Linear Algebra 2

Lecture #15

Orthogonal and orthonormal bases. Gram-Schmidt orthogonalization

Irena Penev

March 6, 2024

- This lecture has four parts:

- This lecture has four parts:
    1. Vector projection

- This lecture has four parts:
  1. Vector projection
  2. Orthogonal and orthonormal sets. Orthogonal and orthonormal bases

- This lecture has four parts:
  1. Vector projection
  2. Orthogonal and orthonormal sets. Orthogonal and orthonormal bases
  3. Coordinate vectors w.r.t. orthogonal and orthonormal bases. Fourier coefficients

- This lecture has four parts:
  1. Vector projection
  2. Orthogonal and orthonormal sets. Orthogonal and orthonormal bases
  3. Coordinate vectors w.r.t. orthogonal and orthonormal bases. Fourier coefficients
  4. Gram-Schmidt orthogonalization

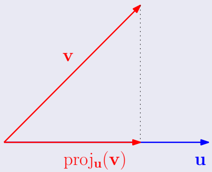
1. Vector projection

1. Vector projection

Definition

Suppose we are given a real or complex vector space V, equipped with a scalar product $\langle \cdot, \cdot \rangle$. For a **non-zero** vector $\mathbf{u} \in V$ and any vector $\mathbf{v} \in V$, the *orthogonal projection* of $\mathbf{v}$ onto $\mathbf{u}$ is the vector

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



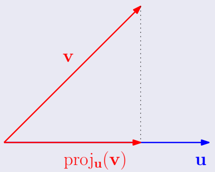
- **Remarks:**

1. Vector projection

Definition

Suppose we are given a real or complex vector space V, equipped with a scalar product $\langle\cdot,\cdot\rangle$. For a **non-zero** vector $\mathbf{u} \in V$ and any vector $\mathbf{v} \in V$, the *orthogonal projection* of $\mathbf{v}$ onto $\mathbf{u}$ is the vector

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



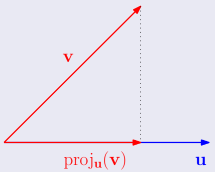
- **Remarks:**
  - Since $\mathbf{u} \neq \mathbf{0}$, r.1 or c.1 guarantees that $\langle \mathbf{u}, \mathbf{u}\rangle > 0$, and so the expression above is well-defined (that is, we are not dividing by zero).

1. Vector projection

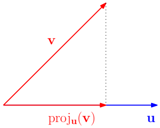
Definition

Suppose we are given a real or complex vector space V, equipped with a scalar product $\langle \cdot, \cdot \rangle$. For a **non-zero** vector $\mathbf{u} \in V$ and any vector $\mathbf{v} \in V$, the *orthogonal projection* of $\mathbf{v}$ onto $\mathbf{u}$ is the vector

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

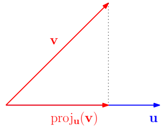


- **Remarks:**
  - Since $\mathbf{u} \neq \mathbf{0}$, r.1 or c.1 guarantees that $\langle \mathbf{u}, \mathbf{u} \rangle > 0$, and so the expression above is well-defined (that is, we are not dividing by zero).
  - $\text{proj}_{\mathbf{u}}(\mathbf{v})$ is a scalar multiple of $\mathbf{u}$.



$$\mathrm{proj}_{\mathbf{u}}(\mathbf{v}) = \frac{\langle \mathbf{v}, \mathbf{u} \rangle}{\langle \mathbf{u}, \mathbf{u} \rangle}\mathbf{u}$$

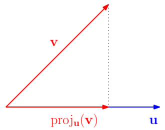
- As the picture suggests, for any scalar $\alpha \neq 0$, the projection of **v** onto α**u** is the same as the projection of **v** onto **u**.



$$\text{proj}_{\mathbf{u}}(\mathbf{v}) = \frac{\langle \mathbf{v}, \mathbf{u} \rangle}{\langle \mathbf{u}, \mathbf{u} \rangle}\mathbf{u}$$

- As the picture suggests, for any scalar $\alpha \neq 0$, the projection of $\mathbf{v}$ onto $\alpha\mathbf{u}$ is the same as the projection of $\mathbf{v}$ onto $\mathbf{u}$.
- Indeed, if V is a complex vector space, then we have that

$$\begin{aligned}
\text{proj}_{\alpha\mathbf{u}}(\mathbf{v}) &= \frac{\langle \mathbf{v}, \alpha\mathbf{u} \rangle}{\langle \alpha\mathbf{u}, \alpha\mathbf{u} \rangle}(\alpha\mathbf{u}) && \text{by definition} \\
&= \frac{\langle \mathbf{v}, \alpha\mathbf{u} \rangle}{\alpha\langle \mathbf{u}, \alpha\mathbf{u} \rangle}(\alpha\mathbf{u}) && \text{by c.3} \\
&= \frac{\overline{\alpha}\langle \mathbf{v}, \mathbf{u} \rangle}{\alpha\overline{\alpha}\langle \mathbf{u}, \mathbf{u} \rangle}(\alpha\mathbf{u}) && \text{by c.3'} \\
&= \frac{\langle \mathbf{v}, \mathbf{u} \rangle}{\langle \mathbf{u}, \mathbf{u} \rangle}\mathbf{u} \\
&= \text{proj}_{\mathbf{u}}(\mathbf{v}) && \text{by definition.}
\end{aligned}$$



$$\mathrm{proj}_{\mathbf{u}}(\mathbf{v}) = \frac{\langle \mathbf{v}, \mathbf{u} \rangle}{\langle \mathbf{u}, \mathbf{u} \rangle}\mathbf{u}$$

