Linear Algebra 2

Lecture #16

The orthogonal complement of a subspace. Orthogonal projection onto a subspace

Irena Penev

March 13, 2024

- This lecture has three parts:

- This lecture has three parts:
  1. The orthogonal complement of a subspace

- This lecture has three parts:
    1. The orthogonal complement of a subspace
    2. Orthogonal projection onto a subspace

- This lecture has three parts:
  1. The orthogonal complement of a subspace
  2. Orthogonal projection onto a subspace
  3. Orthogonal projection onto subspaces of $\mathbb{R}^n$

1. The orthogonal complement of a subspace

1. The orthogonal complement of a subspace

Definition

Let V be a real or complex vector space V, equipped with a scalar product $\langle \cdot, \cdot \rangle$. For a set $A \subseteq V$,[a] the *orthogonal complement* of A, denoted by $A^\perp$, is the set of all vectors in V that are orthogonal to A.

[a] Here, A may or may not be a subspace of V.

- Thus, we have the following:

$$\begin{aligned} A^\perp &= \{\mathbf{v} \in V \mid \mathbf{v} \perp A\} \\ &= \{\mathbf{v} \in V \mid \mathbf{v} \perp \mathbf{a} \ \ \forall \mathbf{a} \in A\} \\ &= \{\mathbf{v} \in V \mid \langle \mathbf{v}, \mathbf{a} \rangle = 0 \ \ \forall \mathbf{a} \in A\}. \end{aligned}$$

Proposition 6.4.1

Let V be a real or complex vector space, equipped with a scalar product $\langle \cdot, \cdot \rangle$. Let $A, B \subseteq V$. Then

- (a) $A^{\perp}$ is a subspace of V;[a]
- (b) if $A \subseteq B$, then $A^{\perp} \supseteq B^{\perp}$.

[a] Note that it is possible that $A = \emptyset$. In this case, we simply get that $A^{\perp} = V$. This is because every vector in V is (vacuously) orthogonal to every vector in the empty set.

Proof (outline).

Proposition 6.4.1

Let V be a real or complex vector space, equipped with a scalar product $\langle \cdot, \cdot \rangle$. Let $A, B \subseteq V$. Then

- (a) $A^{\perp}$ is a subspace of V;[a]
- (b) if $A \subseteq B$, then $A^{\perp} \supseteq B^{\perp}$.

[a] Note that it is possible that $A = \emptyset$. In this case, we simply get that $A^{\perp} = V$. This is because every vector in V is (vacuously) orthogonal to every vector in the empty set.

Proof (outline). For (a), we simply check that $A^{\perp}$ contains $\mathbf{0}$ and is closed under vector addition and scalar multiplication (details: Lecture Notes).

Proposition 6.4.1

Let V be a real or complex vector space, equipped with a scalar product $\langle \cdot, \cdot \rangle$. Let $A, B \subseteq V$. Then

(a) $A^\perp$ is a subspace of V;[a]

(b) if $A \subseteq B$, then $A^\perp \supseteq B^\perp$.

[a] Note that it is possible that $A = \emptyset$. In this case, we simply get that $A^\perp = V$. This is because every vector in V is (vacuously) orthogonal to every vector in the empty set.

Proof (outline). For (a), we simply check that $A^\perp$ contains $\mathbf{0}$ and is closed under vector addition and scalar multiplication (details: Lecture Notes).

Part (b) is "obvious": if $A \subseteq B$, then any vector that is orthogonal to every vector in B is, in particular, orthogonal to every vector in A. $\square$

Proposition 6.4.2

Let V be a real or complex vector space, equipped with a scalar product $\langle \cdot, \cdot \rangle$. Let $\mathbf{u}_1, \dots, \mathbf{u}_k \in V$. Then $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}^\perp = \text{Span}(\mathbf{u}_1, \dots, \mathbf{u}_k)^\perp$.

Proof.

Proposition 6.4.2

Let V be a real or complex vector space, equipped with a scalar product $\langle \cdot, \cdot \rangle$. Let $\mathbf{u}_1, \dots, \mathbf{u}_k \in V$. Then $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}^\perp = \text{Span}(\mathbf{u}_1, \dots, \mathbf{u}_k)^\perp$.

Proof. Since $\{\mathbf{u}_1, \dots, \mathbf{u}_k\} \subseteq \text{Span}(\mathbf{u}_1, \dots, \mathbf{u}_k)$, Prop. 6.4.1(b) guarantees that $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}^\perp \supseteq \text{Span}(\mathbf{u}_1, \dots, \mathbf{u}_k)^\perp$.

Proposition 6.4.2

Let V be a real or complex vector space, equipped with a scalar product $\langle \cdot, \cdot \rangle$. Let $\mathbf{u}_1, \dots, \mathbf{u}_k \in V$. Then $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}^\perp = \text{Span}(\mathbf{u}_1, \dots, \mathbf{u}_k)^\perp$.

Proof. Since $\{\mathbf{u}_1, \dots, \mathbf{u}_k\} \subseteq \text{Span}(\mathbf{u}_1, \dots, \mathbf{u}_k)$, Prop. 6.4.1(b) guarantees that $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}^\perp \supseteq \text{Span}(\mathbf{u}_1, \dots, \mathbf{u}_k)^\perp$.

Let us prove the reverse inclusion.

Proposition 6.4.2

Let V be a real or complex vector space, equipped with a scalar product $\langle \cdot, \cdot \rangle$. Let $\mathbf{u}_1, \dots, \mathbf{u}_k \in V$. Then $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}^\perp = \text{Span}(\mathbf{u}_1, \dots, \mathbf{u}_k)^\perp$.

Proof. Since $\{\mathbf{u}_1, \dots, \mathbf{u}_k\} \subseteq \text{Span}(\mathbf{u}_1, \dots, \mathbf{u}_k)$, Prop. 6.4.1(b) guarantees that $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}^\perp \supseteq \text{Span}(\mathbf{u}_1, \dots, \mathbf{u}_k)^\perp$.

Let us prove the reverse inclusion. Fix $\mathbf{x} \in \{\mathbf{u}_1, \dots, \mathbf{u}_k\}^\perp$. WTS $\mathbf{x} \in \text{Span}(\mathbf{u}_1, \dots, \mathbf{u}_k)^\perp$.

Proposition 6.4.2

Let V be a real or complex vector space, equipped with a scalar product $\langle \cdot, \cdot \rangle$. Let $\mathbf{u}_1, \dots, \mathbf{u}_k \in V$. Then $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}^\perp = \text{Span}(\mathbf{u}_1, \dots, \mathbf{u}_k)^\perp$.

Proof. Since $\{\mathbf{u}_1, \dots, \mathbf{u}_k\} \subseteq \text{Span}(\mathbf{u}_1, \dots, \mathbf{u}_k)$, Prop. 6.4.1(b) guarantees that $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}^\perp \supseteq \text{Span}(\mathbf{u}_1, \dots, \mathbf{u}_k)^\perp$.

Let us prove the reverse inclusion. Fix $\mathbf{x} \in \{\mathbf{u}_1, \dots, \mathbf{u}_k\}^\perp$. WTS $\mathbf{x} \in \text{Span}(\mathbf{u}_1, \dots, \mathbf{u}_k)^\perp$. Fix $\mathbf{u} \in \text{Span}(\mathbf{u}_1, \dots, \mathbf{u}_k)$.

Proposition 6.4.2

Let V be a real or complex vector space, equipped with a scalar product $\langle \cdot, \cdot \rangle$. Let $\mathbf{u}_1, \dots, \mathbf{u}_k \in V$. Then $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}^\perp = \text{Span}(\mathbf{u}_1, \dots, \mathbf{u}_k)^\perp$.

Proof. Since $\{\mathbf{u}_1, \dots, \mathbf{u}_k\} \subseteq \text{Span}(\mathbf{u}_1, \dots, \mathbf{u}_k)$, Prop. 6.4.1(b) guarantees that $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}^\perp \supseteq \text{Span}(\mathbf{u}_1, \dots, \mathbf{u}_k)^\perp$.

Let us prove the reverse inclusion. Fix $\mathbf{x} \in \{\mathbf{u}_1, \dots, \mathbf{u}_k\}^\perp$. WTS $\mathbf{x} \in \text{Span}(\mathbf{u}_1, \dots, \mathbf{u}_k)^\perp$. Fix $\mathbf{u} \in \text{Span}(\mathbf{u}_1, \dots, \mathbf{u}_k)$. Then there exist scalars $\alpha_1, \dots, \alpha_k$ s.t. $\mathbf{u} = \alpha_1\mathbf{u}_1 + \dots + \alpha_k\mathbf{u}_k$.

Proposition 6.4.2

Let V be a real or complex vector space, equipped with a scalar product $\langle \cdot, \cdot \rangle$. Let $\mathbf{u}_1, \dots, \mathbf{u}_k \in V$. Then $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}^\perp = \text{Span}(\mathbf{u}_1, \dots, \mathbf{u}_k)^\perp$.

Proof. Since $\{\mathbf{u}_1, \dots, \mathbf{u}_k\} \subseteq \text{Span}(\mathbf{u}_1, \dots, \mathbf{u}_k)$, Prop. 6.4.1(b) guarantees that $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}^\perp \supseteq \text{Span}(\mathbf{u}_1, \dots, \mathbf{u}_k)^\perp$.

Let us prove the reverse inclusion. Fix $\mathbf{x} \in \{\mathbf{u}_1, \dots, \mathbf{u}_k\}^\perp$. WTS $\mathbf{x} \in \text{Span}(\mathbf{u}_1, \dots, \mathbf{u}_k)^\perp$. Fix $\mathbf{u} \in \text{Span}(\mathbf{u}_1, \dots, \mathbf{u}_k)$. Then there exist scalars $\alpha_1, \dots, \alpha_k$ s.t. $\mathbf{u} = \alpha_1 \mathbf{u}_1 + \dots + \alpha_k \mathbf{u}_k$. But now

$$\begin{aligned}
\langle \mathbf{u}, \mathbf{x} \rangle &= \langle \alpha_1 \mathbf{u}_1 + \dots + \alpha_k \mathbf{u}_k, \mathbf{x} \rangle \\
&= \alpha_1 \langle \mathbf{u}_1, \mathbf{x} \rangle + \dots + \alpha_k \langle \mathbf{u}_k, \mathbf{x} \rangle \\
&\stackrel{(*)}{=} \alpha_1 0 + \dots + \alpha_k 0 \;=\; 0,
\end{aligned}$$

where (*) follows from the fact that $\mathbf{x} \in \{\mathbf{u}_1, \dots, \mathbf{u}_k\}^\perp$. This proves that $\mathbf{x} \perp \mathbf{u}$, and consequently, $\mathbf{x} \in \text{Span}(\mathbf{u}_1, \dots, \mathbf{u}_k)^\perp$. $\square$

- Recall from subsection 3.1.3 of the Lecture Notes (last semester) that if V is a vector space over a field $\mathbb{F}$, and U and W are subspaces of V, then

$$U + W \quad := \quad \{\mathbf{u} + \mathbf{w} \mid \mathbf{u} \in U,\ \mathbf{w} \in W\}$$

  is a subspace of V.
- Moreover, recall from subsection 3.2.6 of the Lecture Notes that if $V = U + W$ and $U \cap W = \{\mathbf{0}\}$, then we say that V is the *direct sum* of U and W, and we write $V = U \oplus W$.

Theorem 6.4.3

Let V be a finite-dimensional real or complex vector space, equipped with a scalar product $\langle \cdot, \cdot \rangle$ and the induced norm $|| \cdot ||$. Let U be a subspace of V. Then $U^\perp$ is a subspace of V, and all the following hold:

a. if $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ is an orthogonal basis of U, and $\{\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{u}_{k+1}, \dots, \mathbf{u}_n\}$ is an extension of that basis to an orthogonal basis of V, then $\{\mathbf{u}_{k+1}, \dots, \mathbf{u}_n\}$ is an orthogonal basis of $U^\perp$;
b. if $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ is an orthonormal basis of U, and $\{\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{u}_{k+1}, \dots, \mathbf{u}_n\}$ is an extension of that basis to an orthonormal basis of V, then $\{\mathbf{u}_{k+1}, \dots, \mathbf{u}_n\}$ is an orthonormal basis of $U^\perp$;
c. $(U^\perp)^\perp = U$;
d. $V = U \oplus U^\perp$, that is, $V = U + U^\perp$ and $U \cap U^\perp = \{\mathbf{0}\}$;
e. $\dim(V) = \dim(U) + \dim(U^\perp)$.

Proof. By Proposition 6.4.1(a), $U^\perp$ is a subspace of V. It remains to prove (a)-(e).

Proof. By Proposition 6.4.1(a), $U^{\perp}$ is a subspace of V. It remains to prove (a)-(e).

Theorem 6.4.3

(a) if $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ is an orthogonal basis of U, and $\{\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{u}_{k+1}, \dots, \mathbf{u}_n\}$ is an extension of that basis to an orthogonal basis of V, then $\{\mathbf{u}_{k+1}, \dots, \mathbf{u}_n\}$ is an orthogonal basis of $U^{\perp}$;

Proof of (a).

Proof. By Proposition 6.4.1(a), $U^\perp$ is a subspace of V. It remains to prove (a)-(e).

Theorem 6.4.3

(a) if $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ is an orthogonal basis of U, and $\{\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{u}_{k+1}, \dots, \mathbf{u}_n\}$ is an extension of that basis to an orthogonal basis of V, then $\{\mathbf{u}_{k+1}, \dots, \mathbf{u}_n\}$ is an orthogonal basis of $U^\perp$;

Proof of (a). Assume that $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ is an orthogonal basis of U, and that $\{\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{u}_{k+1}, \dots, \mathbf{u}_n\}$ is an extension of that basis to an orthogonal basis of V. WTS $\{\mathbf{u}_{k+1}, \dots, \mathbf{u}_n\}$ is an orthogonal basis of $U^\perp$.

Proof. By Proposition 6.4.1(a), $U^{\perp}$ is a subspace of V. It remains to prove (a)-(e).

Theorem 6.4.3

(a) if $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ is an orthogonal basis of U, and $\{\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{u}_{k+1}, \dots, \mathbf{u}_n\}$ is an extension of that basis to an orthogonal basis of V, then $\{\mathbf{u}_{k+1}, \dots, \mathbf{u}_n\}$ is an orthogonal basis of $U^{\perp}$;

Proof of (a). Assume that $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ is an orthogonal basis of U, and that $\{\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{u}_{k+1}, \dots, \mathbf{u}_n\}$ is an extension of that basis to an orthogonal basis of V. WTS $\{\mathbf{u}_{k+1}, \dots, \mathbf{u}_n\}$ is an orthogonal basis of $U^{\perp}$. Clearly, $\{\mathbf{u}_{k+1}, \dots, \mathbf{u}_n\}$ is an orthogonal set of vectors, and so it suffices to show that $\{\mathbf{u}_{k+1}, \dots, \mathbf{u}_n\}$ is in fact a basis of $U^{\perp}$.

Proof. By Proposition 6.4.1(a), $U^\perp$ is a subspace of V. It remains to prove (a)-(e).

Theorem 6.4.3

(a) if $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ is an orthogonal basis of U, and $\{\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{u}_{k+1}, \dots, \mathbf{u}_n\}$ is an extension of that basis to an orthogonal basis of V, then $\{\mathbf{u}_{k+1}, \dots, \mathbf{u}_n\}$ is an orthogonal basis of $U^\perp$;

Proof of (a). Assume that $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ is an orthogonal basis of U, and that $\{\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{u}_{k+1}, \dots, \mathbf{u}_n\}$ is an extension of that basis to an orthogonal basis of V. WTS $\{\mathbf{u}_{k+1}, \dots, \mathbf{u}_n\}$ is an orthogonal basis of $U^\perp$. Clearly, $\{\mathbf{u}_{k+1}, \dots, \mathbf{u}_n\}$ is an orthogonal set of vectors, and so it suffices to show that $\{\mathbf{u}_{k+1}, \dots, \mathbf{u}_n\}$ is in fact a basis of $U^\perp$. We already know that $\{\mathbf{u}_{k+1}, \dots, \mathbf{u}_n\}$ is linearly independent (because it is a subset of the basis $\{\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{u}_{k+1}, \dots, \mathbf{u}_n\}$ of V), and so we need only show that $\text{Span}(\mathbf{u}_{k+1}, \dots, \mathbf{u}_n) = U^\perp$.

Theorem 6.4.3

(a) if $\{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ is an orthogonal basis of U, and $\{\mathbf{u}_1, \ldots, \mathbf{u}_k, \mathbf{u}_{k+1}, \ldots, \mathbf{u}_n\}$ is an extension of that basis to an orthogonal basis of V, then $\{\mathbf{u}_{k+1}, \ldots, \mathbf{u}_n\}$ is an orthogonal basis of $U^\perp$;

Proof of (a) (cont.). Reminder: WTS $\text{Span}(\mathbf{u}_{k+1}, \ldots, \mathbf{u}_n) = U^\perp$.

Theorem 6.4.3

(a) if $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ is an orthogonal basis of U, and $\{\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{u}_{k+1}, \dots, \mathbf{u}_n\}$ is an extension of that basis to an orthogonal basis of V, then $\{\mathbf{u}_{k+1}, \dots, \mathbf{u}_n\}$ is an orthogonal basis of $U^\perp$;

Proof of (a) (cont.). Reminder: WTS $\text{Span}(\mathbf{u}_{k+1}, \dots, \mathbf{u}_n) = U^\perp$.

We first prove that $\text{Span}(\mathbf{u}_{k+1}, \dots, \mathbf{u}_n) \supseteq U^\perp$.

Theorem 6.4.3

(a) if $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ is an orthogonal basis of U, and $\{\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{u}_{k+1}, \dots, \mathbf{u}_n\}$ is an extension of that basis to an orthogonal basis of V, then $\{\mathbf{u}_{k+1}, \dots, \mathbf{u}_n\}$ is an orthogonal basis of $U^\perp$;

Proof of (a) (cont.). Reminder: WTS $\text{Span}(\mathbf{u}_{k+1}, \dots, \mathbf{u}_n) = U^\perp$.

We first prove that $\text{Span}(\mathbf{u}_{k+1}, \dots, \mathbf{u}_n) \supseteq U^\perp$. Fix $\mathbf{x} \in U^\perp$.

Theorem 6.4.3

(a) if $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ is an orthogonal basis of U, and $\{\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{u}_{k+1}, \dots, \mathbf{u}_n\}$ is an extension of that basis to an orthogonal basis of V, then $\{\mathbf{u}_{k+1}, \dots, \mathbf{u}_n\}$ is an orthogonal basis of $U^\perp$;

Proof of (a) (cont.). Reminder: WTS $\text{Span}(\mathbf{u}_{k+1}, \dots, \mathbf{u}_n) = U^\perp$.

