + \cdots + \alpha_n \mathbf{x}_n = \mathbf{0} \ \Rightarrow \ \alpha_1 = \alpha_2 = \cdots = \alpha_n = 0.$$

Definition (Span)

The span of $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n \in \mathcal{V}$ is the set of all vectors that are linear combination of them, i.e.

$$\text{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_n\} = \{\mathbf{v} \in \mathcal{V} \ : \exists \ \alpha_1, \ldots, \alpha_n, \text{ such that } \mathbf{v} = \alpha_1 \mathbf{v}_1, \ldots, \alpha_n \mathbf{v}_n\}$$

Vector spaces contd.

Definition (Linear independence)

A set of vectors, $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$, is **linearly independent** if

$$\alpha_1 \mathbf{x}_1 + \alpha_2 \mathbf{x}_2 + \cdots + \alpha_n \mathbf{x}_n = \mathbf{0} \;\Rightarrow\; \alpha_1 = \alpha_2 = \cdots = \alpha_n = 0.$$

Definition (Span)

The span of $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n \in \mathcal{V}$ is the set of all vectors that are linear combination of them, i.e.

$$\text{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_n\} = \{\mathbf{v} \in \mathcal{V} \;: \exists\; \alpha_1, \ldots, \alpha_n, \text{ such that } \mathbf{v} = \alpha_1 \mathbf{v}_1, \ldots, \alpha_n \mathbf{v}_n\}$$

Definition (Basis)

A set of vectors that is linearly independent and its spans is the whole of $\mathcal{V}$, is said to be a **basis** of $\mathcal{V}$.

Vector spaces contd.

Definition (Dimension)

The **dimension** of a vector space, $\mathcal{V}$, (denoted $\mathbf{dim}(\mathcal{V})$) is the number of vectors in the basis of $\mathcal{V}$.

If dimension of a vector space is finite, it is said to be finite-dimensional. Otherwise it is infinite-dimensional.

Vector spaces contd.

Example

Let $\mathbb{R}_n[x]$ denote the vector space of polynomial of degree at most n over the field of real numbers.

1. Show that it is indeed a vector space!
2. Since every degree n polynomial can be written as $a_0x^0 + a_1x^1 + \cdots + a_nx^n$, what is the dimension of $\mathbb{R}_n[x]$?
3. Show that the family $\{x^i\}_{i=0}^n$ is linearly independent.
4. Hence is the family $\{x^i\}_{i=0}^n$ a basis?

Euclidean spaces

Euclidean spaces

- The Euclidean spaces is when $\mathcal{V} = \mathbb{R}^n$ (most popular vector space)

Euclidean spaces

- The Euclidean spaces is when $\mathcal{V} = \mathbb{R}^n$ (most popular vector space)
- The vectors in this space consist of n-tuples of real numbers, i.e.

$$\mathbf{x} = (x_1, x_2, \dots x_n)$$

Euclidean spaces

- The Euclidean spaces is when $\mathcal{V} = \mathbb{R}^n$ (most popular vector space)
- The vectors in this space consist of n-tuples of real numbers, i.e.

$$\mathbf{x} = (x_1, x_2, \ldots x_n)$$

- For this course and typically in Data Science, vectors will be column vectors, i.e.

$$\mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

Euclidean spaces

- The Euclidean spaces is when $\mathcal{V} = \mathbb{R}^n$ (most popular vector space)
- The vectors in this space consist of n-tuples of real numbers, i.e.

$$\mathbf{x} = (x_1, x_2, \dots x_n)$$

- For this course and typically in Data Science, vectors will be column vectors, i.e.

$$\mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

- Addition and scalar multiplication are defined component-wise on vectors in $\mathbb{R}^n$, i.e.

$$\mathbf{x} + \mathbf{y} = \begin{bmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{bmatrix} \quad \text{and} \quad \alpha\mathbf{x} = \begin{bmatrix} \alpha x_1 \\ \vdots \\ \alpha x_n \end{bmatrix}$$

Euclidean spaces

- The Euclidean spaces is when $\mathcal{V} = \mathbb{R}^n$ (most popular vector space)
- The vectors in this space consist of n-tuples of real numbers, i.e.

$$\mathbf{x} = (x_1, x_2, \dots x_n)$$

- For this course and typically in Data Science, vectors will be column vectors, i.e.

$$\mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

- Addition and scalar multiplication are defined component-wise on vectors in $\mathbb{R}^n$, i.e.

$$\mathbf{x} + \mathbf{y} = \begin{bmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{bmatrix} \quad \text{and} \quad \alpha\mathbf{x} = \begin{bmatrix} \alpha x_1 \\ \vdots \\ \alpha x_n \end{bmatrix}$$

- Euclidean space is used to mathematically represent physical space, with notions such as **distance**, **length**, and **angles**.