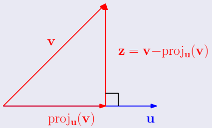
- As the picture suggests, for any scalar $\alpha \neq 0$, the projection of $\mathbf{v}$ onto $\alpha\mathbf{u}$ is the same as the projection of $\mathbf{v}$ onto $\mathbf{u}$.
- Indeed, if V is a complex vector space, then we have that

$$\begin{aligned}
\mathrm{proj}_{\alpha\mathbf{u}}(\mathbf{v}) &= \frac{\langle \mathbf{v}, \alpha\mathbf{u} \rangle}{\langle \alpha\mathbf{u}, \alpha\mathbf{u} \rangle}(\alpha\mathbf{u}) && \text{by definition} \\
&= \frac{\langle \mathbf{v}, \alpha\mathbf{u} \rangle}{\alpha\langle \mathbf{u}, \alpha\mathbf{u} \rangle}(\alpha\mathbf{u}) && \text{by c.3} \\
&= \frac{\overline{\alpha}\langle \mathbf{v}, \mathbf{u} \rangle}{\alpha\overline{\alpha}\langle \mathbf{u}, \mathbf{u} \rangle}(\alpha\mathbf{u}) && \text{by c.3'} \\
&= \frac{\langle \mathbf{v}, \mathbf{u} \rangle}{\langle \mathbf{u}, \mathbf{u} \rangle}\mathbf{u} \\
&= \mathrm{proj}_{\mathbf{u}}(\mathbf{v}) && \text{by definition.}
\end{aligned}$$

- The real case is similar, only without complex conjugates.

Proposition 6.3.1

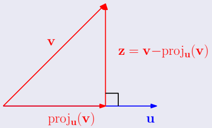
Let V be a real or complex vector space, equipped with a scalar product $\langle \cdot, \cdot \rangle$. Let $\mathbf{u}$ be a non-zero vector in V, let $\mathbf{v}$ be any vector in V, and set $\mathbf{z} := \mathbf{v} - \text{proj}_{\mathbf{u}}(\mathbf{v})$. Then $\mathbf{z} \perp \mathbf{u}$.



Proof.

Proposition 6.3.1

Let V be a real or complex vector space, equipped with a scalar product $\langle \cdot, \cdot \rangle$. Let $\mathbf{u}$ be a non-zero vector in V, let $\mathbf{v}$ be any vector in V, and set $\mathbf{z} := \mathbf{v} - \text{proj}_{\mathbf{u}}(\mathbf{v})$. Then $\mathbf{z} \perp \mathbf{u}$.



Proof. We compute

$$\begin{aligned}
\langle \mathbf{z}, \mathbf{u} \rangle &= \langle \mathbf{v} - \text{proj}_{\mathbf{u}}(\mathbf{v}), \mathbf{u} \rangle = \left\langle \mathbf{v} - \tfrac{\langle \mathbf{v}, \mathbf{u} \rangle}{\langle \mathbf{u}, \mathbf{u} \rangle} \mathbf{u}, \mathbf{u} \right\rangle \\
&\overset{(*)}{=} \langle \mathbf{v}, \mathbf{u} \rangle - \tfrac{\langle \mathbf{v}, \mathbf{u} \rangle}{\langle \mathbf{u}, \mathbf{u} \rangle} \langle \mathbf{u}, \mathbf{u} \rangle \\
&= \langle \mathbf{v}, \mathbf{u} \rangle - \langle \mathbf{v}, \mathbf{u} \rangle = 0,
\end{aligned}$$

where (*) follows from r.2 and r.3 (in the real case) or from c.2 and c.3 (in the complex case). This proves that $\mathbf{z} \perp \mathbf{u}$. □

2. Orthogonal and orthonormal sets. Orthogonal and orthonormal bases

2. Orthogonal and orthonormal sets. Orthogonal and orthonormal bases

Definition

Suppose we are given a real or complex vector space V, equipped with a scalar product $\langle \cdot, \cdot \rangle$ and the induced norm $||\cdot||$.

- An *orthogonal set of vectors* in V is a set of pairwise orthogonal vectors in V.
- An *orthonormal set of vectors* is an orthogonal set of unit vectors (i.e. vectors of length 1).
- An *orthogonal basis* (resp. *orthonormal basis*) of V is an orthogonal (resp. orthonormal) set in V that is also a basis of V.

Proposition 6.3.2

Let V be a real or complex vector space, equipped with a scalar product $\langle \cdot, \cdot \rangle$ and the induced norm $|| \cdot ||$. Then both the following hold:

(a) any **orthogonal** set of **non-zero** vectors in V is linearly independent;

(b) any **orthonormal** set of vectors in V is linearly independent.

Proof.

Proposition 6.3.2

Let V be a real or complex vector space, equipped with a scalar product $\langle \cdot, \cdot \rangle$ and the induced norm $|| \cdot ||$. Then both the following hold:

(a) any **orthogonal** set of **non-zero** vectors in V is linearly independent;

(b) any **orthonormal** set of vectors in V is linearly independent.

Proof. Any orthonormal set of vectors is an orthogonal set of non-zero vectors (because $\mathbf{0}$ is not a unit vector). So, (a) immediately implies (b).

Proposition 6.3.2

Let V be a real or complex vector space, equipped with a scalar product $\langle \cdot, \cdot \rangle$ and the induced norm $|| \cdot ||$. Then both the following hold:

(a) any **orthogonal** set of **non-zero** vectors in V is linearly independent;

(b) any **orthonormal** set of vectors in V is linearly independent.

Proof. Any orthonormal set of vectors is an orthogonal set of non-zero vectors (because $\mathbf{0}$ is not a unit vector). So, (a) immediately implies (b).

It remains to prove (a).

Proposition 6.3.2

(a) any **orthogonal** set of **non-zero** vectors in V is linearly independent;

Proof (continued).

Proposition 6.3.2

(a) any **orthogonal** set of **non-zero** vectors in V is linearly independent;

Proof (continued). Fix an orthogonal set $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ of non-zero vectors in V. WTS this set is linearly independent.