We first prove that $\text{Span}(\mathbf{u}_{k+1}, \dots, \mathbf{u}_n) \supseteq U^\perp$. Fix $\mathbf{x} \in U^\perp$. Then

$$\mathbf{x} \overset{(*)}{=} \sum_{i=1}^{n} \frac{\langle \mathbf{x}, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \mathbf{u}_i \overset{(**)}{=} \sum_{i=k+1}^{n} \frac{\langle \mathbf{x}, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \mathbf{u}_i,$$

where (*) follows from Theorem 6.3.5 (since $\{\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{u}_{k+1}, \dots, \mathbf{u}_n\}$ is an orthogonal basis of V) and (**) follows from the fact that $\mathbf{x} \in U^\perp$ and $\mathbf{u}_1, \dots, \mathbf{u}_k \in U$, we so $\langle \mathbf{x}, \mathbf{u}_i \rangle = 0$ for all $i \in \{1, \dots, k\}$.

Theorem 6.4.3

(a) if $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ is an orthogonal basis of U, and $\{\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{u}_{k+1}, \dots, \mathbf{u}_n\}$ is an extension of that basis to an orthogonal basis of V, then $\{\mathbf{u}_{k+1}, \dots, \mathbf{u}_n\}$ is an orthogonal basis of $U^\perp$;

Proof of (a) (cont.). Reminder: WTS $\text{Span}(\mathbf{u}_{k+1}, \dots, \mathbf{u}_n) = U^\perp$.

We first prove that $\text{Span}(\mathbf{u}_{k+1}, \dots, \mathbf{u}_n) \supseteq U^\perp$. Fix $\mathbf{x} \in U^\perp$. Then

$$\mathbf{x} \overset{(*)}{=} \sum_{i=1}^{n} \frac{\langle \mathbf{x}, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \mathbf{u}_i \overset{(**)}{=} \sum_{i=k+1}^{n} \frac{\langle \mathbf{x}, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \mathbf{u}_i,$$

where (*) follows from Theorem 6.3.5 (since $\{\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{u}_{k+1}, \dots, \mathbf{u}_n\}$ is an orthogonal basis of V) and (**) follows from the fact that $\mathbf{x} \in U^\perp$ and $\mathbf{u}_1, \dots, \mathbf{u}_k \in U$, we so $\langle \mathbf{x}, \mathbf{u}_i \rangle = 0$ for all $i \in \{1, \dots, k\}$.

Thus, $\mathbf{x}$ is a linear combination of the vectors $\mathbf{u}_{k+1}, \dots, \mathbf{u}_n$, and we deduce that $\mathbf{x} \in \text{Span}(\mathbf{u}_{k+1}, \dots, \mathbf{u}_n)$.

Theorem 6.4.3

(a) if $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ is an orthogonal basis of U, and $\{\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{u}_{k+1}, \dots, \mathbf{u}_n\}$ is an extension of that basis to an orthogonal basis of V, then $\{\mathbf{u}_{k+1}, \dots, \mathbf{u}_n\}$ is an orthogonal basis of $U^\perp$;

Proof of (a) (cont.). Reminder: WTS $\text{Span}(\mathbf{u}_{k+1}, \dots, \mathbf{u}_n) = U^\perp$.

We first prove that $\text{Span}(\mathbf{u}_{k+1}, \dots, \mathbf{u}_n) \supseteq U^\perp$. Fix $\mathbf{x} \in U^\perp$. Then

$$\mathbf{x} \overset{(*)}{=} \sum_{i=1}^{n} \frac{\langle \mathbf{x}, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \mathbf{u}_i \overset{(**)}{=} \sum_{i=k+1}^{n} \frac{\langle \mathbf{x}, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \mathbf{u}_i,$$

where (*) follows from Theorem 6.3.5 (since $\{\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{u}_{k+1}, \dots, \mathbf{u}_n\}$ is an orthogonal basis of V) and (**) follows from the fact that $\mathbf{x} \in U^\perp$ and $\mathbf{u}_1, \dots, \mathbf{u}_k \in U$, we so $\langle \mathbf{x}, \mathbf{u}_i \rangle = 0$ for all $i \in \{1, \dots, k\}$.

Thus, $\mathbf{x}$ is a linear combination of the vectors $\mathbf{u}_{k+1}, \dots, \mathbf{u}_n$, and we deduce that $\mathbf{x} \in \text{Span}(\mathbf{u}_{k+1}, \dots, \mathbf{u}_n)$. This proves that $\text{Span}(\mathbf{u}_{k+1}, \dots, \mathbf{u}_n) \supseteq U^\perp$.

Theorem 6.4.3

(a) if $\{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ is an orthogonal basis of U, and $\{\mathbf{u}_1, \ldots, \mathbf{u}_k, \mathbf{u}_{k+1}, \ldots, \mathbf{u}_n\}$ is an extension of that basis to an orthogonal basis of V, then $\{\mathbf{u}_{k+1}, \ldots, \mathbf{u}_n\}$ is an orthogonal basis of $U^\perp$;

Proof of (a) (cont.). Reminder: WTS $\text{Span}(\mathbf{u}_{k+1}, \ldots, \mathbf{u}_n) = U^\perp$.

It remains to show that $\text{Span}(\mathbf{u}_{k+1}, \ldots, \mathbf{u}_n) \subseteq U^\perp$.

Theorem 6.4.3

(a) if $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ is an orthogonal basis of U, and $\{\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{u}_{k+1}, \dots, \mathbf{u}_n\}$ is an extension of that basis to an orthogonal basis of V, then $\{\mathbf{u}_{k+1}, \dots, \mathbf{u}_n\}$ is an orthogonal basis of $U^{\perp}$;

Proof of (a) (cont.). Reminder: WTS $\text{Span}(\mathbf{u}_{k+1}, \dots, \mathbf{u}_n) = U^{\perp}$.

It remains to show that $\text{Span}(\mathbf{u}_{k+1}, \dots, \mathbf{u}_n) \subseteq U^{\perp}$. Fix an arbitrary $\mathbf{x} \in \text{Span}(\mathbf{u}_{k+1}, \dots, \mathbf{u}_n)$. WTS $\mathbf{x} \in U^{\perp}$.

Theorem 6.4.3

(a) if $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ is an orthogonal basis of U, and $\{\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{u}_{k+1}, \dots, \mathbf{u}_n\}$ is an extension of that basis to an orthogonal basis of V, then $\{\mathbf{u}_{k+1}, \dots, \mathbf{u}_n\}$ is an orthogonal basis of $U^\perp$;

Proof of (a) (cont.). Reminder: WTS $\text{Span}(\mathbf{u}_{k+1}, \dots, \mathbf{u}_n) = U^\perp$.

It remains to show that $\text{Span}(\mathbf{u}_{k+1}, \dots, \mathbf{u}_n) \subseteq U^\perp$. Fix an arbitrary $\mathbf{x} \in \text{Span}(\mathbf{u}_{k+1}, \dots, \mathbf{u}_n)$. WTS $\mathbf{x} \in U^\perp$. Fix scalars $\alpha_{k+1}, \dots, \alpha_n$ such that

$$\mathbf{x} = \alpha_{k+1}\mathbf{u}_{k+1} + \dots + \alpha_n\mathbf{u}_n.$$

Theorem 6.4.3

(a) if $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ is an orthogonal basis of U, and $\{\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{u}_{k+1}, \dots, \mathbf{u}_n\}$ is an extension of that basis to an orthogonal basis of V, then $\{\mathbf{u}_{k+1}, \dots, \mathbf{u}_n\}$ is an orthogonal basis of $U^\perp$;

Proof of (a) (cont.). Reminder: WTS $\text{Span}(\mathbf{u}_{k+1}, \dots, \mathbf{u}_n) = U^\perp$.

It remains to show that $\text{Span}(\mathbf{u}_{k+1}, \dots, \mathbf{u}_n) \subseteq U^\perp$. Fix an arbitrary $\mathbf{x} \in \text{Span}(\mathbf{u}_{k+1}, \dots, \mathbf{u}_n)$. WTS $\mathbf{x} \in U^\perp$. Fix scalars $\alpha_{k+1}, \dots, \alpha_n$ such that

$$\mathbf{x} = \alpha_{k+1}\mathbf{u}_{k+1} + \dots + \alpha_n\mathbf{u}_n.$$

Fix any $\mathbf{u} \in U$; we must show that $\mathbf{x} \perp \mathbf{u}$. Since $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ is a basis of U, we know that there exist scalars $\alpha_1, \dots, \alpha_k$ such that

$$\mathbf{u} = \alpha_1\mathbf{u}_1 + \dots + \alpha_k\mathbf{u}_k.$$

Theorem 6.4.3

(a) if $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ is an orthogonal basis of U, and $\{\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{u}_{k+1}, \dots, \mathbf{u}_n\}$ is an extension of that basis to an orthogonal basis of V, then $\{\mathbf{u}_{k+1}, \dots, \mathbf{u}_n\}$ is an orthogonal basis of $U^\perp$;

Proof of (a) (cont.). Reminder: WTS $\text{Span}(\mathbf{u}_{k+1}, \dots, \mathbf{u}_n) = U^\perp$.

It remains to show that $\text{Span}(\mathbf{u}_{k+1}, \dots, \mathbf{u}_n) \subseteq U^\perp$. Fix an arbitrary $\mathbf{x} \in \text{Span}(\mathbf{u}_{k+1}, \dots, \mathbf{u}_n)$. WTS $\mathbf{x} \in U^\perp$. Fix scalars $\alpha_{k+1}, \dots, \alpha_n$ such that

$$\mathbf{x} \quad = \quad \alpha_{k+1}\mathbf{u}_{k+1} + \dots + \alpha_n\mathbf{u}_n.$$

Fix any $\mathbf{u} \in U$; we must show that $\mathbf{x} \perp \mathbf{u}$. Since $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ is a basis of U, we know that there exist scalars $\alpha_1, \dots, \alpha_k$ such that

$$\mathbf{u} \quad = \quad \alpha_1\mathbf{u}_1 + \dots + \alpha_k\mathbf{u}_k.$$

Since $\{\mathbf{u}_1, \dots, \mathbf{u}_k\} \perp \{\mathbf{u}_{k+1}, \dots, \mathbf{u}_n\}$, it readily follows that $\mathbf{x} \perp \mathbf{u}$ (details: Lecture Notes), and we deduce that $\text{Span}(\mathbf{u}_{k+1}, \dots, \mathbf{u}_n) \subseteq U^\perp$. This proves (a).

Theorem 6.4.3

(a) if $\{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ is an orthogonal basis of U, and $\{\mathbf{u}_1, \ldots, \mathbf{u}_k, \mathbf{u}_{k+1}, \ldots, \mathbf{u}_n\}$ is an extension of that basis to an orthogonal basis of V, then $\{\mathbf{u}_{k+1}, \ldots, \mathbf{u}_n\}$ is an orthogonal basis of $U^\perp$;

(b) if $\{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ is an orthonormal basis of U, and $\{\mathbf{u}_1, \ldots, \mathbf{u}_k, \mathbf{u}_{k+1}, \ldots, \mathbf{u}_n\}$ is an extension of that basis to an orthonormal basis of V, then $\{\mathbf{u}_{k+1}, \ldots, \mathbf{u}_n\}$ is an orthonormal basis of $U^\perp$;

Proof of (b). Part (b) follows immediately from part (a).

Theorem 6.4.3

(c) $(U^{\perp})^{\perp} = U$;

(d) $V = U \oplus U^{\perp}$, that is, $V = U + U^{\perp}$ and $U \cap U^{\perp} = \{\mathbf{0}\}$;

(e) $\dim(V) = \dim(U) + \dim(U^{\perp})$.

Proof (continued). It remains to prove (c), (d), and (e).

Theorem 6.4.3

(c) $(U^{\perp})^{\perp} = U$;

(d) $V = U \oplus U^{\perp}$, that is, $V = U + U^{\perp}$ and $U \cap U^{\perp} = \{\mathbf{0}\}$;

(e) $\dim(V) = \dim(U) + \dim(U^{\perp})$.

Proof (continued). It remains to prove (c), (d), and (e).

First, since V is finite-dimensional, so is U.

Theorem 6.4.3

(c) $(U^{\perp})^{\perp} = U$;

(d) $V = U \oplus U^{\perp}$, that is, $V = U + U^{\perp}$ and $U \cap U^{\perp} = \{\mathbf{0}\}$;

(e) $\dim(V) = \dim(U) + \dim(U^{\perp})$.

Proof (continued). It remains to prove (c), (d), and (e).

First, since V is finite-dimensional, so is U. So, by Corollary 6.3.11(a), U has an orthogonal basis $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$.

Theorem 6.4.3

(c) $(U^{\perp})^{\perp} = U$;

(d) $V = U \oplus U^{\perp}$, that is, $V = U + U^{\perp}$ and $U \cap U^{\perp} = \{\mathbf{0}\}$;

(e) $\dim(V) = \dim(U) + \dim(U^{\perp})$.

Proof (continued). It remains to prove (c), (d), and (e).

First, since V is finite-dimensional, so is U. So, by Corollary 6.3.11(a), U has an orthogonal basis $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$. By Corollary 6.3.11(b), the orthogonal basis $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ of U can be extended to an orthogonal basis $\{\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{u}_{k+1}, \dots, \mathbf{u}_n\}$ of V.

Theorem 6.4.3

- (c) $(U^{\perp})^{\perp} = U$;
- (d) $V = U \oplus U^{\perp}$, that is, $V = U + U^{\perp}$ and $U \cap U^{\perp} = \{\mathbf{0}\}$;
- (e) $\dim(V) = \dim(U) + \dim(U^{\perp})$.

Proof (continued). It remains to prove (c), (d), and (e).

First, since V is finite-dimensional, so is U. So, by Corollary 6.3.11(a), U has an orthogonal basis $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$. By Corollary 6.3.11(b), the orthogonal basis $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ of U can be extended to an orthogonal basis $\{\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{u}_{k+1}, \dots, \mathbf{u}_n\}$ of V. By (a), $\{\mathbf{u}_{k+1}, \dots, \mathbf{u}_n\}$ is an orthogonal basis of $U^{\perp}$.

Theorem 6.4.3

(c) $(U^{\perp})^{\perp} = U$;

(d) $V = U \oplus U^{\perp}$, that is, $V = U + U^{\perp}$ and $U \cap U^{\perp} = \{\mathbf{0}\}$;

(e) $\dim(V) = \dim(U) + \dim(U^{\perp})$.

Proof (continued). It remains to prove (c), (d), and (e).

First, since V is finite-dimensional, so is U. So, by Corollary 6.3.11(a), U has an orthogonal basis $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$. By Corollary 6.3.11(b), the orthogonal basis $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ of U can be extended to an orthogonal basis $\{\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{u}_{k+1}, \dots, \mathbf{u}_n\}$ of V. By (a), $\{\mathbf{u}_{k+1}, \dots, \mathbf{u}_n\}$ is an orthogonal basis of $U^{\perp}$. But then $\{\mathbf{u}_{k+1}, \dots, \mathbf{u}_n, \mathbf{u}_1, \dots, \mathbf{u}_k\}$ is an orthogonal basis of V that extends $\{\mathbf{u}_{k+1}, \dots, \mathbf{u}_n\}$,

Theorem 6.4.3

- (c) $(U^{\perp})^{\perp} = U$;
- (d) $V = U \oplus U^{\perp}$, that is, $V = U + U^{\perp}$ and $U \cap U^{\perp} = \{\mathbf{0}\}$;
- (e) $\dim(V) = \dim(U) + \dim(U^{\perp})$.

Proof (continued). It remains to prove (c), (d), and (e).

First, since V is finite-dimensional, so is U. So, by Corollary 6.3.11(a), U has an orthogonal basis $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$. By Corollary 6.3.11(b), the orthogonal basis $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ of U can be extended to an orthogonal basis $\{\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{u}_{k+1}, \dots, \mathbf{u}_n\}$ of V. By (a), $\{\mathbf{u}_{k+1}, \dots, \mathbf{u}_n\}$ is an orthogonal basis of $U^{\perp}$. But then $\{\mathbf{u}_{k+1}, \dots, \mathbf{u}_n, \mathbf{u}_1, \dots, \mathbf{u}_k\}$ is an orthogonal basis of V that extends $\{\mathbf{u}_{k+1}, \dots, \mathbf{u}_n\}$, and so by (a) applied to the vector space $U^{\perp}$, we have that $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ is an orthogonal basis of $(U^{\perp})^{\perp}$.

Theorem 6.4.3

- (c) $(U^\perp)^\perp = U$;
- (d) $V = U \oplus U^\perp$, that is, $V = U + U^\perp$ and $U \cap U^\perp = \{\mathbf{0}\}$;
- (e) $\dim(V) = \dim(U) + \dim(U^\perp)$.

Proof (continued). It remains to prove (c), (d), and (e).

First, since V is finite-dimensional, so is U. So, by Corollary 6.3.11(a), U has an orthogonal basis $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$. By Corollary 6.3.11(b), the orthogonal basis $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ of U can be extended to an orthogonal basis $\{\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{u}_{k+1}, \dots, \mathbf{u}_n\}$ of V. By (a), $\{\mathbf{u}_{k+1}, \dots, \mathbf{u}_n\}$ is an orthogonal basis of $U^\perp$. But then $\{\mathbf{u}_{k+1}, \dots, \mathbf{u}_n, \mathbf{u}_1, \dots, \mathbf{u}_k\}$ is an orthogonal basis of V that extends $\{\mathbf{u}_{k+1}, \dots, \mathbf{u}_n\}$, and so by (a) applied to the vector space $U^\perp$, we have that $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ is an orthogonal basis of $(U^\perp)^\perp$. But now $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ is a basis of both U and $(U^\perp)^\perp$, and it follows that $U = (U^\perp)^\perp$, i.e. (c) holds.

Theorem 6.4.3

- (c) $(U^{\perp})^{\perp} = U$;
- (d) $V = U \oplus U^{\perp}$, that is, $V = U + U^{\perp}$ and $U \cap U^{\perp} = \{\mathbf{0}\}$;
- (e) $\dim(V) = \dim(U) + \dim(U^{\perp})$.

Proof (continued). Further, we have the following:

- $\dim(U) = k$, since $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ is a basis of U;
- $\dim(U^{\perp}) = n - k$, since $\{\mathbf{u}_{k+1}, \dots, \mathbf{u}_n\}$ is a basis of $U^{\perp}$;
- $\dim(V) = n$, since $\{\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{u}_{k+1}, \dots, \mathbf{u}_n\}$ is a basis of V.

It now immediately follows that $\dim(V) = \dim(U) + \dim(U^{\perp})$, i.e. (e) holds.

Theorem 6.4.3

(c) $(U^{\perp})^{\perp} = U$;

(d) $V = U \oplus U^{\perp}$, that is, $V = U + U^{\perp}$ and $U \cap U^{\perp} = \{\mathbf{0}\}$;

(e) $\dim(V) = \dim(U) + \dim(U^{\perp})$.

Proof (continued). Finally, we prove (d).

Theorem 6.4.3

- (c) $(U^{\perp})^{\perp} = U$;
- (d) $V = U \oplus U^{\perp}$, that is, $V = U + U^{\perp}$ and $U \cap U^{\perp} = \{\mathbf{0}\}$;
- (e) $\dim(V) = \dim(U) + \dim(U^{\perp})$.

Proof (continued). Finally, we prove (d).

