

Chapter 4

Representation and Approximation

4.1 Best Approximation

Suppose W is a subspace of a Banach space V . For any $\underline{v} \in V$, consider the problem of finding a vector $\underline{w} \in W$ such that $\|\underline{v} - \underline{w}\|$ is as small as possible.

Definition 4.1.1. *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 $\underline{w}' \in W$.

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

$$\begin{aligned}\underline{v} &= \underline{w} + \underline{e} \\ &= s_1 \underline{w}_1 + \dots + s_n \underline{w}_n + \underline{e},\end{aligned}$$

where \underline{e} is the approximation error.

Finding a best approximation is, in general, rather difficult¹. However, if the norm $\|\cdot\|$ corresponds to the induced norm of an inner product, then one can use orthogonal projection and the problem is greatly simplified. This chapter focuses mainly on computing the best approximation of arbitrary vectors in a Hilbert space.

¹A best approximation exists if W is closed and the Banach space V is reflexive (i.e., it equals its double dual). In addition, it is unique if the Banach space V is strictly convex (i.e., $\|\underline{v} + \underline{w}\| < 2$ for all distinct $\underline{v}, \underline{w} \in V$ such that $\|\underline{v}\| = \|\underline{w}\| = 1$).

Theorem 4.1.2. *Suppose W is a subspace of a Hilbert space V and \underline{v} is a vector in V . Then, we have the following:*

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 has a countable orthogonal basis $\underline{w}_1, \underline{w}_2, \dots$ and is closed, then*

$$\underline{w} = \sum_{i=1}^{\dim(W)} \frac{\langle \underline{v}, \underline{w}_i \rangle}{\|\underline{w}_i\|^2} \underline{w}_i \quad (4.1)$$

exists and equals the best approximation of \underline{v} by vectors in W .

Proof. Let $\underline{w} \in W$ and suppose $\underline{v} - \underline{w}$ is orthogonal to every vector in W . For any $\underline{w}' \in W$, we have $\underline{v} - \underline{w}' = (\underline{v} - \underline{w}) + (\underline{w} - \underline{w}')$ and

$$\begin{aligned} \|\underline{v} - \underline{w}'\|^2 &= \|\underline{v} - \underline{w}\|^2 + 2\operatorname{Re}\langle \underline{v} - \underline{w}, \underline{w} - \underline{w}' \rangle + \|\underline{w} - \underline{w}'\|^2 \\ &= \|\underline{v} - \underline{w}\|^2 + \|\underline{w} - \underline{w}'\|^2 \\ &\geq \|\underline{v} - \underline{w}\|^2. \end{aligned} \quad (4.2)$$

For the converse, we note that, if $\underline{v} - \underline{w}$ is not orthogonal to all vectors in W , then there must be some $\underline{u} \in W$ such that $\langle \underline{v} - \underline{w}, \underline{u} \rangle \neq 0$. Then, we let \underline{w}'' be the projection of $\underline{v} - \underline{w}$ onto \underline{u} . Next, we define $\underline{w}' = \underline{w} + \underline{w}''$ and observe that $\underline{w}' \in W$. Thus, Lemma 3.6.14 implies

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

Thus, \underline{w} is not a best approximation of \underline{v} by vectors in W .

For uniqueness, suppose $\underline{w}, \underline{w}' \in W$ are best approximations of \underline{v} by vectors in W . Then $\|\underline{v} - \underline{w}\| = \|\underline{v} - \underline{w}'\|$ and (4.2) implies that $\|\underline{w} - \underline{w}'\| = 0$. That is, if a best approximation exists then it is unique.

Finally, assume W is closed and $\underline{w}_1, \underline{w}_2, \dots$ is a countable orthogonal basis. Then, for (4.1), let the sequence of partial sums be $\underline{u}_n \triangleq \sum_{i=1}^n \underline{w}_i \langle \underline{v}, \underline{w}_i \rangle / \|\underline{w}_i\|^2$. Next, observe that $\underline{v} - \underline{u}_n$ is orthogonal to \underline{w}_j for $j \in \{1, \dots, n\}$, i.e.,

$$\begin{aligned} \langle \underline{v} - \underline{u}_n, \underline{w}_j \rangle &= \langle \underline{v}, \underline{w}_j \rangle - \left\langle \sum_{i=1}^n \frac{\langle \underline{v}, \underline{w}_i \rangle}{\|\underline{w}_i\|^2} \underline{w}_i, \underline{w}_j \right\rangle \\ &= \langle \underline{v}, \underline{w}_j \rangle - \frac{\langle \underline{v}, \underline{w}_j \rangle}{\|\underline{w}_j\|^2} \langle \underline{w}_j, \underline{w}_j \rangle = 0. \end{aligned}$$

Since $\underline{v} - \underline{u}_n$ is orthogonal to every vector in $W_n = \text{span}\{\underline{w}_1, \dots, \underline{w}_n\}$, we see that \underline{u}_n is the best approximation of \underline{v} by vectors in W_n . If $\dim(W) < \infty$, then this completes the proof.

If $\dim(W) = \infty$, then we must argue that the limit of \underline{u}_n exists and lies in W . The orthogonality of $\underline{w}_1, \dots, \underline{w}_n$ implies $\|\underline{v}\|^2 = \|\underline{v} - \underline{u}_n\|^2 + \|\underline{u}_n\|^2$. From this, we see that $\|\underline{u}_n\|^2 = \sum_{i=1}^n |\langle \underline{v}, \underline{w}_i \rangle|^2 / \|\underline{w}_i\|^2$ is an increasing real sequence upper bounded by $\|\underline{v}\|^2$. It follows that the RHS converges to a finite limit. Thus, we can apply Lemma 3.7.18 to show convergence $\underline{u}_n \rightarrow \underline{w}$. Since W is closed, it follows that $\underline{w} \in W$. By construction, $\underline{v} - \underline{w}$ is orthogonal to \underline{w}_j for $j \in \mathbb{N}$ and, thus, every vector in W . Hence, \underline{w} is the best approximation of \underline{v} by vectors in W . \square

Definition 4.1.3. Whenever the vector \underline{w} in Theorem 4.1.2 exists, it is called the **orthogonal projection** of \underline{v} onto W . If every vector in V has an orthogonal projection onto W , then the mapping $E: V \rightarrow W$, which assigns to each vector in V its orthogonal projection onto W , is called the **orthogonal projection of V onto W** .

One can use Theorem 4.1.14 to verify that this is consistent with the concept of orthogonal projection from Definition 4.1.11. Theorem 4.1.2 also implies the following result, known as Bessel's inequality.

Corollary 4.1.4. Let $\underline{v}_1, \underline{v}_2, \dots$ be a countable orthogonal set of distinct non-zero vectors in an inner-product space V and let $W = \text{span}(\{\underline{v}_1, \underline{v}_2, \dots\})$. If $\underline{v} \in V$, then

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

Moreover, equality holds if and only if

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

Proof. Let the projection of \underline{v} onto the closure of the span of $\underline{v}_1, \underline{v}_2, \dots$ (i.e., \overline{W}) be

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

Then, the error $\underline{u} = \underline{v} - \underline{w}$ satisfies $\langle \underline{u}, \underline{w} \rangle = 0$ and $\|\underline{u}\|^2 = \|\underline{v}\|^2 - \|\underline{w}\|^2$. Noting that $\|\underline{u}\|^2 \geq 0$ and

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

we see that $\|\underline{w}\|^2 \leq \|\underline{v}\|^2$ with equality iff $\underline{u} = \underline{0}$. \square

Problem 4.1.5. Let W be the subspace of \mathbb{R}^2 spanned by the vector $(1, 2)$. Using the standard inner product, let E be the orthogonal projection of \mathbb{R}^2 onto W . Find

1. a formula for $E(x_1, x_2)$
2. the matrix of E in the standard ordered basis, i.e., $E(x_1, x_2) = E\underline{x}$
3. W^\perp
4. an orthonormal basis in which E is represented by the matrix

$$E = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}.$$

4.1.1 Projection Operators

Definition 4.1.6. A function $F: X \rightarrow Y$ 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 4.1.7. Let V be a vector space and $T: V \rightarrow V$ be a linear transformation. If T is idempotent, then T is called a **projection**.

Example 4.1.8. 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 4.1.9. Let V be a vector space and $T: V \rightarrow V$ be a projection operator. Then, the range $\mathcal{R}(T)$ and the $\mathcal{N}(T)$ are disjoint subspaces of V .

Proof. For all $\underline{v} \in V - \{0\}$, we need to prove that \underline{v} is not in both the range and nullspace. Let $\underline{v} \in V$ be in the range of T so that there is a $\underline{w} \in V$ such that $T\underline{w} = \underline{v}$. Then, $T\underline{v} = T^2\underline{w} = T\underline{w} = \underline{v}$ and \underline{v} is not in the null space unless $\underline{v} = \underline{0}$.