Proposition 6.3.2

(a) any **orthogonal** set of **non-zero** vectors in V is linearly independent;

Proof (continued). Fix an orthogonal set $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ of non-zero vectors in V. WTS this set is linearly independent. Fix scalars $\alpha_1, \dots, \alpha_k$ s.t.

$$\alpha_1 \mathbf{u}_1 + \dots + \alpha_k \mathbf{u}_k = \mathbf{0}.$$

WTS $\alpha_1 = \dots = \alpha_k = 0$.

Proposition 6.3.2

(a) any **orthogonal** set of **non-zero** vectors in V is linearly independent;

Proof (continued). Fix an orthogonal set $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ of non-zero vectors in V. WTS this set is linearly independent. Fix scalars $\alpha_1, \dots, \alpha_k$ s.t.

$$\alpha_1 \mathbf{u}_1 + \dots + \alpha_k \mathbf{u}_k \quad = \quad \mathbf{0}.$$

WTS $\alpha_1 = \dots = \alpha_k = 0$. Fix any $i \in \{1, \dots, k\}$. Then

$$\langle \underbrace{\alpha_1 \mathbf{u}_1 + \dots + \alpha_k \mathbf{u}_k}_{=\mathbf{0}}, \mathbf{u}_i \rangle \quad = \quad \langle \mathbf{0}, \mathbf{u}_i \rangle \quad \overset{(*)}{=} \quad 0,$$

where (*) follows from Proposition 6.1.4(c).

Proposition 6.3.2

(a) any **orthogonal** set of **non-zero** vectors in V is linearly independent;

Proof (continued). Reminder: $\langle \alpha_1 \mathbf{u}_1 + \cdots + \alpha_k \mathbf{u}_k, \mathbf{u}_i \rangle = 0$.

Proposition 6.3.2

(a) any **orthogonal** set of **non-zero** vectors in V is linearly independent;

Proof (continued). Reminder: $\langle \alpha_1 \mathbf{u}_1 + \cdots + \alpha_k \mathbf{u}_k, \mathbf{u}_i \rangle = 0$.

On the other hand, note that

$$\langle \alpha_1 \mathbf{u}_1 + \cdots + \alpha_k \mathbf{u}_k, \mathbf{u}_i \rangle \overset{(*)}{=} \alpha_1 \langle \mathbf{u}_1, \mathbf{u}_i \rangle + \cdots + \alpha_k \langle \mathbf{u}_k, \mathbf{u}_i \rangle \overset{(**)}{=} \alpha_i \langle \mathbf{u}_i, \mathbf{u}_i \rangle,$$

where (*) follows from r.2 and r.3 (in the real case) or from c.2 and c.3 (in the complex case), and (**) follows from the fact that $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ is an orthogonal set.

Proposition 6.3.2

(a) any **orthogonal** set of **non-zero** vectors in V is linearly independent;

Proof (continued). Reminder: $\langle \alpha_1 \mathbf{u}_1 + \cdots + \alpha_k \mathbf{u}_k, \mathbf{u}_i \rangle = 0$.

On the other hand, note that

$$\langle \alpha_1 \mathbf{u}_1 + \cdots + \alpha_k \mathbf{u}_k, \mathbf{u}_i \rangle \overset{(*)}{=} \alpha_1 \langle \mathbf{u}_1, \mathbf{u}_i \rangle + \cdots + \alpha_k \langle \mathbf{u}_k, \mathbf{u}_i \rangle \overset{(**)}{=} \alpha_i \langle \mathbf{u}_i, \mathbf{u}_i \rangle,$$

where (*) follows from r.2 and r.3 (in the real case) or from c.2 and c.3 (in the complex case), and (**) follows from the fact that $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ is an orthogonal set.

So,

$$\alpha_i \langle \mathbf{u}_i, \mathbf{u}_i \rangle = 0.$$

Proposition 6.3.2

(a) any **orthogonal** set of **non-zero** vectors in V is linearly independent;

Proof (continued). Reminder: $\langle \alpha_1\mathbf{u}_1 + \cdots + \alpha_k\mathbf{u}_k, \mathbf{u}_i\rangle = 0$.

On the other hand, note that

$$\langle \alpha_1\mathbf{u}_1 + \cdots + \alpha_k\mathbf{u}_k, \mathbf{u}_i\rangle \overset{(*)}{=} \alpha_1\langle \mathbf{u}_1, \mathbf{u}_i\rangle + \cdots + \alpha_k\langle \mathbf{u}_k, \mathbf{u}_i\rangle \overset{(**)}{=} \alpha_i\langle \mathbf{u}_i, \mathbf{u}_i\rangle,$$

where (*) follows from r.2 and r.3 (in the real case) or from c.2 and c.3 (in the complex case), and (**) follows from the fact that $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ is an orthogonal set.

So,

$$\alpha_i\langle \mathbf{u}_i, \mathbf{u}_i\rangle = 0.$$

Since $\mathbf{u}_i \neq \mathbf{0}$, r.1 or c.1 guarantees that $\langle \mathbf{u}_i, \mathbf{u}_i\rangle \neq 0$; consequently, $\alpha_i = 0$.

Proposition 6.3.2

(a) any **orthogonal** set of **non-zero** vectors in V is linearly independent;

Proof (continued). Reminder: $\langle \alpha_1 \mathbf{u}_1 + \cdots + \alpha_k \mathbf{u}_k, \mathbf{u}_i \rangle = 0$.

On the other hand, note that

$$\langle \alpha_1 \mathbf{u}_1 + \cdots + \alpha_k \mathbf{u}_k, \mathbf{u}_i \rangle \overset{(*)}{=} \alpha_1 \langle \mathbf{u}_1, \mathbf{u}_i \rangle + \cdots + \alpha_k \langle \mathbf{u}_k, \mathbf{u}_i \rangle \overset{(**)}{=} \alpha_i \langle \mathbf{u}_i, \mathbf{u}_i \rangle,$$

where (*) follows from r.2 and r.3 (in the real case) or from c.2 and c.3 (in the complex case), and (**) follows from the fact that $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ is an orthogonal set.