Let us first show that $U \cap U^{\perp} = \{\mathbf{0}\}$.

Theorem 6.4.3

- (c) $(U^{\perp})^{\perp} = U$;
- (d) $V = U \oplus U^{\perp}$, that is, $V = U + U^{\perp}$ and $U \cap U^{\perp} = \{\mathbf{0}\}$;
- (e) $\dim(V) = \dim(U) + \dim(U^{\perp})$.

Proof (continued). Finally, we prove (d).

Let us first show that $U \cap U^{\perp} = \{\mathbf{0}\}$. Since U and $U^{\perp}$ are both subspaces of V, they both contain $\mathbf{0}$, and consequently, $\mathbf{0} \in U \cap U^{\perp}$.

Theorem 6.4.3

- (c) $(U^{\perp})^{\perp} = U$;
- (d) $V = U \oplus U^{\perp}$, that is, $V = U + U^{\perp}$ and $U \cap U^{\perp} = \{\mathbf{0}\}$;
- (e) $\dim(V) = \dim(U) + \dim(U^{\perp})$.

Proof (continued). Finally, we prove (d).

Let us first show that $U \cap U^{\perp} = \{\mathbf{0}\}$. Since U and $U^{\perp}$ are both subspaces of V, they both contain $\mathbf{0}$, and consequently, $\mathbf{0} \in U \cap U^{\perp}$.

Now, fix any $\mathbf{u} \in U \cap U^{\perp}$; we must show that $\mathbf{u} = \mathbf{0}$.

Theorem 6.4.3

- (c) $(U^{\perp})^{\perp} = U$;
- (d) $V = U \oplus U^{\perp}$, that is, $V = U + U^{\perp}$ and $U \cap U^{\perp} = \{\mathbf{0}\}$;
- (e) $\dim(V) = \dim(U) + \dim(U^{\perp})$.

Proof (continued). Finally, we prove (d).

Let us first show that $U \cap U^{\perp} = \{\mathbf{0}\}$. Since U and $U^{\perp}$ are both subspaces of V, they both contain $\mathbf{0}$, and consequently, $\mathbf{0} \in U \cap U^{\perp}$.

Now, fix any $\mathbf{u} \in U \cap U^{\perp}$; we must show that $\mathbf{u} = \mathbf{0}$. Since $\mathbf{u} \in U$ and $\mathbf{u} \in U^{\perp}$, we have that $\mathbf{u} \perp \mathbf{u}$, i.e. $\langle \mathbf{u}, \mathbf{u} \rangle = 0$.

Theorem 6.4.3

- (c) $(U^\perp)^\perp = U$;
- (d) $V = U \oplus U^\perp$, that is, $V = U + U^\perp$ and $U \cap U^\perp = \{\mathbf{0}\}$;
- (e) $\dim(V) = \dim(U) + \dim(U^\perp)$.

Proof (continued). Finally, we prove (d).

Let us first show that $U \cap U^\perp = \{\mathbf{0}\}$. Since U and $U^\perp$ are both subspaces of V, they both contain $\mathbf{0}$, and consequently, $\mathbf{0} \in U \cap U^\perp$.

Now, fix any $\mathbf{u} \in U \cap U^\perp$; we must show that $\mathbf{u} = \mathbf{0}$. Since $\mathbf{u} \in U$ and $\mathbf{u} \in U^\perp$, we have that $\mathbf{u} \perp \mathbf{u}$, i.e. $\langle \mathbf{u}, \mathbf{u} \rangle = 0$. But then by the definition of a scalar product, we have that $\mathbf{u} = \mathbf{0}$.

Theorem 6.4.3

- (c) $(U^\perp)^\perp = U$;
- (d) $V = U \oplus U^\perp$, that is, $V = U + U^\perp$ and $U \cap U^\perp = \{\mathbf{0}\}$;
- (e) $\dim(V) = \dim(U) + \dim(U^\perp)$.

Proof (continued). Finally, we prove (d).

Let us first show that $U \cap U^\perp = \{\mathbf{0}\}$. Since U and $U^\perp$ are both subspaces of V, they both contain $\mathbf{0}$, and consequently, $\mathbf{0} \in U \cap U^\perp$.

Now, fix any $\mathbf{u} \in U \cap U^\perp$; we must show that $\mathbf{u} = \mathbf{0}$. Since $\mathbf{u} \in U$ and $\mathbf{u} \in U^\perp$, we have that $\mathbf{u} \perp \mathbf{u}$, i.e. $\langle \mathbf{u}, \mathbf{u} \rangle = 0$. But then by the definition of a scalar product, we have that $\mathbf{u} = \mathbf{0}$. This proves that $U \cap U^\perp = \{\mathbf{0}\}$.

Theorem 6.4.3

- (c) $(U^{\perp})^{\perp} = U$;
- (d) $V = U \oplus U^{\perp}$, that is, $V = U + U^{\perp}$ and $U \cap U^{\perp} = \{\mathbf{0}\}$;
- (e) $\dim(V) = \dim(U) + \dim(U^{\perp})$.

Proof (continued). It remains to show that $V = U + U^{\perp}$.

Theorem 6.4.3

(c) $(U^\perp)^\perp = U$;

(d) $V = U \oplus U^\perp$, that is, $V = U + U^\perp$ and $U \cap U^\perp = \{\mathbf{0}\}$;

(e) $\dim(V) = \dim(U) + \dim(U^\perp)$.

Proof (continued). It remains to show that $V = U + U^\perp$.

It is clear that $U + U^\perp \subseteq V$, and so we need only show that $V \subseteq U + U^\perp$.

Theorem 6.4.3

(c) $(U^{\perp})^{\perp} = U$;

(d) $V = U \oplus U^{\perp}$, that is, $V = U + U^{\perp}$ and $U \cap U^{\perp} = \{\mathbf{0}\}$;

(e) $\dim(V) = \dim(U) + \dim(U^{\perp})$.

Proof (continued). It remains to show that $V = U + U^{\perp}$.

It is clear that $U + U^{\perp} \subseteq V$, and so we need only show that $V \subseteq U + U^{\perp}$.

Fix any $\mathbf{v} \in V$.

Theorem 6.4.3

- (c) $(U^\perp)^\perp = U$;
- (d) $V = U \oplus U^\perp$, that is, $V = U + U^\perp$ and $U \cap U^\perp = \{\mathbf{0}\}$;
- (e) $\dim(V) = \dim(U) + \dim(U^\perp)$.

Proof (continued). It remains to show that $V = U + U^\perp$.

It is clear that $U + U^\perp \subseteq V$, and so we need only show that $V \subseteq U + U^\perp$.

Fix any $\mathbf{v} \in V$. Since $\{\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{u}_{k+1}, \dots, \mathbf{u}_n\}$ is a basis of V, we know that there exist scalars $\alpha_1, \dots, \alpha_n$ such that $\mathbf{v} = \alpha_1 \mathbf{u}_1 + \dots + \alpha_n \mathbf{u}_n$. Set $\mathbf{v}_1 := \alpha_1 \mathbf{u}_1 + \dots + \alpha_k \mathbf{u}_k$ and $\mathbf{v}_2 := \alpha_{k+1} \mathbf{u}_{k+1} + \dots + \alpha_n \mathbf{u}_n$. Then $\mathbf{v} = \mathbf{v}_1 + \mathbf{v}_2$. Since $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ is a basis of U, we see that $\mathbf{v}_1 \in U$, and since $\{\mathbf{u}_{k+1}, \dots, \mathbf{u}_n\}$ is a basis of $U^\perp$, we see that $\mathbf{v}_2 \in U^\perp$. So, $\mathbf{v} = \mathbf{v}_1 + \mathbf{v}_2$ belongs to $U + U^\perp$, and it follows that $V \subseteq U + U^\perp$. This proves (d), and we are done. $\square$

Theorem 6.4.3

Let V be a finite-dimensional real or complex vector space, equipped with a scalar product $\langle \cdot, \cdot \rangle$ and the induced norm $|| \cdot ||$. Let U be a subspace of V. Then $U^\perp$ is a subspace of V, and all the following hold:

a. if $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ is an orthogonal basis of U, and $\{\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{u}_{k+1}, \dots, \mathbf{u}_n\}$ is an extension of that basis to an orthogonal basis of V, then $\{\mathbf{u}_{k+1}, \dots, \mathbf{u}_n\}$ is an orthogonal basis of $U^\perp$;
b. if $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ is an orthonormal basis of U, and $\{\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{u}_{k+1}, \dots, \mathbf{u}_n\}$ is an extension of that basis to an orthonormal basis of V, then $\{\mathbf{u}_{k+1}, \dots, \mathbf{u}_n\}$ is an orthonormal basis of $U^\perp$;
c. $(U^\perp)^\perp = U$;
d. $V = U \oplus U^\perp$, that is, $V = U + U^\perp$ and $U \cap U^\perp = \{\mathbf{0}\}$;
e. $\dim(V) = \dim(U) + \dim(U^\perp)$.

- As a corollary of Theorem 6.4.3(a-b), we obtain the following computationally useful proposition.
  - The proposition is long, and we need to slides to state it.

Proposition 6.4.4

Let V be a finite-dimensional real or complex vector space, equipped with a scalar product $\langle \cdot, \cdot \rangle$ and the induced norm $|| \cdot ||$. Let $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ be any linearly independent set of vectors V, and let $\{\mathbf{v}_1, \dots, \mathbf{v}_k, \mathbf{v}_{k+1}, \dots, \mathbf{v}_n\}$ be an extension of that linearly independent set to a basis of V. Set $U := \text{Span}(\mathbf{v}_1, \dots, \mathbf{v}_k)$.

(a) If the Gram-Schmidt orthogonalization process (version 1) is applied to input vectors $\mathbf{v}_1, \dots, \mathbf{v}_k, \mathbf{v}_{k+1}, \dots, \mathbf{v}_n$ to produce output vectors $\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{u}_{k+1}, \dots, \mathbf{u}_n$, then both the following hold:
  - $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ is an orthogonal basis of U, and $\{\mathbf{u}_{k+1}, \dots, \mathbf{u}_n\}$ is an orthogonal basis of $U^\perp$;
  - $\left\{\frac{\mathbf{u}_1}{||\mathbf{u}_1||}, \dots, \frac{\mathbf{u}_k}{||\mathbf{u}_k||}\right\}$ is an orthonormal basis of U, and $\left\{\frac{\mathbf{u}_{k+1}}{||\mathbf{u}_{k+1}||}, \dots, \frac{\mathbf{u}_n}{||\mathbf{u}_n||}\right\}$ is an orthonormal basis of $U^\perp$.

Proposition 6.4.4

Let V be a finite-dimensional real or complex vector space, equipped with a scalar product $\langle \cdot, \cdot \rangle$ and the induced norm $|| \cdot ||$. Let $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ be any linearly independent set of vectors V, and let $\{\mathbf{v}_1, \dots, \mathbf{v}_k, \mathbf{v}_{k+1}, \dots, \mathbf{v}_n\}$ be an extension of that linearly independent set to a basis of V. Set $U := \text{Span}(\mathbf{v}_1, \dots, \mathbf{v}_k)$.

(b) If the Gram-Schmidt orthogonalization process (version 2) is applied to input vectors $\mathbf{v}_1, \dots, \mathbf{v}_k, \mathbf{v}_{k+1}, \dots, \mathbf{v}_n$ to produce output vectors $\mathbf{z}_1, \dots, \mathbf{z}_k, \mathbf{z}_{k+1}, \dots, \mathbf{z}_n$, then $\{\mathbf{z}_1, \dots, \mathbf{z}_k\}$ is an orthonormal basis of U, and $\{\mathbf{z}_{k+1}, \dots, \mathbf{z}_n\}$ is an orthonormal basis of $U^\perp$.

Proposition 6.4.4

Let V be a finite-dimensional real or complex vector space, equipped with a scalar product $\langle \cdot, \cdot \rangle$ and the induced norm $|| \cdot ||$. Let $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ be any linearly independent set of vectors V, and let $\{\mathbf{v}_1, \dots, \mathbf{v}_k, \mathbf{v}_{k+1}, \dots, \mathbf{v}_n\}$ be an extension of that linearly independent set to a basis of V. Set $U := \text{Span}(\mathbf{v}_1, \dots, \mathbf{v}_k)$.

(b) If the Gram-Schmidt orthogonalization process (version 2) is applied to input vectors $\mathbf{v}_1, \dots, \mathbf{v}_k, \mathbf{v}_{k+1}, \dots, \mathbf{v}_n$ to produce output vectors $\mathbf{z}_1, \dots, \mathbf{z}_k, \mathbf{z}_{k+1}, \dots, \mathbf{z}_n$, then $\{\mathbf{z}_1, \dots, \mathbf{z}_k\}$ is an orthonormal basis of U, and $\{\mathbf{z}_{k+1}, \dots, \mathbf{z}_n\}$ is an orthonormal basis of $U^\perp$.

- This is an easy corollary of Theorem 6.4.3 (details: Lecture Notes).

Example 6.4.5

Consider the following vectors in $\mathbb{R}^4$:

$$\mathbf{a}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{bmatrix}, \quad \mathbf{a}_2 = \begin{bmatrix} 2 \\ 2 \\ 2 \\ 0 \end{bmatrix}, \quad \mathbf{a}_3 = \begin{bmatrix} 0 \\ 3 \\ 3 \\ 3 \end{bmatrix}, \quad \mathbf{a}_4 = \begin{bmatrix} 2 \\ 4 \\ 4 \\ 2 \end{bmatrix}.$$

Compute an orthonormal basis of $U := \text{Span}(\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3, \mathbf{a}_4)$ and an orthonormal basis of $U^\perp$.

Solution.

Example 6.4.5

Consider the following vectors in $\mathbb{R}^4$:

$$\mathbf{a}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{bmatrix}, \quad \mathbf{a}_2 = \begin{bmatrix} 2 \\ 2 \\ 2 \\ 0 \end{bmatrix}, \quad \mathbf{a}_3 = \begin{bmatrix} 0 \\ 3 \\ 3 \\ 3 \end{bmatrix}, \quad \mathbf{a}_4 = \begin{bmatrix} 2 \\ 4 \\ 4 \\ 2 \end{bmatrix}.$$

Compute an orthonormal basis of $U := \mathrm{Span}(\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3, \mathbf{a}_4)$ and an orthonormal basis of $U^\perp$.

Solution. First, we need to find a basis of U and extend it to a basis of $\mathbb{R}^4$. For this, we use Proposition 3.3.21. We consider the standard basis $\mathcal{E}_4 = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4\}$ of $\mathbb{R}^4$, and we form the matrix

$$\begin{aligned} C \quad &:= \quad \left[\begin{array}{cccc|cccc} \mathbf{a}_1 & \mathbf{a}_2 & \mathbf{a}_3 & \mathbf{a}_4 & \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 & \mathbf{e}_4 \end{array} \right] \\ &= \quad \left[\begin{array}{cccc|cccc} 1 & 2 & 0 & 2 & 1 & 0 & 0 & 0 \\ 1 & 2 & 3 & 4 & 0 & 1 & 0 & 0 \\ 1 & 2 & 3 & 4 & 0 & 0 & 1 & 0 \\ 0 & 0 & 3 & 2 & 0 & 0 & 0 & 1 \end{array} \right]. \end{aligned}$$

Solution (continued). Reminder:

$$C := \left[\begin{array}{cccc|cccc} \mathbf{a}_1 & \mathbf{a}_2 & \mathbf{a}_3 & \mathbf{a}_4 & \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 & \mathbf{e}_4 \end{array}\right].$$

Solution (continued). Reminder:

$$C := \left[\begin{array}{cccc:cccc} \mathbf{a}_1 & \mathbf{a}_2 & \mathbf{a}_3 & \mathbf{a}_4 & \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 & \mathbf{e}_4 \end{array}\right].$$

By row reducing, we obtain

$$\mathrm{RREF}(C) = \left[\begin{array}{cccc:cccc} 1 & 2 & 0 & 2 & 0 & 0 & 1 & -1 \\ 0 & 0 & 1 & 2/3 & 0 & 0 & 0 & 1/3 \\ 0 & 0 & 0 & 0 & 1 & 0 & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 \end{array}\right].$$

Solution (continued). Reminder:

$C := \left[\begin{array}{cccc:cccc} \mathbf{a}_1 & \mathbf{a}_2 & \mathbf{a}_3 & \mathbf{a}_4 & \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 & \mathbf{e}_4 \end{array}\right].$

By row reducing, we obtain

$$\mathrm{RREF}(C) = \left[\begin{array}{cccc:cccc} 1 & 2 & 0 & 2 & 0 & 0 & 1 & -1 \\ 0 & 0 & 1 & 2/3 & 0 & 0 & 0 & 1/3 \\ 0 & 0 & 0 & 0 & 1 & 0 & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 \end{array}\right].$$

As we can see, the pivot columns of C are its first, third, fifth, and sixth column.

Solution (continued). Reminder:

$$C := \left[\begin{array}{cccc|cccc} \mathbf{a}_1 & \mathbf{a}_2 & \mathbf{a}_3 & \mathbf{a}_4 & \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 & \mathbf{e}_4 \end{array}\right].$$

By row reducing, we obtain

$$\mathrm{RREF}(C) = \left[\begin{array}{cccc|cccc} 1 & 2 & 0 & 2 & 0 & 0 & 1 & -1 \\ 0 & 0 & 1 & 2/3 & 0 & 0 & 0 & 1/3 \\ 0 & 0 & 0 & 0 & 1 & 0 & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 \end{array}\right].$$

As we can see, the pivot columns of C are its first, third, fifth, and sixth column. So, by Proposition 3.3.21, $\{\mathbf{a}_1, \mathbf{a}_3\}$ is a basis of U, and $\{\mathbf{a}_1, \mathbf{a}_3, \mathbf{e}_1, \mathbf{e}_2\}$ is a basis of $\mathbb{R}^4$ that extends $\{\mathbf{a}_1, \mathbf{a}_3\}$.

