Chapter 3: Inner Products and Norms

We can now classify vectors that do/do not point in the same direction through the ideas of linear dependence/independence. This essentially gives us a yes or no answer to the question. Inner products and norms give a much more quantitative answer to similar questions. Furthermore, inner products and norms lead to further numerical methods and matrix factorizations.

Inner products

The most basic example of an inner product is the usual dot product

$$\vec{v} \cdot \vec{w} = \langle \vec{v}, \vec{w} \rangle = v_1 w_1 + \cdots + v_n w_n = \sum_{i=1}^{n} v_i w_i .$$

for two vectors in $\mathbb{R}^n$. (column vectors). Now, note that

$$\underset{\substack{\nearrow \\ 1\times n}}{} \quad V \cdot W = \underbrace{V^T}_{1\times n} \underbrace{W}_{n\times 1} \quad \leftarrow 1\times 1 - \text{scalar!}.$$

If we take the dot product of a vector with itself we obtain the sum of the squares

$$V \cdot V = v_1^2 + \cdots + v_n^2 .$$

This gives rise to the standard Euclidean norm/length of a vector v:

$$\|v\| = \sqrt{v \cdot v} = \sqrt{v_1^2 + \dots + v_n^2}.$$

We now make this idea of an inner product more precise:

<u>Definition</u>: An inner product on a vector space V is a pairing that takes two elements $\vec{v}, \vec{w} \in V$ and produces a real number $\langle \vec{v}, \vec{w} \rangle \in \mathbb{R}$. The inner product is required to satisfy the following axioms for all $\vec{u}, \vec{v}, \vec{w} \in V$ and $c, d \in \mathbb{R}$:

1) Bilinearity: (linear in each arguement separately)

$$\langle cu + dv, w \rangle = c\langle u, w \rangle + d\langle v, w \rangle$$
$$\langle u, cv + dw \rangle = c\langle u, v \rangle + d\langle u, w \rangle$$

2) Symmetry:

$$\langle v, w \rangle = \langle w, v \rangle$$

3) Positivity:

$$\langle v, v \rangle > 0 \text{ whenever } v \neq 0, \text{ while } \langle 0, 0 \rangle = 0.$$

Vector space + inner product $\Rightarrow$ inner product space.

Ex (Another inner product on $\mathbb{R}^2$)

Set

$$\langle \vec{u}, \vec{v} \rangle = 2u_1v_1 + 5u_2v_2$$

Let's check the inner product axioms:

Bilinearity:

$$\langle c\vec{u}+d\vec{v}, \vec{w} \rangle = 2(cu_1+dv_1)w_1 + 5(cu_2+dv_2)w_2$$
$$= c2u_1w_1 + d2v_1w_1 + c5u_2w_2 + d5v_2w_2$$
$$= c\langle \vec{u}, \vec{w} \rangle + d\langle \vec{v}, \vec{w} \rangle \checkmark$$

(The other calculation is nearly the same)

Symmetry:

$$\langle \vec{u}, \vec{w} \rangle = 2u_1w_1 + 5u_2w_2 = \langle \vec{w}, \vec{u} \rangle \checkmark$$

Positivity: $\langle \vec{v}, \vec{v} \rangle = 2v_1^2 + 5v_2^2 > 0$ if $v \neq 0$ since one of $v_1, v_2 \neq 0$.

Ex This can be generalized. Let $c_1, \dots, c_n$ be positive numbers. Then the pairing

$$\langle \vec{u}, \vec{v} \rangle = \sum_{i=1}^{n} c_i u_i v_i \quad \text{is a weighted}$$

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

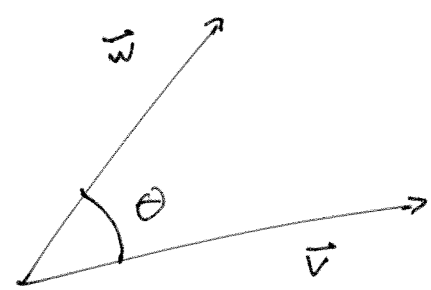
Inequalities

The Cauchy-Schwarz Inequality

Historically, the Cauchy-Schwarz inequality derived itself from the relation

$$\vec{v} \cdot \vec{w} = \|\vec{v}\| \, \|\vec{w}\| \cos\theta$$

equivalent to the law of cosines.



