Linear Algebra

Euclidian Norm, Euclidian Distance, & Angle

Department of Computer Engineering

Sharif University of Technology

Hamid R. Rabiee rabiee@sharif.edu
Maryam Ramezani maryam.ramezani@sharif.edu

Table of contents

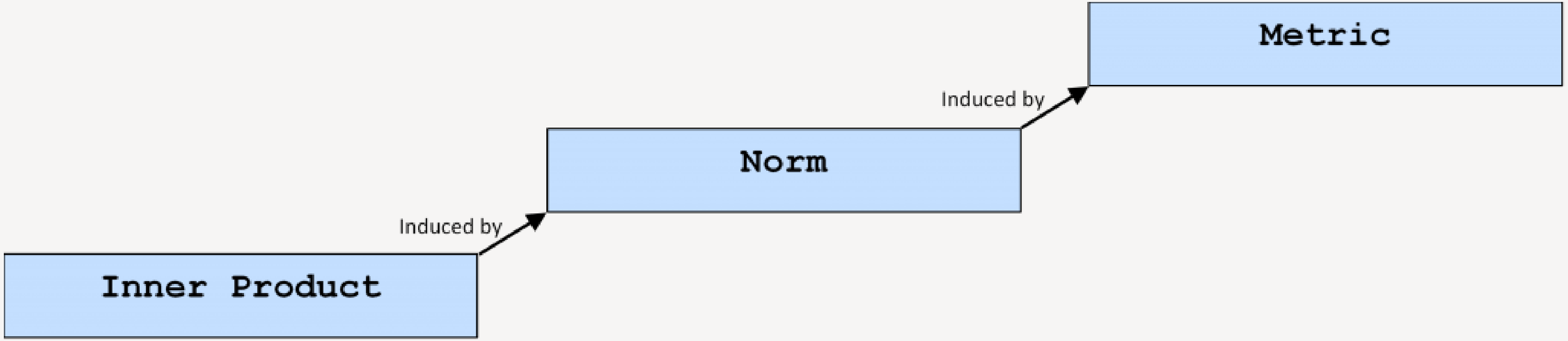
01

Introduction

The reason to use norms

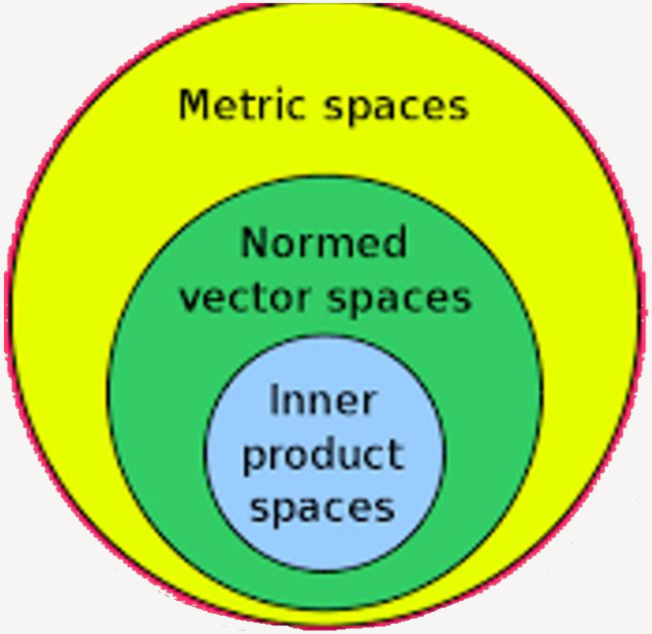
- Machine learning uses vectors, matrices, and tensors as the basic units of representation
- Two reasons to use norms:
  - To estimate how **big** a vector/matrix/tensor is
    - How big is the difference between two tensors is
  - To estimate how **close** one tensor is to another
    - How close is one image to another

Inner Products, Norms and Metrics



Inner Products, Norms and Metrics

- Given an inner product <A , B>, one can obtain a norm doing $|| A ||^2 = <A , A>$
- And given a norm || A ||, one can obtain a metric using the difference vector || A - B||



Inner Products, Norms and Metrics

Vector space	Generalization
metric	metric space
norm	normed
scalar product	inner product space

Euclidean Norm

Definition

Functions closely related to inner products are so-called **norms**. Norms are specific functions that can be interpreted as a distance function between a vector and the origin.

