ECE 586: Vector Space Methods
Lecture 17: Best Approximation

Henry D. Pfister
Duke University

4.1: Best Approximation

Let W be a subspace of a Banach space V and, for any $\underline{v} \in V$, consider finding a vector $\underline{w} \in W$ such that $\|\underline{v} - \underline{w}\|$ is as small as possible.

Definition

The vector $\underline{w} \in W$ is a best approximation of $\underline{v} \in V$ by vectors in W if

$$\|\underline{v} - \underline{w}\| \leq \|\underline{v} - \underline{w}'\|,$$

for all $\underline{w}' \in W$.

Example

If W is spanned by the vectors $\underline{w}_1, \ldots, \underline{w}_n \in V$, then we can write

$$\underline{v} = \underline{w} + \underline{e} = s_1 \underline{w}_1 + \cdots + s_n \underline{w}_n + \underline{e},$$

where $\underline{e} = \underline{v} - \underline{w}$ is the approximation error.

Vector Projection Revisited

Let $\underline{u}, \underline{v}$ be vectors in an inner-product space V with inner product $\langle \cdot, \cdot \rangle$.

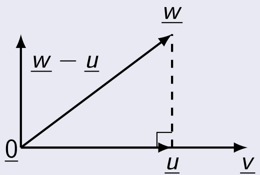
Lemma

If $\langle \underline{w}, \underline{v} \rangle = 0$, *then* $\|\underline{w}+\underline{v}\|^2 = \|\underline{w}\|^2 + 2\,Re\{\langle \underline{w}, \underline{v} \rangle\} + \|\underline{v}\|^2 = \|\underline{w}\|^2 + \|\underline{v}\|^2$.

Definition (Vector Projection)

The projection of $\underline{w}$ onto $\underline{v}$ is defined to be

$$\underline{u} = \frac{\langle \underline{w}, \underline{v} \rangle}{\|\underline{v}\|^2} \underline{v}$$



Lemma

Let $\underline{u}$ *be the projection of* $\underline{w}$ *onto* $\underline{v}$. *If* $\langle \underline{w}, \underline{v} \rangle \neq 0$, *then* $\|\underline{w} - \underline{u}\| < \|\underline{w}\|$.

Proof.

$\langle \underline{w} - \underline{u}, \underline{u} \rangle = 0$ implies $\|\underline{w}\|^2 = \|(\underline{w} - \underline{u}) + \underline{u}\|^2 = \|\underline{w} - \underline{u}\|^2 + \|\underline{u}\|^2$. □

4.1: Orthogonal Projection

In an arbitrary Banach space, finding a best approximation can be hard.

For the induced norm of a Hilbert space, orthogonal projection simplifies this!

Theorem (Projection)

Suppose W is a subspace of a Hilbert space V and $\underline{v} \in V$. Then,

1. *The vector $\underline{w} \in W$ is a best approximation of $\underline{v} \in V$ by vectors in W if and only if $\underline{v} - \underline{w}$ is orthogonal to every vector in W.*
2. *If a best approximation of $\underline{v} \in V$ by vectors in W exists, it is unique.*
3. *If W is a closed subspace with a countable orthogonal basis $\underline{w}_1, \underline{w}_2, \ldots,$ then the best approximation of $\underline{v}$ by vectors in W is*

$$\underline{w} = \sum_{i=1}^{\dim(W)} \frac{\langle \underline{v}, \underline{w}_i \rangle}{\|\underline{w}_i\|^2} \underline{w}_i.$$

Note: the implied linear mapping $T : V \to W$ defined by $T(\underline{v}) = \underline{w}$ is called the orthogonal projection of V onto W.

Proof in live session.

Orthogonal Projection Example

Example

For the standard inner product space $V = \mathbb{R}^3$, let W be the subspace spanned by

$$\begin{aligned} \underline{v}_1 &= (2,2,1), \\ \underline{v}_2 &= (3,6,0). \end{aligned}$$

Then, the Gram-Schmidt process generates the orthogonal basis

$$\begin{aligned} \underline{w}_1 &= (2,2,1), \\ \underline{w}_2 &= (-1,2,-2) \end{aligned}$$

and the orthogonal projection of $\underline{v} \in V$ onto W is defined by

$$\begin{aligned} T\underline{v} &= \sum_{i=1}^{2} \frac{\langle \underline{v}, \underline{w}_i \rangle}{\|\underline{w}_i\|^2} \underline{w}_i \\ &= \frac{1}{9}\langle \underline{v}, (2,2,1)\rangle(2,2,1) + \frac{1}{9}\langle \underline{v}, (-1,2,-2)\rangle(-1,2,-2). \end{aligned}$$