Linear maps

Linear maps

Definition (Linear map)

A **linear map** is a function $\mathcal{T} : \mathcal{V} \to \mathcal{U}$, where $\mathcal{V}$ and $\mathcal{U}$ are vector spaces, that satisfies

1. $\mathcal{T}(\mathbf{x}+\mathbf{y}) = \mathcal{T}(\mathbf{x}) + \mathcal{T}(\mathbf{y})$ for all $\mathbf{x}, \mathbf{y} \in \mathcal{V}$
2. $\mathcal{T}(\alpha\mathbf{x}) = \alpha\mathcal{T}(\mathbf{x})$ for all $\mathbf{x} \in \mathcal{V}, \ \alpha \in \mathbb{R}$

Linear maps

Definition (Linear map)

A **linear map** is a function $\mathcal{T} : \mathcal{V} \to \mathcal{U}$, where $\mathcal{V}$ and $\mathcal{U}$ are vector spaces, that satisfies

1. $\mathcal{T}(\mathbf{x} + \mathbf{y}) = \mathcal{T}(\mathbf{x}) + \mathcal{T}(\mathbf{y})$ for all $\mathbf{x}, \mathbf{y} \in \mathcal{V}$
2. $\mathcal{T}(\alpha \mathbf{x}) = \alpha \mathcal{T}(\mathbf{x})$ for all $\mathbf{x} \in \mathcal{V}, \ \alpha \in \mathbb{R}$

- A composition of linear maps, $\mathcal{T}$ and $\mathcal{S}$ will be denoted as $\mathcal{T} \circ \mathcal{S}$
- A linear map from $\mathcal{V}$ to itself is called a linear operator

Linear maps

Definition (Linear map)

A **linear map** is a function $\mathcal{T} : \mathcal{V} \to \mathcal{U}$, where $\mathcal{V}$ and $\mathcal{U}$ are vector spaces, that satisfies

1. $\mathcal{T}(\mathbf{x}+\mathbf{y}) = \mathcal{T}(\mathbf{x}) + \mathcal{T}(\mathbf{y})$ for all $\mathbf{x}, \mathbf{y} \in \mathcal{V}$
2. $\mathcal{T}(\alpha\mathbf{x}) = \alpha\mathcal{T}(\mathbf{x})$ for all $\mathbf{x} \in \mathcal{V},\ \alpha \in \mathbb{R}$

- A composition of linear maps, $\mathcal{T}$ and $\mathcal{S}$ will be denoted as $\mathcal{T} \circ \mathcal{S}$
- A linear map from $\mathcal{V}$ to itself is called a linear operator
- A linear map preserves addition and scalar multiplication, hence reflect the structure of vector spaces $\Rightarrow$ a **homomorphism** of vector spaces
- An invertible homomorphism (where the inverse is also a homomorphism) is called an **isomorphism**

Linear maps

Definition (Linear map)

A **linear map** is a function $\mathcal{T} : \mathcal{V} \to \mathcal{U}$, where $\mathcal{V}$ and $\mathcal{U}$ are vector spaces, that satisfies

1. $\mathcal{T}(\mathbf{x} + \mathbf{y}) = \mathcal{T}(\mathbf{x}) + \mathcal{T}(\mathbf{y})$ for all $\mathbf{x}, \mathbf{y} \in \mathcal{V}$
2. $\mathcal{T}(\alpha \mathbf{x}) = \alpha \mathcal{T}(\mathbf{x})$ for all $\mathbf{x} \in \mathcal{V}, \ \alpha \in \mathbb{R}$

- A composition of linear maps, $\mathcal{T}$ and $\mathcal{S}$ will be denoted as $\mathcal{T} \circ \mathcal{S}$
- A linear map from $\mathcal{V}$ to itself is called a linear operator
- A linear map preserves addition and scalar multiplication, hence reflect the structure of vector spaces $\Rightarrow$ a **homomorphism** of vector spaces
- An invertible homomorphism (where the inverse is also a homomorphism) is called an **isomorphism**
- $\mathcal{V}$ and $\mathcal{W}$ are **isomorphic**, when there is an isomorphism from $\mathcal{V}$ to $\mathcal{W}$
- Every real n-dimensional vector space is isomorphic to $\mathbb{R}^n$

The matrix of a linear map

- In the finite-dimensional setting, matrices are used to represent and manipulate **vectors** and **linear maps** on a computer

The matrix of a linear map

- In the finite-dimensional setting, matrices are used to represent and manipulate **vectors** and **linear maps** on a computer

Definition (Matrix of a linear map)

Suppose $\mathbf{v}_1, \ldots, \mathbf{v}_n$ and $\mathbf{u}_1, \ldots, \mathbf{u}_n$ are bases of $\mathcal{V}$ and $\mathcal{U}$ respectively, and $\mathcal{T} : \mathcal{V} \to \mathcal{U}$ is a linear map. Then the **matrix** of $\mathcal{T}$, with entries A_{ij} for $i = 1, \ldots, m$ and $j = 1, \ldots, n$ is defined by

$$\mathcal{T}(\mathbf{v}_j) = A_{1j}\mathbf{u}_1 + A_{2j}\mathbf{u}_2 + \cdots + A_{mj}\mathbf{u}_m,$$

i.e. the jth column of $\mathbf{A}$ is the coordinates of $\mathcal{T}(\mathbf{v}_j)$ in the chosen basis for $\mathcal{U}$.