Let \underline{v} be in the null space of T , then $T\underline{v} = \underline{0}$. But, $T\underline{v} = \underline{v}$ for all \underline{v} in the range. Therefore, \underline{v} is not in the range unless $\underline{v} = \underline{0}$. From this, we see that only $\underline{0} \in V$ is in both the range and nullspace. Therefore, they are disjoint subspaces. \square

Example 4.1.10. Consider the linear transform $T: V \rightarrow V$ defined by $T = I - P$, where P is a projection. It is easy to verify that T is a projection operator because

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

Notice also that $P(I - P)\underline{v} = \underline{0}$ implies that $\mathcal{R}(T) \subseteq \mathcal{N}(P)$ and $T\underline{v} = \underline{v}$ for $\underline{v} \in \mathcal{N}(P)$ implies $\mathcal{N}(P) \subseteq \mathcal{R}(T)$. Therefore, $\mathcal{R}(T) = \mathcal{N}(P)$ and $I - P$ is a projection onto $\mathcal{N}(P)$.

Definition 4.1.11. Let V be an inner-product space and $P: V \rightarrow V$ be a projection operator. If $\mathcal{R}(P) \perp \mathcal{N}(P)$, then P is called a **orthogonal projection**.

Example 4.1.12. Let V be an inner-product space and $P: V \rightarrow V$ be an orthogonal projection. Since $P\underline{v} = \underline{0}$ iff $(I - P)\underline{v} = \underline{v}$ and $I - P$ is a projection, we have $\mathcal{N}(P) = \mathcal{R}(I - P)$. Then, $\underline{v} = P\underline{v} + (I - P)\underline{v}$ defines an orthogonal decomposition of \underline{v} because $P\underline{v} \in \mathcal{R}(P)$ is orthogonal to $(I - P)\underline{v} \in \mathcal{N}(P)$. In addition, $V = \mathcal{R}(P) \oplus \mathcal{N}(P)$ and hence $\mathcal{N}(P) = \mathcal{R}(P)^\perp$.

Theorem 4.1.13. 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. First, we observe that

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

Then, if $P = P^H$, we have $P = P^2 = P^H P$ and the RHS above is zero. Thus, the LHS is zero and we have $\mathcal{R}(P) \perp \mathcal{N}(P)$. Conversely, if $\mathcal{R}(P) \perp \mathcal{N}(P)$, then

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

By symmetry, we also have $\langle \underline{u}, P\underline{v} \rangle = \langle P\underline{u}, P\underline{v} \rangle$ and together these imply that

$$\underline{v}^H P\underline{u} = \langle P\underline{u}, \underline{v} \rangle = \langle \underline{u}, P\underline{v} \rangle = \underline{v}^H P^H \underline{u}$$

for all $\underline{u}, \underline{v} \in V$. Thus, $P = P^H$. □

Theorem 4.1.14. Suppose W is a closed subspace of a separable Hilbert space V and let E denote the orthogonal projection of V on W . Then, E is an idempotent linear transformation of V onto W , $E\underline{w}' = \underline{0}$ iff $\underline{w}' \in W^\perp$, and

$$V = W \oplus W^\perp.$$

Proof. Let \underline{v} be any vector in V . Since $E\underline{v}$ is the best approximation of \underline{v} by vectors in W , it follows that $\underline{v} \in W$ implies $E\underline{v} = \underline{v}$. Therefore, $E(E\underline{v}) = E\underline{v}$ for any $\underline{v} \in V$ since $E\underline{v} \in W$. That is, $E^2 = E$ and E is idempotent.

To show that E is a linear transformation, let $\underline{w}_1, \underline{w}_2, \dots$ be a countable orthonormal basis for W (whose existence follows from Theorem 3.7.17). Using part 3 of Theorem 4.1.2, we find that

$$\begin{aligned} E(s_1\underline{v}_1 + \underline{v}_2) &= \sum_{i=1}^{\dim(W)} \langle s_1\underline{v}_1 + \underline{v}_2, \underline{w}_i \rangle \underline{w}_i \\ &= s_1 \sum_{i=1}^{\dim(W)} \langle \underline{v}_1, \underline{w}_i \rangle \underline{w}_i + \sum_{i=1}^{\dim(W)} \langle \underline{v}_2, \underline{w}_i \rangle \underline{w}_i \\ &= s_1 E\underline{v}_1 + E\underline{v}_2. \end{aligned}$$

Therefore, E is a linear transformation. It also follows that $E\underline{w}' = \underline{0}$ iff $\underline{w}' \in W^\perp$ because W^\perp can be defined by the fact that $\langle \underline{w}', \underline{w}_i \rangle = 0$ for $i \in \mathbb{N}$.

Again, let $\underline{v} \in V$ and recall that (by Theorem 4.1.2) $E\underline{v}$ is the unique vector in W such that $\underline{v} - E\underline{v}$ is in W^\perp . Therefore, the equation $\underline{v} = E\underline{v} + (\underline{v} - E\underline{v})$ gives a unique decomposition of \underline{v} into $E\underline{v} \in W$ and $\underline{v} - E\underline{v} \in W^\perp$. This unique decomposition implies that V is the direct sum of W and W^\perp . Lastly, one finds from the definition of W^\perp that

$$W \cap W^\perp = \{\underline{u} \in W \mid \langle \underline{u}, \underline{w} \rangle = 0 \forall \underline{w} \in W\} \subseteq \{\underline{u} \in W \mid \langle \underline{u}, \underline{u} \rangle = 0\} = \{\underline{0}\}. \quad \square$$

Corollary 4.1.15. *Let W be a closed subspace of a separable Hilbert space V and E be the orthogonal projection of V on W . Then $I - E$ is the orthogonal projection of V on W^\perp .*

Proof. This follows directly from the orthogonal decomposition in Theorem 4.1.14. One can also verify that $I - E$ is an idempotent linear transformation of V with range W^\perp and nullspace W . From Definition 4.1.11, we see that $I - E$ is an orthogonal projection. \square

Example 4.1.16. *Let $V = \mathbb{C}^n$ be the standard n -dimensional complex Hilbert space. Let $U \in \mathbb{C}^{n \times m}$ be a matrix whose columns $\underline{u}_1, \dots, \underline{u}_m$ form an orthonormal set in V . Then, the best approximation of $\underline{v} \in V$ by vectors in $\mathcal{R}(U)$ (as defined by (4.1)) can also be written as*

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

4.2 Computing Approximations in Hilbert Spaces

4.2.1 Normal Equations

Suppose V is a Hilbert space the subspace W is spanned by $\underline{w}_1, \dots, \underline{w}_n \in V$. Consider the situation where the sequence $\underline{w}_1, \dots, \underline{w}_n$ is linearly independent, but not orthogonal. In this case, it is not possible to apply (4.1) directly. It is nevertheless possible to obtain a similar expression for the best approximation of \underline{v} by vectors in W . Theorem 4.1.2 asserts that $\hat{\underline{v}} \in W$ is a best approximation of $\underline{v} \in V$ by vectors in W if and only if $\underline{v} - \hat{\underline{v}}$ is orthogonal to every vector in W . This implies that

$$\langle \underline{v} - \hat{\underline{v}}, \underline{w}_j \rangle = \left\langle \underline{v} - \sum_{i=1}^n s_i \underline{w}_i, \underline{w}_j \right\rangle = 0$$

or, equivalently,

$$\sum_{i=1}^n s_i \langle \underline{w}_i, \underline{w}_j \rangle = \langle \underline{v}, \underline{w}_j \rangle$$

for $j = 1, \dots, n$. These conditions yield a system of n linear equations in n unknowns, which can be written in the matrix form

$$\begin{bmatrix} \langle \underline{w}_1, \underline{w}_1 \rangle & \langle \underline{w}_2, \underline{w}_1 \rangle & \cdots & \langle \underline{w}_n, \underline{w}_1 \rangle \\ \langle \underline{w}_1, \underline{w}_2 \rangle & \langle \underline{w}_2, \underline{w}_2 \rangle & \cdots & \langle \underline{w}_n, \underline{w}_2 \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle \underline{w}_1, \underline{w}_n \rangle & \langle \underline{w}_2, \underline{w}_n \rangle & \cdots & \langle \underline{w}_n, \underline{w}_n \rangle \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \\ \vdots \\ s_n \end{bmatrix} = \begin{bmatrix} \langle \underline{v}, \underline{w}_1 \rangle \\ \langle \underline{v}, \underline{w}_2 \rangle \\ \vdots \\ \langle \underline{v}, \underline{w}_n \rangle \end{bmatrix}.$$

We can rewrite this matrix equation as

$$G \underline{s} = \underline{t}$$

where

$$\underline{t}^T = (\langle \underline{v}, \underline{w}_1 \rangle, \langle \underline{v}, \underline{w}_2 \rangle, \dots, \langle \underline{v}, \underline{w}_n \rangle)$$

is the **cross-correlation vector**, and

$$\underline{s}^T = (s_1, s_2, \dots, s_n)$$

is the vector of coefficients. Equations of this form are collectively known as the **normal equations**.