4.1.1: Orthogonal Projection onto an Orthonormal Set

Let $V = \mathbb{C}^n$ be the standard n-dimensional complex Hilbert space and $U \in \mathbb{C}^{n \times m}$ be a matrix with orthonormal columns $\underline{u}_1, \ldots, \underline{u}_m$:

$$U = \begin{bmatrix} | & | & & | \\ \underline{u}_1 & \underline{u}_2 & \cdots & \underline{u}_m \\ | & | & & | \end{bmatrix}$$

Then, the best approximation of $\underline{v} \in V$ by vectors in $\mathcal{R}(U)$, given by

$$\underline{w} = \sum_{i=1}^{m} \frac{\langle \underline{v}, \underline{u}_i \rangle}{\left\| \underline{u}_i \right\|^2} \underline{u}_i,$$

can also be written as

$$\underline{w} = UU^H \underline{v} = \sum_{i=1}^{m} \underline{u}_i (\underline{u}_i^H \underline{v}).$$

4.1.1: What is a Projection? (1)

Definition

A function $F\colon X \to Y$ with $Y \subseteq X$ is idempotent if $F(F(x)) = F(x)$. When F is a linear transformation, this reduces to $F^2 = F \cdot F = F$.

Definition

Let V be a vector space and $T\colon V \to V$ be a linear transformation. If T is idempotent, then T is called a projection because $T\underline{v} = \underline{v}$ if $\underline{v} \in \mathcal{R}(T)$.

Example

The idempotent matrix A is a projection onto the first two coordinates.

$$A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

Theorem

Let V be a vector space and $T\colon V \to V$ be a (linear) projection operator. Then, the range $\mathcal{R}(T)$ and the $\mathcal{N}(T)$ are disjoint subspaces of V.

Proof.

For any non-zero $\underline{v} \in \mathcal{R}(T)$, there is a non-zero $\underline{w} \in V$ such that $\underline{v} = T\underline{w}$. Thus, $T\underline{v} = T^2\underline{w} = T\underline{w} = \underline{v}$ and T projects onto $\mathcal{R}(T)$. Since $\underline{v} \neq \underline{0}$, we have $\underline{v} \notin \mathcal{N}(T)$. Thus, a non-zero vector cannot be in both subspaces. $\square$

Example

Consider the linear transform $T\colon V \to V$ defined by $T = I - P$, where P is a projection. It is easy to verify that T is a projection operator because

$$T^2 = (I-P)(I-P) = I - P - P + P^2 = I - P = T.$$

In fact, T is a projection onto $\mathcal{R}(T) = \mathcal{R}(I-P) = \mathcal{N}(P)$ because $P\underline{v} = \underline{0}$ (i.e., $\underline{v} \in \mathcal{N}(P)$) if and only if $(I-P)\underline{v} = \underline{v}$ (i.e., $\underline{v} \in \mathcal{R}(T)$).

4.1.1: Orthogonal Projection Operators

Definition

Let V be an inner-product space and $P\colon V \to V$ be a (linear) projection operator. If $\mathcal{R}(P) \perp \mathcal{N}(P)$, then P is called an orthogonal projection.

Example

Let $P\colon V \to V$ be an orthogonal projection. Since $P\underline{v} = \underline{0}$ iff $(I - P)\underline{v} = \underline{v}$, $\mathcal{N}(P) = \mathcal{R}(I - P)$. Thus, $P\underline{v} \in \mathcal{R}(P)$ is orthogonal to $(I - P)\underline{v} \in \mathcal{N}(P)$.

Theorem

For $V = F^n$ with the standard inner product, a projection matrix P is an orthogonal projection matrix if and only if it is Hermitian (i.e. $P^H = P$).

Proof.

If $P = P^H$, then $P = P^2 = P^H P$ and $\mathcal{R}(P) \perp \mathcal{N}(P)$ follows from

$$\langle P\underline{u}, (I - P)\underline{v}\rangle = \underline{v}^H (I - P)^H P\underline{u} = \underline{v}^H (P - P^H P)\underline{u} = 0.$$

If $\mathcal{R}(P) \perp \mathcal{N}(P)$, then $P = P^H$ because $\langle P\underline{u}, \underline{v}\rangle = \langle \underline{u}, P\underline{v}\rangle$ for all $\underline{u}, \underline{v}$ by

$$\langle P\underline{u}, \underline{v}\rangle = \langle P\underline{u}, P\underline{v} + (I - P)\underline{v}\rangle = \langle P\underline{u}, P\underline{v}\rangle + \langle P\underline{u}, (I - P)\underline{v}\rangle = \langle P\underline{u}, P\underline{v}\rangle. \quad \square$$

Next Steps

- To continue studying after this video –
  - Try the required reading: Course Notes EF 4.1 - 4.1.1
  - Or the recommended reading: LADR 6C
  - Also, look at the problems in Assignment 7