2: Vector space geometry

- norms
- inner products
- orthogonality

Transpose operation

Real vectors

- let x be a real n-vector, i.e., $x \in \mathbb{R}^n$
- the **transpose** operator turns a column-vector into a row-vector (and vice-versa)

$$x = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \qquad x^T = \begin{pmatrix} 1 & 2 & 3 \end{pmatrix}$$

- the set of row-vectors with n elements is denoted by $\mathbb{R}^{1\times n}$

Complex vectors

- for complex n-vectors, the **conjugate transpose** is the analogous operation
- transpose and then conjugate ($\bar{y}$) each element

$$z = \begin{pmatrix} 2 + 7\mathbf{i} \\ -1 - 4\mathbf{i} \end{pmatrix}, \qquad z^* = \begin{pmatrix} 2 - 7\mathbf{i} & -1 + 4\mathbf{i} \end{pmatrix}$$

Vector norms

- the **Euclidean** norm of a vector $z \in \mathbb{R}^n$

$$\|z\|_2 = \sqrt{z^T z} = \sqrt{\langle z, z \rangle}$$

- $\|z\|_2$ measures the length of a vector (from the origin)

- $\|z - x\|_2$ measures the distance between vectors z and x, often denoted

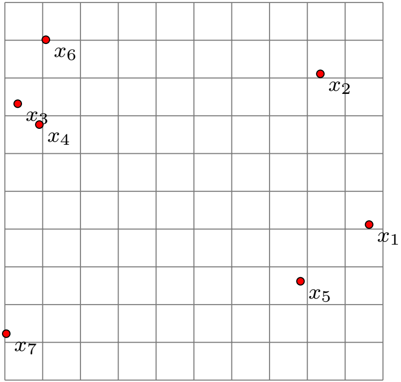
$$\mathbf{dist}(z, x) = \|z - x\|_2$$

- the Euclidean norm is **induced** by the **inner product** on $\mathbb{R}^n$

- extends to a norm on $\mathbb{C}^n$ via conjugate transposition

$$\|w\|_2 = \sqrt{w^* w}$$

Nearest neighbor



given the set of vectors $\mathcal{X} = \{x_1, \ldots, x_7\}$, x_j is the nearest neighbor of $z \in \mathcal{X}$ if

$$\|z - x_j\| \leq \|z - x_k\| \text{ for all } k = 1, \ldots, 7$$

Beyond $\|\cdot\|_2$

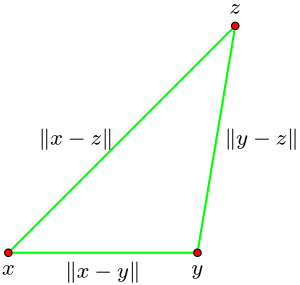
- a **norm** is **any** function $f : V \to \mathbb{R}$ that satisfies
  - **Positive-definiteness:** $f(x) \geq 0$ for all $x \in V$ with $f(x) = 0$ if and only if $x = 0$
  - **Triangle Inequality:** $f(x + y) \leq f(x) + f(y)$ for all $x, y \in V$
  - **Homogeneity:** $f(\lambda x) = |\lambda| f(x)$ for all $\lambda \in \mathbb{C}$ and $x \in V$

  when satisfied, we replace f with $\|\cdot\|$

- a norm that is derived from an inner product is called an **induced norm**

- together a vector space and a norm are called a **normed vector space** $(V, \|\cdot\|)$

Triangle inequality



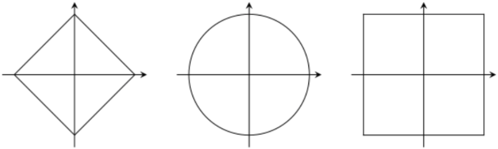
- looking at the figure, it is clear that $\|x - z\| \le \|x - y\| + \|y - z\|$
- the triangle inequality tells us

$$\|x - z\| = \|(x - y) + (y - z)\| \le \|(x - y)\| + \|(y - z)\|$$

Non-Euclidean norms

ℓ_p-norms

$$\|x\|_p = (|x_1|^p + |x_2|^p + \ldots + |x_n|^p)^{\frac{1}{p}}, \quad p \geq 1$$