So,

$$\alpha_i \langle \mathbf{u}_i, \mathbf{u}_i \rangle = 0.$$

Since $\mathbf{u}_i \neq \mathbf{0}$, r.1 or c.1 guarantees that $\langle \mathbf{u}_i, \mathbf{u}_i \rangle \neq 0$; consequently, $\alpha_i = 0$. Since $i \in \{1, \dots, k\}$ was chosen arbitrarily, it follows that $\alpha_1 = \cdots = \alpha_k = 0$, and we are done. □

Proposition 6.3.2

Let V be a real or complex vector space, equipped with a scalar product $\langle \cdot, \cdot \rangle$ and the induced norm $|| \cdot ||$. Then both the following hold:

a) any **orthogonal** set of **non-zero** vectors in V is linearly independent;

b) any **orthonormal** set of vectors in V is linearly independent.

Proposition 6.3.3

Let V be a real or complex vector space, equipped with a scalar product $\langle \cdot, \cdot \rangle$ and the induced norm $|| \cdot ||$. Let $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ be an **orthogonal** set of vectors in V. Then all the following hold:

a. for all scalars $\alpha_1, \dots, \alpha_k$, we have that $\{\alpha_1 \mathbf{u}_1, \dots, \alpha_k \mathbf{u}_k\}$ is an **orthogonal** set of vectors;
b. if vectors $\mathbf{u}_1, \dots, \mathbf{u}_k$ are all non-zero, then $\left\{ \frac{\mathbf{u}_1}{||\mathbf{u}_1||}, \dots, \frac{\mathbf{u}_k}{||\mathbf{u}_k||} \right\}$ is an **orthonormal** set of vectors, and consequently, an orthonormal basis of $\text{Span}(\mathbf{u}_1, \dots, \mathbf{u}_k)$;
c. if $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ is an orthogonal basis of V, then $\left\{ \frac{\mathbf{u}_1}{||\mathbf{u}_1||}, \dots, \frac{\mathbf{u}_k}{||\mathbf{u}_k||} \right\}$ is an orthonormal basis of V.

- Proof: Lecture Notes.

Proposition 6.3.4

Let V be a finite-dimensional real or complex vector space, equipped with a scalar product $\langle \cdot, \cdot \rangle$ and the induced norm $|| \cdot ||$. Set $n := \dim(V)$. Then both the following hold:

(a) any orthogonal set of n non-zero vectors in V is an orthogonal basis of V;

(b) any orthonormal set of n vectors in V is an orthonormal basis of V.

Proof.

Proposition 6.3.4

Let V be a finite-dimensional real or complex vector space, equipped with a scalar product $\langle \cdot, \cdot \rangle$ and the induced norm $|| \cdot ||$. Set $n := \dim(V)$. Then both the following hold:

(a) any orthogonal set of n non-zero vectors in V is an orthogonal basis of V;

(b) any orthonormal set of n vectors in V is an orthonormal basis of V.

Proof. By Proposition 6.3.2, any orthogonal set of non-zero vectors is linearly independent, and by Corollary 3.2.20(a), any linearly independent set of size n in an n-dimensional vector space is a basis of that vector space. This proves (a).

Proposition 6.3.4

Let V be a finite-dimensional real or complex vector space, equipped with a scalar product $\langle \cdot, \cdot \rangle$ and the induced norm $|| \cdot ||$. Set $n := \dim(V)$. Then both the following hold:

(a) any orthogonal set of n non-zero vectors in V is an orthogonal basis of V;

(b) any orthonormal set of n vectors in V is an orthonormal basis of V.

Proof. By Proposition 6.3.2, any orthogonal set of non-zero vectors is linearly independent, and by Corollary 3.2.20(a), any linearly independent set of size n in an n-dimensional vector space is a basis of that vector space. This proves (a).

Part (b) follows from (a), since any orthonormal set of vectors is, in particular, an orthogonal set of non-zero vectors (because $\mathbf{0}$ is not a unit vector). □

3. Coordinate vectors w.r.t. orthogonal and orthonormal bases. Fourier coefficients

3. Coordinate vectors w.r.t. orthogonal and orthonormal bases. Fourier coefficients

- If we have an orthogonal basis of a real or complex vector space (equipped with a scalar product and the norm induced by it), then every vector in that vector space can be expressed as a linear combination of those basis vectors in a particularly nice way, that is, we have a convenient formula for the coefficients in front of the basis vectors (see Theorem 6.3.5, next slide).

3. Coordinate vectors w.r.t. orthogonal and orthonormal bases. Fourier coefficients

- If we have an orthogonal basis of a real or complex vector space (equipped with a scalar product and the norm induced by it), then every vector in that vector space can be expressed as a linear combination of those basis vectors in a particularly nice way, that is, we have a convenient formula for the coefficients in front of the basis vectors (see Theorem 6.3.5, next slide).
- If our basis is orthonormal, then we get an even nicer formula for the coefficients (see Corollary 6.3.6, next slide).
    - The formula from Corollary 6.3.6 follows immediately from the one for Theorem 6.3.5.
    - The coefficients from Corollary 6.3.6 are called the "Fourier coefficients."

Theorem 6.3.5

Let V be a real or complex vector space, equipped with a scalar product $\langle \cdot, \cdot \rangle$. Let $\mathcal{B} = \{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ be an **orthogonal** basis of V. Then for all $\mathbf{v} \in V$, we have that

$$\mathbf{v} = \sum_{i=1}^{n} \text{proj}_{\mathbf{u}_i}(\mathbf{v}) = \sum_{i=1}^{n} \frac{\langle \mathbf{v}, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \mathbf{u}_i,$$

and consequently, $\left[\ \mathbf{v}\ \right]_{\mathcal{B}} = \left[\ \frac{\langle \mathbf{v}, \mathbf{u}_1 \rangle}{\langle \mathbf{u}_1, \mathbf{u}_1 \rangle} \quad \dots \quad \frac{\langle \mathbf{v}, \mathbf{u}_n \rangle}{\langle \mathbf{u}_n, \mathbf{u}_n \rangle}\ \right]^T$.