Solution (continued). Reminder:

$$C := \left[\begin{array}{cccc|cccc} \mathbf{a}_1 & \mathbf{a}_2 & \mathbf{a}_3 & \mathbf{a}_4 & \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 & \mathbf{e}_4 \end{array}\right].$$

By row reducing, we obtain

$$\mathrm{RREF}(C) = \left[\begin{array}{cccc|cccc} 1 & 2 & 0 & 2 & 0 & 0 & 1 & -1 \\ 0 & 0 & 1 & 2/3 & 0 & 0 & 0 & 1/3 \\ 0 & 0 & 0 & 0 & 1 & 0 & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 \end{array}\right].$$

As we can see, the pivot columns of C are its first, third, fifth, and sixth column. So, by Proposition 3.3.21, $\{\mathbf{a}_1, \mathbf{a}_3\}$ is a basis of U, and $\{\mathbf{a}_1, \mathbf{a}_3, \mathbf{e}_1, \mathbf{e}_2\}$ is a basis of $\mathbb{R}^4$ that extends $\{\mathbf{a}_1, \mathbf{a}_3\}$. By applying the Gram-Schmidt orthogonalization process (version 2) to the vectors $\mathbf{a}_1, \mathbf{a}_3, \mathbf{e}_1, \mathbf{e}_2$, we obtain the following vectors (next slide):

Solution (continued).

$$\mathbf{z}_1 = \begin{bmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \\ 0 \end{bmatrix}, \qquad \mathbf{z}_2 = \begin{bmatrix} -2/\sqrt{15} \\ 1/\sqrt{15} \\ 1/\sqrt{15} \\ 3/\sqrt{15} \end{bmatrix},$$

$$\mathbf{z}_3 = \begin{bmatrix} 2/\sqrt{10} \\ -1/\sqrt{10} \\ -1/\sqrt{10} \\ 2/\sqrt{10} \end{bmatrix}, \qquad \mathbf{z}_4 = \begin{bmatrix} 0 \\ 1/\sqrt{2} \\ -1/\sqrt{2} \\ 0 \end{bmatrix}.$$

Solution (continued).

$$\mathbf{z}_1 = \begin{bmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \\ 0 \end{bmatrix}, \qquad \mathbf{z}_2 = \begin{bmatrix} -2/\sqrt{15} \\ 1/\sqrt{15} \\ 1/\sqrt{15} \\ 3/\sqrt{15} \end{bmatrix},$$

$$\mathbf{z}_3 = \begin{bmatrix} 2/\sqrt{10} \\ -1/\sqrt{10} \\ -1/\sqrt{10} \\ 2/\sqrt{10} \end{bmatrix}, \qquad \mathbf{z}_4 = \begin{bmatrix} 0 \\ 1/\sqrt{2} \\ -1/\sqrt{2} \\ 0 \end{bmatrix}.$$

By Proposition 6.4.4(b), $\{\mathbf{z}_1, \mathbf{z}_2\}$ is an orthonormal basis of U, whereas $\{\mathbf{z}_3, \mathbf{z}_4\}$ is an orthonormal basis of $U^{\perp}$. $\square$

Example 6.4.5

Consider the following vectors in $\mathbb{R}^4$:

$$\mathbf{a}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{bmatrix}, \quad \mathbf{a}_2 = \begin{bmatrix} 2 \\ 2 \\ 2 \\ 0 \end{bmatrix}, \quad \mathbf{a}_3 = \begin{bmatrix} 0 \\ 3 \\ 3 \\ 3 \end{bmatrix}, \quad \mathbf{a}_4 = \begin{bmatrix} 2 \\ 4 \\ 4 \\ 2 \end{bmatrix}.$$

Compute an orthonormal basis of $U := \text{Span}(\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3, \mathbf{a}_4)$ and an orthonormal basis of $U^\perp$.

- **Remark:** We could also have applied the Gram-Schmidt orthogonalization process (version 1) to the vectors $\mathbf{a}_1, \mathbf{a}_3, \mathbf{e}_1, \mathbf{e}_2$, and then normalized the output vectors. We would have gotten the same vectors $\mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_3, \mathbf{z}_4$ as above. Proposition 6.4.4(a) would then imply that $\{\mathbf{z}_1, \mathbf{z}_2\}$ is an orthonormal basis of U, whereas $\{\mathbf{z}_3, \mathbf{z}_4\}$ is an orthonormal basis of $U^\perp$.

2. Orthogonal projection onto a subspace

2. Orthogonal projection onto a subspace

Theorem 6.5.1

Let V be a finite-dimensional real or complex vector space, equipped with a scalar product $\langle \cdot, \cdot \rangle$ and the induced norm $||\cdot||$. Let U be a subspace of V, and let $\mathbf{x} \in V$. Then there exists a unique vector $\mathbf{x}_U \in U$ that has the property that

$$||\mathbf{x} - \mathbf{x}_U|| = \min_{\mathbf{u} \in U} ||\mathbf{x} - \mathbf{u}||.$$

Moreover, if $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ is an orthogonal basis of U, then this vector $\mathbf{x}_U$ is given by the formula

$$\mathbf{x}_U = \sum_{i=1}^{k} \text{proj}_{\mathbf{u}_i}(\mathbf{x}) = \sum_{i=1}^{k} \frac{\langle \mathbf{x}, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \mathbf{u}_i.$$

- **Terminology/Notation:** The vector $\mathbf{x}_U$ from Theorem 6.5.1 is called the *orthogonal projection* of $\mathbf{x}$ onto U.

Theorem 6.5.1

Let V be a finite-dimensional real or complex vector space, equipped with a scalar product $\langle \cdot, \cdot \rangle$ and the induced norm $|| \cdot ||$. Let U be a subspace of V, and let $\mathbf{x} \in V$. Then there exists a unique vector $\mathbf{x}_U \in U$ that has the property that

$$||\mathbf{x} - \mathbf{x}_U|| \quad = \quad \min_{\mathbf{u} \in U} ||\mathbf{x} - \mathbf{u}||.$$

Moreover, if $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ is an orthogonal basis of U, then this vector $\mathbf{x}_U$ is given by the formula

$$\mathbf{x}_U \quad = \quad \sum_{i=1}^{k} \text{proj}_{\mathbf{u}_i}(\mathbf{x}) \quad = \quad \sum_{i=1}^{k} \frac{\langle \mathbf{x}, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \mathbf{u}_i.$$

- **Remarks:**

Theorem 6.5.1

Let V be a finite-dimensional real or complex vector space, equipped with a scalar product $\langle \cdot, \cdot \rangle$ and the induced norm $||\cdot||$. Let U be a subspace of V, and let $\mathbf{x} \in V$. Then there exists a unique vector $\mathbf{x}_U \in U$ that has the property that

$$||\mathbf{x} - \mathbf{x}_U|| \quad = \quad \min_{\mathbf{u} \in U} ||\mathbf{x} - \mathbf{u}||.$$

Moreover, if $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ is an orthogonal basis of U, then this vector $\mathbf{x}_U$ is given by the formula

$$\mathbf{x}_U \quad = \quad \sum_{i=1}^{k} \text{proj}_{\mathbf{u}_i}(\mathbf{x}) \quad = \quad \sum_{i=1}^{k} \frac{\langle \mathbf{x}, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \mathbf{u}_i.$$

- **Remarks:**
  - If $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ is an **orthonormal** basis of U, then

$$\mathbf{x}_U \quad = \quad \sum_{i=1}^{k} \text{proj}_{\mathbf{u}_i}(\mathbf{x}) \quad = \quad \sum_{i=1}^{k} \langle \mathbf{x}, \mathbf{u}_i \rangle \, \mathbf{u}_i.$$

Theorem 6.5.1

Let V be a finite-dimensional real or complex vector space, equipped with a scalar product $\langle \cdot, \cdot \rangle$ and the induced norm $||\cdot||$. Let U be a subspace of V, and let $\mathbf{x} \in V$. Then there exists a unique vector $\mathbf{x}_U \in U$ that has the property that

$$||\mathbf{x} - \mathbf{x}_U|| \quad = \quad \min_{\mathbf{u} \in U} ||\mathbf{x} - \mathbf{u}||.$$

Moreover, if $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ is an orthogonal basis of U, then this vector $\mathbf{x}_U$ is given by the formula

$$\mathbf{x}_U \quad = \quad \sum_{i=1}^{k} \mathrm{proj}_{\mathbf{u}_i}(\mathbf{x}) \quad = \quad \sum_{i=1}^{k} \frac{\langle \mathbf{x}, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \mathbf{u}_i.$$

- **Remarks:**
  - If $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ is an **orthonormal** basis of U, then
    $$\mathbf{x}_U \quad = \quad \sum_{i=1}^{k} \mathrm{proj}_{\mathbf{u}_i}(\mathbf{x}) \quad = \quad \sum_{i=1}^{k} \langle \mathbf{x}, \mathbf{u}_i \rangle \, \mathbf{u}_i.$$
  - If $\mathbf{x} \in U$, then $\mathbf{x}_U = \mathbf{x}$, since in this case, the expression $||\mathbf{x} - \mathbf{u}||$ (for $\mathbf{u} \in U$) is minimized for $\mathbf{u} = \mathbf{x}$.

Theorem 6.5.1

Let V be a finite-dimensional real or complex vector space, equipped with a scalar product $\langle \cdot, \cdot \rangle$ and the induced norm $||\cdot||$. Let U be a subspace of V, and let $\mathbf{x} \in V$. Then there exists a unique vector $\mathbf{x}_U \in U$ that has the property that

$$||\mathbf{x} - \mathbf{x}_U|| = \min_{\mathbf{u} \in U} ||\mathbf{x} - \mathbf{u}||.$$

Moreover, if $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ is an orthogonal basis of U, then this vector $\mathbf{x}_U$ is given by the formula

$$\mathbf{x}_U = \sum_{i=1}^{k} \text{proj}_{\mathbf{u}_i}(\mathbf{x}) = \sum_{i=1}^{k} \frac{\langle \mathbf{x}, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \mathbf{u}_i.$$

- Now let's prove the theorem!

Proof.

Proof. Using Corollary 6.3.11, we fix an orthogonal basis $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ of U, and we extend it to an orthogonal basis $\{\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{u}_{k+1}, \dots, \mathbf{u}_n\}$ of V.

Proof. Using Corollary 6.3.11, we fix an orthogonal basis $\{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ of U, and we extend it to an orthogonal basis $\{\mathbf{u}_1, \ldots, \mathbf{u}_k, \mathbf{u}_{k+1}, \ldots, \mathbf{u}_n\}$ of V. By Theorem 6.4.3(a), $\{\mathbf{u}_{k+1}, \ldots, \mathbf{u}_n\}$ is an orthogonal basis of $U^\perp$.

Proof. Using Corollary 6.3.11, we fix an orthogonal basis $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ of U, and we extend it to an orthogonal basis $\{\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{u}_{k+1}, \dots, \mathbf{u}_n\}$ of V. By Theorem 6.4.3(a), $\{\mathbf{u}_{k+1}, \dots, \mathbf{u}_n\}$ is an orthogonal basis of $U^{\perp}$. Set

$$\mathbf{u}^* \quad := \quad \sum_{i=1}^{k} \frac{\langle \mathbf{x}, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \mathbf{u}_i.$$

Proof. Using Corollary 6.3.11, we fix an orthogonal basis $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ of U, and we extend it to an orthogonal basis $\{\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{u}_{k+1}, \dots, \mathbf{u}_n\}$ of V. By Theorem 6.4.3(a), $\{\mathbf{u}_{k+1}, \dots, \mathbf{u}_n\}$ is an orthogonal basis of $U^\perp$. Set

$$\mathbf{u}^* \quad := \quad \sum_{i=1}^{k} \frac{\langle \mathbf{x}, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \mathbf{u}_i.$$

- So, $\mathbf{u}^*$ is defined via the formula from the statement of the theorem.

Proof. Using Corollary 6.3.11, we fix an orthogonal basis $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ of U, and we extend it to an orthogonal basis $\{\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{u}_{k+1}, \dots, \mathbf{u}_n\}$ of V. By Theorem 6.4.3(a), $\{\mathbf{u}_{k+1}, \dots, \mathbf{u}_n\}$ is an orthogonal basis of $U^\perp$. Set

$$\mathbf{u}^* \quad := \quad \sum_{i=1}^{k} \frac{\langle \mathbf{x}, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \mathbf{u}_i.$$

- So, $\mathbf{u}^*$ is defined via the formula from the statement of the theorem.
- The reason we call it $\mathbf{u}^*$ rather than $\mathbf{x}_U$ is because we have not proven the existence and uniqueness of $\mathbf{x}_U$ yet.

Proof. Using Corollary 6.3.11, we fix an orthogonal basis $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ of U, and we extend it to an orthogonal basis $\{\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{u}_{k+1}, \dots, \mathbf{u}_n\}$ of V. By Theorem 6.4.3(a), $\{\mathbf{u}_{k+1}, \dots, \mathbf{u}_n\}$ is an orthogonal basis of $U^\perp$. Set

$$\mathbf{u}^* \quad := \quad \sum_{i=1}^{k} \frac{\langle \mathbf{x}, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \mathbf{u}_i.$$

- So, $\mathbf{u}^*$ is defined via the formula from the statement of the theorem.
- The reason we call it $\mathbf{u}^*$ rather than $\mathbf{x}_U$ is because we have not proven the existence and uniqueness of $\mathbf{x}_U$ yet.
- However, this is just a minor stylistic matter!

Proof. Using Corollary 6.3.11, we fix an orthogonal basis $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ of U, and we extend it to an orthogonal basis $\{\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{u}_{k+1}, \dots, \mathbf{u}_n\}$ of V. By Theorem 6.4.3(a), $\{\mathbf{u}_{k+1}, \dots, \mathbf{u}_n\}$ is an orthogonal basis of $U^\perp$. Set

$$\mathbf{u}^* \quad := \quad \sum_{i=1}^{k} \frac{\langle \mathbf{x}, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \mathbf{u}_i.$$

- So, $\mathbf{u}^*$ is defined via the formula from the statement of the theorem.
- The reason we call it $\mathbf{u}^*$ rather than $\mathbf{x}_U$ is because we have not proven the existence and uniqueness of $\mathbf{x}_U$ yet.
- However, this is just a minor stylistic matter!

Since $\mathbf{u}^*$ is a linear combination of the vectors $\mathbf{u}_1, \dots, \mathbf{u}_k$, which form a basis of U, we see that $\mathbf{u}^* \in U$.

Proof. Using Corollary 6.3.11, we fix an orthogonal basis $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ of U, and we extend it to an orthogonal basis $\{\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{u}_{k+1}, \dots, \mathbf{u}_n\}$ of V. By Theorem 6.4.3(a), $\{\mathbf{u}_{k+1}, \dots, \mathbf{u}_n\}$ is an orthogonal basis of $U^\perp$. Set

$$\mathbf{u}^* \quad := \quad \sum_{i=1}^{k} \frac{\langle \mathbf{x}, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \mathbf{u}_i.$$

- So, $\mathbf{u}^*$ is defined via the formula from the statement of the theorem.
- The reason we call it $\mathbf{u}^*$ rather than $\mathbf{x}_U$ is because we have not proven the existence and uniqueness of $\mathbf{x}_U$ yet.
- However, this is just a minor stylistic matter!

Since $\mathbf{u}^*$ is a linear combination of the vectors $\mathbf{u}_1, \dots, \mathbf{u}_k$, which form a basis of U, we see that $\mathbf{u}^* \in U$.

Now, fix any $\mathbf{u} \in U$. We must show that $||\mathbf{x} - \mathbf{u}^*|| \leq ||\mathbf{x} - \mathbf{u}||$, and that equality holds iff $\mathbf{u}^* = \mathbf{u}$. Clearly, this is sufficient to prove the theorem.

Proof (continued). Reminder: $\mathbf{u}^* := \sum_{i=1}^{k} \frac{\langle \mathbf{x}, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \mathbf{u}_i$; $\mathbf{u} \in U$; WTS $||\mathbf{x} - \mathbf{u}^*|| \leq ||\mathbf{x} - \mathbf{u}||$ and that equality holds iff $\mathbf{u}^* = \mathbf{u}$.

Proof (continued). Reminder: $\mathbf{u}^* := \sum_{i=1}^{k} \frac{\langle \mathbf{x}, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \mathbf{u}_i$; $\mathbf{u} \in U$; WTS $||\mathbf{x} - \mathbf{u}^*|| \leq ||\mathbf{x} - \mathbf{u}||$ and that equality holds iff $\mathbf{u}^* = \mathbf{u}$.

Let us first prove that $(\mathbf{u}^* - \mathbf{u}) \perp (\mathbf{x} - \mathbf{u}^*)$.

Proof (continued). Reminder: $\mathbf{u}^* := \sum_{i=1}^{k} \frac{\langle \mathbf{x}, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \mathbf{u}_i$; $\mathbf{u} \in U$; WTS $||\mathbf{x} - \mathbf{u}^*|| \leq ||\mathbf{x} - \mathbf{u}||$ and that equality holds iff $\mathbf{u}^* = \mathbf{u}$.

Let us first prove that $(\mathbf{u}^* - \mathbf{u}) \perp (\mathbf{x} - \mathbf{u}^*)$. Since $\mathbf{u}^*, \mathbf{u} \in U$, and since U is a subspace of V, it is clear that $\mathbf{u}^* - \mathbf{u} \in U$.

Proof (continued). Reminder: $\mathbf{u}^* := \sum_{i=1}^{k} \frac{\langle \mathbf{x}, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \mathbf{u}_i$; $\mathbf{u} \in U$; WTS $||\mathbf{x} - \mathbf{u}^*|| \leq ||\mathbf{x} - \mathbf{u}||$ and that equality holds iff $\mathbf{u}^* = \mathbf{u}$.

Let us first prove that $(\mathbf{u}^* - \mathbf{u}) \perp (\mathbf{x} - \mathbf{u}^*)$. Since $\mathbf{u}^*, \mathbf{u} \in U$, and since U is a subspace of V, it is clear that $\mathbf{u}^* - \mathbf{u} \in U$. So, it suffices to show that $\mathbf{x} - \mathbf{u}^* \in U^{\perp}$.

Proof (continued). Reminder: $\mathbf{u}^* := \sum_{i=1}^{k} \frac{\langle \mathbf{x}, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \mathbf{u}_i$; $\mathbf{u} \in U$; WTS $||\mathbf{x} - \mathbf{u}^*|| \leq ||\mathbf{x} - \mathbf{u}||$ and that equality holds iff $\mathbf{u}^* = \mathbf{u}$.

Let us first prove that $(\mathbf{u}^* - \mathbf{u}) \perp (\mathbf{x} - \mathbf{u}^*)$. Since $\mathbf{u}^*, \mathbf{u} \in U$, and since U is a subspace of V, it is clear that $\mathbf{u}^* - \mathbf{u} \in U$. So, it suffices to show that $\mathbf{x} - \mathbf{u}^* \in U^\perp$.

By Theorem 6.3.5, we have that

$$\mathbf{x} = \sum_{i=1}^{n} \frac{\langle \mathbf{x}, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \mathbf{u}_i,$$

Proof (continued). Reminder: $\mathbf{u}^* := \sum_{i=1}^{k} \frac{\langle \mathbf{x}, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \mathbf{u}_i$; $\mathbf{u} \in U$; WTS $||\mathbf{x} - \mathbf{u}^*|| \leq ||\mathbf{x} - \mathbf{u}||$ and that equality holds iff $\mathbf{u}^* = \mathbf{u}$.

Let us first prove that $(\mathbf{u}^* - \mathbf{u}) \perp (\mathbf{x} - \mathbf{u}^*)$. Since $\mathbf{u}^*, \mathbf{u} \in U$, and since U is a subspace of V, it is clear that $\mathbf{u}^* - \mathbf{u} \in U$. So, it suffices to show that $\mathbf{x} - \mathbf{u}^* \in U^\perp$.