The matrix of a linear map

- In the finite-dimensional setting, matrices are used to represent and manipulate **vectors** and **linear maps** on a computer

Definition (Matrix of a linear map)

Suppose $\mathbf{v}_1, \ldots, \mathbf{v}_n$ and $\mathbf{u}_1, \ldots, \mathbf{u}_n$ are bases of $\mathcal{V}$ and $\mathcal{U}$ respectively, and $\mathcal{T} : \mathcal{V} \to \mathcal{U}$ is a linear map. Then the **matrix** of $\mathcal{T}$, with entries A_{ij} for $i = 1, \ldots, m$ and $j = 1, \ldots, n$ is defined by

$$\mathcal{T}(\mathbf{v}_j) = A_{1j}\mathbf{u}_1 + A_{2j}\mathbf{u}_2 + \cdots + A_{mj}\mathbf{u}_m,$$

i.e. the jth column of $\mathbf{A}$ is the coordinates of $\mathcal{T}(\mathbf{v}_j)$ in the chosen basis for $\mathcal{U}$.

- Conversely, every matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$ induces a linear map $\mathcal{T} : \mathbb{R}^n \to \mathbb{R}^m$ given by

$$\mathcal{T}(\mathbf{x}) = \mathbf{A}\mathbf{x}$$

The matrix of a linear map

- In the finite-dimensional setting, matrices are used to represent and manipulate **vectors** and **linear maps** on a computer

Definition (Matrix of a linear map)

Suppose $\mathbf{v}_1, \ldots, \mathbf{v}_n$ and $\mathbf{u}_1, \ldots, \mathbf{u}_n$ are bases of $\mathcal{V}$ and $\mathcal{U}$ respectively, and $\mathcal{T} : \mathcal{V} \to \mathcal{U}$ is a linear map. Then the **matrix** of $\mathcal{T}$, with entries A_{ij} for $i = 1, \ldots, m$ and $j = 1, \ldots, n$ is defined by

$$\mathcal{T}(\mathbf{v}_j) = A_{1j}\mathbf{u}_1 + A_{2j}\mathbf{u}_2 + \cdots + A_{mj}\mathbf{u}_m,$$

i.e. the jth column of $\mathbf{A}$ is the coordinates of $\mathcal{T}(\mathbf{v}_j)$ in the chosen basis for $\mathcal{U}$.

- Conversely, every matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$ induces a linear map $\mathcal{T} : \mathbb{R}^n \to \mathbb{R}^m$ given by

$$\mathcal{T}(\mathbf{x}) = \mathbf{A}\mathbf{x}$$

- And the matrix of this map with respect to the standard bases of $\mathbb{R}^n$ and $\mathbb{R}^m$ is simply $\mathbf{A}$

The matrix of a linear map

- If $\mathbf{A} \in \mathbb{R}^{m \times n}$, its transpose, $\mathbf{A}^T \in \mathbb{R}^{n \times m}$ is $\left(\mathbf{A}^T\right)_{ij} = A_{ji}$ for each (i, j)
- $\Rightarrow$ the columns of $\mathbf{A}$ become the **rows** of $\mathbf{A}^T$, and the rows of $\mathbf{A}$ become the **columns** of $\mathbf{A}^T$

The matrix of a linear map

- If $\mathbf{A} \in \mathbb{R}^{m\times n}$, its transpose, $\mathbf{A}^T \in \mathbb{R}^{n\times m}$ is $\left(\mathbf{A}^T\right)_{ij} = A_{ji}$ for each (i,j)
- $\Rightarrow$ the columns of $\mathbf{A}$ become the **rows** of $\mathbf{A}^T$, and the rows of $\mathbf{A}$ become the **columns** of $\mathbf{A}^T$
- Some nicealgebraic properties of the **transpose**:
  1. $\left(\mathbf{A}^T\right)^T = \mathbf{A}$
  2. $(\mathbf{A}+\mathbf{B})^T = \mathbf{A}^T + \mathbf{B}^T$
  3. $(\alpha\mathbf{A})^T = \alpha\mathbf{A}^T$
  4. $(\mathbf{AB})^T = \mathbf{B}^T\mathbf{A}^T$

Nullspace & range space

- Some of the most important subspaces are those induced by **linear maps**

Nullspace & range space

- Some of the most important subspaces are those induced by **linear maps**

Definition (Nullspace and range)

If $\mathcal{T} : \mathcal{V} \to \mathcal{U}$ is a linear map, we define the **nullspace** of $\mathcal{T}$ as

$$\text{null}(\mathcal{T}) = \{\mathbf{v} \in \mathcal{V} \mid \mathcal{T}(\mathbf{v}) = \mathbf{0}\},$$

and the **range** of $\mathcal{T}$ as

$$\text{range}(\mathcal{T}) = \{\mathbf{u} \in \mathcal{U} \mid \exists \mathbf{v} \in \mathcal{V} \text{ such that } \mathcal{T}(\mathbf{v}) = \mathbf{u}\}$$

Nullspace & range space

- Some of the most important subspaces are those induced by **linear maps**

Definition (Nullspace and range)

If $\mathcal{T} : \mathcal{V} \rightarrow \mathcal{U}$ is a linear map, we define the **nullspace** of $\mathcal{T}$ as

$$\text{null}(\mathcal{T}) = \{\mathbf{v} \in \mathcal{V} \mid \mathcal{T}(\mathbf{v}) = \mathbf{0}\},$$

and the **range** of $\mathcal{T}$ as

$$\text{range}(\mathcal{T}) = \{\mathbf{u} \in \mathcal{U} \mid \exists \mathbf{v} \in \mathcal{V} \text{ such that } \mathcal{T}(\mathbf{v}) = \mathbf{u}\}$$

