Session 4: Norms and inner-products

Optimization and Computational Linear Algebra for Data Science

Léo Miolane

Contents

Norms and inner-products

Questions

Questions

Questions

Orthogonality

Definition

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. $x \perp \mathrm{Span}(v_1, \ldots, v_k)$.

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

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

Proof

Orthogonal projection

Picture

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

Orthogonal projection and distance

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:

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

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 $(v_1, \ldots, v_k)$ 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

Consequence

Consequence

Corollary

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

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

Proof of Cauchy-Schwarz inequality

Cauchy-Schwarz inequality

Theorem

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 $\alpha \in \mathbb{R}$.

Proof

Proof

Questions?

Questions?

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. $A A^{\mathsf{T}} = \mathrm{Id}_n$

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: $\|Ax\| = \|x\|$.