By Theorem 6.3.5, we have that

$$\mathbf{x} \quad = \quad \sum_{i=1}^{n} \frac{\langle \mathbf{x}, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \mathbf{u}_i,$$

and it follows that

$$\mathbf{x} - \mathbf{u}^* \quad = \quad \sum_{i=k+1}^{n} \frac{\langle \mathbf{x}, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \mathbf{u}_i.$$

Proof (continued). Reminder: $\mathbf{u}^* := \sum_{i=1}^{k} \frac{\langle \mathbf{x},\mathbf{u}_i \rangle}{\langle \mathbf{u}_i,\mathbf{u}_i \rangle}\mathbf{u}_i$; $\mathbf{u} \in U$; WTS $||\mathbf{x} - \mathbf{u}^*|| \leq ||\mathbf{x} - \mathbf{u}||$ and that equality holds iff $\mathbf{u}^* = \mathbf{u}$.

Let us first prove that $(\mathbf{u}^* - \mathbf{u}) \perp (\mathbf{x} - \mathbf{u}^*)$. Since $\mathbf{u}^*, \mathbf{u} \in U$, and since U is a subspace of V, it is clear that $\mathbf{u}^* - \mathbf{u} \in U$. So, it suffices to show that $\mathbf{x} - \mathbf{u}^* \in U^\perp$.

By Theorem 6.3.5, we have that

$$\mathbf{x} = \sum_{i=1}^{n} \frac{\langle \mathbf{x},\mathbf{u}_i \rangle}{\langle \mathbf{u}_i,\mathbf{u}_i \rangle}\mathbf{u}_i,$$

and it follows that

$$\mathbf{x} - \mathbf{u}^* = \sum_{i=k+1}^{n} \frac{\langle \mathbf{x},\mathbf{u}_i \rangle}{\langle \mathbf{u}_i,\mathbf{u}_i \rangle}\mathbf{u}_i.$$

So, $\mathbf{x} - \mathbf{u}^*$ is a linear combination of the vectors $\mathbf{u}_{k+1}, \dots, \mathbf{u}_n$; since those $n-k$ vectors form a basis of $U^\perp$, it follows that $\mathbf{x} - \mathbf{u}^* \in U^\perp$.

Proof (continued). Reminder: $\mathbf{u}^* := \sum_{i=1}^{k} \frac{\langle \mathbf{x},\mathbf{u}_i \rangle}{\langle \mathbf{u}_i,\mathbf{u}_i \rangle}\mathbf{u}_i$; $\mathbf{u} \in U$; WTS $||\mathbf{x}-\mathbf{u}^*|| \leq ||\mathbf{x}-\mathbf{u}||$ and that equality holds iff $\mathbf{u}^* = \mathbf{u}$.

Let us first prove that $(\mathbf{u}^* - \mathbf{u}) \perp (\mathbf{x} - \mathbf{u}^*)$. Since $\mathbf{u}^*, \mathbf{u} \in U$, and since U is a subspace of V, it is clear that $\mathbf{u}^* - \mathbf{u} \in U$. So, it suffices to show that $\mathbf{x} - \mathbf{u}^* \in U^\perp$.

By Theorem 6.3.5, we have that

$$\mathbf{x} \;=\; \sum_{i=1}^{n} \frac{\langle \mathbf{x},\mathbf{u}_i \rangle}{\langle \mathbf{u}_i,\mathbf{u}_i \rangle}\mathbf{u}_i,$$

and it follows that

$$\mathbf{x} - \mathbf{u}^* \;=\; \sum_{i=k+1}^{n} \frac{\langle \mathbf{x},\mathbf{u}_i \rangle}{\langle \mathbf{u}_i,\mathbf{u}_i \rangle}\mathbf{u}_i.$$

So, $\mathbf{x} - \mathbf{u}^*$ is a linear combination of the vectors $\mathbf{u}_{k+1}, \dots, \mathbf{u}_n$; since those $n-k$ vectors form a basis of $U^\perp$, it follows that $\mathbf{x} - \mathbf{u}^* \in U^\perp$. This proves that $(\mathbf{u}^* - \mathbf{u}) \perp (\mathbf{x} - \mathbf{u}^*)$.

Proof (continued). Reminder: $\mathbf{u}^* := \sum_{i=1}^{k} \frac{\langle \mathbf{x}, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \mathbf{u}_i$; $\mathbf{u} \in U$; WTS $||\mathbf{x} - \mathbf{u}^*|| \leq ||\mathbf{x} - \mathbf{u}||$ and that equality holds iff $\mathbf{u}^* = \mathbf{u}$.

Now that we have shown that vectors $\mathbf{u}^* - \mathbf{u}$ and $\mathbf{x} - \mathbf{u}^*$ are orthogonal to each other, we can apply the Pythagorean theorem to them, as follows:

$$\begin{aligned} ||\mathbf{x} - \mathbf{u}||^2 &= ||(\mathbf{x} - \mathbf{u}^*) + (\mathbf{u}^* - \mathbf{u})||^2 \\ &\overset{(*)}{=} ||\mathbf{x} - \mathbf{u}^*||^2 + ||\mathbf{u}^* - \mathbf{u}||^2 \\ &\geq ||\mathbf{x} - \mathbf{u}^*||^2, \end{aligned}$$

where (*) follows from the Pythagorean theorem.

Proof (continued). Reminder: $\mathbf{u}^* := \sum_{i=1}^{k} \frac{\langle \mathbf{x}, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \mathbf{u}_i$; $\mathbf{u} \in U$; WTS $||\mathbf{x} - \mathbf{u}^*|| \leq ||\mathbf{x} - \mathbf{u}||$ and that equality holds iff $\mathbf{u}^* = \mathbf{u}$.

Now that we have shown that vectors $\mathbf{u}^* - \mathbf{u}$ and $\mathbf{x} - \mathbf{u}^*$ are orthogonal to each other, we can apply the Pythagorean theorem to them, as follows:

$$\begin{aligned} ||\mathbf{x} - \mathbf{u}||^2 &= ||(\mathbf{x} - \mathbf{u}^*) + (\mathbf{u}^* - \mathbf{u})||^2 \\ &\stackrel{(*)}{=} ||\mathbf{x} - \mathbf{u}^*||^2 + ||\mathbf{u}^* - \mathbf{u}||^2 \\ &\geq ||\mathbf{x} - \mathbf{u}^*||^2, \end{aligned}$$

where (*) follows from the Pythagorean theorem. Consequently, we have that $||\mathbf{x} - \mathbf{u}^*|| \leq ||\mathbf{x} - \mathbf{u}||$. Moreover, the inequality above is an equality iff $||\mathbf{u}^* - \mathbf{u}|| = 0$, i.e. iff $\mathbf{u}^* = \mathbf{u}$. This completes the argument. □

Theorem 6.5.1

Let V be a finite-dimensional real or complex vector space, equipped with a scalar product $\langle \cdot, \cdot \rangle$ and the induced norm $|| \cdot ||$. Let U be a subspace of V, and let $\mathbf{x} \in V$. Then there exists a unique vector $\mathbf{x}_U \in U$ that has the property that

$$||\mathbf{x} - \mathbf{x}_U|| \quad = \quad \min_{\mathbf{u} \in U} ||\mathbf{x} - \mathbf{u}||.$$

Moreover, if $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ is an orthogonal basis of U, then this vector $\mathbf{x}_U$ is given by the formula

$$\mathbf{x}_U \quad = \quad \sum_{i=1}^{k} \text{proj}_{\mathbf{u}_i}(\mathbf{x}) \quad = \quad \sum_{i=1}^{k} \frac{\langle \mathbf{x}, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \mathbf{u}_i.$$

- **Terminology/Notation:** The vector $\mathbf{x}_U$ from Theorem 6.5.1 is called the *orthogonal projection* of $\mathbf{x}$ onto U.

Corollary 6.5.2

Let V be a finite-dimensional real or complex vector space, equipped with a scalar product $\langle \cdot, \cdot \rangle$ and the induced norm $|| \cdot ||$. Let $\mathbf{u}$ be any non-zero vector in V, and set $U := \text{Span}(\mathbf{u})$.[a] Then for every $\mathbf{x} \in V$, we have that

$$\mathbf{x}_U \quad = \quad \text{proj}_{\mathbf{u}}(\mathbf{x}) \quad = \quad \tfrac{\langle \mathbf{x}, \mathbf{u} \rangle}{\langle \mathbf{u}, \mathbf{u} \rangle} \mathbf{u}.$$

[a] So, U is a one-dimensional subspace of V.

Proof.

Corollary 6.5.2

Let V be a finite-dimensional real or complex vector space, equipped with a scalar product $\langle \cdot, \cdot \rangle$ and the induced norm $|| \cdot ||$. Let $\mathbf{u}$ be any non-zero vector in V, and set $U := \text{Span}(\mathbf{u})$.[a] Then for every $\mathbf{x} \in V$, we have that

$$\mathbf{x}_U \quad = \quad \text{proj}_{\mathbf{u}}(\mathbf{x}) \quad = \quad \frac{\langle \mathbf{x}, \mathbf{u} \rangle}{\langle \mathbf{u}, \mathbf{u} \rangle} \mathbf{u}.$$

[a] So, U is a one-dimensional subspace of V.

Proof. Clearly, $\{\mathbf{u}\}$ is an orthogonal basis of U. So, the result follows immediately from Theorem 6.5.1. □

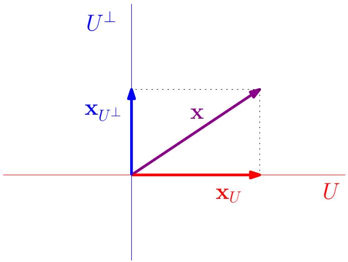
Corollary 6.5.3

Let V be a finite-dimensional real or complex vector space, equipped with a scalar product $\langle \cdot, \cdot \rangle$ and the induced norm $|| \cdot ||$. Let U be a subspace of V, and let $\mathbf{x} \in V$. Then

$$\mathbf{x} = \mathbf{x}_U + \mathbf{x}_{U^\perp}.$$

Moreover, this is the unique way of expressing $\mathbf{x}$ as a sum of a vector in U and a vector in $U^\perp$.[a]

[a] This means that for all $\mathbf{y} \in U$ and $\mathbf{z} \in U^\perp$, if $\mathbf{x} = \mathbf{y} + \mathbf{z}$, then $\mathbf{y} = \mathbf{x}_U$ and $\mathbf{z} = \mathbf{x}_{U^\perp}$.



Corollary 6.5.3

Let V be a finite-dimensional real or complex vector space, equipped with a scalar product $\langle \cdot, \cdot \rangle$ and the induced norm $|| \cdot ||$. Let U be a subspace of V, and let $\mathbf{x} \in V$. Then

$$\mathbf{x} = \mathbf{x}_U + \mathbf{x}_{U^\perp}.$$

Moreover, this is the unique way of expressing $\mathbf{x}$ as a sum of a vector in U and a vector in $U^\perp$.[a]

[a] This means that for all $\mathbf{y} \in U$ and $\mathbf{z} \in U^\perp$, if $\mathbf{x} = \mathbf{y} + \mathbf{z}$, then $\mathbf{y} = \mathbf{x}_U$ and $\mathbf{z} = \mathbf{x}_{U^\perp}$.

Proof.

Corollary 6.5.3

Let V be a finite-dimensional real or complex vector space, equipped with a scalar product $\langle \cdot, \cdot \rangle$ and the induced norm $|| \cdot ||$. Let U be a subspace of V, and let $\mathbf{x} \in V$. Then

$$\mathbf{x} = \mathbf{x}_U + \mathbf{x}_{U^\perp}.$$

Moreover, this is the unique way of expressing $\mathbf{x}$ as a sum of a vector in U and a vector in $U^\perp$.[a]

[a] This means that for all $\mathbf{y} \in U$ and $\mathbf{z} \in U^\perp$, if $\mathbf{x} = \mathbf{y} + \mathbf{z}$, then $\mathbf{y} = \mathbf{x}_U$ and $\mathbf{z} = \mathbf{x}_{U^\perp}$.

Proof. By Corollary 6.3.11, U has an orthogonal basis $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$, and moreover, this basis can be extended to an orthogonal basis $\{\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{u}_{k+1}, \dots, \mathbf{u}_n\}$ of V. By Theorem 6.4.3(a), we have that $\{\mathbf{u}_{k+1}, \dots, \mathbf{u}_n\}$ is an orthogonal basis of $U^\perp$.

Corollary 6.5.3

Let V be a finite-dimensional real or complex vector space, equipped with a scalar product $\langle \cdot, \cdot \rangle$ and the induced norm $|| \cdot ||$. Let U be a subspace of V, and let $\mathbf{x} \in V$. Then

$$\mathbf{x} \quad = \quad \mathbf{x}_U + \mathbf{x}_{U^\perp}.$$

Moreover, this is the unique way of expressing $\mathbf{x}$ as a sum of a vector in U and a vector in $U^\perp$.[a]

[a] This means that for all $\mathbf{y} \in U$ and $\mathbf{z} \in U^\perp$, if $\mathbf{x} = \mathbf{y} + \mathbf{z}$, then $\mathbf{y} = \mathbf{x}_U$ and $\mathbf{z} = \mathbf{x}_{U^\perp}$.

Proof. By Corollary 6.3.11, U has an orthogonal basis $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$, and moreover, this basis can be extended to an orthogonal basis $\{\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{u}_{k+1}, \dots, \mathbf{u}_n\}$ of V. By Theorem 6.4.3(a), we have that $\{\mathbf{u}_{k+1}, \dots, \mathbf{u}_n\}$ is an orthogonal basis of $U^\perp$. Now, by Theorem 6.3.5, we have that

$$\mathbf{x}_U \quad = \quad \sum_{i=1}^{k} \frac{\langle \mathbf{x}, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \mathbf{u}_i \qquad \text{and} \qquad \mathbf{x}_{U^\perp} \quad = \quad \sum_{i=k+1}^{n} \frac{\langle \mathbf{x}, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \mathbf{u}_i.$$

Solution (continued). On the other hand, by Theorem 6.3.5, we have that

$$\mathbf{x} \quad = \quad \sum_{i=1}^{n} \frac{\langle \mathbf{x}, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \mathbf{u}_i.$$

Solution (continued). On the other hand, by Theorem 6.3.5, we have that

$$\mathbf{x} \;=\; \sum_{i=1}^{n} \frac{\langle \mathbf{x}, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \mathbf{u}_i.$$

Consequently,

$$\mathbf{x} \;=\; \sum_{i=1}^{n} \frac{\langle \mathbf{x}, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \mathbf{u}_i \;=\; \left(\sum_{i=1}^{k} \frac{\langle \mathbf{x}, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \mathbf{u}_i \right) + \left(\sum_{i=k+1}^{n} \frac{\langle \mathbf{x}, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \mathbf{u}_i \right) \;=\; \mathbf{x}_U + \mathbf{x}_{U^\perp}.$$

Solution (continued). On the other hand, by Theorem 6.3.5, we have that

$$\mathbf{x} = \sum_{i=1}^{n} \frac{\langle \mathbf{x}, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \mathbf{u}_i.$$

Consequently,

$$\mathbf{x} = \sum_{i=1}^{n} \frac{\langle \mathbf{x}, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \mathbf{u}_i = \Big(\sum_{i=1}^{k} \frac{\langle \mathbf{x}, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \mathbf{u}_i \Big) + \Big(\sum_{i=k+1}^{n} \frac{\langle \mathbf{x}, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \mathbf{u}_i \Big) = \mathbf{x}_U + \mathbf{x}_{U^\perp}.$$

It remains to prove the uniqueness part of the corollary.

Solution (continued). On the other hand, by Theorem 6.3.5, we have that

$$\mathbf{x} \;=\; \sum_{i=1}^{n} \frac{\langle \mathbf{x}, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \mathbf{u}_i.$$

Consequently,

$$\mathbf{x} \;=\; \sum_{i=1}^{n} \frac{\langle \mathbf{x}, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \mathbf{u}_i \;=\; \Big(\sum_{i=1}^{k} \frac{\langle \mathbf{x}, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \mathbf{u}_i \Big) + \Big(\sum_{i=k+1}^{n} \frac{\langle \mathbf{x}, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \mathbf{u}_i \Big) \;=\; \mathbf{x}_U + \mathbf{x}_{U^\perp}.$$

It remains to prove the uniqueness part of the corollary. So, suppose that $\mathbf{y} \in U$ and $\mathbf{z} \in U^\perp$ are such that $\mathbf{x} = \mathbf{y} + \mathbf{z}$. WTS $\mathbf{y} = \mathbf{x}_U$ and $\mathbf{z} = \mathbf{x}_{U^\perp}$.

Solution (continued). On the other hand, by Theorem 6.3.5, we have that

$$\mathbf{x} = \sum_{i=1}^{n} \frac{\langle \mathbf{x}, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \mathbf{u}_i.$$

Consequently,

$$\mathbf{x} = \sum_{i=1}^{n} \frac{\langle \mathbf{x}, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \mathbf{u}_i = \Big(\sum_{i=1}^{k} \frac{\langle \mathbf{x}, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \mathbf{u}_i \Big) + \Big(\sum_{i=k+1}^{n} \frac{\langle \mathbf{x}, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \mathbf{u}_i \Big) = \mathbf{x}_U + \mathbf{x}_{U^\perp}.$$

It remains to prove the uniqueness part of the corollary. So, suppose that $\mathbf{y} \in U$ and $\mathbf{z} \in U^\perp$ are such that $\mathbf{x} = \mathbf{y} + \mathbf{z}$. WTS $\mathbf{y} = \mathbf{x}_U$ and $\mathbf{z} = \mathbf{x}_{U^\perp}$. We have that

$$\mathbf{x}_U + \mathbf{x}_{U^\perp} = \mathbf{x} = \mathbf{y} + \mathbf{z},$$

and consequently,

$$\mathbf{x}_U - \mathbf{y} = \mathbf{z} - \mathbf{x}_{U^\perp}.$$

Solution (continued). On the other hand, by Theorem 6.3.5, we have that

$$\mathbf{x} = \sum_{i=1}^{n} \frac{\langle \mathbf{x}, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \mathbf{u}_i.$$

Consequently,

$$\mathbf{x} = \sum_{i=1}^{n} \frac{\langle \mathbf{x}, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \mathbf{u}_i = \left(\sum_{i=1}^{k} \frac{\langle \mathbf{x}, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \mathbf{u}_i \right) + \left(\sum_{i=k+1}^{n} \frac{\langle \mathbf{x}, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \mathbf{u}_i \right) = \mathbf{x}_U + \mathbf{x}_{U^\perp}.$$

It remains to prove the uniqueness part of the corollary. So, suppose that $\mathbf{y} \in U$ and $\mathbf{z} \in U^\perp$ are such that $\mathbf{x} = \mathbf{y} + \mathbf{z}$. WTS $\mathbf{y} = \mathbf{x}_U$ and $\mathbf{z} = \mathbf{x}_{U^\perp}$. We have that

$$\mathbf{x}_U + \mathbf{x}_{U^\perp} = \mathbf{x} = \mathbf{y} + \mathbf{z},$$

and consequently,

$$\mathbf{x}_U - \mathbf{y} = \mathbf{z} - \mathbf{x}_{U^\perp}.$$

But $\mathbf{x}_U - \mathbf{y} \in U$ and $\mathbf{z} - \mathbf{x}_{U^\perp} \in U^\perp$.