Definition 4.2.1. The $n \times n$ matrix

$$G = \begin{bmatrix} \langle \underline{w}_1, \underline{w}_1 \rangle & \langle \underline{w}_2, \underline{w}_1 \rangle & \cdots & \langle \underline{w}_n, \underline{w}_1 \rangle \\ \langle \underline{w}_1, \underline{w}_2 \rangle & \langle \underline{w}_2, \underline{w}_2 \rangle & \cdots & \langle \underline{w}_n, \underline{w}_2 \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle \underline{w}_1, \underline{w}_n \rangle & \langle \underline{w}_2, \underline{w}_n \rangle & \cdots & \langle \underline{w}_n, \underline{w}_n \rangle \end{bmatrix} \quad (4.3)$$

is called the **Gramian** matrix. Since $g_{ij} = \langle \underline{w}_j, \underline{w}_i \rangle$, it follows that the Gramian is a Hermitian symmetric matrix, i.e., $G^H = G$.

Definition 4.2.2. A matrix $M \in F^{n \times n}$ is **positive-semidefinite** if $M^H = M$ and $\underline{v}^H M \underline{v} \geq 0$ for all $\underline{v} \in F^n - \{0\}$. If the inequality is strict, M is **positive-definite**.

An important aspect of positive-definite matrices is that they are always invertible. This follows from noting that $M \underline{v} = 0$ for $\underline{v} \neq 0$ implies that $\underline{v}^H M \underline{v} = 0$ and contradicts the definition of positive definite.

Theorem 4.2.3. A Gramian matrix G is always positive-semidefinite. It is positive-definite if and only if the vectors $\underline{w}_1, \dots, \underline{w}_n$ are linearly independent.

Proof. Since $g_{ij} = \langle \underline{w}_j, \underline{w}_i \rangle$, the conjugation property of the inner product implies $G^H = G$. Using $\underline{v} = (v_1, \dots, v_n)^T \in F^n$, we can write

$$\begin{aligned} \underline{v}^H G \underline{v} &= \sum_{i=1}^n \sum_{j=1}^n \bar{v}_i g_{ij} v_j = \sum_{i=1}^n \sum_{j=1}^n \bar{v}_i \langle \underline{w}_j, \underline{w}_i \rangle v_j \\ &= \sum_{i=1}^n \sum_{j=1}^n \langle v_j \underline{w}_j, v_i \underline{w}_i \rangle = \left\langle \sum_{j=1}^n v_j \underline{w}_j, \sum_{i=1}^n v_i \underline{w}_i \right\rangle \\ &= \left\| \sum_{i=1}^n v_i \underline{w}_i \right\|^2 \geq 0. \end{aligned} \quad (4.4)$$

That is, $\underline{v}^H G \underline{v} \geq 0$ for all $\underline{v} \in F^n$.

Suppose that G is not positive-definite. Then, there exists $\underline{v} \in F^n - \{0\}$ such that $\underline{v}^H G \underline{v} = 0$. By (4.4), this implies that

$$\sum_{i=1}^n v_i \underline{w}_i = 0$$

and hence the sequence of vectors $\underline{w}_1, \dots, \underline{w}_n$ is not linearly independent.

Conversely, if G is positive-definite then $\underline{v}^H G \underline{v} > 0$ and

$$\left\| \sum_{i=1}^n v_i \underline{w}_i \right\| > 0$$

for all $\underline{v} \in F^n - \{0\}$. Thus, the vectors $\underline{w}_1, \dots, \underline{w}_n$ are linearly independent. \square

4.2.2 Orthogonality Principle

Theorem 4.2.4. *Let $\underline{w}_1, \dots, \underline{w}_n$ be vectors in an inner-product space V and denote the span of $\underline{w}_1, \dots, \underline{w}_n$ by W . For any vector $\underline{v} \in V$, the norm of the error vector*

$$\underline{e} = \underline{v} - \sum_{i=1}^n s_i \underline{w}_i \quad (4.5)$$

*is minimized when the error vector \underline{e} is orthogonal to every vector in W . If $\hat{\underline{v}}$ denotes the **least-squares** approximation of \underline{v} then*

$$\langle \underline{v} - \hat{\underline{v}}, \underline{w}_j \rangle = 0$$

for $j = 1, \dots, n$.

Proof. Minimizing $\|\underline{e}\|^2$ over \underline{s} , where \underline{e} is given by (4.5) requires minimizing

$$\begin{aligned} J(\underline{s}) &= \left\langle \underline{v} - \sum_{i=1}^n s_i \underline{w}_i, \underline{v} - \sum_{j=1}^n s_j \underline{w}_j \right\rangle \\ &= \langle \underline{v}, \underline{v} \rangle - \sum_{i=1}^n \langle s_i \underline{w}_i, \underline{v} \rangle - \sum_{j=1}^n \langle \underline{v}, s_j \underline{w}_j \rangle + \sum_{i=1}^n \sum_{j=1}^n \langle s_i \underline{w}_i, s_j \underline{w}_j \rangle \\ &= \langle \underline{v}, \underline{v} \rangle - \sum_{i=1}^n s_i \langle \underline{w}_i, \underline{v} \rangle - \sum_{j=1}^n \bar{s}_j \langle \underline{v}, \underline{w}_j \rangle + \sum_{i=1}^n \sum_{j=1}^n s_i \bar{s}_j \langle \underline{w}_i, \underline{w}_j \rangle. \end{aligned}$$

To take the derivative with respect to $\underline{s} \in \mathbb{C}^n$, we use the decomposition $\underline{s} = \underline{a} + j\underline{b}$, with $\underline{a}, \underline{b} \in \mathbb{R}^n$, and define the differential operators

$$\frac{\partial}{\partial \underline{s}} \triangleq \frac{1}{2} \left(\frac{\partial}{\partial \underline{a}} - j \frac{\partial}{\partial \underline{b}} \right) \quad \frac{\partial}{\partial \bar{\underline{s}}} \triangleq \frac{1}{2} \left(\frac{\partial}{\partial \underline{a}} + j \frac{\partial}{\partial \underline{b}} \right),$$

where $\partial/\partial \underline{a} = (\partial/\partial a_1, \dots, \partial/\partial a_n)^T$. Since $J(\underline{s})$ is real function of \underline{s} , a stationary

point of J can be found by setting either derivative to $\underline{0}$. Choosing \underline{s} , gives

$$\begin{aligned} \frac{\partial}{\partial \underline{s}} J(\underline{s}) &= - \begin{bmatrix} \langle \underline{v}, \underline{w}_1 \rangle \\ \langle \underline{v}, \underline{w}_2 \rangle \\ \vdots \\ \langle \underline{v}, \underline{w}_n \rangle \end{bmatrix} + \begin{bmatrix} \langle \underline{w}_1, \underline{w}_1 \rangle & \langle \underline{w}_2, \underline{w}_1 \rangle & \cdots & \langle \underline{w}_n, \underline{w}_1 \rangle \\ \langle \underline{w}_1, \underline{w}_2 \rangle & \langle \underline{w}_2, \underline{w}_2 \rangle & \cdots & \langle \underline{w}_n, \underline{w}_2 \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle \underline{w}_1, \underline{w}_n \rangle & \langle \underline{w}_2, \underline{w}_n \rangle & \cdots & \langle \underline{w}_n, \underline{w}_n \rangle \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \\ \vdots \\ s_n \end{bmatrix} \\ &= \underline{0}. \end{aligned}$$

In matrix form, this yields the familiar equation

$$G\underline{s} = \underline{t}.$$

To ensure that this extremum is in fact a minimum, one can compute the 2nd derivative to show that the Hessian is G . Since G is a positive-semidefinite matrix, the extremum is indeed a minimum.

This implies that $\|\underline{e}\|^2$ is minimized if and only if $G\underline{s} = \underline{t}$. That is, $\|\underline{e}\|^2$ is minimized if and only if $\underline{v} - \hat{\underline{v}}$ is orthogonal to every vector in W . \square

Note that it is also possible to prove this theorem using the Cauchy-Schwarz inequality or the projection theorem.

