Session 5: Orthogonal Matrices

Optimization and Computational Linear Algebra for Data Science

Marylou Gabrié (based on material by Léo Miolane)

Contents

1. Gram-Schmidt algorithm

Gram-Schmidt algorithm: purpose

The Gram-Schmidt process takes as

- **Input:** a *linearly independent* family $(x_1, \dots, x_k)$ of $\mathbb{R}^n$.
- **Output:** an *orthonormal basis* $(v_1, \dots v_k)$ of $\mathrm{Span}(x_1, \dots, x_k)$.

Consequence

Every subspace of $\mathbb{R}^n$ admits an orthonormal basis.

1.2 Gram-Schmidt construction

The Gram-Schmidt process constructs $v_1, v_2, \ldots, v_k$ in this order, such that for all $i \in \{1, \ldots, k\}$:

$$\mathcal{H}_i : \begin{cases} (v_1, \ldots, v_i) \text{ is an orthonormal family} \\ \mathrm{Span}(v_1, \ldots, v_i) = \mathrm{Span}(x_1, \ldots, x_i). \end{cases}$$

Iterative construction of the v_i's

Iterative construction of the v_i's

Iterative construction of the v_i's

2. Orthogonal matrices

2.1 Orthogonal matrices definition

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

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

$\mathbb{R}^n$ Orthogonal matrices = rotations

2.3 Orthonormal bases

- Let $(a_1, a_2, \cdots a_n)$ an orthonormal basis of $\mathbb{R}^n$, and A the $\mathbb{R}^{n\times n}$ matrix collecting the basis vectors in its columns.
- Consider $x = (x_1, x_2, \cdots, x_n)$ in $\mathbb{R}^n$, where $x_1, \cdots x_n$ are the coordinates in the canonical basis of $\mathbb{R}^n$.

Proposition

The coefficients of x in the $(a_1, a_2, \cdots a_n)$ basis are given by

$$x' = A^\top x$$

Proof.

Questions?

Questions?

Questions?

Session 6 - Eigenvalues & eigenvectors (preview)

Introduction

1.1 Definition

Definition

Let $A \in \mathbb{R}^{n \times n}$. A **non-zero** vector $v \in \mathbb{R}^n$ is said to be an *eigenvector* of A is there exists $\lambda \in \mathbb{R}$ such that

$$Av = \lambda v.$$

The scalar λ is called the eigenvalue (of A) associated to v.

Examples: I_d? matrix A with $\ker(A) \neq \{0\}$?

Example: diagonal matrices

Matrix with no eigenvalues/vectors

Example: orthogonal projection

1.3 Eigenspaces

Definition

If $\lambda \in \mathbb{R}$ is an eigenvalue of $A \in \mathbb{R}^{n \times n}$, the set

$$E_\lambda(A) = \{x \in \mathbb{R}^n \mid Ax = \lambda x\}$$

is called the eigenspace of A associated to λ. The dimension of $E_\lambda(A)$ is called the multiplicity of the eigenvalue λ.

Examples: Eigenvalue 1 for I_d? Eigenvalue 0 for $\ker(A)$?

Questions?

Questions?

Questions?