- The nullspace and range of a linear map are always **subspaces** of its domain and codomain, respectively

Nullspace & range space

- Some of the most important subspaces are those induced by **linear maps**

Definition (Nullspace and range)

If $\mathcal{T} \ : \ \mathcal{V} \to \mathcal{U}$ is a linear map, we define the **nullspace** of $\mathcal{T}$ as

$$\text{null}(\mathcal{T}) = \{\mathbf{v} \in \mathcal{V} \mid \mathcal{T}(\mathbf{v}) = \mathbf{0}\},$$

and the **range** of $\mathcal{T}$ as

$$\text{range}(\mathcal{T}) = \{\mathbf{u} \in \mathcal{U} \mid \exists \mathbf{v} \in \mathcal{V} \text{ such that } \ \mathcal{T}(\mathbf{v}) = \mathbf{u}\}$$

- The nullspace and range of a linear map are always **subspaces** of its domain and codomain, respectively
- The **columnspace** of a matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$ is the span of its columns (considered as vectors in $\mathbb{R}^m$), denoted range$(\mathbf{A})$
- The **rowspace** of a matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$ is the span of its rows (considered as vectors in $\mathbb{R}^n$), denoted range $\left(\mathbf{A}^T\right)$

Nullspace & range space

- Some of the most important subspaces are those induced by **linear maps**

Definition (Nullspace and range)

If $\mathcal{T} : \mathcal{V} \to \mathcal{U}$ is a linear map, we define the **nullspace** of $\mathcal{T}$ as

$$\text{null}(\mathcal{T}) = \{\mathbf{v} \in \mathcal{V} \mid \mathcal{T}(\mathbf{v}) = \mathbf{0}\},$$

and the **range** of $\mathcal{T}$ as

$$\text{range}(\mathcal{T}) = \{\mathbf{u} \in \mathcal{U} \mid \exists \mathbf{v} \in \mathcal{V} \text{ such that } \mathcal{T}(\mathbf{v}) = \mathbf{u}\}$$

- The nullspace and range of a linear map are always **subspaces** of its domain and codomain, respectively
- The **columnspace** of a matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$ is the span of its columns (considered as vectors in $\mathbb{R}^m$), denoted range$(\mathbf{A})$
- The **rowspace** of a matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$ is the span of its rows (considered as vectors in $\mathbb{R}^n$), denoted range $\left(\mathbf{A}^T\right)$
- The **rank** satisfies: $\text{rank}(\mathbf{A}) = \mathbf{dim}\,(\text{range}(\mathbf{A})) = \mathbf{dim}\left(\text{range}\left(\mathbf{A}^T\right)\right)$

Metric spaces

Metric spaces

- Metrics generalize the notion of distance from Euclidean space (although metric spaces need not be vector spaces)

Metric spaces

- Metrics generalize the notion of distance from Euclidean space (although metric spaces need not be vector spaces)

Definition (Metric)

A metric on a set $\mathcal{X}$ is a function $d(\cdot,\cdot) : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ that $\forall \mathbf{x}, \mathbf{y} \in \mathcal{X}$ satisfies

(a) $d(\mathbf{x}, \mathbf{y}) \geq 0$ for all $\mathbf{x}$ and $\mathbf{y}$ (*nonnegativity*)

(b) $d(\mathbf{x}, \mathbf{y}) = 0$ if and only if $\mathbf{x} = \mathbf{y}$ (*definiteness*)

(c) $d(\mathbf{x}, \mathbf{y}) = d(\mathbf{y}, \mathbf{x})$ (*symmetry*)

(d) $d(\mathbf{x}, \mathbf{y}) \leq d(\mathbf{x}, \mathbf{z}) + d(\mathbf{z}, \mathbf{y})$ (*triangle inequality*)

Metric spaces

- Metrics generalize the notion of distance from Euclidean space (although metric spaces need not be vector spaces)

Definition (Metric)

A metric on a set $\mathcal{X}$ is a function $d(\cdot,\cdot) : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ that $\forall \mathbf{x}, \mathbf{y} \in \mathcal{X}$ satisfies

(a)	$d(\mathbf{x}, \mathbf{y}) \geq 0$ for all $\mathbf{x}$ and $\mathbf{y}$	(*nonnegativity*)
(b)	$d(\mathbf{x}, \mathbf{y}) = 0$ if and only if $\mathbf{x} = \mathbf{y}$	(*definiteness*)
(c)	$d(\mathbf{x}, \mathbf{y}) = d(\mathbf{y}, \mathbf{x})$	(*symmetry*)
(d)	$d(\mathbf{x}, \mathbf{y}) \leq d(\mathbf{x}, \mathbf{z}) + d(\mathbf{z}, \mathbf{y})$	(*triangle inequality*)

- A pseudo-metric satisfies (a), (c) and (d) but not necessarily (b)

Metric spaces

- Metrics generalize the notion of distance from Euclidean space (although metric spaces need not be vector spaces)

Definition (Metric)

A metric on a set $\mathcal{X}$ is a function $d(\cdot,\cdot) : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ that $\forall \mathbf{x}, \mathbf{y} \in \mathcal{X}$ satisfies