4.3 Approximation for Systems of Linear Equations

4.3.1 Matrix Representation

For finite-dimensional vector spaces, least-squares (i.e., best approximation) problems have natural matrix representations. Suppose $V = F^m$ and $\underline{w}_1, \underline{w}_2, \dots, \underline{w}_n \in V$ are column vectors. Then, the approximation vector is given by

$$\hat{\underline{v}} = \sum_{i=1}^n s_i \underline{w}_i$$

In matrix form, we have

$$\hat{\underline{v}} = A\underline{s},$$

where $A = [\underline{w}_1 \cdots \underline{w}_n]$. The optimization problem can then be reformulated as follows. Determine $\underline{s} \in F^n$ such that

$$\|\underline{e}\|^2 = \|\underline{v} - \hat{\underline{v}}\|^2 = \|\underline{v} - A\underline{s}\|^2$$

is minimized. Note that this occurs when the error vector is orthogonal to every vector in W , i.e.,

$$\langle \underline{e}, \underline{w}_j \rangle = \langle \underline{v} - \hat{\underline{v}}, \underline{w}_j \rangle = \langle \underline{v} - A\underline{s}, \underline{w}_j \rangle = 0$$

for $j = 1, \dots, n$.

4.3.2 Standard Inner Products

When $\|\cdot\|$ is the norm induced by the standard inner product, these conditions can be expressed as

$$\begin{bmatrix} \underline{w}_1^H \\ \vdots \\ \underline{w}_n^H \end{bmatrix} (\underline{v} - A\underline{s}) = \underline{0}.$$

Using the definition of A , we obtain

$$A^H A \underline{s} = A^H \underline{v}.$$

The matrix $A^H A$ is the Gramian G defined in (4.3). The vector $A^H \underline{v}$ is the cross correlation vector \underline{t} .

When the vectors $\underline{w}_1, \dots, \underline{w}_n$ are linearly independent, the Gramian matrix is positive definite and hence invertible. The optimal solution for the least-squares problem is therefore given by

$$\underline{s} = (A^H A)^{-1} A^H \underline{v} = G^{-1} \underline{t}.$$

The matrix $(A^H A)^{-1} A^H$ is often called the **pseudoinverse**.

The best approximation of $\underline{v} \in V$ by vectors in W is equal to

$$\hat{\underline{v}} = A \underline{s} = A (A^H A)^{-1} A^H \underline{v}.$$

The matrix $P = A (A^H A)^{-1} A^H$ is called the **projection matrix** for the range of A . It defines an orthogonal projection onto the range of A (i.e., the subspace spanned by the columns of A).

4.3.3 Generalized Inner Products

We can also consider the case of a general inner product. Recall that an inner product on V is completely determined by the values

$$h_{ji} = \langle \underline{e}_i, \underline{e}_j \rangle$$

and that it can be expressed in terms of the matrix H (where $[H]_{j,i} = h_{j,i}$) as

$$\langle \underline{v}, \underline{w} \rangle = \underline{w}^H H \underline{v}.$$

Minimizing $\|\underline{e}\|^2 = \|\underline{v} - A\underline{s}\|^2$ and using the orthogonality principle lead to the matrix equation

$$A^H H A \underline{s} = A^H H \underline{v}.$$

When the vectors $\underline{w}_1, \dots, \underline{w}_n$ are linearly independent, the optimal solution is given by

$$\underline{s} = (A^H H A)^{-1} A^H H \underline{v}.$$

4.3.4 Minimum Error

Let $\hat{\underline{v}} \in W$ be the best approximation of \underline{v} by vectors in W . Again, we can write

$$\underline{v} = \hat{\underline{v}} + \underline{e},$$

where $\underline{e} \in W^\perp$ is the minimum achievable error. The squared norm of the minimum error is given implicitly by

$$\|\underline{v}\|^2 = \|\hat{\underline{v}} + \underline{e}\|^2 = \langle \hat{\underline{v}} + \underline{e}, \hat{\underline{v}} + \underline{e} \rangle = \langle \hat{\underline{v}}, \hat{\underline{v}} \rangle + \langle \underline{e}, \underline{e} \rangle = \|\hat{\underline{v}}\|^2 + \|\underline{e}\|^2.$$

We can then find an explicit expression for the approximation error,

$$\begin{aligned} \|\underline{e}\|^2 &= \|\underline{v}\|^2 - \|\hat{\underline{v}}\|^2 = \underline{v}^H H \underline{v} - \hat{\underline{v}}^H H \hat{\underline{v}} \\ &= \underline{v}^H H \underline{v} - \underline{s}^H A^H H A \underline{s} \\ &= \underline{v}^H H \underline{v} - \underline{v}^H H A (A^H H A)^{-1} A^H H \underline{v} \\ &= \underline{v}^H \left(H - H A (A^H H A)^{-1} A^H H \right) \underline{v}. \end{aligned}$$

4.4 Applications and Examples in Signal Processing

4.4.1 Linear Regression

Let $(x_1, y_1), \dots, (x_n, y_n)$ be a collection of points in \mathbb{R}^2 . A **linear regression** problem consists in finding scalars a and b such that

$$y_i \approx ax_i + b$$

for $i = 1, \dots, n$. Define the error component e_i by $e_i = y_i - ax_i - b$, then

$$\begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = a \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} + b \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} + \begin{bmatrix} e_1 \\ \vdots \\ e_n \end{bmatrix} = \begin{bmatrix} x_1 & 1 \\ \vdots & \vdots \\ x_n & 1 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} + \begin{bmatrix} e_1 \\ \vdots \\ e_n \end{bmatrix}.$$

In vector form, we can rewrite this equation as

$$\underline{y} = A\underline{s} + \underline{e},$$

where $\underline{y} = (y_1, \dots, y_n)^T$, $\underline{s} = (a, b)^T$, $\underline{e} = (e_1, \dots, e_n)^T$, and

$$A = \begin{bmatrix} x_1 & 1 \\ \vdots & \vdots \\ x_n & 1 \end{bmatrix}.$$

This equation has a form analog to the matrix representation of a least-squares problems. Consider the goal of minimizing $\|\underline{e}\|^2$. The line that minimizes the sums of the squares of the *vertical* distances between the data abscissas and the line is then given by

$$\underline{s} = (A^H A)^{-1} A^H \underline{y}.$$

4.4.2 Linear Minimum Mean-Squared Error Estimation

Let Y, X_1, \dots, X_n be a set of zero-mean random variables. The goal of the linear minimum mean-squared error (LMMSE) estimation problem is to find coefficients s_1, \dots, s_n such that

$$\hat{Y} = s_1 X_1 + \dots + s_n X_n$$

minimizes the MSE $E[|Y - \hat{Y}|^2]$. Using the inner product defined by

$$\langle X, Y \rangle = E[X\bar{Y}], \quad (4.6)$$

we can compute the linear minimum mean-squared estimate \hat{Y} using

$$G\underline{s} = \underline{t},$$

where

$$G = \begin{bmatrix} E[X_1\bar{X}_1] & E[X_2\bar{X}_1] & \cdots & E[X_n\bar{X}_1] \\ E[X_1\bar{X}_2] & E[X_2\bar{X}_2] & \cdots & E[X_n\bar{X}_2] \\ \vdots & \vdots & \ddots & \vdots \\ E[X_1\bar{X}_n] & E[X_2\bar{X}_n] & \cdots & E[X_n\bar{X}_n] \end{bmatrix}$$

and

$$\underline{t} = \begin{bmatrix} E[Y\bar{X}_1] \\ E[Y\bar{X}_2] \\ \vdots \\ E[Y\bar{X}_n] \end{bmatrix}.$$

If the matrix G is invertible, the minimum mean-squared error is given by

$$\|Y - \hat{Y}\|^2 = E[Y\bar{Y}] - \underline{t}^H G^{-1} \underline{t}.$$

4.4.3 The Wiener Filter

Suppose that the sequence of zero-mean random variables $\{X[t]\}$ is wide-sense stationary, and consider the FIR filter

$$\begin{aligned} Y[t] &= \sum_{k=0}^{K-1} h[k]X[t-k] \\ &= \begin{bmatrix} X[t] & \cdots & X[t-K+1] \end{bmatrix} \begin{bmatrix} h[0] \\ \vdots \\ h[K-1] \end{bmatrix} = (\underline{X}[t])^T \underline{h}. \end{aligned}$$

The goal is to design this filter in such a way that its output is as close as possible to a desired sequence $\{Z[t]\}$. In particular, we want to minimize the mean-squared error

$$\|Z[t] - Y[t]\|^2 = E[|Z[t] - Y[t]|^2].$$

By the orthogonality principle, the mean-squared error is minimized when the error is orthogonal to the data; that is, for $j = 0, 1, \dots, K - 1$, we have

$$\left\langle Z[t] - \sum_{k=0}^{K-1} h[k]X[t-k], X[t-j] \right\rangle = 0,$$

or, equivalently, we can write

$$\langle Z[t], X[t-j] \rangle = \sum_{k=0}^{K-1} h[k] \langle X[t-k], X[t-j] \rangle.$$

Using (4.6), we obtain

$$\mathbb{E} [Z[t]\bar{X}[t-j]] = \sum_{k=0}^{K-1} h[k] \mathbb{E} [X[t-k]\bar{X}[t-j]]. \quad (4.7)$$

where $j = 1, \dots, K - 1$.