Solution (continued). On the other hand, by Theorem 6.3.5, we have that

$$\mathbf{x} \;=\; \sum_{i=1}^{n} \frac{\langle \mathbf{x}, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \mathbf{u}_i.$$

Consequently,

$$\mathbf{x} \;=\; \sum_{i=1}^{n} \frac{\langle \mathbf{x}, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \mathbf{u}_i \;=\; \Big(\sum_{i=1}^{k} \frac{\langle \mathbf{x}, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \mathbf{u}_i \Big) + \Big(\sum_{i=k+1}^{n} \frac{\langle \mathbf{x}, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \mathbf{u}_i \Big) \;=\; \mathbf{x}_U + \mathbf{x}_{U^\perp}.$$

It remains to prove the uniqueness part of the corollary. So, suppose that $\mathbf{y} \in U$ and $\mathbf{z} \in U^\perp$ are such that $\mathbf{x} = \mathbf{y} + \mathbf{z}$. WTS $\mathbf{y} = \mathbf{x}_U$ and $\mathbf{z} = \mathbf{x}_{U^\perp}$. We have that

$$\mathbf{x}_U + \mathbf{x}_{U^\perp} \;=\; \mathbf{x} \;=\; \mathbf{y} + \mathbf{z},$$

and consequently,

$$\mathbf{x}_U - \mathbf{y} \;=\; \mathbf{z} - \mathbf{x}_{U^\perp}.$$

But $\mathbf{x}_U - \mathbf{y} \in U$ and $\mathbf{z} - \mathbf{x}_{U^\perp} \in U^\perp$. Since $U \cap U^\perp = \{\mathbf{0}\}$ (by Theorem 6.4.3(d)), it follows that $\mathbf{x}_U - \mathbf{y} = \mathbf{z} - \mathbf{x}_{U^\perp} = \mathbf{0}$,

Solution (continued). On the other hand, by Theorem 6.3.5, we have that

$$\mathbf{x} \;=\; \sum_{i=1}^{n} \frac{\langle \mathbf{x}, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \mathbf{u}_i.$$

Consequently,

$$\mathbf{x} \;=\; \sum_{i=1}^{n} \frac{\langle \mathbf{x}, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \mathbf{u}_i \;=\; \Big(\sum_{i=1}^{k} \frac{\langle \mathbf{x}, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \mathbf{u}_i \Big) + \Big(\sum_{i=k+1}^{n} \frac{\langle \mathbf{x}, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \mathbf{u}_i \Big) \;=\; \mathbf{x}_U + \mathbf{x}_{U^\perp}.$$

It remains to prove the uniqueness part of the corollary. So, suppose that $\mathbf{y} \in U$ and $\mathbf{z} \in U^\perp$ are such that $\mathbf{x} = \mathbf{y} + \mathbf{z}$. WTS $\mathbf{y} = \mathbf{x}_U$ and $\mathbf{z} = \mathbf{x}_{U^\perp}$. We have that

$$\mathbf{x}_U + \mathbf{x}_{U^\perp} \;=\; \mathbf{x} \;=\; \mathbf{y} + \mathbf{z},$$

and consequently,

$$\mathbf{x}_U - \mathbf{y} \;=\; \mathbf{z} - \mathbf{x}_{U^\perp}.$$

But $\mathbf{x}_U - \mathbf{y} \in U$ and $\mathbf{z} - \mathbf{x}_{U^\perp} \in U^\perp$. Since $U \cap U^\perp = \{\mathbf{0}\}$ (by Theorem 6.4.3(d)), it follows that $\mathbf{x}_U - \mathbf{y} = \mathbf{z} - \mathbf{x}_{U^\perp} = \mathbf{0}$, and consequently, $\mathbf{y} = \mathbf{x}_U$ and $\mathbf{z} = \mathbf{x}_{U^\perp}$. $\square$

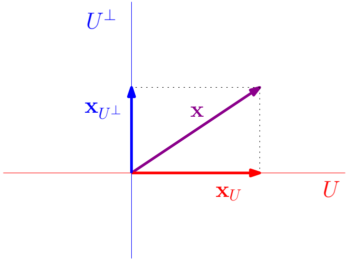
Corollary 6.5.3

Let V be a finite-dimensional real or complex vector space, equipped with a scalar product $\langle\cdot,\cdot\rangle$ and the induced norm $||\cdot||$. Let U be a subspace of V, and let $\mathbf{x} \in V$. Then

$$\mathbf{x} \quad = \quad \mathbf{x}_U + \mathbf{x}_{U^\perp}.$$

Moreover, this is the unique way of expressing $\mathbf{x}$ as a sum of a vector in U and a vector in $U^\perp$.[a]

[a] This means that for all $\mathbf{y} \in U$ and $\mathbf{z} \in U^\perp$, if $\mathbf{x} = \mathbf{y} + \mathbf{z}$, then $\mathbf{y} = \mathbf{x}_U$ and $\mathbf{z} = \mathbf{x}_{U^\perp}$.



- Suppose that V is a finite-dimensional real or complex vector space, equipped with a scalar product $\langle \cdot, \cdot \rangle$ and the induced norm $|| \cdot ||$, and suppose that U is a subspace of V.

- Suppose that V is a finite-dimensional real or complex vector space, equipped with a scalar product $\langle \cdot, \cdot \rangle$ and the induced norm $|| \cdot ||$, and suppose that U is a subspace of V.
- We can then define the function $\text{proj}_U : V \to V$ by setting $\text{proj}_U(\mathbf{x}) = \mathbf{x}_U$ for all $\mathbf{x} \in V$ (where $\mathbf{x}_U$ is the orthogonal projection of $\mathbf{x}$ onto U, as in Theorem 6.5.1).

- Suppose that V is a finite-dimensional real or complex vector space, equipped with a scalar product $\langle \cdot, \cdot \rangle$ and the induced norm $|| \cdot ||$, and suppose that U is a subspace of V.
- We can then define the function $\text{proj}_U : V \to V$ by setting $\text{proj}_U(\mathbf{x}) = \mathbf{x}_U$ for all $\mathbf{x} \in V$ (where $\mathbf{x}_U$ is the orthogonal projection of $\mathbf{x}$ onto U, as in Theorem 6.5.1).
- Clearly, $\text{proj}_U(\mathbf{u}) = \mathbf{u}$ for all $\mathbf{u} \in U$.

- Suppose that V is a finite-dimensional real or complex vector space, equipped with a scalar product $\langle \cdot, \cdot \rangle$ and the induced norm $|| \cdot ||$, and suppose that U is a subspace of V.
- We can then define the function $\text{proj}_U : V \to V$ by setting $\text{proj}_U(\mathbf{x}) = \mathbf{x}_U$ for all $\mathbf{x} \in V$ (where $\mathbf{x}_U$ is the orthogonal projection of $\mathbf{x}$ onto U, as in Theorem 6.5.1).
- Clearly, $\text{proj}_U(\mathbf{u}) = \mathbf{u}$ for all $\mathbf{u} \in U$.
- Moreover, we have that $\text{Im}\big(\text{proj}_U\big) = U$ and $\text{proj}_U[U] = U$.

- Suppose that V is a finite-dimensional real or complex vector space, equipped with a scalar product $\langle \cdot, \cdot \rangle$ and the induced norm $|| \cdot ||$, and suppose that U is a subspace of V.
- We can then define the function $\text{proj}_U : V \to V$ by setting $\text{proj}_U(\mathbf{x}) = \mathbf{x}_U$ for all $\mathbf{x} \in V$ (where $\mathbf{x}_U$ is the orthogonal projection of $\mathbf{x}$ onto U, as in Theorem 6.5.1).
- Clearly, $\text{proj}_U(\mathbf{u}) = \mathbf{u}$ for all $\mathbf{u} \in U$.
- Moreover, we have that $\text{Im}\big(\text{proj}_U\big) = U$ and $\text{proj}_U[U] = U$.
- Using the formula from Theorem 6.5.1, we can easily see that the function proj_U is linear.

- Suppose that V is a finite-dimensional real or complex vector space, equipped with a scalar product $\langle \cdot, \cdot \rangle$ and the induced norm $|| \cdot ||$, and suppose that U is a subspace of V.
- We can then define the function $\text{proj}_U : V \to V$ by setting $\text{proj}_U(\mathbf{x}) = \mathbf{x}_U$ for all $\mathbf{x} \in V$ (where $\mathbf{x}_U$ is the orthogonal projection of $\mathbf{x}$ onto U, as in Theorem 6.5.1).
- Clearly, $\text{proj}_U(\mathbf{u}) = \mathbf{u}$ for all $\mathbf{u} \in U$.
- Moreover, we have that $\text{Im}\big(\text{proj}_U\big) = U$ and $\text{proj}_U[U] = U$.
- Using the formula from Theorem 6.5.1, we can easily see that the function proj_U is linear.
- Indeed, if $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ is any orthogonal basis of U (such a basis exists by Corollary 6.3.11), then the following hold (next two slides):

- for all $\mathbf{x}, \mathbf{y} \in V$, we have that

$$\begin{aligned}
\text{proj}_U(\mathbf{x}+\mathbf{y}) &\overset{(*)}{=} \sum_{i=1}^{k} \frac{\langle \mathbf{x}+\mathbf{y}, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \mathbf{u}_i \\
&\overset{(**)}{=} \sum_{i=1}^{k} \frac{\langle \mathbf{x}, \mathbf{u}_i \rangle + \langle \mathbf{y}, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \mathbf{u}_i \\
&= \Big(\sum_{i=1}^{k} \frac{\langle \mathbf{x}, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \mathbf{u}_i \Big) + \Big(\sum_{i=1}^{k} \frac{\langle \mathbf{y}, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \mathbf{u}_i \Big) \\
&\overset{(*)}{=} \text{proj}_U(\mathbf{x}) + \text{proj}_U(\mathbf{y}),
\end{aligned}$$

  where both instances of (*) follow from Theorem 6.5.1, and (**) follows from r.2 or c.2;

- for all $\mathbf{x} \in V$ and scalars α, we have that

$$\begin{aligned}
\text{proj}_U(\alpha\mathbf{x}) &\overset{(*)}{=} \sum_{i=1}^{k} \frac{\langle \alpha\mathbf{x}, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \mathbf{u}_i \\
&\overset{(**)}{=} \sum_{i=1}^{k} \frac{\alpha\langle \mathbf{x}, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \mathbf{u}_i \\
&= \alpha \sum_{i=1}^{k} \frac{\langle \mathbf{x}, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \mathbf{u}_i \\
&\overset{(*)}{=} \alpha \text{proj}_U(\mathbf{x}),
\end{aligned}$$

where both instances of (*) follow from Theorem 6.5.1, and (**) follows from r.3 or c.3.

3. Orthogonal projection onto subspaces of $\mathbb{R}^n$

3. Orthogonal projection onto subspaces of $\mathbb{R}^n$

- In the remainder of this lecture, we assume that $\mathbb{R}^n$ is equipped with the standard scalar product $\cdot$ and the induced norm $||\cdot||$.

3. Orthogonal projection onto subspaces of $\mathbb{R}^n$

- In the remainder of this lecture, we assume that $\mathbb{R}^n$ is equipped with the standard scalar product $\cdot$ and the induced norm $||\cdot||$.
- Now, suppose that U is a subspace of $\mathbb{R}^n$.

3. Orthogonal projection onto subspaces of $\mathbb{R}^n$

- In the remainder of this lecture, we assume that $\mathbb{R}^n$ is equipped with the standard scalar product $\cdot$ and the induced norm $|| \cdot ||$.
- Now, suppose that U is a subspace of $\mathbb{R}^n$.
  - As we saw, $\text{proj}_U : \mathbb{R}^n \to \mathbb{R}^n$ is linear (and its image is U).

3. Orthogonal projection onto subspaces of $\mathbb{R}^n$

- In the remainder of this lecture, we assume that $\mathbb{R}^n$ is equipped with the standard scalar product $\cdot$ and the induced norm $||\cdot||$.
- Now, suppose that U is a subspace of $\mathbb{R}^n$.
  - As we saw, $\text{proj}_U : \mathbb{R}^n \to \mathbb{R}^n$ is linear (and its image is U).
  - Since proj_U is linear, it has a standard matrix.

3. Orthogonal projection onto subspaces of $\mathbb{R}^n$

- In the remainder of this lecture, we assume that $\mathbb{R}^n$ is equipped with the standard scalar product $\cdot$ and the induced norm $||\cdot||$.
- Now, suppose that U is a subspace of $\mathbb{R}^n$.
  - As we saw, $\text{proj}_U : \mathbb{R}^n \to \mathbb{R}^n$ is linear (and its image is U).
  - Since proj_U is linear, it has a standard matrix.
    - Note that this matrix belongs to $\mathbb{R}^{n \times n}$.

3. Orthogonal projection onto subspaces of $\mathbb{R}^n$

- In the remainder of this lecture, we assume that $\mathbb{R}^n$ is equipped with the standard scalar product $\cdot$ and the induced norm $||\cdot||$.
- Now, suppose that U is a subspace of $\mathbb{R}^n$.
  - As we saw, $\text{proj}_U : \mathbb{R}^n \to \mathbb{R}^n$ is linear (and its image is U).
  - Since proj_U is linear, it has a standard matrix.
    - Note that this matrix belongs to $\mathbb{R}^{n \times n}$.
  - Our goal is to we give formulas for the standard matrices of orthogonal projections onto various subspaces of $\mathbb{R}^n$.

- In section 3.3 of the Lecture Notes (last semester), we defined the row space of a matrix A to be the span of the rows of A.

- In section 3.3 of the Lecture Notes (last semester), we defined the row space of a matrix A to be the span of the rows of A.
- Proposition 3.3.1(b) states that

$$\text{Row}(A) \quad = \quad \{\mathbf{u}^T \mid \mathbf{u} \in \text{Col}(A^T)\}.$$

- In section 3.3 of the Lecture Notes (last semester), we defined the row space of a matrix A to be the span of the rows of A.
- Proposition 3.3.1(b) states that

$$\mathrm{Row}(A) \;\; = \;\; \{\mathbf{u}^T \mid \mathbf{u} \in \mathrm{Col}(A^T)\}.$$

- In what follows, it will be convenient to slightly modify the definition of the row space, as follows:

$$\mathrm{Row}(A) \;\; := \;\; \mathrm{Col}(A^T).$$

- In section 3.3 of the Lecture Notes (last semester), we defined the row space of a matrix A to be the span of the rows of A.
- Proposition 3.3.1(b) states that

$$\mathrm{Row}(A) \quad = \quad \{\mathbf{u}^T \mid \mathbf{u} \in \mathrm{Col}(A^T)\}.$$

- In what follows, it will be convenient to slightly modify the definition of the row space, as follows:

$$\mathrm{Row}(A) \quad := \quad \mathrm{Col}(A^T).$$

- So, we (re)defined the row space of a matrix to be the span of the transposes of its rows.

- For example, for the matrix

$$A = \begin{bmatrix} 1 & 2 & 1 & 2 \\ 2 & 3 & 2 & 3 \\ 3 & 4 & 3 & 4 \end{bmatrix},$$

we have that

$$A^T = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 1 & 2 & 3 \\ 2 & 3 & 4 \end{bmatrix},$$

and consequently,

$$\text{Row}(A) = \text{Span}\left(\begin{bmatrix} 1 \\ 2 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ 3 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 3 \\ 4 \\ 3 \\ 4 \end{bmatrix} \right).$$

- For example, for the matrix

$$A = \begin{bmatrix} 1 & 2 & 1 & 2 \\ 2 & 3 & 2 & 3 \\ 3 & 4 & 3 & 4 \end{bmatrix},$$

we have that

$$A^T = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 1 & 2 & 3 \\ 2 & 3 & 4 \end{bmatrix},$$

and consequently,

$$\text{Row}(A) = \text{Span}\left(\begin{bmatrix} 1 \\ 2 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ 3 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 3 \\ 4 \\ 3 \\ 4 \end{bmatrix} \right).$$

- If this change of definition bothers you, then every time you see $\text{Row}(\square)$, mentally replace it with $\text{Col}(\square^T)$.

Theorem 6.6.1

Let $A \in \mathbb{R}^{n \times m}$. Then $\text{Row}(A)^\perp = \text{Nul}(A)$ and $\text{Row}(A) = \text{Nul}(A)^\perp$.

Proof.

Theorem 6.6.1

Let $A \in \mathbb{R}^{n \times m}$. Then $\text{Row}(A)^\perp = \text{Nul}(A)$ and $\text{Row}(A) = \text{Nul}(A)^\perp$.

Proof. In view of Theorem 6.4.3(c), it suffices to show that $\text{Row}(A)^\perp = \text{Nul}(A)$.

- Indeed, by Theorem 6.4.3(c), we have that

$$(\text{Row}(A)^\perp)^\perp \ = \ \text{Row}(A).$$

- So, if $\text{Row}(A)^\perp = \text{Nul}(A)$, then

$$\text{Nul}(A)^\perp \ = \ (\text{Row}(A)^\perp)^\perp = \text{Row}(A).$$

Theorem 6.6.1

Let $A \in \mathbb{R}^{n \times m}$. Then $\text{Row}(A)^\perp = \text{Nul}(A)$ and $\text{Row}(A) = \text{Nul}(A)^\perp$.

Proof. In view of Theorem 6.4.3(c), it suffices to show that $\text{Row}(A)^\perp = \text{Nul}(A)$.

- Indeed, by Theorem 6.4.3(c), we have that

$$(\text{Row}(A)^\perp)^\perp \;=\; \text{Row}(A).$$

- So, if $\text{Row}(A)^\perp = \text{Nul}(A)$, then

$$\text{Nul}(A)^\perp \;=\; (\text{Row}(A)^\perp)^\perp = \text{Row}(A).$$

Set

$$A \;=\; \begin{bmatrix} \mathbf{a}_1^T \\ \vdots \\ \mathbf{a}_n^T \end{bmatrix},$$

so that $\text{Row}(A) = \text{Span}(\mathbf{a}_1, \dots, \mathbf{a}_n)$.

Proof (continued). Now, for all vectors $\mathbf{x} \in \mathbb{R}^m$:

$$\begin{aligned}
\mathbf{x} \in \text{Nul}(A) \quad &\iff \quad A\mathbf{x} = \mathbf{0} \\
&\iff \quad \begin{bmatrix} \mathbf{a}_1^T \\ \vdots \\ \mathbf{a}_n^T \end{bmatrix} \mathbf{x} = \mathbf{0} \\
&\iff \quad \begin{bmatrix} \mathbf{a}_1 \cdot \mathbf{x} \\ \vdots \\ \mathbf{a}_n \cdot \mathbf{x} \end{bmatrix} = \mathbf{0} \\
&\iff \quad \mathbf{a}_i \cdot \mathbf{x} = 0 \;\; \forall i \in \{1, \dots, n\} \\
&\iff \quad \mathbf{a}_i \perp \mathbf{x} \;\; \forall i \in \{1, \dots, n\} \\
&\iff \quad \mathbf{x} \in \{\mathbf{a}_1, \dots, \mathbf{a}_n\}^\perp \\
&\overset{(*)}{\iff} \quad \mathbf{x} \in \text{Span}(\mathbf{a}_1, \dots, \mathbf{a}_n)^\perp \\
&\iff \quad \mathbf{x} \in \text{Row}(A)^\perp,
\end{aligned}$$

where (*) follows from the fact that $\{\mathbf{a}_1, \dots, \mathbf{a}_m\}^\perp = \text{Span}(\mathbf{a}_1, \dots, \mathbf{a}_m)^\perp$ (by Proposition 6.4.2). This proves that $\text{Nul}(A) = \text{Row}(A)^\perp$, and we are done. $\square$

Theorem 6.6.1

Let $A \in \mathbb{R}^{n \times m}$. Then $\mathrm{Row}(A)^\perp = \mathrm{Nul}(A)$ and $\mathrm{Row}(A) = \mathrm{Nul}(A)^\perp$.