The inequality comes from $|\cos\theta| \le 1$ so that

$$|\vec{v} \cdot \vec{w}| \le \|\vec{v}\| \, \|\vec{w}\|.$$

This result is actually much more general:

<u>Theorem</u> Every inner product satisfies the Cauchy-Schwarz inequality:

$$|\langle \vec{v}, \vec{w} \rangle| \le \|\vec{v}\| \, \|\vec{w}\|.$$

Note: $\|\vec{v}\| = \sqrt{\langle \vec{v}, \vec{v} \rangle}$ is defined in terms of the specific inner product used.

Note: Equality holds if and only if the vectors are parallel.

Given any inner product we can use the quotient

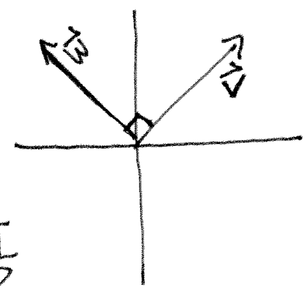
$$\cos\theta = \frac{\langle \vec{v}, \vec{w} \rangle}{\|\vec{v}\| \, \|\vec{w}\|}$$

to define the "angle" between vectors. The reason for the quotes on angle is the following: Consider the dot product and the inner product we saw before

$$\vec{v} \cdot \vec{w} = v_1 w_1 + v_2 w_2$$

$$\langle \vec{v}, \vec{w} \rangle = 2 v_1 w_1 + 5 v_2 w_2$$

Choose $\vec{v} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, $\vec{w} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$



$$\vec{v} \cdot \vec{w} = 0 \Rightarrow \cos\theta = 0 \Rightarrow \theta = \frac{\pi}{2}$$

$\langle \vec{v}, \vec{w} \rangle = 3$

$\|\vec{v}\| = \sqrt{2+5} = \sqrt{7}$ $\quad \Rightarrow \cos\theta = \frac{3}{7} \neq 0$!

$\|\vec{w}\| = \sqrt{7}$

The "angle" depends on the inner product.

Othogonality

Definition Two elements $\vec{v}, \vec{w} \in V$ of an inner product space V are called orthogonal if their inner product vanishes:

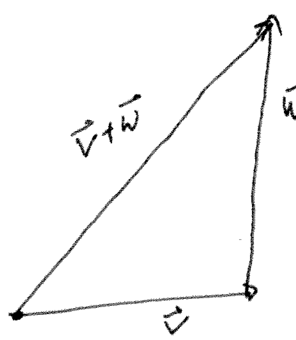
$$\langle \vec{v}, \vec{w} \rangle = 0.$$

Note: othogonal vectors cannot be linearly dependent.
This will come up again later.

Ex

$$\left\langle \begin{pmatrix} 5 \\ 2 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \end{pmatrix} \right\rangle = 2 \cdot 5(-1) + 5 \cdot 2 \cdot 1 = 0 \checkmark$$

The Triangle Inequality



Theorem
The norm associated with an inner product satisfies the triangle in quality

$$\| \vec{v} + \vec{w} \| \leq \| \vec{v} \| + \| \vec{w} \| \quad \text{for all } \vec{v}, \vec{w} \in V.$$