Definition

For $v \in V$, we define the eculidean norm of v, denoted $||v||$, by:

$$||v|| = \sqrt{< v, v >}$$

Euclidean Norm

Note

- Euclidean Norm (2-norm, l_2 norm, length)

$$||x|| = \sqrt{x_1^2 + x_2^2 + \ldots + x_n^2} = \sqrt{x^T x}$$

  - A vector whose length is 1 is called a unit vector
  - **Normalizing**: divide a non-zero vector by its length which is a unit vector in the same direction of original vector
  - It is a nonnegative scalar
  - In $\mathbb{R}^2$ follows from the Pythagorean Theorem.
  - What about $\mathbb{R}^3$?
  - What is the shape of $||x||_2 = 1$?

Euclidean Norm

Example

Norm of $P_n(x)$ in the term of inner product $\langle p_n(x), q_n(x)\rangle = \int_0^1 p_n(x)q_n(x)dx$:

$$||P_n(x)|| = \sqrt{\int_0^1 P_n^2(x)dx}$$

02

Inequalities

Chebyshev Inequality

Theorem 1

Suppose that k of the numbers $|x_1|, |x_2|, \dots, |x_n|$ are $\geq a$ then k of the numbers $x_1^2, x_2^2, \dots, x_n^2\ are \geq a^2$

So $\left||x|\right|^2 = x_1^2 + x_2^2 + \ \dots + x_n^2 \geq ka^2\ so\ we\ have\ k \leq \frac{\left||x|\right|^2}{a^2}$

Number of x_i with $|x_i| \geq a$ is no more than $\frac{\left||x|\right|^2}{a^2}$

Question

- What happens when $\frac{\left||x|\right|^2}{a^2} \geq n$?
- No entry of a vector can be larger in magnitude than the norm of the vector. (why?)

Answers

- **When** $\frac{||x||^2}{a^2} \geq n$:

  The bound $k \leq \frac{||x||^2}{a^2}$ becomes **non-informative**, because we already know $k \leq n$.

- **Why no entry can exceed the norm (**$| x_i | \leq \| x \|_2$**):**

  Since $\| x \|_2^2 = \sum_{j=1}^{n} x_j^2 \geq x_i^2$,taking square roots gives $\| x \|_2 \geq | x_i |$.

Cauchy-Schwartz Inequality

Theorem 2

For two n-vectors $a\ and\ b$, $|a^T b| \le\ ||a||\ ||b||$

Written out:

$$|a_1 b_1 + \cdots + a_n b_n| \le (a_1^2 + \ \ldots + a_n^2)^{\frac{1}{2}} (b_1^2 + \ \ldots + b_n^2)^{\frac{1}{2}}$$

$$\left(\sum_{i=1}^{n} x_i y_i\right)^2 \le \left(\sum_{i=1}^{n} x_i^2\right)\left(\sum_{i=1}^{n} y_i^2\right)$$

Cauchy-Schwartz Inequality - Proof

It is clearly true if either $a\ or\ b\ is\ 0$.
So, assume $\alpha = ||a||$ and $\beta = ||b||$ are non-zero
We have

$$0 \leq \left||\beta a - \alpha b|\right|^2$$
$$= \left||\beta a|\right|^2 - 2\,(\beta a)^T(\alpha b) + \left||\alpha b|\right|^2$$
$$= \beta^2\left||a|\right|^2 - 2\,\beta\alpha(a^T b) + \alpha^2\left||b|\right|^2$$
$$= 2\left||a|\right|^2\left||b|\right|^2 - 2\left||a|\right|\left||b|\right|(a^T b)$$

Divide by 2 $\left||a|\right|\left||b|\right|$ to get $a^T b \leq \left||a|\right|\left||b|\right|$
Apply to $-a, b$ to get other half of Cauchy-Schwartz inequality.

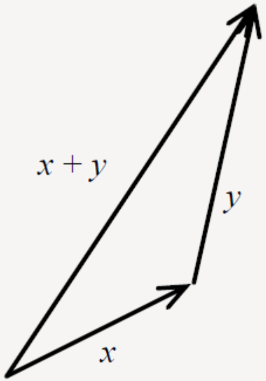
Cauchy-Schwarz inequality holds with equality when one of the vectors is a multiple of the other
If and only if a and b are linear dependent