(a)	$d(\mathbf{x}, \mathbf{y}) \geq 0$ for all $\mathbf{x}$ and $\mathbf{y}$	(*nonnegativity*)
(b)	$d(\mathbf{x}, \mathbf{y}) = 0$ if and only if $\mathbf{x} = \mathbf{y}$	(*definiteness*)
(c)	$d(\mathbf{x}, \mathbf{y}) = d(\mathbf{y}, \mathbf{x})$	(*symmetry*)
(d)	$d(\mathbf{x}, \mathbf{y}) \leq d(\mathbf{x}, \mathbf{z}) + d(\mathbf{z}, \mathbf{y})$	(*triangle inequality*)

- Metrics allow limits to be defined for mathematical objects other than reals

Metric spaces

- Metrics generalize the notion of distance from Euclidean space (although metric spaces need not be vector spaces)

Definition (Metric)

A metric on a set $\mathcal{X}$ is a function $d(\cdot, \cdot) : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ that $\forall \mathbf{x}, \mathbf{y} \in \mathcal{X}$ satisfies

(a)	$d(\mathbf{x}, \mathbf{y}) \geq 0$ for all $\mathbf{x}$ and $\mathbf{y}$	(*nonnegativity*)
(b)	$d(\mathbf{x}, \mathbf{y}) = 0$ if and only if $\mathbf{x} = \mathbf{y}$	(*definiteness*)
(c)	$d(\mathbf{x}, \mathbf{y}) = d(\mathbf{y}, \mathbf{x})$	(*symmetry*)
(d)	$d(\mathbf{x}, \mathbf{y}) \leq d(\mathbf{x}, \mathbf{z}) + d(\mathbf{z}, \mathbf{y})$	(*triangle inequality*)

- Metrics allow limits to be defined for mathematical objects other than reals

Example

1. $\{x_n\} \subseteq \mathcal{S}$ converges to x if for any $\epsilon > 0,\ \exists N \in \mathbb{N} \mid d(x_n, x) < \epsilon\ \forall\ n \geq N$
2. Euclidean distance: $d_E(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|_2$
3. Loss/cost functions: (a) MSE, (b) MAE, (c) Cross Entropy, etc

Normed spaces

Normed spaces

- Norms generalize the notion of length from Euclidean space

Normed spaces

- Norms generalize the notion of length from Euclidean space

Definition (Vector norm)

A norm on a real vector space $\mathcal{V}$ is a function $\|\cdot\| \ : \ \mathcal{V} \to \mathbb{R}$ that for all $\mathbf{x}, \mathbf{y} \in \mathcal{V}$ and $\lambda \in \mathbb{R}$ satisfies

(a)	$\|\mathbf{x}\| \geq 0$	(*nonnegativity*)
(b)	$\|\mathbf{x}\| = 0$ if and only if $\mathbf{x} = \mathbf{0}$	(*definitiveness*)
(c)	$\|\lambda \mathbf{x}\| = \|\lambda\| \|\mathbf{x}\|$	(*homogeneiety*)
(d)	$\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$	(*triangle inequality*)

Normed spaces

- Norms generalize the notion of length from Euclidean space

Definition (Vector norm)

A norm on a real vector space $\mathcal{V}$ is a function $\|\cdot\| : \mathcal{V} \to \mathbb{R}$ that for all $\mathbf{x}, \mathbf{y} \in \mathcal{V}$ and $\lambda \in \mathbb{R}$ satisfies

(a)	$\|\mathbf{x}\| \geq 0$	(*nonnegativity*)
(b)	$\|\mathbf{x}\| = 0$ if and only if $\mathbf{x} = \mathbf{0}$	(*definitiveness*)
(c)	$\|\lambda \mathbf{x}\| = \lvert\lambda\rvert \|\mathbf{x}\|$	(*homogeniety*)
(d)	$\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$	(*triangle inequality*)

- A vector space endowed with a norm is called a **normed (vector) space**

Normed spaces

- Norms generalize the notion of length from Euclidean space

Definition (Vector norm)

A norm on a real vector space $\mathcal{V}$ is a function $\|\cdot\| \; : \; \mathcal{V} \to \mathbb{R}$ that for all $\mathbf{x}, \mathbf{y} \in \mathcal{V}$ and $\lambda \in \mathbb{R}$ satisfies

(a) $\|\mathbf{x}\| \geq 0$ (*nonnegativity*)
(b) $\|\mathbf{x}\| = 0$ if and only if $\mathbf{x} = \mathbf{0}$ (*definitiveness*)
(c) $\|\lambda \mathbf{x}\| = |\lambda| \|\mathbf{x}\|$ (*homogeniety*)
(d) $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$ (*triangle inequality*)

- A vector space endowed with a norm is called a **normed (vector) space**
- Note that any norm on $\mathcal{V}$ induces a distance metric on $\mathcal{V}$:

$$d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|$$

Normed spaces

- Norms generalize the notion of length from Euclidean space

Definition (Vector norm)

A norm on a real vector space $\mathcal{V}$ is a function $\|\cdot\| \;:\; \mathcal{V} \to \mathbb{R}$ that for all $\mathbf{x}, \mathbf{y} \in \mathcal{V}$ and $\lambda \in \mathbb{R}$ satisfies

(a)	$\|\mathbf{x}\| \geq 0$	(*nonnegativity*)
(b)	$\|\mathbf{x}\| = 0$ if and only if $\mathbf{x} = \mathbf{0}$	(*definitiveness*)
(c)	$\|\lambda\mathbf{x}\| = \vert\lambda\vert\|\mathbf{x}\|$	(*homogeniety*)
(d)	$\|\mathbf{x}+\mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$	(*triangle inequality*)