Proof:

$$\|\vec{v}+\vec{w}\|^2 = \langle \vec{v}+\vec{w}, \vec{v}+\vec{w}\rangle$$
$$= \langle \vec{v}+\vec{w}, \vec{v}\rangle + \langle \vec{v}+\vec{w}, \vec{w}\rangle$$
$$= \langle \vec{v}, \vec{v}\rangle + \langle \vec{w}, \vec{v}\rangle + \langle \vec{v}, \vec{w}\rangle + \langle \vec{w}, \vec{w}\rangle$$
$$= \langle \vec{v}, \vec{v}\rangle + 2\langle \vec{w}, \vec{v}\rangle + \langle \vec{w}, \vec{w}\rangle$$
$$\leq \langle \vec{v}, \vec{v}\rangle + 2\|\vec{w}\|\|\vec{v}\| + \langle \vec{w}, \vec{w}\rangle$$
$$= \|\vec{v}\|^2 + 2\|\vec{w}\|\|\vec{v}\| + \|\vec{w}\|^2$$
$$= (\|\vec{v}\| + \|\vec{w}\|)^2.$$

A square-root gives the result.

Norms: So far we use the inner product to define a norm $\|\vec{v}\| = \sqrt{\langle v, v\rangle}$. What is a norm? Are norms always attached to inner products? In this section we answer those questions.

Definition: A norm on the vector space V assigns a real number $\|\vec{v}\|$ to each vector $\vec{v} \in V$ subject to the following axioms for every $\vec{v}, \vec{w} \in V$ and $c \in \mathbb{R}$.

a) Positivity: $\|\vec{v}\| \geq 0$, with $\|\vec{v}\| = 0$ if and only if $\vec{v} = 0$.
b) Homogeneity: $\|c\vec{v}\| = |c| \, \|\vec{v}\|$
c) Triangle inequality: $\|\vec{v}+\vec{w}\| \leq \|\vec{v}\| + \|\vec{w}\|$

This definition assumes the existence of the triangle inequality. This indicates that a norm doesn't need an inner product.

Ex Consider the 1-norm

$$\|\vec{v}\|_1 = |v_1| + \dots + |v_n|$$

Ex The ∞-norm or max norm

$$\|\vec{v}\|_\infty = \max\{|v_1|, \dots, |v_n|\}.$$

can be obtained as the limit as $p \to \infty$.

Ex More generally the p-norm

$$\|\vec{v}\|_p = \left(\sum_{i=1}^{n} |v_i|^p\right)^{1/p}$$

- A norm measures the size of a vector $\vec{v}$.
- Given two vectors $\vec{v}$ and $\vec{w}$ the norm of the difference $\|\vec{v}-\vec{w}\|$ measures how close the two vectors are to one-another.

Definition An element $\vec{v} \in V$ such that

$$\|\vec{v}\| = 1$$

is called a <u>unit vector</u>.

- Given a vector $\vec{v} \in V$, the vector

$$\vec{u} = \frac{\vec{v}}{\|\vec{v}\|} \quad \text{is a unit vector.} \quad (\vec{v} \neq 0).$$

Check:

$$\|\vec{u}\| = \left\| \frac{\vec{v}}{\|\vec{v}\|} \right\| \overset{\text{homogeneity}}{=} \frac{1}{\|\vec{v}\|} \|\vec{v}\| = 1.$$

<u>Ex</u> Scale $\vec{v} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ so that $\|\vec{u}\|_2 = 1$

set $$\vec{u} = \frac{1}{\|\vec{v}\|} \vec{v} = \frac{1}{\sqrt{1^2+1^2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}.$$

<u>Positive Definite Matrices</u>:

An $n \times n$ matrix is said to be positive definite if it is symmetric, $K^T = K$, and satisfies the condition

$$\vec{x}^T K \vec{x} > 0 \quad \text{for all } \vec{x} \neq 0, \ \vec{x} \in \mathbb{R}^n.$$

In this case we write $K > 0$.

<u>Remark</u>: This does not tell you anything about the positivity of the matrix entries

Q: How to check if a matrix is positive definite?

We will arrive a better characterizations later but for now we check directly.