Triangle Inequality

Theorem 3

Consider a triangle in two or three dimensions:

$$||x+y|| \leq ||x|| + ||y||$$



Verification of triangle inequality:

$$||\boldsymbol{x}+\boldsymbol{y}||^2 = ||\boldsymbol{x}||^2 + ||\boldsymbol{y}||^2 + \mathbf{2}\ \underline{\boldsymbol{x}^T\boldsymbol{y}}$$

$$\leq ||\boldsymbol{x}||^2 + ||\boldsymbol{y}||^2 + \mathbf{2}\ \underline{||\boldsymbol{x}||\,||\boldsymbol{y}||}$$

Cauchy-Schwartz Inequality

$$= \left(||\boldsymbol{x}|| + ||\boldsymbol{y}||\right)^2$$

$$\Rightarrow ||\boldsymbol{x}+\boldsymbol{y}|| \leq ||\boldsymbol{x}|| + ||\boldsymbol{y}||$$

03

Euclidean Norm

Vector Norm Properties

Important Properties

1. Absolute Homogenity / Linearity:

$$||\alpha x|| = |\alpha|\ ||x||$$

2. Subadditivity / Triangle Inequality:

$$||x + y|| \leq ||x|| + ||y||$$

3. Positive definiteness / Point separating:

$$if\ ||x|| = 0\ then\ x = 0$$

$$(from\ 1\ \&\ 3): For\ every\ x, ||x|| = 0\ iff\ x = 0$$

4. Non-negativity:

$$||x|| \geq 0$$

Norm of sum

Theorem 4

If x and y are vectors:

$$||x+y|| = \sqrt{||x||^2 + 2\,x^T y + ||y||^2}$$

Proof:

$$||x+y||^2 = (x+y)^T(x+y)$$

$$= x^T x + x^T y + y^T x + y^T y$$

$$= ||x||^2 + 2\,x^T y + ||y||^2$$

Inner product and norm

Theorem 5

Take any inner product $\langle \cdot,\cdot \rangle$ and define $f(x) = \sqrt{\langle x, x \rangle}$. Then f is a norm.

Note

Every inner product gives rise to a norm, but not every norm comes from an inner product. (Think about norm 2 and norm max)

Proof

- **Positive definiteness**

  Inner product properties give:

  $\langle x, x\rangle \geq 0$ for all x, so $f(x) = \sqrt{\langle x, x\rangle} \geq 0.$

  $\langle x, x\rangle = 0$ iff $x = 0.$

  Hence $f(x) = 0 \iff x = 0.$

- **Homogeneity (absolute scalability)**

  For any scalar α,

  $f(\alpha x) = \sqrt{\langle \alpha x, \alpha x\rangle} = \sqrt{\alpha\,\bar{\alpha}\ \langle x, x\rangle} = \sqrt{|\,\alpha\,|^2\ \langle x, x\rangle} = |\,\alpha\,| \sqrt{\langle x, x\rangle} = |\,\alpha\,|\ f(x).$

  (For real spaces, $\alpha\,\bar{\alpha} = \alpha^2$.)

Proof (cont')

- **Triangle inequality**

  Use Cauchy–Schwarz:

  $$|\langle x, y\rangle| \leq f(x)\, f(y).$$

  Now expand:

  $$f(x+y)^2 = \langle x+y, x+y\rangle = \langle x, x\rangle + \langle y, y\rangle + \langle x, y\rangle + \langle y, x\rangle$$

  In real spaces, $\langle y, x\rangle = \langle x, y\rangle$,so

  $$f(x+y)^2 = f(x)^2 + f(y)^2 + 2\langle x, y\rangle \leq f(x)^2 + f(y)^2 + 2|\langle x, y\rangle|$$

  $$\leq f(x)^2 + f(y)^2 + 2f(x)f(y) = \big(f(x) + f(y)\big)^2$$

  Taking square roots (both sides ≥ 0):

  $$f(x+y) \leq f(x) + f(y).$$

- All three norm axioms hold, so f is a norm. □

Norm of block vectors

Note

Suppose a,b,c are vectors:

$$\left\|\begin{bmatrix} a \\ b \\ c \end{bmatrix}\right\|^2 = a^T a + b^T b + c^T c = \|a\|^2 + \|b\|^2 + \|c\|^2$$

So, we have

$$\left\|\begin{bmatrix} a \\ b \\ c \end{bmatrix}\right\| = \sqrt{\|a\|^2 + \|b\|^2 + \|c\|^2} = \left\|\begin{bmatrix} \|a\| \\ \|b\| \\ \|c\| \end{bmatrix}\right\|$$

(Parse RHS very carefully!)