- A vector space endowed with a norm is called a **normed (vector) space**
- Note that any norm on $\mathcal{V}$ induces a distance metric on $\mathcal{V}$:
$$d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|$$
- This satisfies axioms for metrics, hence any **normed space** is also a **metric space** (since axioms for norms $\Rightarrow$ axioms for metrics)

Vector Norms

- This course will focus on a few specific norms on $\mathbb{R}^n$, i.e. p-norms

Vector Norms

- This course will focus on a few specific norms on $\mathbb{R}^n$, i.e. p-norms
- For $\mathbf{x} \in \mathbb{R}^n$, the p-norm is defined as $\|\mathbf{x}\|_p := \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p}$

Vector Norms

- This course will focus on a few specific norms on $\mathbb{R}^n$, i.e. p-norms
- For $\mathbf{x} \in \mathbb{R}^n$, the p-norm is defined as $\|\mathbf{x}\|_p := \left(\sum_{i=1}^n |x_i|^p\right)^{1/p}$

Example

(1)	ℓ_2-norm:	$\|\mathbf{x}\|_2 := \sqrt{\sum_{i=1}^p x_i^2}$	(Euclidean norm)
(2)	ℓ_1-norm:	$\|\mathbf{x}\|_1 := \sum_{i=1}^p \lvert x_i \rvert$	(Manhattan norm)
(3)	ℓ_∞-norm:	$\|\mathbf{x}\|_\infty := \max_{i=1,\ldots,p} \lvert x_i \rvert$	(Chebyshev norm)

Vector Norms

- This course will focus on a few specific norms on $\mathbb{R}^n$, i.e. p-norms
- For $\mathbf{x} \in \mathbb{R}^n$, the p-norm is defined as $\|\mathbf{x}\|_p := \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p}$

Example

(1)	ℓ_2-norm:	$\|\mathbf{x}\|_2 := \sqrt{\sum_{i=1}^{p} x_i^2}$	(Euclidean norm)
(2)	ℓ_1-norm:	$\|\mathbf{x}\|_1 := \sum_{i=1}^{p} \|x_i\|$	(Manhattan norm)
(3)	ℓ_∞-norm:	$\|\mathbf{x}\|_\infty := \max_{i=1,\ldots,p} \|x_i\|$	(Chebyshev norm)

- $p \geq 1$ is required for the general definition of the p-norm because the triangle inequality fails to hold if $p < 1$

Vector Norms

- This course will focus on a few specific norms on $\mathbb{R}^n$, i.e. p-norms
- For $\mathbf{x} \in \mathbb{R}^n$, the p-norm is defined as $\|\mathbf{x}\|_p := \left(\sum_{i=1}^n |x_i|^p\right)^{1/p}$

Example

(1)	ℓ_2-norm:	$\|\mathbf{x}\|_2 := \sqrt{\sum_{i=1}^p x_i^2}$	(Euclidean norm)
(2)	ℓ_1-norm:	$\|\mathbf{x}\|_1 := \sum_{i=1}^p \|x_i\|$	(Manhattan norm)
(3)	ℓ_∞-norm:	$\|\mathbf{x}\|_\infty := \max_{i=1,\dots,p} \|x_i\|$	(Chebyshev norm)

- $p \geq 1$ is required for the general definition of the p-norm because the triangle inequality fails to hold if $p < 1$
- In a finite-dimensional vector space $\mathcal{V}$, all norms on $\mathcal{V}$ are equivalent, i.e. for two norms $\|\cdot\|_A$ and $\|\cdot\|_B$ there exist constants $\alpha, \beta > 0$ such that

$$\alpha\|\mathbf{x}\|_A \leq \|\mathbf{x}\|_B \leq \beta\|\mathbf{x}\|_A \quad \text{for all } \mathbf{x} \in \mathcal{V}$$

Vector norms contd.

Definition (Quasi-norm)

A quasi-norm satisfies all the norm properties except (d) triangle inequality, which is replaced by $\|\mathbf{x}+\mathbf{y}\| \leq c\,(\|\mathbf{x}\|+\|\mathbf{y}\|)$ for a constant $c \geq 1$.

Vector norms contd.

Definition (Quasi-norm)

A quasi-norm satisfies all the norm properties except (d) triangle inequality, which is replaced by $\|\mathbf{x}+\mathbf{y}\| \leq c\,(\|\mathbf{x}\|+\|\mathbf{y}\|)$ for a constant $c \geq 1$.

Definition (Semi(pseudo)-norm)

A semi(pseudo)-norm satisfies all the norm properties except (b) definiteness.

Vector norms contd.

Definition (Quasi-norm)

A quasi-norm satisfies all the norm properties except (d) triangle inequality, which is replaced by $\|\mathbf{x}+\mathbf{y}\| \leq c\,(\|\mathbf{x}\| + \|\mathbf{y}\|)$ for a constant $c \geq 1$.

Definition (Semi(pseudo)-norm)

A semi(pseudo)-norm satisfies all the norm properties except (b) definiteness.

