# Session 4: Norms, Inner Products and Orthogonality

Optimization and Computational Linear Algebra for Data Science

Marylou Gabrié (based on material by Léo Miolane)

# Contents

# 1. Norms

# 1.1 Introduction: the Euclidean norm

### Definition

We define the Euclidean norm of $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ as:

$$\|x\|_2 = \sqrt{x_1^2 + \cdots + x_n^2}.$$

# 1.1 Introduction: the Euclidean norm

**Definition**

We define the Euclidean norm of $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ as:

$$\|x\|_2 = \sqrt{x_1^2 + \cdots + x_n^2}.$$

**Observations.**

# 1.2 General norms

Let $V$ be a vector space.

**Definition**

A norm $\|\cdot\|$ on $V$ is a function from $V$ to $\mathbb{R}_{\geq 0}$ that verifies:

1. *Homogeneity*: $\|\alpha v\| = |\alpha| \times \|v\|$ for all $\alpha \in \mathbb{R}$ and $v \in V$.
2. *Positive definiteness*: if $\|v\| = 0$ for some $v \in V$, then $v = 0$.
3. *Triangular inequality*: $\|u + v\| \leq \|u\| + \|v\|$ for all $u, v \in V$.

# Other examples

- The $\ell_1$ norm

$$\|x\|_1 \stackrel{\text{def}}{=} \sum_{i=1}^{n} |x_i| \;=\; |x_1| + \cdots + |x_n|.$$

# Other examples

- The infinity-norm

$$\|x\|_\infty \stackrel{\text{def}}{=} \max(|x_1|, \ldots, |x_n|).$$

# Exercise: Balls drawing

For each of the norms $\|\cdot\|_2$, $\|\cdot\|_1$ and $\|\cdot\|_\infty$, draw the «ball»:

$$B = \{x \in \mathbb{R}^2 \mid \|x\| \leq 1\}.$$

# Exercise: Balls drawing

For each of the norms $\|\cdot\|_2$, $\|\cdot\|_1$ and $\|\cdot\|_\infty$, draw the «ball»:

$$B = \{x \in \mathbb{R}^2 \mid \|x\| \leq 1\}.$$

# Exercise: Balls drawing

For each of the norms $\|\cdot\|_2$, $\|\cdot\|_1$ and $\|\cdot\|_\infty$, draw the «ball»:

$$B = \{x \in \mathbb{R}^2 \mid \|x\| \leq 1\}.$$

# 2. Inner products

# 2.1 The Euclidean dot product

### Definition

We define the Euclidean dot product of two vectors $x$ and $y$ of $\mathbb{R}^n$ as:

$$x \cdot y \; = \; \sum_{i=1}^{n} x_i \, y_i \; = \; x_1 y_1 + \cdots + x_n y_n.$$

# 2.2 Inner product

Let $V$ be a vector space.

**Definition**

An inner product on $V$ is a function $\langle \cdot, \cdot \rangle$ from $V \times V$ to $\mathbb{R}$ that verifies the following points:

1. *Symmetry*: $\langle u, v \rangle = \langle v, u \rangle$ for all $u, v \in V$.
2. *Linearity*: $\langle u + v, w \rangle = \langle u, w \rangle + \langle v, w \rangle$ and $\langle \alpha v, w \rangle = \alpha \langle v, w \rangle$ for all $u, v, w \in V$ and $\alpha \in \mathbb{R}$.
3. *Positive definiteness*: $\langle v, v \rangle \geq 0$ with equality if and only if $v = 0$.

# Other example

If $V$ is the set of all random variables (on a probability space $\Omega$) that have a finite second moment, then

$$\langle X, Y \rangle \stackrel{\text{def}}{=} \mathbb{E}[XY]$$

is an inner product on $V$.

# 2.3 Norm induced by an inner product

### Proposition

If $\langle \cdot, \cdot \rangle$ is an inner product on $V$ then

$$\|v\| \stackrel{\text{def}}{=} \sqrt{\langle v, v \rangle}$$

is a norm on $V$. We say that the norm $\| \cdot \|$ is induced by the inner product $\langle \cdot, \cdot \rangle$.

# Example

Consider again the set $V$ of all random variables (on a probability space $\Omega$) that have a finite second moment, with the inner product:

$$\langle X, Y \rangle \stackrel{\text{def}}{=} \mathbb{E}[XY].$$

# 2.4 Cauchy-Schwarz inequality

### Theorem (Cauchy-Schwarz inequality)

*Let $\|\cdot\|$ be the norm induced by the inner product $\langle \cdot, \cdot \rangle$ on the vector space $V$. Then for all $x, y \in V$:*

$$|\langle x, y \rangle| \leq \|x\| \, \|y\|. \tag{1}$$

*Moreover, there is equality in* (1) *if and only if $x$ and $y$ are linearly dependent, i.e. $x = \alpha y$* or *$y = \alpha x$ for some $\alpha \in \mathbb{R}$.*

# Examples

For random variable, with the inner product $\langle X, Y \rangle = \mathbb{E}[XY]$:

# 2.5 Applications in data science

- Measure distances / strengths
  - e.g. Nearest neighbors
  - e.g. Regularization
- Measure angles / correlations

# 3. Orthogonality

# 3.1 Definitions: Orthogonality

Let $V$ be a vector space, and $\langle \cdot, \cdot \rangle$ be an inner product on V.

Definition

- We say that vectors $x$ and $y$ are *orthogonal* if $\langle x, y \rangle = 0$. We write then $x \perp y$.
- We say that a vector $x$ is orthogonal to a set of vectors $A$ if $x$ is orthogonal to all the vectors in $A$. We write then $x \perp A$.