For this specific case where the normal equations are defined in terms of the expectation operator, these equations are called the **Wiener-Hopf** equations. The Gramian of the Wiener-Hopf equations can be expressed in a more familiar form using the autocorrelation matrix. Recall that $\{X[t]\}$ is a wide-sense stationary process. As such, we have

$$R_{xx}(j-k) = R_{xx}(j,k) = \mathbb{E} [X[t-k]\bar{X}[t-j]] = \langle X[t-k], X[t-j] \rangle.$$

Also define

$$R_{zx}(j) = \mathbb{E} [Z[t]\bar{X}[t-j]] = \langle Z[t], X[t-j] \rangle.$$

Using this notation, we can rewrite (4.7) as

$$R_{zx} = \begin{bmatrix} R_{zx}(0) \\ R_{zx}(1) \\ \vdots \\ R_{zx}(K-1) \end{bmatrix} = R_{xx} \begin{bmatrix} h[0] \\ h[1] \\ \vdots \\ h[K-1] \end{bmatrix}$$

where the $K \times K$ autocorrelation matrix is given by

$$R_{xx} = \begin{bmatrix} R_{xx}[0] & \bar{R}_{xx}[1] & \cdots & \bar{R}_{xx}[K-1] \\ R_{xx}[1] & R_{xx}[0] & \cdots & \bar{R}_{xx}[K-2] \\ \vdots & \vdots & \ddots & \vdots \\ R_{xx}[K-1] & R_{xx}[K-2] & \cdots & R_{xx}[0] \end{bmatrix}.$$

Note that the matrix R_{xx} is Toeplitz, i.e., all the elements on a diagonal are equal. Assuming that R_{xx} is invertible, the optimal filter taps are then given by

$$\underline{h} = R_{xx}^{-1} R_{zx}.$$

The minimum mean-squared error is given by

$$\begin{aligned} \|Z - Y\|^2 &= \|Z\|^2 - \|Y\|^2 \\ &= E[Z\bar{Z}] - E[\underline{h}^H \bar{X} X^T \underline{h}] \\ &= E[Z\bar{Z}] - \underline{h}^H R_{xx} \underline{h} \\ &= E[Z\bar{Z}] - R_{zx}^H \underline{h}, \end{aligned}$$

where t can be ignored because the processes are WSS.

4.4.4 LMMSE Filtering in Practice

While theoretical treatments of optimal filtering often assume one has well-defined random variables with known statistics, this is rarely the case in practice. Yet, there is a very close connection between Wiener filtering and natural data driven approaches. Consider the problem from the previous section and let $x[1], x[2], \dots, x[N]$ and $z[1], z[2], \dots, z[N]$ be realizations of the random processes.

As an application, one can think of the $x[t]$ sequence as the received samples in a wireless communication system and the $z[t]$ sequence as a *pilot sequence* (i.e., known to both the transmitter and receiver). It is assumed the transmitted sequence has been convolved with an unknown LTI system. This type of degradation is known as intersymbol interference (ISI) and the goal is to find a linear filter $h[0], h[1], \dots, h[K-1]$ that removes as much ISI as possible. A suitable cost function for this goal is

$$J(\underline{h}) = \sum_{t=K}^N \lambda^{N-t} \left| z[t] - \sum_{k=0}^{K-1} h[k] x[t-k] \right|^2,$$

where the exponential weighting factor λ emphasizes the most recently received symbols because, in reality, the channel conditions are changing with time.

Using the vector $\underline{z} = [z[K] \ z[K+1] \ \cdots \ z[N]]$ and the matrix

$$A = \begin{bmatrix} x[K] & x[K-1] & \cdots & x[1] \\ x[K+1] & x[K] & \cdots & x[2] \\ \vdots & \vdots & \ddots & \vdots \\ x[N] & x[N-1] & \cdots & x[N-K+1] \end{bmatrix},$$

we can rewrite this cost function as

$$J(\underline{h}) = (\underline{A}\underline{h} - \underline{z})^H \Lambda (\underline{A}\underline{h} - \underline{z}),$$

where Λ is a diagonal matrix with diagonal entries $[\lambda^{N-K} \ \lambda^{N-K+1} \ \cdots \ \lambda^0]$. Using the orthogonality principle, one finds that the optimal solution is given by the normal equation

$$A^H \Lambda A \underline{h} = A^H \Lambda \underline{z}.$$

To see the connection with Wiener filtering, the key observation is that the matrix $A^H \Lambda A$ and the vector $A^H \Lambda \underline{z}$ are sample-average estimates of the correlation matrix and cross-correlation vector. This is because, for large N and λ close to 1, we have

$$[A^H \Lambda A]_{ij} = \sum_{t=K}^N \lambda^{N-t} x[t-j+1] \bar{x}[t-i+1] \approx \frac{R_{xx}(i-j)}{1-\lambda}$$

and

$$[A^H \Lambda \underline{z}]_i = \sum_{t=K}^N \lambda^{N-t} z[t] \bar{x}[t-i+1] \approx \frac{R_{zx}(i)}{1-\lambda}.$$

Another benefit of this approach is that, as each new sample arrives, the solution \underline{h} can be updated with low complexity. Consider the matrix $G_N = A^H \Lambda A$ and vector $\underline{b}_N = A^H \Lambda \underline{z}$ as a function of N . Then, $G_{N+1} = \lambda G_N + \underline{u}^H \underline{u}$ and $\underline{b}_{N+1} = \lambda \underline{b}_N + z[N+1] \underline{u}^H$, where

$$\underline{u} = [x[N+1] \ x[N] \ \cdots \ x[N-K+2]]. \quad (4.8)$$

The updated solution vector $\underline{h}_{N+1} = G_{N+1}^{-1} \underline{b}_{N+1}$ can be computed efficiently using the Sherman-Morrison matrix inversion formula.

4.5 Dual Approximation

4.5.1 Minimum-Norm Solutions

In many cases, one is interested in finding the minimum-norm vector that satisfies some feasibility constraints. For example, an underdetermined system of linear equations has an infinite number of solutions. But, in practice, it often makes sense to prefer the minimum-norm solution over other solutions. Finding this solution is very similar to finding the best approximation.

Let V be a Hilbert space and $\underline{w}_1, \underline{w}_2, \dots, \underline{w}_n$ be a set of linearly independent vectors in V . For any $\underline{v} \in V$, consider finding the scalars s_1, s_2, \dots, s_n that minimize

$$\left\| \underline{v} - \sum_{i=1}^n s_i \underline{w}_i \right\|.$$

The answer is clearly given by the best approximation of \underline{v} by vectors in the span of $\underline{w}_1, \underline{w}_2, \dots, \underline{w}_n$. The orthogonality principle tells us that s_1, s_2, \dots, s_n must satisfy

$$\begin{bmatrix} \langle \underline{w}_1, \underline{w}_1 \rangle & \langle \underline{w}_2, \underline{w}_1 \rangle & \cdots & \langle \underline{w}_n, \underline{w}_1 \rangle \\ \langle \underline{w}_1, \underline{w}_2 \rangle & \langle \underline{w}_2, \underline{w}_2 \rangle & \cdots & \langle \underline{w}_n, \underline{w}_2 \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle \underline{w}_1, \underline{w}_n \rangle & \langle \underline{w}_2, \underline{w}_n \rangle & \cdots & \langle \underline{w}_n, \underline{w}_n \rangle \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \\ \vdots \\ s_n \end{bmatrix} = \begin{bmatrix} \langle \underline{v}, \underline{w}_1 \rangle \\ \langle \underline{v}, \underline{w}_2 \rangle \\ \vdots \\ \langle \underline{v}, \underline{w}_n \rangle \end{bmatrix}. \quad (4.9)$$

The same problem can also be posed in a different manner.