The norm of a stacked vector is the norm of the vector formed from the norms of sub-vectors.

04

Euclidean Metric (Distance)

Metric Properties

Important Properties

Let V be a real vector space over $\mathbb{R}$. A function $V \times V \rightarrow \mathbb{R}$ is called metric or **distance function** on V, and (V, R) a metric space, if for all $u, v, w \in V$ the following holds true:

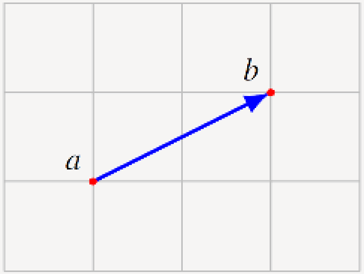
(i) $d(v, w) \geq 0$ and $d(v, w) = 0$ if and only if $v = w$;

(ii) $d(v, w) = d(v, w)$;

(iii) $d(v, w) \leq d(v, u) + d(u, w)$.

Euclidean Distance

- Distance between two n-vectors shows the vectors are “close” or “nearby” or “far”.



- Distance:

$$dist(a, b) = ||a - b||$$

Comparing Norm and Distance

Norm

(Normed Linear Space)

1. $\|x - y\| \geq 0$
2. $\|x - y\| = 0 \Rightarrow x = y$
3. $\|\lambda(x - y)\| = |\lambda| \, \|x - y\|$

Distance Function

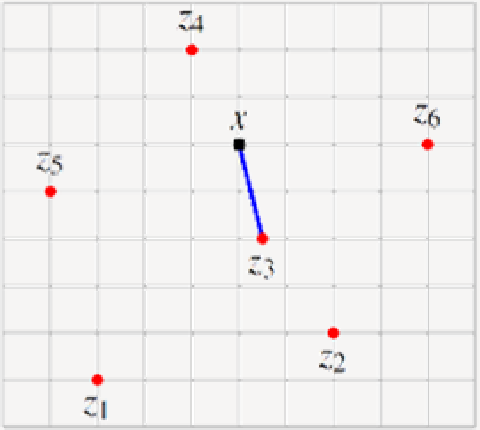
(Metric Space)

1. $dist(x, y) \geq 0$
2. $dist(x, y) = 0 \Rightarrow x = y$
3. $dist(x, y) = dist(y, x)$

ML Application

Feature Distance and Nearest Neighbors:

- if x, y are feature vectors for two entities, $||x - y||$ is the feature distance
- if $z_1, z_2, \dots, z_m$ is a list of vectors, z_j is the nearest neighbor of x if:
- $||x - z_j|| \leq ||x - z_i||, \quad i = 1, 2, \dots, m$



05

Angle

Angle

Definition

Angle between two non-zero vectors a, b is defined as:

$$\angle(a, b) = \arccos\left(\frac{a^T b}{||a||\,||b||}\right)$$

$\angle(a, b)$ is the number in $[0, \pi]$ that satisfies:

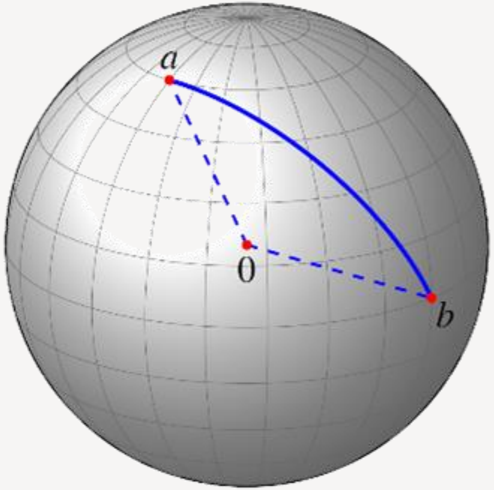
$$a^T b = ||a||\ ||b|| \cos(\angle(a, b))$$

Coincides with ordinary angle between vectors in 2D and 3D

Application

Spherical distance:

- $if\ a, b\ are\ on\ sphere\ with\ radius\ R, distance\ along\ the\ sphere\ is\ R\ \angle(a, b)$



Resources

- Chapter 1: Advanced Linear and Matrix Algebra, Nathaniel Johnston
- Chapter 6: Linear Algebra David Cherney
- Linear Algebra and Optimization for Machine Learning
- Introduction to Applied Linear Algebra Vectors, Matrices, and Least Squares