Theorem 6.6.1

Let $A \in \mathbb{R}^{n \times m}$. Then $\text{Row}(A)^{\perp} = \text{Nul}(A)$ and $\text{Row}(A) = \text{Nul}(A)^{\perp}$.

Corollary 6.6.2

Let $A \in \mathbb{R}^{n \times m}$. Then all the following hold:

(a) $\text{Nul}(A^T A) = \text{Nul}(A)$;

(b) $\text{Row}(A^T A) = \text{Row}(A)$;

(c) $\text{rank}(A^T A) = \text{rank}(A)$.

Proof.

Theorem 6.6.1

Let $A \in \mathbb{R}^{n\times m}$. Then $\text{Row}(A)^\perp = \text{Nul}(A)$ and $\text{Row}(A) = \text{Nul}(A)^\perp$.

Corollary 6.6.2

Let $A \in \mathbb{R}^{n\times m}$. Then all the following hold:

a. $\text{Nul}(A^T A) = \text{Nul}(A)$;
b. $\text{Row}(A^T A) = \text{Row}(A)$;
c. $\text{rank}(A^T A) = \text{rank}(A)$.

Proof. We first prove (a).

Theorem 6.6.1

Let $A \in \mathbb{R}^{n \times m}$. Then $\text{Row}(A)^\perp = \text{Nul}(A)$ and $\text{Row}(A) = \text{Nul}(A)^\perp$.

Corollary 6.6.2

Let $A \in \mathbb{R}^{n \times m}$. Then all the following hold:

(a) $\text{Nul}(A^T A) = \text{Nul}(A)$;

(b) $\text{Row}(A^T A) = \text{Row}(A)$;

(c) $\text{rank}(A^T A) = \text{rank}(A)$.

Proof. We first prove (a). Note that $A^T A \in \mathbb{R}^{m \times m}$, and that both $\text{Nul}(A)$ and $\text{Nul}(A^T A)$ are subspaces of $\mathbb{R}^m$.

Theorem 6.6.1

Let $A \in \mathbb{R}^{n \times m}$. Then $\text{Row}(A)^{\perp} = \text{Nul}(A)$ and $\text{Row}(A) = \text{Nul}(A)^{\perp}$.

Corollary 6.6.2

Let $A \in \mathbb{R}^{n \times m}$. Then all the following hold:

(a) $\text{Nul}(A^T A) = \text{Nul}(A)$;

(b) $\text{Row}(A^T A) = \text{Row}(A)$;

(c) $\text{rank}(A^T A) = \text{rank}(A)$.

Proof. We first prove (a). Note that $A^T A \in \mathbb{R}^{m \times m}$, and that both $\text{Nul}(A)$ and $\text{Nul}(A^T A)$ are subspaces of $\mathbb{R}^m$. Now, fix any $\mathbf{x} \in \mathbb{R}^m$. WTS $\mathbf{x} \in \text{Nul}(A^T A)$ iff $\mathbf{x} \in \text{Nul}(A)$.

Theorem 6.6.1

Let $A \in \mathbb{R}^{n \times m}$. Then $\text{Row}(A)^\perp = \text{Nul}(A)$ and $\text{Row}(A) = \text{Nul}(A)^\perp$.

Corollary 6.6.2

Let $A \in \mathbb{R}^{n \times m}$. Then all the following hold:

(a) $\text{Nul}(A^T A) = \text{Nul}(A)$;

(b) $\text{Row}(A^T A) = \text{Row}(A)$;

(c) $\text{rank}(A^T A) = \text{rank}(A)$.

Proof. We first prove (a). Note that $A^T A \in \mathbb{R}^{m \times m}$, and that both $\text{Nul}(A)$ and $\text{Nul}(A^T A)$ are subspaces of $\mathbb{R}^m$. Now, fix any $\mathbf{x} \in \mathbb{R}^m$. WTS $\mathbf{x} \in \text{Nul}(A^T A)$ iff $\mathbf{x} \in \text{Nul}(A)$.

Suppose first that $\mathbf{x} \in \text{Nul}(A)$.

Theorem 6.6.1

Let $A \in \mathbb{R}^{n \times m}$. Then $\text{Row}(A)^\perp = \text{Nul}(A)$ and $\text{Row}(A) = \text{Nul}(A)^\perp$.

Corollary 6.6.2

Let $A \in \mathbb{R}^{n \times m}$. Then all the following hold:

(a) $\text{Nul}(A^T A) = \text{Nul}(A)$;

(b) $\text{Row}(A^T A) = \text{Row}(A)$;

(c) $\text{rank}(A^T A) = \text{rank}(A)$.

Proof. We first prove (a). Note that $A^T A \in \mathbb{R}^{m \times m}$, and that both $\text{Nul}(A)$ and $\text{Nul}(A^T A)$ are subspaces of $\mathbb{R}^m$. Now, fix any $\mathbf{x} \in \mathbb{R}^m$. WTS $\mathbf{x} \in \text{Nul}(A^T A)$ iff $\mathbf{x} \in \text{Nul}(A)$.

Suppose first that $\mathbf{x} \in \text{Nul}(A)$. Then $A\mathbf{x} = \mathbf{0}$, and consequently, $A^T A\mathbf{x} = \mathbf{0}$. So, $\mathbf{x} \in \text{Nul}(A^T A)$.

Theorem 6.6.1

Let $A \in \mathbb{R}^{n \times m}$. Then $\text{Row}(A)^{\perp} = \text{Nul}(A)$ and $\text{Row}(A) = \text{Nul}(A)^{\perp}$.

Corollary 6.6.2

Let $A \in \mathbb{R}^{n \times m}$. Then all the following hold:

- (a) $\text{Nul}(A^T A) = \text{Nul}(A)$;
- (b) $\text{Row}(A^T A) = \text{Row}(A)$;
- (c) $\text{rank}(A^T A) = \text{rank}(A)$.

Proof (continued). Suppose, conversely, that $\mathbf{x} \in \text{Nul}(A^T A)$.

Theorem 6.6.1

Let $A \in \mathbb{R}^{n\times m}$. Then $\text{Row}(A)^\perp = \text{Nul}(A)$ and $\text{Row}(A) = \text{Nul}(A)^\perp$.

Corollary 6.6.2

Let $A \in \mathbb{R}^{n\times m}$. Then all the following hold:

(a) $\text{Nul}(A^TA) = \text{Nul}(A)$;

(b) $\text{Row}(A^TA) = \text{Row}(A)$;

(c) $\text{rank}(A^TA) = \text{rank}(A)$.

Proof (continued). Suppose, conversely, that $\mathbf{x} \in \text{Nul}(A^TA)$. Then $A^TA\mathbf{x} = \mathbf{0}$, and it follows that $\mathbf{x}^TA^TA\mathbf{x} = \mathbf{0}$.

Theorem 6.6.1

Let $A \in \mathbb{R}^{n \times m}$. Then $\text{Row}(A)^\perp = \text{Nul}(A)$ and $\text{Row}(A) = \text{Nul}(A)^\perp$.

Corollary 6.6.2

Let $A \in \mathbb{R}^{n \times m}$. Then all the following hold:

a. $\text{Nul}(A^T A) = \text{Nul}(A)$;
b. $\text{Row}(A^T A) = \text{Row}(A)$;
c. $\text{rank}(A^T A) = \text{rank}(A)$.

Proof (continued). Suppose, conversely, that $\mathbf{x} \in \text{Nul}(A^T A)$. Then $A^T A\mathbf{x} = \mathbf{0}$, and it follows that $\mathbf{x}^T A^T A\mathbf{x} = \mathbf{0}$. But note that

$$\mathbf{x}^T A^T A\mathbf{x} \;=\; (A\mathbf{x})^T(A\mathbf{x}) \;=\; (A\mathbf{x}) \cdot (A\mathbf{x}) \;=\; ||A\mathbf{x}||^2;$$

Theorem 6.6.1

Let $A \in \mathbb{R}^{n \times m}$. Then $\text{Row}(A)^{\perp} = \text{Nul}(A)$ and $\text{Row}(A) = \text{Nul}(A)^{\perp}$.

Corollary 6.6.2

Let $A \in \mathbb{R}^{n \times m}$. Then all the following hold:

(a) $\text{Nul}(A^T A) = \text{Nul}(A)$;

(b) $\text{Row}(A^T A) = \text{Row}(A)$;

(c) $\text{rank}(A^T A) = \text{rank}(A)$.

Proof (continued). Suppose, conversely, that $\mathbf{x} \in \text{Nul}(A^T A)$. Then $A^T A\mathbf{x} = \mathbf{0}$, and it follows that $\mathbf{x}^T A^T A\mathbf{x} = \mathbf{0}$. But note that

$$\mathbf{x}^T A^T A\mathbf{x} \quad = \quad (A\mathbf{x})^T (A\mathbf{x}) \quad = \quad (A\mathbf{x}) \cdot (A\mathbf{x}) \quad = \quad ||A\mathbf{x}||^2;$$

consequently, $||A\mathbf{x}||^2 = 0$.

Theorem 6.6.1

Let $A \in \mathbb{R}^{n \times m}$. Then $\text{Row}(A)^{\perp} = \text{Nul}(A)$ and $\text{Row}(A) = \text{Nul}(A)^{\perp}$.

Corollary 6.6.2

Let $A \in \mathbb{R}^{n \times m}$. Then all the following hold:

(a) $\text{Nul}(A^T A) = \text{Nul}(A)$;

(b) $\text{Row}(A^T A) = \text{Row}(A)$;

(c) $\text{rank}(A^T A) = \text{rank}(A)$.

Proof (continued). Suppose, conversely, that $\mathbf{x} \in \text{Nul}(A^T A)$. Then $A^T A\mathbf{x} = \mathbf{0}$, and it follows that $\mathbf{x}^T A^T A\mathbf{x} = \mathbf{0}$. But note that

$$\mathbf{x}^T A^T A\mathbf{x} = (A\mathbf{x})^T (A\mathbf{x}) = (A\mathbf{x}) \cdot (A\mathbf{x}) = ||A\mathbf{x}||^2;$$

consequently, $||A\mathbf{x}||^2 = 0$. It follows that $||A\mathbf{x}|| = 0$, and therefore, $A\mathbf{x} = \mathbf{0}$, i.e. $\mathbf{x} \in \text{Nul}(A)$. This proves (a).

Theorem 6.6.1

Let $A \in \mathbb{R}^{n \times m}$. Then $\text{Row}(A)^\perp = \text{Nul}(A)$ and $\text{Row}(A) = \text{Nul}(A)^\perp$.

Corollary 6.6.2

Let $A \in \mathbb{R}^{n \times m}$. Then all the following hold:

(a) $\text{Nul}(A^T A) = \text{Nul}(A)$;

(b) $\text{Row}(A^T A) = \text{Row}(A)$;

(c) $\text{rank}(A^T A) = \text{rank}(A)$.

Proof (continued). For (b), we observe that

$$\begin{aligned}
\text{Row}(A^T A) &= \text{Nul}(A^T A)^\perp && \text{by Theorem 6.6.1} \\
&= \text{Nul}(A)^\perp && \text{by (a)} \\
&= \text{Row}(A) && \text{by Theorem 6.6.1.}
\end{aligned}$$

Theorem 6.6.1

Let $A \in \mathbb{R}^{n \times m}$. Then $\text{Row}(A)^\perp = \text{Nul}(A)$ and $\text{Row}(A) = \text{Nul}(A)^\perp$.

Corollary 6.6.2

Let $A \in \mathbb{R}^{n \times m}$. Then all the following hold:

- (a) $\text{Nul}(A^T A) = \text{Nul}(A)$;
- (b) $\text{Row}(A^T A) = \text{Row}(A)$;
- (c) $\text{rank}(A^T A) = \text{rank}(A)$.

Proof (continued). Finally, for (c), we have the following:

$$\begin{aligned} \text{rank}(A^T A) &= \dim\big(\text{Row}(A^T A)\big) && \text{by Theorem 3.3.9} \\ &= \dim\big(\text{Row}(A)\big) && \text{by (b)} \\ &= \text{rank}(A) && \text{by Theorem 3.3.9.} \end{aligned}$$

This completes the argument. □

Theorem 6.6.3

Let $A \in \mathbb{R}^{n \times m}$ be a matrix of rank m (i.e. A is a matrix of full column rank). Then the matrix $A(A^T A)^{-1}A^T$ is the standard matrix of orthogonal projection onto $\text{Col}(A)$, that is, for all $\mathbf{x} \in \mathbb{R}^n$, the orthogonal projection of $\mathbf{x}$ onto $C := \text{Col}(A)$ is given by

$$\mathbf{x}_C = A(A^T A)^{-1}A^T\mathbf{x}.$$

Proof.

Theorem 6.6.3

Let $A \in \mathbb{R}^{n \times m}$ be a matrix of rank m (i.e. A is a matrix of full column rank). Then the matrix $A(A^TA)^{-1}A^T$ is the standard matrix of orthogonal projection onto $\mathrm{Col}(A)$, that is, for all $\mathbf{x} \in \mathbb{R}^n$, the orthogonal projection of $\mathbf{x}$ onto $C := \mathrm{Col}(A)$ is given by

$$\mathbf{x}_C \quad = \quad A(A^TA)^{-1}A^T\mathbf{x}.$$

Proof. Fix $\mathbf{x} \in \mathbb{R}^n$. We must first check that the expression $A(A^TA)^{-1}A^T\mathbf{x}$ is defined and belongs to $C = \mathrm{Col}(A)$.

Theorem 6.6.3

Let $A \in \mathbb{R}^{n \times m}$ be a matrix of rank m (i.e. A is a matrix of full column rank). Then the matrix $A(A^TA)^{-1}A^T$ is the standard matrix of orthogonal projection onto $\text{Col}(A)$, that is, for all $\mathbf{x} \in \mathbb{R}^n$, the orthogonal projection of $\mathbf{x}$ onto $C := \text{Col}(A)$ is given by

$$\mathbf{x}_C = A(A^TA)^{-1}A^T\mathbf{x}.$$

Proof. Fix $\mathbf{x} \in \mathbb{R}^n$. We must first check that the expression $A(A^TA)^{-1}A^T\mathbf{x}$ is defined and belongs to $C = \text{Col}(A)$.

First, note that $A^TA \in \mathbb{R}^{m \times m}$, and that by Corollary 6.6.2(a), we have that $\text{rank}(A^TA) = \text{rank}(A) = m$.

Theorem 6.6.3

Let $A \in \mathbb{R}^{n\times m}$ be a matrix of rank m (i.e. A is a matrix of full column rank). Then the matrix $A(A^TA)^{-1}A^T$ is the standard matrix of orthogonal projection onto $\text{Col}(A)$, that is, for all $\mathbf{x} \in \mathbb{R}^n$, the orthogonal projection of $\mathbf{x}$ onto $C := \text{Col}(A)$ is given by

$$\mathbf{x}_C = A(A^TA)^{-1}A^T\mathbf{x}.$$

Proof. Fix $\mathbf{x} \in \mathbb{R}^n$. We must first check that the expression $A(A^TA)^{-1}A^T\mathbf{x}$ is defined and belongs to $C = \text{Col}(A)$.

First, note that $A^TA \in \mathbb{R}^{m\times m}$, and that by Corollary 6.6.2(a), we have that $\text{rank}(A^TA) = \text{rank}(A) = m$. So, by the Invertible Matrix Theorem, A^TA is invertible, and we see that $(A^TA)^{-1}$ is defined and belongs to $\mathbb{R}^{m\times m}$.

Theorem 6.6.3

Let $A \in \mathbb{R}^{n \times m}$ be a matrix of rank m (i.e. A is a matrix of full column rank). Then the matrix $A(A^T A)^{-1} A^T$ is the standard matrix of orthogonal projection onto $\text{Col}(A)$, that is, for all $\mathbf{x} \in \mathbb{R}^n$, the orthogonal projection of $\mathbf{x}$ onto $C := \text{Col}(A)$ is given by

$$\mathbf{x}_C \quad = \quad A(A^T A)^{-1} A^T \mathbf{x}.$$

Proof. Fix $\mathbf{x} \in \mathbb{R}^n$. We must first check that the expression $A(A^T A)^{-1} A^T \mathbf{x}$ is defined and belongs to $C = \text{Col}(A)$.

First, note that $A^T A \in \mathbb{R}^{m \times m}$, and that by Corollary 6.6.2(a), we have that $\text{rank}(A^T A) = \text{rank}(A) = m$. So, by the Invertible Matrix Theorem, $A^T A$ is invertible, and we see that $(A^T A)^{-1}$ is defined and belongs to $\mathbb{R}^{m \times m}$. Since $A \in \mathbb{R}^{n \times m}$, $(A^T A)^{-1} \in \mathbb{R}^{m \times m}$, and $A^T \in \mathbb{R}^{m \times n}$, we see that $A(A^T A)^{-1} A^T \in \mathbb{R}^{n \times n}$;

Theorem 6.6.3

Let $A \in \mathbb{R}^{n \times m}$ be a matrix of rank m (i.e. A is a matrix of full column rank). Then the matrix $A(A^TA)^{-1}A^T$ is the standard matrix of orthogonal projection onto $\text{Col}(A)$, that is, for all $\mathbf{x} \in \mathbb{R}^n$, the orthogonal projection of $\mathbf{x}$ onto $C := \text{Col}(A)$ is given by

$$\mathbf{x}_C = A(A^TA)^{-1}A^T\mathbf{x}.$$

Proof. Fix $\mathbf{x} \in \mathbb{R}^n$. We must first check that the expression $A(A^TA)^{-1}A^T\mathbf{x}$ is defined and belongs to $C = \text{Col}(A)$.

First, note that $A^TA \in \mathbb{R}^{m \times m}$, and that by Corollary 6.6.2(a), we have that $\text{rank}(A^TA) = \text{rank}(A) = m$. So, by the Invertible Matrix Theorem, A^TA is invertible, and we see that $(A^TA)^{-1}$ is defined and belongs to $\mathbb{R}^{m \times m}$. Since $A \in \mathbb{R}^{n \times m}$, $(A^TA)^{-1} \in \mathbb{R}^{m \times m}$, and $A^T \in \mathbb{R}^{m \times n}$, we see that $A(A^TA)^{-1}A^T \in \mathbb{R}^{n \times n}$; since $\mathbf{x} \in \mathbb{R}^n$, we see that $A(A^TA)^{-1}A^T\mathbf{x}$ is defined and belongs to $\mathbb{R}^n$.

