**CSE 840: Computational Foundations of Artificial Intelligence** Feb 10, 2025

Lecture 7

*Instructor: Vishnu Boddeti* *Scribe: Xinrui Yu*

# 1 Introduction

Metric spaces measure distances and lengths. The inner product extends this idea to vector spaces by defining angles and magnitudes.

# 2 Inner Product and Hilbert Spaces

**Definition 1** *Consider a vector space $V$. A mapping $< \cdot, \cdot >: V \times V \to F$ is called an inner product if it satisfies the following properties:*

- *(P1):* $< x_1 + x_2, y > = < x_1, y > + < x_2, y >$
- *(P1):* $< \lambda x, y > = \lambda < x, y > (\lambda \in F)$
- *(P3):* $< x, y > = \overline{< y, x >}$
- *(P4):*$< x, x > \geq 0$
- *(P5):* $< x, x > = 0 \Leftrightarrow x = 0$

Examples:

- Euclidiean inner product on $\mathbb{R}^n$: $< x, y > = \sum_{i=1}^{n} x_i y_i$.
- On $\mathbb{C}^n$, $\langle x, y \rangle = \sum x_i \overline{y_i}$
- $\mathcal{C}([a, b])$: $< f, g > = \int_a^b f(t)g(t)dt$ is an inner product (but space would not be complete)

**Definition 2** *A vector space with a norm is called a **normed space**. If every Cauchy sequence in the space converges, then it is called a **Banach space**. A vector space with an inner product is called a **pre-Hilbert space**. If it is additionally complete, then $V$ is called a **Hilbert Space**.*

Consider a vector space with an inner product $< \cdot, \cdot >$. Define $\| \cdot \| : V \to \mathbb{R}$ as $\|x\| := \sqrt{< x, x >}$. Then $\| \cdot \|$ is a norm on $V$, the norm is induced by $< \cdot, \cdot >$. In general, the other way does not work.

Consider a vector space $V$ with norm $\| \cdot \|$. Then $d : V \times V \to \mathbb{R}$, $d(x, y) := \|x - y\|$ is a metric on $V$, the metric is induced by the norm. In general, the other direction does not work.

# 3 Orthogonal Basis and Projection

**Definition 3** *Consider a pre-Hilbert spave $V$. Two vectors $v_1, v_2 \in V$ are called **orthogonal** if* $< v_1, v_2 >= 0$

Notation: $v_1 \perp v_2$ Two sets $v_1, v_2 \subset V$ are called orthogonal if $\forall v_1 \in V_1, v_2 \in V_2 :< v_1, v_2 >= 0$

Vectors are called orthogonal if additionally the two vectors have norm of 1.

A set of vectors $v_1, v_2, \cdots, v_n$ is called orthonormal is any two vectors are orthonormal. For a set $S \subseteq V$ we define its orthogonal complement $S^\perp$ as follows:

$$S^\perp := \{v \in V | v \perp s, \forall s \in S\}$$

# 4 Orthogonal Projection

**Definition 4** *$A \in \mathcal{L}(V)$ is called a **projection** if $A^2 = A$.*

**Theorem 5** *Let $U$ be a finite-dim subspace of a pre-Hilber-space $H$. Then there exists a linear projection $P_U : H \to U$, and $Ker(P_U) = U^\perp$. $P_U$ is then called the **orthogonal projection** of $H$ on $U$.*

Construction: Let $v_1, \cdots, v_n$ be an orthogonal basis of $U$. Define $P_U : V \to U$ by $P_U(w) = \sum_{i=1}^n \frac{<w,v_i>}{\|v_i\|} v_i$

**Remark 6** *In an orthonormal basis $v_1, \cdots, v_n$, the representation of a vector is given by*

$$v = \sum_{i=1}^n < v, u_i > u_i$$

# 5 Gram-Schmidt Orthogonalization

It is a procedure that takes any basis $v_1, \cdots, v_n$ of a finite-dim vector space and transforms it into another basis $u_1, \cdots, u_n$ that is orthonormal.

Intuition: iterative procedure

Step 1: $u_1 = \frac{v_1}{\|v_1\|}$, $U_1 = span\{u_1\}$

Step k: Assume that we already have $u_1, u_2, \cdots, u_{k-1}$.

- Project $v_k$ on $U_{k-1}$ and keep "the rest":

$$\tilde{u}_k = v_k - P_{U_{k-1}}(v_k)$$

- Normalize:

$$u_k = \frac{\tilde{u}_k}{\|\tilde{u}_k\|}$$

In practice, use Householder reflection for a numerically stable orthogonalization.

## 6 Orthogonal Matrices

**Definition 7** *Let $Q \in \mathbb{R}^{n \times n}$ be a matrix with orthonormal column vectors (w.r.t Euclidean inner product). Then $Q$ is called an orthogonal matrix.*

*If $Q \in \mathbb{C}^n$ and the columns are orthonormal (w.r.t the standard inner product on $\mathbb{C}$), then it is called unitary.*

Examples:

- Identity:
$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$
- Reflection:
$$\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$
- Permutation:
$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$
- Rotation:
$$\begin{bmatrix} cos\theta & -sin\theta \\ sin\theta & cos\theta \end{bmatrix}$$
- Rotation in $\mathbb{R}^3$:
$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & cos\theta & -sin\theta \\ 0 & sin\theta & cos\theta \end{bmatrix}$$
- General rotation can be written as a product of "elementary" rotation.

Properties of orthogonal matrix $Q$:

- columns are orthogonal $\Leftrightarrow$ rows are orthogonal.
- $Q$ is always invertible, and $Q^{-1} = Q^T$
- $Q$ realizes an isometry: $\forall v \in V : \|Qv\| = \|v\|$.
- $Q$ preserves angles: $< Qu, Qv >=< u, v >, \forall u, v \in V$
- $|det(Q)| = 1$

The respective properties also holds for unitary matrices $U$. $(U^{-1} = \bar{U}^T$

**Theorem 8** *Let $S \in \mathcal{L}(V)$ for a real vector space $V$. Then the following are equivalent:*

- *$S$ is an isometry: $\|Sv\| = \|v\|, v \in V$.*

- *There exists an orthonormal basis of $V$ such that the matrix of $S$ has the following form:*

$$A = \begin{bmatrix} B_1 & 0 & 0 & \cdots & 0 \\ 0 & B_2 & 0 & \cdots & 0 \\ 0 & 0 & B_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & B_n \end{bmatrix}$$

  *where each $B$ is a block, they are either a real number with value 1 or -1, or a $2 \times 2$ rotation matrix.*