$p = 1$ $p = 2$ $p = \infty$

- $\|x\|_1 = |x_1| + |x_2| + \ldots |x_n|$
- $\|x\|_2 = |x_1|^2 + |x_2|^2 + \ldots |x_n|^2$
- $\|x\|_\infty = \max \; \{|x_1|, |x_2|, \ldots, |x_n|\}$

Inner product on $\mathbb{R}^n$

- given $x, y \in \mathbb{R}^n$, the **inner product** (or dot product) is

$$x^T y = \sum_{i=1}^{n} x_i y_i$$

- often written as $\langle x, y \rangle$

- later we will generalize to vector fields beyond $\mathbb{R}^n$ and use $\langle \cdot, \cdot \rangle_V$

Use cases

- **summation:** to sum the elements in x, take an inner product with the $\mathbf{1}$ vector

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

Use cases

- **summation:** to sum the elements in x, take an inner product with the $\mathbf{1}$ vector

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

- **average:** to compute the average of all n elements

$$\mathbf{ave}(x) = \left(\frac{\mathbf{1}}{n}\right)^T x = (x_1 + x_2 + \cdots + x_n)/n$$

Use cases

- **summation:** to sum the elements in x, take an inner product with the $\mathbf{1}$ vector

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

- **average:** to compute the average of all n elements

$$\mathbf{ave}(x) = \left(\frac{\mathbf{1}}{n}\right)^T x = (x_1 + x_2 + \cdots + x_n)/n$$

- **sum of squares:** the inner product of a vector with itself

$$x^T x = x_1^2 + \ldots, + x_n^2$$

Use cases

- **summation:** to sum the elements in x, take an inner product with the $\mathbf{1}$ vector

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

- **average:** to compute the average of all n elements

$$\mathbf{ave}(x) = \left(\frac{\mathbf{1}}{n}\right)^T x = (x_1 + x_2 + \cdots + x_n)/n$$

- **sum of squares:** the inner product of a vector with itself

$$x^T x = x_1^2 + \ldots, + x_n^2$$

- **pick element i:** retrieve the i^{th}-element of x

$$e_i^T x = x_i$$

Use cases

- **summation:** to sum the elements in x, take an inner product with the $\mathbf{1}$ vector

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

- **average:** to compute the average of all n elements

$$\mathbf{ave}(x) = \left(\frac{\mathbf{1}}{n}\right)^T x = (x_1 + x_2 + \cdots + x_n)/n$$

- **sum of squares:** the inner product of a vector with itself

$$x^T x = x_1^2 + \ldots, + x_n^2$$

- **pick element i:** retrieve the i^{th}-element of x

$$e_i^T x = x_i$$

- **differencing:** $(e_i - e_j)^T x = x_i - x_j$

Inner products

Given a **vector space** V, an **inner product** is a function $f : V \times V \to \mathbb{R}$, denoted by $\langle \cdot, \cdot \rangle$, that satisfies the following four conditions:

1. **Conjugate symmetry:** $\langle x, y \rangle = \overline{\langle y, x \rangle}$
2. **Positive definiteness:** $\langle x, x \rangle \geq 0$ with $\langle x, x \rangle = 0$ if and only if $x = 0$
3. **Linearity (in the second argument):** $\langle x, \alpha y + \beta z \rangle = \alpha \langle x, y \rangle + \beta \langle x, z \rangle$

Notes

- some communities use the opposite convention
- V and $\langle \cdot, \cdot \rangle$ together form an **inner product space**

Examples on $\mathbb{R}^n$

1. $\langle x, y\rangle = x_1y_1 + \cdots + x_ny_n$

2. $\langle x, y\rangle = \gamma_1x_1y_1 + \cdots + \gamma_nx_ny_n$ where $\gamma_i \geq 0$

3. $\langle x, y\rangle = \langle Dx, Dy\rangle$ where D is an invertible matrix and $\langle\cdot,\cdot\rangle$ is defined as in 1

Beyond finite-dimensional vector spaces

- if $V = C([0, 1])$, i.e., vector space of continuous functions on $[0, 1]$, then

$$\langle f, g\rangle = \int_0^1 f(x)g(x)\mathrm{d}x$$

  defines an inner product

Inner products induce norms

given any inner product $\langle x, y \rangle$, a norm can be defined

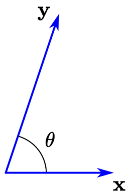
$$\|x\| = \sqrt{\langle x, x \rangle}$$

- not all norms are constructed this way
- $\|x\|_2$ clearly is
- there is no inner product associated to $\|\cdot\|_\infty$ or $\|\cdot\|_1$
- an **inner product space** is a vector space V over F with an inner product