Theorem 6.6.3

Let $A \in \mathbb{R}^{n \times m}$ be a matrix of rank m (i.e. A is a matrix of full column rank). Then the matrix $A(A^T A)^{-1} A^T$ is the standard matrix of orthogonal projection onto $\text{Col}(A)$, that is, for all $\mathbf{x} \in \mathbb{R}^n$, the orthogonal projection of $\mathbf{x}$ onto $C := \text{Col}(A)$ is given by

$$\mathbf{x}_C = A(A^T A)^{-1} A^T \mathbf{x}.$$

Proof. Fix $\mathbf{x} \in \mathbb{R}^n$. We must first check that the expression $A(A^T A)^{-1} A^T \mathbf{x}$ is defined and belongs to $C = \text{Col}(A)$.

First, note that $A^T A \in \mathbb{R}^{m \times m}$, and that by Corollary 6.6.2(a), we have that $\text{rank}(A^T A) = \text{rank}(A) = m$. So, by the Invertible Matrix Theorem, $A^T A$ is invertible, and we see that $(A^T A)^{-1}$ is defined and belongs to $\mathbb{R}^{m \times m}$. Since $A \in \mathbb{R}^{n \times m}$, $(A^T A)^{-1} \in \mathbb{R}^{m \times m}$, and $A^T \in \mathbb{R}^{m \times n}$, we see that $A(A^T A)^{-1} A^T \in \mathbb{R}^{n \times n}$; since $\mathbf{x} \in \mathbb{R}^n$, we see that $A(A^T A)^{-1} A^T \mathbf{x}$ is defined and belongs to $\mathbb{R}^n$. Meanwhile, $(A^T A)^{-1} A^T \mathbf{x}$ is a vector in $\mathbb{R}^m$, and so (next slide):

Theorem 6.6.3

Let $A \in \mathbb{R}^{n\times m}$ be a matrix of rank m (i.e. A is a matrix of full column rank). Then the matrix $A(A^TA)^{-1}A^T$ is the standard matrix of orthogonal projection onto $\mathrm{Col}(A)$, that is, for all $\mathbf{x} \in \mathbb{R}^n$, the orthogonal projection of $\mathbf{x}$ onto $C := \mathrm{Col}(A)$ is given by

$$\mathbf{x}_C = A(A^TA)^{-1}A^T\mathbf{x}.$$

Proof (continued).

$$A(A^TA)^{-1}A^T\mathbf{x} = \underbrace{A}_{\in\mathbb{R}^{n\times m}}\Big(\underbrace{(A^TA)^{-1}A^T\mathbf{x}}_{\in\mathbb{R}^m}\Big)$$

is a linear combination of the columns of A. By definition, this means that $A(A^TA)^{-1}A^T\mathbf{x} \in \mathrm{Col}(A) = C$.

Theorem 6.6.3

Let $A \in \mathbb{R}^{n \times m}$ be a matrix of rank m (i.e. A is a matrix of full column rank). Then the matrix $A(A^T A)^{-1}A^T$ is the standard matrix of orthogonal projection onto $\text{Col}(A)$, that is, for all $\mathbf{x} \in \mathbb{R}^n$, the orthogonal projection of $\mathbf{x}$ onto $C := \text{Col}(A)$ is given by

$$\mathbf{x}_C = A(A^T A)^{-1}A^T\mathbf{x}.$$

Proof (continued). In view of Corollary 6.5.3, it is now enough to prove that $(\mathbf{x} - A(A^T A)^{-1}A^T\mathbf{x}) \in C^\perp$, for it will then follow that $\mathbf{x}_C = A(A^T A)^{-1}A^T\mathbf{x}$, which is what we need to show.

Theorem 6.6.3

Let $A \in \mathbb{R}^{n \times m}$ be a matrix of rank m (i.e. A is a matrix of full column rank). Then the matrix $A(A^T A)^{-1} A^T$ is the standard matrix of orthogonal projection onto $\text{Col}(A)$, that is, for all $\mathbf{x} \in \mathbb{R}^n$, the orthogonal projection of $\mathbf{x}$ onto $C := \text{Col}(A)$ is given by

$$\mathbf{x}_C \;=\; A(A^T A)^{-1} A^T \mathbf{x}.$$

Proof (continued). In view of Corollary 6.5.3, it is now enough to prove that $(\mathbf{x} - A(A^T A)^{-1} A^T \mathbf{x}) \in C^\perp$, for it will then follow that $\mathbf{x}_C = A(A^T A)^{-1} A^T \mathbf{x}$, which is what we need to show.

- Indeed, if we can show that $(\mathbf{x} - A(A^T A)^{-1} A^T \mathbf{x}) \in C^\perp$, then we get that

  $$\mathbf{x} \;=\; \underbrace{A(A^T A)^{-1} A^T \mathbf{x}}_{\in C} + \big(\underbrace{\mathbf{x} - A(A^T A)^{-1} A^T \mathbf{x}}_{\in C^\perp}\big),$$

  which (by Corollary 6.5.3) implies that $\mathbf{x}_C = A(A^T A)^{-1} A^T \mathbf{x}$ and $\mathbf{x}_{C^\perp} = \mathbf{x} - A(A^T A)^{-1} A^T \mathbf{x}$.

Theorem 6.6.3

Let $A \in \mathbb{R}^{n \times m}$ be a matrix of rank m (i.e. A is a matrix of full column rank). Then the matrix $A(A^T A)^{-1} A^T$ is the standard matrix of orthogonal projection onto $\text{Col}(A)$, that is, for all $\mathbf{x} \in \mathbb{R}^n$, the orthogonal projection of $\mathbf{x}$ onto $C := \text{Col}(A)$ is given by

$$\mathbf{x}_C \;=\; A(A^T A)^{-1} A^T \mathbf{x}.$$

Proof (continued). But note that

$$C^\perp \;=\; \text{Col}(A)^\perp \;=\; \text{Row}(A^T)^\perp \;\stackrel{(*)}{=}\; \text{Nul}(A^T),$$

where (*) follows from Theorem 6.6.1.

Theorem 6.6.3

Let $A \in \mathbb{R}^{n \times m}$ be a matrix of rank m (i.e. A is a matrix of full column rank). Then the matrix $A(A^TA)^{-1}A^T$ is the standard matrix of orthogonal projection onto $\text{Col}(A)$, that is, for all $\mathbf{x} \in \mathbb{R}^n$, the orthogonal projection of $\mathbf{x}$ onto $C := \text{Col}(A)$ is given by

$$\mathbf{x}_C = A(A^TA)^{-1}A^T\mathbf{x}.$$

Proof (continued). But note that

$$C^\perp = \text{Col}(A)^\perp = \text{Row}(A^T)^\perp \stackrel{(*)}{=} \text{Nul}(A^T),$$

where (*) follows from Theorem 6.6.1. So, it in fact suffices to show that the vector $\mathbf{x} - A(A^TA)^{-1}A^T\mathbf{x}$ belongs to $\text{Nul}(A^T)$.

Theorem 6.6.3

Let $A \in \mathbb{R}^{n \times m}$ be a matrix of rank m (i.e. A is a matrix of full column rank). Then the matrix $A(A^T A)^{-1} A^T$ is the standard matrix of orthogonal projection onto $\mathrm{Col}(A)$, that is, for all $\mathbf{x} \in \mathbb{R}^n$, the orthogonal projection of $\mathbf{x}$ onto $C := \mathrm{Col}(A)$ is given by

$$\mathbf{x}_C \;=\; A(A^T A)^{-1} A^T \mathbf{x}.$$

Proof (continued). But note that

$$C^\perp \;=\; \mathrm{Col}(A)^\perp \;=\; \mathrm{Row}(A^T)^\perp \;\overset{(*)}{=}\; \mathrm{Nul}(A^T),$$

where (*) follows from Theorem 6.6.1. So, it in fact suffices to show that the vector $\mathbf{x} - A(A^T A)^{-1} A^T \mathbf{x}$ belongs to $\mathrm{Nul}(A^T)$. For this, we compute:

$$A^T\Big(\mathbf{x} - A(A^T A)^{-1} A^T \mathbf{x}\Big) \;=\; A^T \mathbf{x} - \underbrace{A^T A(A^T A)^{-1}}_{=I_m} A^T \mathbf{x} \;=\; \mathbf{0}.$$

Theorem 6.6.3

Let $A \in \mathbb{R}^{n\times m}$ be a matrix of rank m (i.e. A is a matrix of full column rank). Then the matrix $A(A^TA)^{-1}A^T$ is the standard matrix of orthogonal projection onto $\text{Col}(A)$, that is, for all $\mathbf{x} \in \mathbb{R}^n$, the orthogonal projection of $\mathbf{x}$ onto $C := \text{Col}(A)$ is given by

$$\mathbf{x}_C \quad = \quad A(A^TA)^{-1}A^T\mathbf{x}.$$

Proof (continued). But note that

$$C^\perp \quad = \quad \text{Col}(A)^\perp \quad = \quad \text{Row}(A^T)^\perp \quad \overset{(*)}{=} \quad \text{Nul}(A^T),$$

where (*) follows from Theorem 6.6.1. So, it in fact suffices to show that the vector $\mathbf{x} - A(A^TA)^{-1}A^T\mathbf{x}$ belongs to $\text{Nul}(A^T)$. For this, we compute:

$$A^T\Big(\mathbf{x} - A(A^TA)^{-1}A^T\mathbf{x}\Big) \quad = \quad A^T\mathbf{x} - \underbrace{A^TA(A^TA)^{-1}}_{=I_m}A^T\mathbf{x} \quad = \quad \mathbf{0}.$$

This proves that $\mathbf{x} - A(A^TA)^{-1}A^T\mathbf{x} \in \text{Nul}(A^T)$, and we are done. $\square$

Theorem 6.6.3

Let $A \in \mathbb{R}^{n \times m}$ be a matrix of rank m (i.e. A is a matrix of full column rank). Then the matrix $A(A^T A)^{-1} A^T$ is the standard matrix of orthogonal projection onto $\mathrm{Col}(A)$, that is, for all $\mathbf{x} \in \mathbb{R}^n$, the orthogonal projection of $\mathbf{x}$ onto $C := \mathrm{Col}(A)$ is given by

$$\mathbf{x}_C = A(A^T A)^{-1} A^T \mathbf{x}.$$

Theorem 6.6.3

Let $A \in \mathbb{R}^{n \times m}$ be a matrix of rank m (i.e. A is a matrix of full column rank). Then the matrix $A(A^T A)^{-1} A^T$ is the standard matrix of orthogonal projection onto $\text{Col}(A)$, that is, for all $\mathbf{x} \in \mathbb{R}^n$, the orthogonal projection of $\mathbf{x}$ onto $C := \text{Col}(A)$ is given by

$$\mathbf{x}_C \quad = \quad A(A^T A)^{-1} A^T \mathbf{x}.$$

Corollary 6.6.4

Let $\mathbf{a}$ be a non-zero vector in $\mathbb{R}^n$. Then the standard matrix of projection onto the line $L := \text{Span}(\mathbf{a})$ is the matrix

$$\mathbf{a}(\mathbf{a}^T\mathbf{a})^{-1}\mathbf{a}^T \quad = \quad \mathbf{a}(\mathbf{a} \cdot \mathbf{a})^{-1}\mathbf{a}^T \quad = \quad \tfrac{1}{\mathbf{a} \cdot \mathbf{a}}\mathbf{a}\mathbf{a}^T.$$

Consequently, for every vector $\mathbf{x} \in \mathbb{R}^n$, we have that

$$\mathbf{x}_L \quad = \quad \text{proj}_L(\mathbf{x}) \quad = \quad \tfrac{1}{\mathbf{a} \cdot \mathbf{a}}\mathbf{a}\mathbf{a}^T\mathbf{x}.$$

Proof. This is a special case of Theorem 6.6.3 for $A = \left[\ \mathbf{a} \ \right]$. $\square$

Theorem 6.6.5

Let U be a subspace of $\mathbb{R}^n$, and let $P \in \mathbb{R}^{n \times n}$ be the standard matrix of proj_U. Then $I_n - P$ is the standard matrix of $\text{proj}_{U^\perp}$, that is, for all $\mathbf{x} \in \mathbb{R}^n$, the orthogonal projection of $\mathbf{x}$ onto $U^\perp$ is given by $\mathbf{x}_{U^\perp} = (I_n - P)\mathbf{x}$.

Proof.

Theorem 6.6.5

Let U be a subspace of $\mathbb{R}^n$, and let $P \in \mathbb{R}^{n \times n}$ be the standard matrix of proj_U. Then $I_n - P$ is the standard matrix of $\text{proj}_{U^\perp}$, that is, for all $\mathbf{x} \in \mathbb{R}^n$, the orthogonal projection of $\mathbf{x}$ onto $U^\perp$ is given by $\mathbf{x}_{U^\perp} = (I_n - P)\mathbf{x}$.

Proof. We observe that for all $\mathbf{x} \in \mathbb{R}^n$, we have that

$$(I_n - P)\mathbf{x} \ = \ I_n\mathbf{x} - P\mathbf{x} \ \overset{(*)}{=} \ \mathbf{x} - \mathbf{x}_U \ \overset{(**)}{=} \ \mathbf{x}_{U^\perp},$$

where (*) follows from the fact that P is the standard matrix of proj_U, and (**) follows from Corollary 6.5.3. So, $I_n - P$ is indeed the standard matrix of $\text{proj}_{U^\perp}$. □

- Theorem 6.6.5: if $P \in \mathbb{R}^{n \times n}$ is the standard matrix of proj_U, then $I_n - P$ is the standard matrix of $\text{proj}_{U^\perp}$.

- Theorem 6.6.5: if $P \in \mathbb{R}^{n\times n}$ is the standard matrix of proj_U, then $I_n - P$ is the standard matrix of $\mathrm{proj}_{U^\perp}$.

Corollary 6.6.6

Let $A \in \mathbb{R}^{n\times m}$ be a matrix of rank n (i.e. A is a matrix of full row rank). Then the matrix $I_m - A^T(AA^T)^{-1}A$ is the standard matrix of orthogonal projection onto $N := \mathrm{Nul}(A)$, that is, for all $\mathbf{x} \in \mathbb{R}^m$, the orthogonal projection of $\mathbf{x}$ onto N is given by
$\mathbf{x}_N = (I_m - A^T(AA^T)^{-1}A)\mathbf{x}$.

Proof.

- Theorem 6.6.5: if $P \in \mathbb{R}^{n\times n}$ is the standard matrix of proj_U, then $I_n - P$ is the standard matrix of $\text{proj}_{U^\perp}$.

Corollary 6.6.6

Let $A \in \mathbb{R}^{n\times m}$ be a matrix of rank n (i.e. A is a matrix of full row rank). Then the matrix $I_m - A^T(AA^T)^{-1}A$ is the standard matrix of orthogonal projection onto $N := \text{Nul}(A)$, that is, for all $\mathbf{x} \in \mathbb{R}^m$, the orthogonal projection of $\mathbf{x}$ onto N is given by $\mathbf{x}_N = (I_m - A^T(AA^T)^{-1}A)\mathbf{x}$.

Proof. First, note that

$$\text{Nul}(A) \overset{(*)}{=} \text{Row}(A)^\perp = \text{Col}(A^T)^\perp.$$

where (*) follows from Theorem 6.6.1.

- Theorem 6.6.5: if $P \in \mathbb{R}^{n\times n}$ is the standard matrix of proj_U, then $I_n - P$ is the standard matrix of $\text{proj}_{U^\perp}$.

Corollary 6.6.6

Let $A \in \mathbb{R}^{n\times m}$ be a matrix of rank n (i.e. A is a matrix of full row rank). Then the matrix $I_m - A^T(AA^T)^{-1}A$ is the standard matrix of orthogonal projection onto $N := \text{Nul}(A)$, that is, for all $\mathbf{x} \in \mathbb{R}^m$, the orthogonal projection of $\mathbf{x}$ onto N is given by
$\mathbf{x}_N = (I_m - A^T(AA^T)^{-1}A)\mathbf{x}$.

Proof. First, note that

$$\text{Nul}(A) \overset{(*)}{=} \text{Row}(A)^\perp = \text{Col}(A^T)^\perp.$$

where (*) follows from Theorem 6.6.1. Note further that $A^T \in \mathbb{R}^{m\times n}$ and that $\text{rank}(A^T) = \text{rank}(A) = n$, i.e. A^T has full column rank.

- Theorem 6.6.5: if $P \in \mathbb{R}^{n \times n}$ is the standard matrix of proj_U, then $I_n - P$ is the standard matrix of $\text{proj}_{U^\perp}$.

Corollary 6.6.6

Let $A \in \mathbb{R}^{n \times m}$ be a matrix of rank n (i.e. A is a matrix of full row rank). Then the matrix $I_m - A^T(AA^T)^{-1}A$ is the standard matrix of orthogonal projection onto $N := \text{Nul}(A)$, that is, for all $\mathbf{x} \in \mathbb{R}^m$, the orthogonal projection of $\mathbf{x}$ onto N is given by
$\mathbf{x}_N = (I_m - A^T(AA^T)^{-1}A)\mathbf{x}$.

Proof. First, note that

$$\text{Nul}(A) \stackrel{(*)}{=} \text{Row}(A)^\perp = \text{Col}(A^T)^\perp.$$

where (*) follows from Theorem 6.6.1. Note further that $A^T \in \mathbb{R}^{m \times n}$ and that $\text{rank}(A^T) = \text{rank}(A) = n$, i.e. A^T has full column rank. So, by Theorem 6.6.3, the standard matrix of orthogonal projection onto $\text{Col}(A^T)$ is $A^T(AA^T)^{-1}A$.

- Theorem 6.6.5: if $P \in \mathbb{R}^{n\times n}$ is the standard matrix of proj_U, then $I_n - P$ is the standard matrix of $\text{proj}_{U^\perp}$.

Corollary 6.6.6

Let $A \in \mathbb{R}^{n\times m}$ be a matrix of rank n (i.e. A is a matrix of full row rank). Then the matrix $I_m - A^T(AA^T)^{-1}A$ is the standard matrix of orthogonal projection onto $N := \text{Nul}(A)$, that is, for all $\mathbf{x} \in \mathbb{R}^m$, the orthogonal projection of $\mathbf{x}$ onto N is given by
$\mathbf{x}_N = (I_m - A^T(AA^T)^{-1}A)\mathbf{x}$.

Proof. First, note that

$$\text{Nul}(A) \overset{(*)}{=} \text{Row}(A)^\perp = \text{Col}(A^T)^\perp.$$

where (*) follows from Theorem 6.6.1. Note further that $A^T \in \mathbb{R}^{m\times n}$ and that $\text{rank}(A^T) = \text{rank}(A) = n$, i.e. A^T has full column rank. So, by Theorem 6.6.3, the standard matrix of orthogonal projection onto $\text{Col}(A^T)$ is $A^T(AA^T)^{-1}A$. Finally, by Theorem 6.6.5, the standard matrix of orthogonal projection onto $\text{Col}(A^T)^\perp = \text{Nul}(A)$ is $I_m - A^T(AA^T)^{-1}A$. □