Theorem 4.5.1. *Let V be a Hilbert space and $\underline{w}_1, \underline{w}_2, \dots, \underline{w}_n$ be a set of linearly independent vectors in V . The **dual approximation** problem is to find the vector $\underline{w} \in V$ of minimum-norm that satisfies $\langle \underline{w}, \underline{w}_i \rangle = c_i$ for $i = 1, \dots, n$. This vector is given by*

$$\underline{w} = \sum_{i=1}^n s_i \underline{w}_i, \quad (4.10)$$

where the coefficients s_1, s_2, \dots, s_n can be found by solving (4.9) with $\langle \underline{v}, \underline{w}_i \rangle = c_i$.

Proof. Let $W = \text{span}(\underline{w}_1, \underline{w}_2, \dots, \underline{w}_n)$ and notice that the subset

$$A = \{ \underline{u} \in V \mid \langle \underline{u}, \underline{w}_i \rangle = c_i \forall i = 1, \dots, n \}$$

is simply the orthogonal complement W^\perp translated by some fixed vector $\underline{v} \in A$. This is because, for all $\underline{x} \in W^\perp$, we have $\langle \underline{x}, \underline{w}_i \rangle = 0$ and $\langle \underline{v} - \underline{x}, \underline{w}_i \rangle = c_i$.

Thus, there is a one-to-one correspondence between $\underline{u} \in A$ and $\underline{x} \in W^\perp$ given by $\underline{u} = \underline{v} - \underline{x}$ and this implies

$$\min_{\underline{u} \in A} \|\underline{u}\| = \min_{\underline{x} \in W^\perp} \|\underline{v} - \underline{x}\|.$$

Since the optimizer \underline{x}^* of the RHS expression is the orthogonal projection of \underline{v} onto W^\perp , the resulting error vector $\underline{v} - \underline{x}^*$ must be contained in $(W^\perp)^\perp = W$. With this knowledge, we let $\underline{w} = \underline{v} - \underline{x}^*$ and use (4.9) to solve for s_1, \dots, s_n because we know the RHS satisfies $\langle \underline{v}, \underline{w}_i \rangle = c_i$. Since $\underline{w}_1, \dots, \underline{w}_n$ are linearly independent, the Gramian is full rank and this linear system has a unique solution. \square

4.5.2 Underdetermined Linear Systems

Let $A \in \mathbb{C}^{m \times n}$ with $m < n$ be the matrix representation of an underdetermined system of linearly independent equations and $\underline{v} \in \mathbb{C}^m$ be any column vector. Then, the dual approximation theorem can be applied to solve the problem

$$\min_{\underline{s}: A\underline{s} = \underline{v}} \|\underline{s}\|.$$

To see this as a dual approximation, we can rewrite the constraint as $(A^H)^H \underline{s} = \underline{v}$ so that the columns of A^H define the vectors $\underline{w}_1, \dots, \underline{w}_n$ in Theorem 4.5.1. Then, the theorem concludes that the minimum norm solution lies in $\mathcal{R}(A^H)$ (i.e., the column space of A^H). Using (4.10), one gathers that $\hat{\underline{s}} = A^H \underline{t}$ for some \underline{t} with $A(A^H \underline{t}) = \underline{v}$. If the rows of A are linearly independent, then the columns of A^H are linearly independent and $(AA^H)^{-1}$ exists. In this case, the solution $\hat{\underline{s}}$ can be obtained in closed form and is given by

$$\hat{\underline{s}} = A^H (AA^H)^{-1} \underline{v}.$$

4.6 Projection onto Convex Sets

So far, we have focused on the projection of vectors onto subspaces. In this section, similar results are obtained for the projection of vectors onto convex sets.

Definition 4.6.1. *Let V be a vector space. The subset $A \subseteq V$ is called a **convex set** if, for all $\underline{a}_1, \underline{a}_2 \in A$ and $\lambda \in (0, 1)$, we have $\lambda \underline{a}_1 + (1 - \lambda) \underline{a}_2 \in A$. The set is **strictly convex** if, for all $\underline{a}_1, \underline{a}_2 \in \bar{A}$ and $\lambda \in (0, 1)$, we have $\lambda \underline{a}_1 + (1 - \lambda) \underline{a}_2 \in A^\circ$.*

this to $\underline{x} = \underline{v} - \underline{u}_n$ and $\underline{y} = \underline{v} - \underline{u}_m$ gives

$$\begin{aligned} \|\underline{u}_m - \underline{u}_n\|^2 &= \|(\underline{v} - \underline{u}_n) - (\underline{v} - \underline{u}_m)\|^2 \\ &= 2\|\underline{v} - \underline{u}_n\|^2 + 2\|\underline{v} - \underline{u}_m\|^2 - \|(\underline{v} - \underline{u}_n) + (\underline{v} - \underline{u}_m)\|^2 \\ &= 2\|\underline{v} - \underline{u}_n\|^2 + 2\|\underline{v} - \underline{u}_m\|^2 - 4\left\|\underline{v} - \frac{\underline{u}_n + \underline{u}_m}{2}\right\|^2 \\ &\leq 2\|\underline{v} - \underline{u}_n\|^2 + 2\|\underline{v} - \underline{u}_m\|^2 - 4d^2 \end{aligned}$$

because the convexity of A implies $\frac{\underline{u}_n + \underline{u}_m}{2} \in A$ and therefore $\left\|\underline{v} - \frac{\underline{u}_n + \underline{u}_m}{2}\right\|^2 \geq d^2$. Since the limit of the RHS (as $m, n \rightarrow \infty$) equals 0, we find that the sequence \underline{u}_n is Cauchy and therefore the limit \underline{u}^* must exist. Since $\underline{u}_n \in A$ and A is closed, we also see that $\underline{u}^* \in A$. Therefore, the infimum is achieved as a minimum.

Uniqueness can be seen by assuming instead that $\underline{u}_m, \underline{u}_n$ are two elements in A which are both at a distance d from \underline{v} . Then, the above derivation shows that $\|\underline{u}_m - \underline{u}_n\|^2 \leq 0$. Therefore, they are the same point. \square

Remark 4.6.6. *The same result holds for norm projections in many other Banach spaces including L^p and ℓ^p for $1 < p < \infty$. In general, it is required that the Banach space be strictly convex (for uniqueness) and reflexive (for existence).*

Earlier in this chapter, we studied the equivalence between the orthogonality and Hilbert-space projections onto subspaces. The following result can be seen as a generalization of that result to Hilbert-space projections onto convex sets.

Theorem 4.6.7. *For any $\underline{v} \notin A$, a necessary and sufficient condition for $\underline{u}^* = P_A(\underline{v})$ is that $\operatorname{Re}\langle \underline{v} - \underline{u}^*, \underline{u} - \underline{u}^* \rangle \leq 0$ for all $\underline{u} \in A$.*

Proof. Let $\underline{u}^* = P_A(\underline{v})$ be the unique projection of \underline{v} onto A . For all $\underline{u} \in A$ and any $\alpha \in (0, 1)$, observe that $\underline{u}' = (1 - \alpha)\underline{u}^* + \alpha\underline{u} = \underline{u}^* + \alpha(\underline{u} - \underline{u}^*) \in A$ due to convexity. The optimality of \underline{u}^* implies that

$$\begin{aligned} \|\underline{v} - \underline{u}^*\|^2 &\leq \|\underline{v} - \underline{u}'\|^2 \\ &\leq \|\underline{v} - \underline{u}^* - \alpha(\underline{u} - \underline{u}^*)\|^2 \\ &= \|\underline{v} - \underline{u}^*\|^2 + \alpha^2\|\underline{u} - \underline{u}^*\|^2 - 2\alpha\operatorname{Re}\langle \underline{v} - \underline{u}^*, \underline{u} - \underline{u}^* \rangle. \end{aligned}$$

Thus, $\operatorname{Re}\langle \underline{v} - \underline{u}^*, \underline{u} - \underline{u}^* \rangle \leq \frac{\alpha}{2}\|\underline{u} - \underline{u}^*\|^2$. One can establish necessity by taking the limit as $\alpha \rightarrow 0$. For sufficiency, we assume $\operatorname{Re}\langle \underline{v} - \underline{u}^*, \underline{u} - \underline{u}^* \rangle \leq 0$ and we

write

$$\begin{aligned} \|\underline{v} - \underline{u}\|^2 - \|\underline{v} - \underline{u}^*\|^2 &= \|(\underline{v} - \underline{u}^*) - (\underline{u} - \underline{u}^*)\|^2 - \|\underline{v} - \underline{u}^*\|^2 \\ &= \|\underline{v} - \underline{u}^*\|^2 + \|\underline{u} - \underline{u}^*\|^2 - 2\operatorname{Re}\langle \underline{v} - \underline{u}^*, \underline{u} - \underline{u}^* \rangle - \|\underline{v} - \underline{u}^*\|^2 \\ &\geq 0. \end{aligned}$$