Relationship to angle

the angle, θ, between two non-zero vectors, x and y is

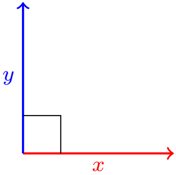
$$\theta = \arccos\left(\frac{x^T y}{\|x\|_2 \|y\|_2}\right)$$



- $x^T y = \|x\|_2 \|y\|_2 \cos\theta$
- by default, θ is measured in radians, i.e., $360^\circ = 2\pi$ radians
- $\|x\|_2 = \sqrt{x_1^2 + x_2^2}$ – more on norms later

Orthogonality

vectors x and y are said to be orthogonal if $\langle x, y \rangle = 0$



- this happens when $\cos\theta = 0$
- often written as $x \perp y$
- orthogonality refers to a pair (or set) of vectors - not an individual vector
- a set of vectors $\{v_1, v_2, \ldots, v_k\}$ are **orthonormal** if

$$v_i \perp v_j = 0 \quad \text{for all } i \neq j \quad \text{and} \quad \|v_i\|_2 = 1 \quad \text{for } i = 1, \ldots, k$$

Cauchy-Schwarz inequality

the Cauchy-Schwarz inequality on $\mathbb{R}^n$ relates inner products to norms

- for any $x, y \in \mathbb{R}^n$

$$|x^T y| \leq \|x\|_2 \|y\|_2$$

- from the properties that define an inner product, this implies

$$0 \leq \frac{|x^T y|}{\|x\|_2 \|y\|_2} \leq 1$$

- for general inner product spaces $(V, \langle \cdot, \cdot \rangle)$

$$|\langle u, v \rangle| \leq \|u\| \|v\|$$

Angles between vectors

- if $\theta = 0$ then x and y are **aligned**

$$x^T y = \|x\|_2 \|y\|_2 \quad \Longleftrightarrow \quad y = \alpha x, \quad \text{for some } \alpha \geq 0 \quad (x \neq 0)$$

- if $\theta = \pi$ then x and y are **opposed**

$$x^T y = -\|x\|_2 \|y\|_2 \quad \Longleftrightarrow \quad y = -\alpha x, \quad \text{for some } \alpha \geq 0 \quad (x \neq 0)$$

- if $\theta = \frac{\pi}{2}$ then x and y are **orthogonal**: $x^T y = 0$

Sign of an inner product

- $x^T y > 0 \iff \theta$ is **acute**

- $x^T y < 0 \iff \theta$ is **obtuse**

Proof of Cauchy-Schwarz inequality

we will now prove that $|x^T y| \leq \|x\|_2 \|y\|_2$ for any $x, y \in \mathbb{R}^n$

Proof

if $x = 0$ or $y = 0$ result is immediate

assume that $\|x\|_2 = \|y\|_2 = 1$, then

$$\|x - y\|_2^2 = x^T x + y^T y - 2x^T y \geq 0 \quad \Longrightarrow \quad x^T y \leq 1$$

for any x, y we have that

$$\left\langle \frac{x}{\|x\|_2}, \frac{y}{\|y\|_2} \right\rangle \leq 1 \quad \Longrightarrow \quad x^T y \leq \|x\|_2 \|y\|_2$$

to show $-1 \leq x^T y$, repeat above arguments for $\|x + y\|_2^2$:

$$\|x + y\|_2^2 = x^T x + y^T y + 2x^T y \geq 0.$$

$\square$

Triangle inequality

- the Cauchy-Schwarz inequality can be used to verify the **triangle inequality**:

$$\begin{aligned}\|x+y\|_2^2 &= \|x\|_2^2 + 2x^Ty + \|y\|_2^2 \\ &\leq \|x\|_2^2 + 2\|x\|_2\|y\|_2 + \|y\|_2^2 \\ &= (\|x\|_2 + \|y\|_2)^2\end{aligned}$$

$$\Longrightarrow \quad \|x+y\|_2 \leq \|x\|_2 + \|y\|_2$$

Further topics

- equivalence of norms
- Pythagorean theorem
- parallelogram law
- inner products and norms for matrices