Theorem 6.3.5

Let V be a real or complex vector space, equipped with a scalar product $\langle \cdot, \cdot \rangle$. Let $\mathcal{B} = \{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ be an **orthogonal** basis of V. Then for all $\mathbf{v} \in V$, we have that

$$\mathbf{v} \;=\; \sum_{i=1}^{n} \text{proj}_{\mathbf{u}_i}(\mathbf{v}) \;=\; \sum_{i=1}^{n} \frac{\langle \mathbf{v}, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \mathbf{u}_i,$$

and consequently, $\left[\ \mathbf{v}\ \right]_{\mathcal{B}} = \left[\ \frac{\langle \mathbf{v}, \mathbf{u}_1 \rangle}{\langle \mathbf{u}_1, \mathbf{u}_1 \rangle} \quad \dots \quad \frac{\langle \mathbf{v}, \mathbf{u}_n \rangle}{\langle \mathbf{u}_n, \mathbf{u}_n \rangle}\ \right]^T.$

Corollary 6.3.6

Let V be a real or complex vector space, equipped with a scalar product $\langle \cdot, \cdot \rangle$ and the induced norm $|| \cdot ||$. Let $\mathcal{B} = \{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ be an **orthonormal** basis of V. Then for all $\mathbf{v} \in V$, we have that

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

and consequently, $\left[\ \mathbf{v}\ \right]_{\mathcal{B}} = \left[\ \langle \mathbf{v}, \mathbf{u}_1 \rangle \quad \dots \quad \langle \mathbf{v}, \mathbf{u}_n \rangle\ \right]^T.$

Theorem 6.3.5

Let V be a real or complex vector space, equipped with a scalar product $\langle \cdot, \cdot \rangle$. Let $\mathcal{B} = \{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ be an **orthogonal** basis of V. Then for all $\mathbf{v} \in V$, we have that

$$\mathbf{v} \;=\; \sum_{i=1}^{n} \text{proj}_{\mathbf{u}_i}(\mathbf{v}) \;=\; \sum_{i=1}^{n} \frac{\langle \mathbf{v}, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \mathbf{u}_i,$$

and consequently, $\left[\ \mathbf{v}\ \right]_{\mathcal{B}} = \left[\ \frac{\langle \mathbf{v}, \mathbf{u}_1 \rangle}{\langle \mathbf{u}_1, \mathbf{u}_1 \rangle} \quad \dots \quad \frac{\langle \mathbf{v}, \mathbf{u}_n \rangle}{\langle \mathbf{u}_n, \mathbf{u}_n \rangle}\ \right]^T$.

Proof.

Theorem 6.3.5

Let V be a real or complex vector space, equipped with a scalar product $\langle\cdot,\cdot\rangle$. Let $\mathcal{B} = \{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ be an **orthogonal** basis of V. Then for all $\mathbf{v} \in V$, we have that

$$\mathbf{v} = \sum_{i=1}^{n} \mathrm{proj}_{\mathbf{u}_i}(\mathbf{v}) = \sum_{i=1}^{n} \frac{\langle \mathbf{v}, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \mathbf{u}_i,$$

and consequently, $\left[\ \mathbf{v}\ \right]_{\mathcal{B}} = \left[\ \frac{\langle \mathbf{v}, \mathbf{u}_1 \rangle}{\langle \mathbf{u}_1, \mathbf{u}_1 \rangle} \quad \dots \quad \frac{\langle \mathbf{v}, \mathbf{u}_n \rangle}{\langle \mathbf{u}_n, \mathbf{u}_n \rangle}\ \right]^T$.

Proof. The second statement follows from the first and from the definition of a coordinate vector.

Theorem 6.3.5

Let V be a real or complex vector space, equipped with a scalar product $\langle \cdot, \cdot \rangle$. Let $\mathcal{B} = \{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ be an **orthogonal** basis of V. Then for all $\mathbf{v} \in V$, we have that

$$\mathbf{v} \;=\; \sum_{i=1}^{n} \text{proj}_{\mathbf{u}_i}(\mathbf{v}) \;=\; \sum_{i=1}^{n} \frac{\langle \mathbf{v}, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \mathbf{u}_i,$$

and consequently, $\left[\ \mathbf{v}\ \right]_{\mathcal{B}} = \left[\ \frac{\langle \mathbf{v}, \mathbf{u}_1 \rangle}{\langle \mathbf{u}_1, \mathbf{u}_1 \rangle} \quad \dots \quad \frac{\langle \mathbf{v}, \mathbf{u}_n \rangle}{\langle \mathbf{u}_n, \mathbf{u}_n \rangle}\ \right]^T$.

Proof. The second statement follows from the first and from the definition of a coordinate vector. It remains to prove the first statement.

Theorem 6.3.5

Let V be a real or complex vector space, equipped with a scalar product $\langle \cdot, \cdot \rangle$. Let $\mathcal{B} = \{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ be an **orthogonal** basis of V. Then for all $\mathbf{v} \in V$, we have that

$$\mathbf{v} = \sum_{i=1}^{n} \text{proj}_{\mathbf{u}_i}(\mathbf{v}) = \sum_{i=1}^{n} \frac{\langle \mathbf{v}, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \mathbf{u}_i,$$

and consequently, $\begin{bmatrix} \mathbf{v} \end{bmatrix}_{\mathcal{B}} = \begin{bmatrix} \frac{\langle \mathbf{v}, \mathbf{u}_1 \rangle}{\langle \mathbf{u}_1, \mathbf{u}_1 \rangle} & \dots & \frac{\langle \mathbf{v}, \mathbf{u}_n \rangle}{\langle \mathbf{u}_n, \mathbf{u}_n \rangle} \end{bmatrix}^T$.

Proof. The second statement follows from the first and from the definition of a coordinate vector. It remains to prove the first statement.

Fix a vector $\mathbf{v} \in V$.