Example

- The ℓ_q-norm becomes a quasi-norm when $q \in (0,1)$ with $c = 2^{1/q} - 1$
- The **total variation norm** (TV-norm) defined (in 1D): $\|\mathbf{x}\|_{\mathrm{TV}} := \sum_{i=1}^{p-1} |x_{i+1} - x_i|$ is a **semi-norm** since it fails to satisfy (b); e.g., $\mathbf{x} = (1,1,\ldots,1)^T$ has $\|\mathbf{x}\|_{\mathrm{TV}} = 0$ even though $\mathbf{x} \neq \mathbf{0}$.

Vector norms contd.

Definition (“ℓ_0-norm”)

$\|\mathbf{x}\|_0 = |\{i : x_i \neq 0\}| \quad \equiv \lim_{q\to 0}\|\mathbf{x}\|_q^q$

Vector norms contd.

Definition ("ℓ_0-norm")

$\|\mathbf{x}\|_0 = |\{i : x_i \neq 0\}| \quad \equiv \lim_{q\to 0} \|\mathbf{x}\|_q^q$

- The "ℓ_0-norm" counts the non-zero components of $\mathbf{x}$

Vector norms contd.

Definition (“ℓ_0-norm”)

$\|\mathbf{x}\|_0 = |\{i : x_i \neq 0\}| \quad \equiv \lim_{q \to 0} \|\mathbf{x}\|_q^q$

- The “ℓ_0-norm” counts the non-zero components of $\mathbf{x}$
- It is not a norm – it does not satisfy norm properties (c) and (d)
- $\Rightarrow$ it is also neither a **quasi**- nor a **semi-norm**

Vector norms contd.

Definition ("ℓ_0-norm")

$\|\mathbf{x}\|_0 = |\{i : x_i \neq 0\}| \quad \equiv \lim_{q\to 0}\|\mathbf{x}\|_q^q$

- The "ℓ_0-norm" counts the non-zero components of $\mathbf{x}$
- It is not a norm – it does not satisfy norm properties (c) and (d)
- $\Rightarrow$ it is also neither a **quasi**- nor a **semi-norm**

Example (Compressed sensing)

Find $\underset{\mathbf{x}\in\mathbb{R}^p}{\arg\min} \; \|\mathbf{x}\|_0$ subject to: $\mathbf{A}\mathbf{x} = \mathbf{y}$.

Vector norms contd.

Definition ("ℓ_0-norm")

$\|\mathbf{x}\|_0 = |\{i : x_i \neq 0\}| \quad \equiv \lim_{q\to 0}\|\mathbf{x}\|_q^q$

- The "ℓ_0-norm" counts the non-zero components of $\mathbf{x}$
- It is not a norm – it does not satisfy norm properties (c) and (d)
- $\Rightarrow$ it is also neither a **quasi**- nor a **semi-norm**

Example (Compressed sensing)

Find $\quad \underset{\mathbf{x}\in\mathbb{R}^p}{\arg\min}\ \|\mathbf{x}\|_0$ subject to: $\mathbf{A}\mathbf{x} = \mathbf{y}$.

Example (Sparse approximation)

Find $\quad \underset{\mathbf{x}\in\mathbb{R}^p}{\arg\min}\ \|\mathbf{A}\mathbf{x} - \mathbf{y}\|_2^2$ subject to: $\|\mathbf{x}\|_0 \leq s$.

Vector norms contd.

Definition (Dual norm)

Let $\|\cdot\|$ be a norm in $\mathbb{R}^p$, then the **dual norm** denoted by $\|\cdot\|^*$ is defined:

$$\|\mathbf{x}\|^* = \sup_{\|\mathbf{y}\|\leq 1} \mathbf{x}^T\mathbf{y}, \quad \text{for all } \mathbf{x}, \mathbf{y} \in \mathbb{R}^p$$

Vector norms contd.

Definition (Dual norm)

Let $\|\cdot\|$ be a norm in $\mathbb{R}^p$, then the **dual norm** denoted by $\|\cdot\|^*$ is defined:

$$\|\mathbf{x}\|^* = \sup_{\|\mathbf{y}\|\leq 1} \mathbf{x}^T\mathbf{y}, \quad \text{for all } \mathbf{x}, \mathbf{y} \in \mathbb{R}^p$$

- The **dual** of the *dual norm* is the original (primal) norm, i.e., $\|\mathbf{x}\|^{**} = \|\mathbf{x}\|$

Vector norms contd.

Definition (Dual norm)

Let $\| \cdot \|$ be a norm in $\mathbb{R}^p$, then the **dual norm** denoted by $\| \cdot \|^*$ is defined:

$$\|\mathbf{x}\|^* = \sup_{\|\mathbf{y}\| \leq 1} \mathbf{x}^T \mathbf{y}, \quad \text{for all } \mathbf{x}, \mathbf{y} \in \mathbb{R}^p$$

- The **dual** of the *dual norm* is the original (primal) norm, i.e., $\|\mathbf{x}\|^{**} = \|\mathbf{x}\|$
- Hölder's inequality $\Rightarrow$ $\| \cdot \|_q$ is a dual norm of $\| \cdot \|_r$ when $\frac{1}{q} + \frac{1}{r} = 1$.

Vector norms contd.