Let $K = \begin{pmatrix} 4 & -2 \\ -2 & 3 \end{pmatrix}$ and $\vec{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2$

$$\vec{x}^T K x = (x_1 \;\; x_2) \begin{pmatrix} 4 & -2 \\ -2 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

$$= (x_1 \;\; x_2) \begin{pmatrix} 4x_1 - 2x_2 \\ -2x_1 + 3x_2 \end{pmatrix} = 4x_1^2 - 2x_2 x_1 - 2x_2 x_1 + 3x_2^2$$

$$= 4x_1^2 - 4x_2 x_1 + 3x_2^2$$

write as sum of squares

$$= 4\left(x_1^2 - x_2 x_1 + \tfrac{3}{4} x_2^2\right)$$

$(a+b)^2 = a^2 + 2ab + b^2$

if $a = x_1$

$b = \frac{x_2}{2}$

$$= 4\left(x_1^2 - x_2 x_1 + \tfrac{1}{4} x_2^2 + \tfrac{1}{2} x_2^2\right)$$

$$= 4\left(x_1^2 - \tfrac{1}{2} x_2^2\right) + 2x_2^2 > 0.$$

The Cholesky Factorization:

We use the theorem:

Theorem:

A symmetrix matrix is positive definite iff it is regular and has all positive pivots.

We combine all our factorizations:

A regular $\Rightarrow$ $A = LU$ $\Rightarrow$ $A = LDV$

A symmetric $\Rightarrow$ $A = LDV = LDL^T$

A has positive pivots $\Rightarrow$

$$D = \begin{pmatrix} d_{11} & & & \\ & d_{22} & & \\ & & \ddots & \\ & & & d_{nn} \end{pmatrix}$$

d_{ii} are the pivots, so
$d_{ii} > 0$ for $i = 1, \dots, n$.

Write

$$S = \begin{pmatrix} \sqrt{d_{11}} & & & \\ & \sqrt{d_{22}} & & \\ & & \ddots & \\ & & & \sqrt{d_{nn}} \end{pmatrix}$$ so that

$$S^2 = SS^T = D$$

We have factored

$$A = LSS^TL^T.$$

Define $M = LS$, $M^T = S^TL^T$ $\Rightarrow$ $A = MM^T$

This is the <u>Cholesky factorization of A</u>.

Note: M is lower triangular.

Ex Find the cholesky factorization of

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & -2 \\ 1 & -2 & 14 \end{pmatrix}.$$

Step 1 Find $A = LU$

$-r_1 + r_2 \to r_2$
$-r_1 + r_3 \to r_3$

$$\begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & -3 \\ 0 & -3 & 13 \end{pmatrix}$$

$3r_2 + r_3 \to r_3$

$$\underbrace{\begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & -3 & 1 \end{pmatrix}}_{= L}, \underbrace{\begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & -3 \\ 0 & 0 & 4 \end{pmatrix}}_{= \mathcal{U}.}$$

All positive!

Step 2 Find $A = LDV$.

Recall: D is just the diagonal matrix consisting of the diagonal of $\mathcal{U}$,

$$D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 4 \end{pmatrix}.$$

$$A = \underbrace{\begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & -3 & 1 \end{pmatrix}}_{= L} \underbrace{\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 4 \end{pmatrix}}_{= D} \underbrace{\begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & -3 \\ 0 & 0 & 1 \end{pmatrix}}_{= L^T}$$

Step 3 Find "$S = \sqrt{D}$" and $M = LS$

$$S = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}, \quad M = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & -3 & 2 \end{pmatrix}, \quad L = MM^T!$$

Applications of Cholesky:

- Linear systems of the form $A\vec{x} = \vec{b}$ with A being symmetric are very common.

We will see these soon. We computed

$$A \longmapsto LU \quad \text{requires} \quad \begin{aligned} &\approx \tfrac{1}{3}n^3 \text{ additions} \\ &\approx \tfrac{1}{3}n^3 \text{ multiplications} \\ &\approx \tfrac{2}{3}n^3 \text{ operations} \end{aligned}$$

$$A \longmapsto MM^T \quad \text{requires} \quad \approx \tfrac{1}{3}n^3 \text{ operations}$$

through some optimized algorithms. While LU works, Cholesky is faster.