Theorem 6.3.5

Let V be a real or complex vector space, equipped with a scalar product $\langle \cdot, \cdot \rangle$. Let $\mathcal{B} = \{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ be an **orthogonal** basis of V. Then for all $\mathbf{v} \in V$, we have that

$$\mathbf{v} \;=\; \sum_{i=1}^{n} \text{proj}_{\mathbf{u}_i}(\mathbf{v}) \;=\; \sum_{i=1}^{n} \frac{\langle \mathbf{v}, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \mathbf{u}_i,$$

and consequently, $\left[\ \mathbf{v}\ \right]_{\mathcal{B}} = \left[\ \frac{\langle \mathbf{v}, \mathbf{u}_1 \rangle}{\langle \mathbf{u}_1, \mathbf{u}_1 \rangle} \quad \dots \quad \frac{\langle \mathbf{v}, \mathbf{u}_n \rangle}{\langle \mathbf{u}_n, \mathbf{u}_n \rangle}\ \right]^T$.

Proof. The second statement follows from the first and from the definition of a coordinate vector. It remains to prove the first statement.

Fix a vector $\mathbf{v} \in V$. By definition, for all $i \in \{1, \dots, n\}$, we have that $\text{proj}_{\mathbf{u}_i}(\mathbf{v}) = \frac{\langle \mathbf{v}, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \mathbf{u}_i$. So, it suffices to show that

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

Proof (continued). Reminder: WTS $\mathbf{v} = \sum_{i=1}^{n} \frac{\langle \mathbf{v}, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \mathbf{u}_i$.

Proof (continued). Reminder: WTS $\mathbf{v} = \sum_{i=1}^{n} \frac{\langle \mathbf{v}, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \mathbf{u}_i$.

Since $\mathbf{v} \in V$ and $\{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ is a basis of V, there exist scalars $\alpha_1, \dots, \alpha_n$ s.t.

$$\mathbf{v} \;=\; \sum_{i=1}^{n} \alpha_i \mathbf{u}_i.$$

Proof (continued). Reminder: WTS $\mathbf{v} = \sum_{i=1}^{n} \frac{\langle \mathbf{v}, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \mathbf{u}_i$.

Since $\mathbf{v} \in V$ and $\{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ is a basis of V, there exist scalars $\alpha_1, \dots, \alpha_n$ s.t.

$$\mathbf{v} \;=\; \sum_{i=1}^{n} \alpha_i \mathbf{u}_i.$$

Now, fix any index $j \in \{1, \dots, n\}$. We then have that

$$\langle \mathbf{v}, \mathbf{u}_j \rangle \;=\; \Big\langle \sum_{i=1}^{n} \alpha_i \mathbf{u}_i, \mathbf{u}_j \Big\rangle \;\overset{(*)}{=}\; \sum_{i=1}^{n} \alpha_i \langle \mathbf{u}_i, \mathbf{u}_j \rangle \;\overset{(**)}{=}\; \alpha_j \langle \mathbf{u}_j, \mathbf{u}_j \rangle,$$

where (*) follows from r.2 and r.3 (in the real case) or from c.2 and c.3 (in the complex case), and (**) follows from the fact that $\mathbf{u}_1, \dots, \mathbf{u}_n$ are pairwise orthogonal.

Proof (continued). Reminder: WTS $\mathbf{v} = \sum_{i=1}^{n} \frac{\langle \mathbf{v}, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \mathbf{u}_i$.

Since $\mathbf{v} \in V$ and $\{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ is a basis of V, there exist scalars $\alpha_1, \dots, \alpha_n$ s.t.

$$\mathbf{v} \;=\; \sum_{i=1}^{n} \alpha_i \mathbf{u}_i.$$

Now, fix any index $j \in \{1, \dots, n\}$. We then have that

$$\langle \mathbf{v}, \mathbf{u}_j \rangle \;=\; \Big\langle \sum_{i=1}^{n} \alpha_i \mathbf{u}_i, \mathbf{u}_j \Big\rangle \;\overset{(*)}{=}\; \sum_{i=1}^{n} \alpha_i \langle \mathbf{u}_i, \mathbf{u}_j \rangle \;\overset{(**)}{=}\; \alpha_j \langle \mathbf{u}_j, \mathbf{u}_j \rangle,$$

where (*) follows from r.2 and r.3 (in the real case) or from c.2 and c.3 (in the complex case), and (**) follows from the fact that $\mathbf{u}_1, \dots, \mathbf{u}_n$ are pairwise orthogonal. Since $\mathbf{u}_j \neq \mathbf{0}$ (because $\{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ is a basis of V), r.1 or c.1 guarantees that $\langle \mathbf{u}_j, \mathbf{u}_j \rangle \neq 0$, and we deduce that

$$\alpha_j \;=\; \frac{\langle \mathbf{v}, \mathbf{u}_j \rangle}{\langle \mathbf{u}_j, \mathbf{u}_j \rangle}.$$

Proof (continued). Reminder: WTS $\mathbf{v} = \sum_{i=1}^{n} \frac{\langle \mathbf{v}, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \mathbf{u}_i$.

Since $\mathbf{v} \in V$ and $\{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ is a basis of V, there exist scalars $\alpha_1, \dots, \alpha_n$ s.t.

$$\mathbf{v} \;=\; \sum_{i=1}^{n} \alpha_i \mathbf{u}_i.$$

Now, fix any index $j \in \{1, \dots, n\}$. We then have that

$$\langle \mathbf{v}, \mathbf{u}_j \rangle \;=\; \Big\langle \sum_{i=1}^{n} \alpha_i \mathbf{u}_i, \mathbf{u}_j \Big\rangle \;\stackrel{(*)}{=}\; \sum_{i=1}^{n} \alpha_i \langle \mathbf{u}_i, \mathbf{u}_j \rangle \;\stackrel{(**)}{=}\; \alpha_j \langle \mathbf{u}_j, \mathbf{u}_j \rangle,$$

where (*) follows from r.2 and r.3 (in the real case) or from c.2 and c.3 (in the complex case), and (**) follows from the fact that $\mathbf{u}_1, \dots, \mathbf{u}_n$ are pairwise orthogonal. Since $\mathbf{u}_j \neq \mathbf{0}$ (because $\{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ is a basis of V), r.1 or c.1 guarantees that $\langle \mathbf{u}_j, \mathbf{u}_j \rangle \neq 0$, and we deduce that

$$\alpha_j \;=\; \frac{\langle \mathbf{v}, \mathbf{u}_j \rangle}{\langle \mathbf{u}_j, \mathbf{u}_j \rangle}.$$