Example 1

i) $\|\cdot\|_2$ is **dual** of $\|\cdot\|_2$ (i.e., $\|\cdot\|_2$ is *self-dual*):
$\sup\{\mathbf{z}^T\mathbf{x} \mid \|\mathbf{x}\|_2 \leq 1\} = \|\mathbf{z}\|_2$.

ii) $\|\cdot\|_1$ is **dual** of $\|\cdot\|_\infty$, (and *vice versa*): $\sup\{\mathbf{z}^T\mathbf{x} \mid \|\mathbf{x}\|_\infty \leq 1\} = \|\mathbf{z}\|_1$.

Vector norms contd.

Example 1

i) $\|\cdot\|_2$ is **dual** of $\|\cdot\|_2$ (i.e., $\|\cdot\|_2$ is *self-dual*):
$\sup\{\mathbf{z}^T\mathbf{x} \mid \|\mathbf{x}\|_2 \leq 1\} = \|\mathbf{z}\|_2$.

ii) $\|\cdot\|_1$ is **dual** of $\|\cdot\|_\infty$, (and *vice versa*): $\sup\{\mathbf{z}^T\mathbf{x} \mid \|\mathbf{x}\|_\infty \leq 1\} = \|\mathbf{z}\|_1$.

Example 2

What is the **dual norm** of $\|\cdot\|_q$ for $q = 1 + 1/\log(p)$?

Vector norms contd.

Example 1

i) $\|\cdot\|_2$ is **dual** of $\|\cdot\|_2$ (i.e., $\|\cdot\|_2$ is *self-dual*):
$\sup\{\mathbf{z}^T\mathbf{x} \mid \|\mathbf{x}\|_2 \le 1\} = \|\mathbf{z}\|_2$.

ii) $\|\cdot\|_1$ is **dual** of $\|\cdot\|_\infty$, (and *vice versa*): $\sup\{\mathbf{z}^T\mathbf{x} \mid \|\mathbf{x}\|_\infty \le 1\} = \|\mathbf{z}\|_1$.

Example 2

What is the **dual norm** of $\|\cdot\|_q$ for $q = 1 + 1/\log(p)$?

Solution

By Hölder's inequality, $\|\cdot\|_r$ is the **dual norm** of $\|\cdot\|_q$ if $\frac{1}{q} + \frac{1}{r} = 1$. Therefore, $r = 1 + \log(p)$ for $q = 1 + 1/\log(p)$.

Inner product space

Inner product space

Definition (Inner product)

An **inner product** on a real vector space $\mathcal{V}$ is a function $\langle \cdot, \cdot\cdot \rangle \ : \ \mathcal{V} \times \mathcal{V} \to \mathbb{R}$ such that for all $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathcal{V}$ and $\alpha, \beta \in \mathbb{R}$ it satisfies

1. $\langle \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{y}, \mathbf{x} \rangle,$ (symmetry)
2. $\langle (\alpha \mathbf{x} + \beta \mathbf{y}), \mathbf{z} \rangle = \alpha \langle \mathbf{x}, \mathbf{z} \rangle + \beta \langle \mathbf{y}, \mathbf{z} \rangle,$ (linearity)
3. $\langle \mathbf{x}, \mathbf{x} \rangle \geq 0$ (positive definiteness)

Inner product space

Definition (Inner product)

An **inner product** on a real vector space $\mathcal{V}$ is a function $\langle \cdot, \cdot\cdot \rangle \; : \; \mathcal{V} \times \mathcal{V} \to \mathbb{R}$ such that for all $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathcal{V}$ and $\alpha, \beta \in \mathbb{R}$ it satisfies

1. $\langle \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{y}, \mathbf{x} \rangle,$ (symmetry)
2. $\langle (\alpha\mathbf{x} + \beta\mathbf{y}), \mathbf{z} \rangle = \alpha\langle \mathbf{x}, \mathbf{z} \rangle + \beta\langle \mathbf{y}, \mathbf{z} \rangle,$ (linearity)
3. $\langle \mathbf{x}, \mathbf{x} \rangle \geq 0$ (positive definiteness)

- The **inner product** (also dot product) of any two vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ is defined as $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^T\mathbf{y} = \sum_i^n x_i y_i$

Inner product space

Definition (Inner product)

An **inner product** on a real vector space $\mathcal{V}$ is a function $\langle \cdot, \cdot\cdot \rangle \ : \ \mathcal{V} \times \mathcal{V} \to \mathbb{R}$ such that for all $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathcal{V}$ and $\alpha, \beta \in \mathbb{R}$ it satisfies

1. $\langle \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{y}, \mathbf{x} \rangle,$ (symmetry)
2. $\langle (\alpha \mathbf{x} + \beta \mathbf{y}), \mathbf{z} \rangle = \alpha \langle \mathbf{x}, \mathbf{z} \rangle + \beta \langle \mathbf{y}, \mathbf{z} \rangle,$ (linearity)
3. $\langle \mathbf{x}, \mathbf{x} \rangle \geq 0$ (positive definiteness)

- The **inner product** (also dot product) of any two vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ is defined as $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^T \mathbf{y} = \sum_i^n x_i y_i$
- Important relations involving the inner product:
  - Hölder's inequality: $|\langle \mathbf{x}, \mathbf{y} \rangle| \leq \|\mathbf{x}\|_q \|\mathbf{y}\|_r$, where $r > 1$ and $\frac{1}{q} + \frac{1}{r} = 1$
  - Cauchy-Schwarz is a special case of Hölder's inequality $(q = r = 2)$