Thus, $\|\underline{v} - \underline{u}\|^2 \geq \|\underline{v} - \underline{u}^*\|^2$ for all $\underline{u} \in A$ and $\underline{u}^* = P_A(\underline{v})$. \square

4.6.1 Projection Properties and Examples

Let A be a closed convex subset of a Hilbert space V over \mathbb{R} . By drawing a simple picture (e.g., see Figure 4.1), one can see that projecting \underline{v} onto A is an operation that is translation invariant. Specifically, this means that translating the set A and the vector \underline{v} by the same vector \underline{v}_0 results in an output that is also translated by \underline{v}_0 . Mathematically, this means that, for all $\underline{v}, \underline{v}_0 \in V$, the projection onto V satisfies

$$\begin{aligned} P_{A+\underline{v}_0}(\underline{v} + \underline{v}_0) &= \arg \min_{\underline{u} \in A+\underline{v}_0} \|\underline{u} - \underline{v} - \underline{v}_0\| \\ &= \underline{v}_0 + \arg \min_{\underline{u}' \in A} \|(\underline{u}' + \underline{v}_0) - \underline{v} - \underline{v}_0\| \\ &= \underline{v}_0 + \arg \min_{\underline{u}' \in A} \|\underline{u}' - \underline{v}\| \\ &= \underline{v}_0 + P_A(\underline{v}). \end{aligned}$$

This also leads to the following trick. If a projection is easy when the set is centered, then one can: (i) translate the problem so that the set is centered, (ii) project onto the centered set, and (iii) translate back.

Using the best approximation theorem, it is easy to verify that the orthogonal projection of $\underline{v} \in V$ onto a one-dimensional subspace $W = \operatorname{span}(\underline{w})$ is given by

$$P_W(\underline{v}) = \frac{\langle \underline{v}, \underline{w} \rangle}{\|\underline{w}\|^2} \underline{w}.$$

A **hyperplane** is a closed subspace of $U \subset V$ that satisfies a single linear equality of the form $\langle \underline{v}, \underline{w} \rangle = 0$ for all $\underline{v} \in U$. Such a subspace is said to have co-dimension one (e.g., if $\dim(V) = n$, then $\dim(U) = n - 1$). Equivalently, U can be seen as the orthogonal complement of a one-dimensional subspace (e.g., $U = W^\perp$). Thus, we can write

$$P_U(\underline{v}) = P_{W^\perp}(\underline{v}) = \underline{v} - \frac{\langle \underline{v}, \underline{w} \rangle}{\|\underline{w}\|^2} \underline{w}.$$

Similarly, a linear equality such as $\langle \underline{v}, \underline{w} \rangle = c$, where \underline{v}_0 is any vector in V satisfying $\langle \underline{v}_0, \underline{w} \rangle = c$, defines an **affine hyperplane**. This is the shifted subspace $U + \underline{v}_0$ of co-dimension one because

$$\langle \underline{v}, \underline{w} \rangle = \langle \underline{u} + \underline{v}_0, \underline{w} \rangle = \langle \underline{u}, \underline{w} \rangle + \langle \underline{v}_0, \underline{w} \rangle = 0 + c = c.$$

Thus, we can project onto $U + \underline{v}_0$ by translating, projecting, and then translating back. This gives

$$P_{U+\underline{v}_0}(\underline{v}) = \left((\underline{v} - \underline{v}_0) - \frac{\langle \underline{v} - \underline{v}_0, \underline{w} \rangle}{\|\underline{w}\|^2} \underline{w} \right) + \underline{v}_0 = \underline{v} - \frac{\langle \underline{v}, \underline{w} \rangle - c}{\|\underline{w}\|^2} \underline{w},$$

which does not depend on the choice of \underline{v}_0 .

Next, let H be the subset of $\underline{v} \in V$ satisfying the linear inequality $\langle \underline{v}, \underline{w} \rangle \geq c$. Then, H is a closed convex set known as a **half space**. For any $\underline{v} \in H$, we have $P_H(\underline{v}) = \underline{v}$ and, for any $\underline{v} \notin H$, we have $P_H(\underline{v}) = P_{U+\underline{v}_0}(\underline{v})$ because the closest point must lie on the separating hyperplane and achieve the inequality with equality. For any $\underline{v} \in H$, one can put these together to see that

$$P_H(\underline{v}) = \begin{cases} \underline{v} & \text{if } \langle \underline{v}, \underline{w} \rangle \geq c \\ \underline{v} - \frac{\langle \underline{v}, \underline{w} \rangle - c}{\|\underline{w}\|^2} \underline{w} & \text{if } \langle \underline{v}, \underline{w} \rangle < c. \end{cases} \quad (4.11)$$

Theorem 4.6.8. *Let V be a Hilbert space over \mathbb{R} and $A \subset V$ be a closed convex set. For any $\underline{v} \notin A$, there is an affine hyperplane $U' = \{\underline{u} \in V \mid \langle \underline{u}, \underline{w} \rangle = c\}$ (defined by $\underline{w} \in V$ and $c \in \mathbb{R}$) such that $\langle \underline{v}, \underline{w} \rangle < c$ and $\langle \underline{u}, \underline{w} \rangle \geq c$ for all $\underline{u} \in A$.*

Proof. Let $\underline{u}^* = P_A(\underline{v})$ be the orthogonal projection of \underline{v} onto A and define $\underline{w} = \underline{u}^* - \underline{v}$ and $c = \langle \underline{u}^*, \underline{w} \rangle$. From Theorem 4.6.7, we see that $\langle \underline{v} - \underline{u}^*, \underline{u} - \underline{u}^* \rangle \leq 0$ for all $\underline{u} \in A$. Thus, for all $\underline{u} \in A$, we have

$$\begin{aligned} \langle \underline{u}, \underline{w} \rangle &= \langle \underline{w}, \underline{u} \rangle = \langle \underline{u}^* - \underline{v}, \underline{u} \rangle = -\langle \underline{v} - \underline{u}^*, \underline{u} \rangle \\ &\stackrel{(a)}{\geq} -\langle \underline{v} - \underline{u}^*, \underline{u}^* \rangle = \langle \underline{u}^* - \underline{v}, \underline{u}^* \rangle \\ &= \langle \underline{w}, \underline{u}^* \rangle = \langle \underline{u}^*, \underline{w} \rangle = c, \end{aligned}$$

where (a) follows from $\langle \underline{v} - \underline{u}^*, \underline{u} - \underline{u}^* \rangle \leq 0$ for all $\underline{u} \in A$. For $\langle \underline{v}, \underline{w} \rangle$, we observe that together, $\underline{u}^* \in A$ and $\underline{v} \notin A$, imply that

$$\begin{aligned} 0 &< \|\underline{u}^* - \underline{v}\|^2 = \langle \underline{u}^* - \underline{v}, \underline{u}^* - \underline{v} \rangle \\ &= \langle \underline{u}^*, \underline{u}^* - \underline{v} \rangle - \langle \underline{v}, \underline{u}^* - \underline{v} \rangle = c - \langle \underline{v}, \underline{u}^* - \underline{v} \rangle, \end{aligned}$$

which shows that $\langle \underline{v}, \underline{w} \rangle < c$ and completes the proof. \square

Corollary 4.6.9 (Farkas' Lemma). *Let $B \in \mathbb{R}^{m \times n}$ be a matrix and $\underline{v} \in \mathbb{R}^m$ be a vector. Then, exactly one of the following two conditions hold:*

1. *There exists $\underline{s} \in \mathbb{R}^n$ such that $B\underline{s} = \underline{v}$ and $\underline{s} \geq \underline{0}$.*
2. *There exists $\underline{w} \in \mathbb{R}^m$ such that $\underline{w}^T B \geq \underline{0}$ and $\underline{w}^T \underline{v} < 0$.*

Proof. Consider the closed convex set $A = \{B\underline{s} \mid \underline{s} \geq \underline{0}\}$, where $\underline{s} \geq \underline{0}$ denotes $s_i \geq 0$ for $i = 1, \dots, n$. If $\underline{v} \in A$, then we see that the first condition holds.

If $\underline{v} \notin A$, then we apply Theorem 4.6.8 using the standard Hilbert space \mathbb{R}^m to see that there exists $\underline{w} \in \mathbb{R}^m$ and $c \in \mathbb{R}$ such that $\underline{w}^T \underline{v} < c$ and $\underline{w}^T \underline{u} \geq c$ for all $\underline{u} \in A$. Next, we observe that A is a cone (i.e., for all $\alpha \geq 0$ and $\underline{u} \in A$, we have $\alpha \underline{u} \in A$). Thus, if there is any $\underline{u} \in A$ such that $\underline{w}^T \underline{u} < 0$, then $\underline{w}^T \underline{u}$ can take any non-positive value for some $\underline{u} \in A$. This gives a contradiction and shows that $\underline{w}^T \underline{u} \geq 0$ for all $\underline{u} \in A$. Since $\underline{0} \in A$, $c > 0$ would also give a contradiction and it follows that $c \leq 0$ and $\underline{w}^T \underline{v} < 0$.