Since $j \in \{1, \dots, n\}$ was chosen arbitrarily, we now deduce that

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

□

Theorem 6.3.5

Let V be a real or complex vector space, equipped with a scalar product $\langle \cdot, \cdot \rangle$. Let $\mathcal{B} = \{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ be an **orthogonal** basis of V. Then for all $\mathbf{v} \in V$, we have that

$$\mathbf{v} = \sum_{i=1}^{n} \text{proj}_{\mathbf{u}_i}(\mathbf{v}) = \sum_{i=1}^{n} \frac{\langle \mathbf{v}, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \mathbf{u}_i,$$

and consequently, $\left[\ \mathbf{v}\ \right]_{\mathcal{B}} = \left[\ \frac{\langle \mathbf{v}, \mathbf{u}_1 \rangle}{\langle \mathbf{u}_1, \mathbf{u}_1 \rangle} \quad \dots \quad \frac{\langle \mathbf{v}, \mathbf{u}_n \rangle}{\langle \mathbf{u}_n, \mathbf{u}_n \rangle}\ \right]^T.$

Theorem 6.3.5

Let V be a real or complex vector space, equipped with a scalar product $\langle \cdot, \cdot \rangle$. Let $\mathcal{B} = \{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ be an **orthogonal** basis of V. Then for all $\mathbf{v} \in V$, we have that

$$\mathbf{v} \quad = \quad \sum_{i=1}^{n} \text{proj}_{\mathbf{u}_i}(\mathbf{v}) \quad = \quad \sum_{i=1}^{n} \frac{\langle \mathbf{v}, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \mathbf{u}_i,$$

and consequently, $\left[\ \mathbf{v}\ \right]_{\mathcal{B}} = \left[\ \frac{\langle \mathbf{v}, \mathbf{u}_1 \rangle}{\langle \mathbf{u}_1, \mathbf{u}_1 \rangle} \quad \dots \quad \frac{\langle \mathbf{v}, \mathbf{u}_n \rangle}{\langle \mathbf{u}_n, \mathbf{u}_n \rangle}\ \right]^T$.

Corollary 6.3.6

Let V be a real or complex vector space, equipped with a scalar product $\langle \cdot, \cdot \rangle$ and the induced norm $|| \cdot ||$. Let $\mathcal{B} = \{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ be an **orthonormal** basis of V. Then for all $\mathbf{v} \in V$, we have that

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

and consequently, $\left[\ \mathbf{v}\ \right]_{\mathcal{B}} = \left[\ \langle \mathbf{v}, \mathbf{u}_1 \rangle \quad \dots \quad \langle \mathbf{v}, \mathbf{u}_n \rangle\ \right]^T$.

4. Gram-Schmidt orthogonalization

4. Gram-Schmidt orthogonalization

- Our goal is to describe the "Gram-Schmidt orthogonalization process," which gives a recipe for transforming an arbitrary basis of a real or complex vector space (equipped with a scalar product and the norm induced by it) into an orthogonal (and even orthonormal) basis.

4. Gram-Schmidt orthogonalization
    - Our goal is to describe the “Gram-Schmidt orthogonalization process,” which gives a recipe for transforming an arbitrary basis of a real or complex vector space (equipped with a scalar product and the norm induced by it) into an orthogonal (and even orthonormal) basis.
    - There are in fact two different (but similar) versions of the Gram-Schmidt orthogonalization process.

4. Gram-Schmidt orthogonalization

- Our goal is to describe the "Gram-Schmidt orthogonalization process," which gives a recipe for transforming an arbitrary basis of a real or complex vector space (equipped with a scalar product and the norm induced by it) into an orthogonal (and even orthonormal) basis.
- There are in fact two different (but similar) versions of the Gram-Schmidt orthogonalization process.
    - The first version first produces an orthogonal basis, and then (optionally) produces an orthonormal basis.

4. Gram-Schmidt orthogonalization

- Our goal is to describe the “Gram-Schmidt orthogonalization process,” which gives a recipe for transforming an arbitrary basis of a real or complex vector space (equipped with a scalar product and the norm induced by it) into an orthogonal (and even orthonormal) basis.
- There are in fact two different (but similar) versions of the Gram-Schmidt orthogonalization process.
    - The first version first produces an orthogonal basis, and then (optionally) produces an orthonormal basis.
    - The second version produces an orthonormal basis directly.

4. Gram-Schmidt orthogonalization

- Our goal is to describe the "Gram-Schmidt orthogonalization process," which gives a recipe for transforming an arbitrary basis of a real or complex vector space (equipped with a scalar product and the norm induced by it) into an orthogonal (and even orthonormal) basis.
- There are in fact two different (but similar) versions of the Gram-Schmidt orthogonalization process.
    - The first version first produces an orthogonal basis, and then (optionally) produces an orthonormal basis.
    - The second version produces an orthonormal basis directly.
- We first describe the first version, we give a numerical example, and we prove the correctness of the process.

4. Gram-Schmidt orthogonalization

- Our goal is to describe the “Gram-Schmidt orthogonalization process,” which gives a recipe for transforming an arbitrary basis of a real or complex vector space (equipped with a scalar product and the norm induced by it) into an orthogonal (and even orthonormal) basis.
- There are in fact two different (but similar) versions of the Gram-Schmidt orthogonalization process.
  - The first version first produces an orthogonal basis, and then (optionally) produces an orthonormal basis.
  - The second version produces an orthonormal basis directly.
- We first describe the first version, we give a numerical example, and we prove the correctness of the process.
- Then we describe the second version.
  - The proof of correctness is similar to the proof of the first, and we omit it.
  - A numerical example is given in the Lecture Notes.