**Exercise:** If $x$ is orthogonal to $v_1, \ldots, v_k$ then $x$ is orthogonal to any linear combination of these vectors i.e. $x \perp \mathrm{Span}(v_1, \ldots, v_k)$.

# Orthogonal & orthonormal families

### Definition

We say that a family of vectors $(v_1, \dots, v_k)$ is:

- *orthogonal* if the vectors $v_1, \dots, v_n$ are pairwise orthogonal, i.e. $\langle v_i, v_j \rangle = 0$ for all $i \neq j$.
- *orthonormal* if it is orthogonal and if all the $v_i$ have unit norm: $\|v_1\| = \cdots = \|v_k\| = 1$.

# Coordinates in an orthonormal basis

**Proposition**

A vector space of finite dimension admits an orthonormal basis.

**Proposition**

Assume that $\dim(V) = n$ and let $(v_1, \dots, v_n)$ be an **orthonormal** basis of $V$. Then the coordinates of a vector $x \in V$ in the basis $(v_1, \dots, v_n)$ are $(\langle v_1, x\rangle, \dots, \langle v_n, x\rangle)$:

$$x = \langle v_1, x\rangle v_1 + \cdots + \langle v_n, x\rangle v_n.$$

# Coordinates in an orthonormal basis

**Remark.** Let $x, y$ in $V$ with coordinates $x = (\alpha_1, \cdots, \alpha_n)$ and $y = (\beta_1, \cdots, \beta_n)$ in an othononormal basis $(v_1, \cdots, v_n)$.

- $\langle x, y \rangle =$

- $\|x, y\| =$

# 3.2 Pythagorean Theorem

**Theorem (Pythagorean theorem)**

*Let* $\|\cdot\|$ *be the norm induced by* $\langle \cdot, \cdot \rangle$. *For all* $x, y \in V$ *we have*

$$x \perp y \iff \|x+y\|^2 = \|x\|^2 + \|y\|^2.$$

**Proof.**

$\square$

# Application to random variables

For random variables with finite second moment, with the inner product $\langle X, Y \rangle = \mathbb{E}[XY]$:

# Picture

From now, $\langle \cdot, \cdot \rangle$ denotes the Euclidean dot product, and $\| \cdot \|$ the Euclidean norm. What is the vector of $S$ that is *the closest* to $x$ ?

# 3.3 Orthogonal projection

From now, $\langle \cdot, \cdot \rangle$ denotes the Euclidean dot product, and $\| \cdot \|$ the Euclidean norm.

**Definition**

Let $S$ be a subspace of $\mathbb{R}^n$. The **orthogonal projection** of a vector $x$ onto $S$ is defined as the vector $P_S(x)$ in $S$ that minimizes the distance to $x$:

$$P_S(x) \stackrel{\text{def}}{=} \operatorname*{arg\,min}_{y \in S} \|x - y\|.$$

The **distance of $x$ to the subspace $S$** is then defined as

$$d(x, S) \stackrel{\text{def}}{=} \min_{y \in S} \|x - y\| = \|x - P_S(x)\|.$$

# Computing orthogonal projections

## Proposition

Let $S$ be a subspace of $\mathbb{R}^n$ and let $(v_1, \dots, v_k)$ be an **orthonormal basis** of $S$. Then for all $x \in \mathbb{R}^n$,

$$P_S(x) = \langle v_1, x \rangle v_1 + \cdots + \langle v_k, x \rangle v_k.$$

# Proof

# Consequences

Let $V = \begin{pmatrix} | & & | \\ v_1 & \cdots & v_k \\ | & & | \end{pmatrix}$ gather the orthonormal basis-vectors of the subspace $S$

### Proposition

The orthogonal projection is given by $P_S(x) = VV^\top x$.

- $P_S$ is a linear transform.
- $VV^\top$ is its matrix.

# Consequences

### Corollary

For all $x \in \mathbb{R}^n$,

- $x - P_S(x)$ is orthogonal to $S$.
- $\|P_S(x)\| \leq \|x\|$.

Prove it in the homework!

# 3.4 Orthogonal complement

Let $S$ be a subspace of $V$.

### Definition

We define the orthogonal complement of $S$ as

$$S^{\perp} = \{\, x \in V \mid x \perp S \,\}$$

### Properties

- $S^{\perp}$ is a subspace of $V$.
- $\dim(S^{\perp}) = \dim(V) - \dim(S)$

**Exercise.** Prove it!

# Questions?

# Questions?

# 4. Orthogonal matrices (preview)

# Orthogonal matrices

### Definition

A matrix $A \in \mathbb{R}^{n \times n}$ is called an *orthogonal matrix* if its columns are an orthonormal family.

# A proposition

### Proposition

Let $A \in \mathbb{R}^{n\times n}$. The following points are equivalent:

1. $A$ is orthogonal.
2. $A^\mathsf{T} A = \mathrm{Id}_n$.
3. $AA^\mathsf{T} = \mathrm{Id}_n$

# Orthogonal matrices & norm

### Proposition

Let $A \in \mathbb{R}^{n\times n}$ be an orthogonal matrix. Then $A$ preserves the dot product in the sense that for all $x, y \in \mathbb{R}^n$,

$$\langle Ax, Ay\rangle = \langle x, y\rangle.$$

In particular if we take $x = y$ we see that $A$ preserves the Euclidean norm: $\|Ax\| = \|x\|$.

# Questions?

# Questions?

# Questions?