This establishes that either the first or second condition must hold. But, if they both hold, then we get the contradiction

$$0 > \underline{w}^T \underline{v} = \underline{w}^T (B\underline{s}) = (\underline{w}^T B)\underline{s} \geq 0,$$

where the last step holds because $\underline{w}^T B \geq \underline{0}$ and $\underline{s} \geq \underline{0}$. Thus, exactly one holds. \square

Theorem 4.6.10. *Let V be Hilbert space over \mathbb{R} and $A \subset V$ be a closed convex set. Then, A equals the intersection of a set of half spaces.*

Proof. Let \mathcal{H} be the set of all half spaces in V and let $\mathcal{G} = \{H \in \mathcal{H} \mid A \subseteq H\}$ be the subset of half spaces containing A . For example, consider the half spaces defined by tangent planes passing through points on the boundary of A . Let $B = \bigcap_{H \in \mathcal{G}} H$ be the intersection of all the half spaces in \mathcal{G} . Since each half space contains A , it is clear that $A \subseteq B$. To show that $A = B$, we will show that $(x \notin A) \rightarrow (x \notin B)$. If $x \notin A$, then Theorem 4.6.8 shows that there is an affine hyperplane that separates x from A and an associated half space G that contains A but not x . Since G contains A , it follows that $G \in \mathcal{G}$. But, $x \notin G$ implies $x \notin B$ because B is the intersection of the half spaces in \mathcal{G} . This completes the proof. \square

4.6.2 Minimum Distance Between Two Convex Sets

Now, consider the smallest distance between two disjoint closed convex sets $A, B \subseteq V$. In this case, a unique solution may exist but a some things can go wrong. If the two sets are not strictly convex (e.g., consider two squares), then it is clearly possible for their to multiple pairs of points that achieve the minimum distance. Even if the two sets are strictly convex, one may find that the infimum is achieved as the points wander off to infinity. For example, consider the strictly convex hyperbolic sets $A = \{(x, y) | x^2 - y^2 \geq 1, x > 0\}$ and $B = \{(x, y) | y^2 - x^2 \geq 1, y \geq 0\}$. These two sets share the line $x = y > 0$ as an asymptote, so their infimal distance is 0.

To understand this behavior, we first note that the distance $f(\underline{u}, \underline{v}) = \|\underline{u} - \underline{v}\|$ is a convex function on the convex product set $A \times B$. It follows that any local minimum value is a global minimum value distance.

Theorem 4.6.11. *Let V be a Hilbert space and consider the infimal distance*

$$d = \inf_{\underline{u} \in A, \underline{v} \in B} \|\underline{u} - \underline{v}\|$$

between two disjoint closed convex sets $A, B \subseteq V$. If either set is compact, then the infimum is achieved. If the infimum is achieved and either set is strictly convex, then the minimizing points $\underline{u}^, \underline{v}^*$ are unique.*

Proof. Consider any sequence $(\underline{u}_1, \underline{v}_1), (\underline{u}_2, \underline{v}_2), \dots \in A \times B$ that satisfies

$$\lim_{n \rightarrow \infty} \|\underline{u}_n - \underline{v}_n\| = d.$$

If B is compact, then there is a subsequence \underline{v}_{n_j} that converges to some $\underline{v}^* \in B$. Since $\|P_A(\underline{v}_n) - \underline{v}_n\| \leq \|\underline{u}_n - \underline{v}_n\|$, we can replace \underline{u}_{n_j} by $P_A(\underline{v}_{n_j})$ and still achieve the infimum. The continuity of P_A also shows that $\underline{u}_{n_j} \rightarrow \underline{u}^* = P_A(\underline{v}^*)$ and this implies $\|\underline{u}^* - \underline{v}^*\| = d$. Also, $P_B(\underline{u}^*) = \underline{v}^*$ because \underline{v}^* is the unique closest point in B to \underline{u}^* . Notice that \underline{v}^* may not be unique (due to the subsequence construction) and, thus, the pair $(\underline{u}^*, \underline{v}^*)$ is not unique in general.

Since $\|\underline{u} - \underline{v}\|$ is a convex function on the convex product set $A \times B$, there is a (possibly empty) convex set of minimizers

$$M = \{(\underline{u}, \underline{v}) \in A \times B \mid \|\underline{u} - \underline{v}\| = d\}.$$

Also, each component of the $(\underline{u}, \underline{v})$ points in M must lie on the boundary of its set because otherwise one could reduce the smallest distance by moving one point along the minimum distance line towards the boundary. Now, suppose that (i) A is strictly convex and (ii) M contains more than one pair of minimizers. Then, condition (ii) implies that there must be two boundary points $\underline{u}_1, \underline{u}_2 \in \partial A$ such that $\alpha \underline{u}_1 + (1 - \alpha) \underline{u}_2 \in \partial A$ for $\alpha \in [0, 1]$. But this contradicts condition (i) and shows that, if A is strictly convex, then there is at most one pair $(\underline{u}^*, \underline{v}^*) \in M$ of minimizing points. \square

Remark 4.6.12. Finding the minimum distance between two disjoint closed convex sets $A, B \subseteq V$ is a classic problem that is solved nicely by the idea of alternating minimization. Let $\underline{v}_0 \in B$ be an arbitrary initial point and define

$$\begin{aligned}\underline{u}_{n+1} &= \arg \min_{\underline{u} \in A} \|\underline{u} - \underline{v}_n\| \\ \underline{v}_{n+1} &= \arg \min_{\underline{v} \in B} \|\underline{u}_{n+1} - \underline{v}\|.\end{aligned}$$

Notice that the sequence $d_n = \|\underline{u}_n - \underline{v}_n\|$ is non-increasing and must therefore have a limit. By adapting the previous proof, one can show that, if either set is compact, then the sequence $(\underline{u}_n, \underline{v}_n)$ converges to a pair of vectors that minimize the distance.

Theorem 4.6.13. Let V be a Hilbert space over \mathbb{R} and A, B be disjoint closed convex subsets of V . If either set is compact, then there is an affine hyperplane $\{\underline{a} \in V \mid \langle \underline{a}, \underline{w} \rangle = c\}$ (defined by $\underline{w} \in V$ and $c \in \mathbb{R}$) such that $\langle \underline{u}, \underline{w} \rangle > c$ for all $\underline{u} \in A$ and $\langle \underline{u}, \underline{w} \rangle < c$ for all $\underline{u} \in B$.

Proof. Applying Theorem 4.6.11 gives a pair of points $(\underline{u}^*, \underline{v}^*) \in A \times B$ that minimize the distance and satisfy $\underline{u}^* = P_A(\underline{v}^*)$ and $\underline{v}^* = P_B(\underline{u}^*)$. Applying Theorem 4.6.8 to $P_A(\underline{v}^*)$ shows that $\langle \underline{u}, \underline{u}^* - \underline{v}^* \rangle \geq \langle \underline{u}^*, \underline{u}^* - \underline{v}^* \rangle$ for all $\underline{u} \in A$. Similarly, applying Theorem 4.6.8 to $P_B(\underline{u}^*)$ shows that $\langle \underline{v}, \underline{v}^* - \underline{u}^* \rangle \geq \langle \underline{v}^*, \underline{v}^* - \underline{u}^* \rangle$ for all $\underline{v} \in B$. Negating this gives $\langle \underline{v}, \underline{u}^* - \underline{v}^* \rangle \leq \langle \underline{v}^*, \underline{u}^* - \underline{v}^* \rangle$. Now, we observe that $\langle \underline{u}^*, \underline{u}^* - \underline{v}^* \rangle - \langle \underline{v}^*, \underline{u}^* - \underline{v}^* \rangle = \|\underline{u}^* - \underline{v}^*\|^2 > 0$ because $\underline{u}^* \neq \underline{v}^*$. Thus, we can choose $\underline{w} = \underline{u}^* - \underline{v}^*$ and $c = \frac{1}{2}(\langle \underline{u}^*, \underline{u}^* - \underline{v}^* \rangle + \langle \underline{v}^*, \underline{u}^* - \underline{v}^* \rangle)$ to guarantee that $\langle \underline{u}, \underline{w} \rangle > c$ for all $\underline{u} \in A$ and $\langle \underline{u}, \underline{w} \rangle < c$ for all $\underline{u} \in B$. \square