Gram-Schmidt orthogonalization process (version 1)

Let V be a real or complex vector space, equipped with a scalar product $\langle \cdot, \cdot \rangle$ and the induced norm $|| \cdot ||$, and let $\mathbf{v}_1, \dots, \mathbf{v}_k$ be linearly independent vectors in V. For all $\ell \in \{1, \dots, k\}$, set

$$\mathbf{u}_\ell \quad := \quad \mathbf{v}_\ell - \sum_{i=1}^{\ell-1} \text{proj}_{\mathbf{u}_i}(\mathbf{v}_\ell) \quad = \quad \mathbf{v}_\ell - \sum_{i=1}^{\ell-1} \frac{\langle \mathbf{v}_\ell, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \mathbf{u}_i.$$

Then $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ is an orthogonal basis of $\text{Span}(\mathbf{v}_1, \dots, \mathbf{v}_k)$, and $\left\{ \frac{\mathbf{u}_1}{||\mathbf{u}_1||}, \dots, \frac{\mathbf{u}_k}{||\mathbf{u}_k||} \right\}$ is an orthonormal basis of $\text{Span}(\mathbf{v}_1, \dots, \mathbf{v}_k)$.

- The sequence $\mathbf{u}_1, \dots, \mathbf{u}_k$ is obtained (recursively) as follows:
  - $\mathbf{u}_1 := \mathbf{v}_1$;
  - $\mathbf{u}_2 := \mathbf{v}_2 - \text{proj}_{\mathbf{u}_1}(\mathbf{v}_2)$;
  - $\mathbf{u}_3 := \mathbf{v}_3 - \Big(\text{proj}_{\mathbf{u}_1}(\mathbf{v}_3) + \text{proj}_{\mathbf{u}_2}(\mathbf{v}_3) \Big)$;

    $\vdots$
  - $\mathbf{u}_k := \mathbf{v}_k - \Big(\text{proj}_{\mathbf{u}_1}(\mathbf{v}_k) + \text{proj}_{\mathbf{u}_2}(\mathbf{v}_k) + \dots + \text{proj}_{\mathbf{u}_{k-1}}(\mathbf{v}_k) \Big)$.

Example 6.3.8

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

$$\mathbf{v}_1 = \begin{bmatrix} 3 \\ 4 \\ -4 \\ 3 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} -5 \\ 10 \\ 2 \\ 11 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} 8 \\ 19 \\ 11 \\ -2 \end{bmatrix}.$$

Set $U := \text{Span}(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$. Using the Gram-Schmidt orthogonalization process (version 1):

(a) compute an orthogonal basis of U (w.r.t. the standard scalar product $\cdot$ in $\mathbb{R}^4$).

(b) compute an orthonormal basis of U (w.r.t. the standard scalar product $\cdot$ in $\mathbb{R}^4$ and the norm $|| \cdot ||$ induced by it).

- **Remark:** To see that $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ really are linearly independent, we compute

$$\text{RREF}\Big(\begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3 \end{bmatrix}\Big) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix},$$

and we deduce that $\text{rank}\Big(\begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3 \end{bmatrix}\Big) = 3$, i.e. $\begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3 \end{bmatrix}$ has full column rank. So, by Theorem 3.3.12(a), vectors $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ are linearly independent.

Example 6.3.8

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

$$\mathbf{v}_1 = \begin{bmatrix} 3 \\ 4 \\ -4 \\ 3 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} -5 \\ 10 \\ 2 \\ 11 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} 8 \\ 19 \\ 11 \\ -2 \end{bmatrix}.$$

Set $U := \text{Span}(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$. Using the Gram-Schmidt orthogonalization process (version 1):

(a) compute an orthogonal basis of U (w.r.t. the standard scalar product $\cdot$ in $\mathbb{R}^4$).

(b) compute an orthonormal basis of U (w.r.t. the standard scalar product $\cdot$ in $\mathbb{R}^4$ and the norm $||\cdot||$ induced by it).

- Solution: On the board.

- Let's now prove the correctness of the Gram-Schmidt orthogonalization process!

- Let's now prove the correctness of the Gram-Schmidt orthogonalization process!
- We begin with a technical proposition.

- Let's now prove the correctness of the Gram-Schmidt orthogonalization process!
- We begin with a technical proposition.

Proposition 6.3.7

Let V be a real or complex vector space, equipped with a scalar product $\langle \cdot, \cdot \rangle$. Let $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ be an orthogonal set of non-zero vectors in V. Let $\mathbf{v} \in V$, and set $\mathbf{y} := \sum_{i=1}^{k} \text{proj}_{\mathbf{u}_i}(\mathbf{v}) = \sum_{i=1}^{k} \frac{\langle \mathbf{v}, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \mathbf{u}_i$ and $\mathbf{z} := \mathbf{v} - \mathbf{y}$. Then all the following hold:

a. $\{\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{z}\}$ is an orthogonal set of vectors;
b. $\mathbf{z} = \mathbf{0}$ iff $\mathbf{v} \in \text{Span}(\mathbf{u}_1, \dots, \mathbf{u}_k)$;
c. $\text{Span}(\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{v}) = \text{Span}(\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{z})$.

Proof.

- Let's now prove the correctness of the Gram-Schmidt orthogonalization process!
- We begin with a technical proposition.

Proposition 6.3.7

Let V be a real or complex vector space, equipped with a scalar product $\langle \cdot, \cdot \rangle$. Let $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ be an orthogonal set of non-zero vectors in V. Let $\mathbf{v} \in V$, and set $\mathbf{y} := \sum_{i=1}^{k} \text{proj}_{\mathbf{u}_i}(\mathbf{v}) = \sum_{i=1}^{k} \frac{\langle \mathbf{v}, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \mathbf{u}_i$ and $\mathbf{z} := \mathbf{v} - \mathbf{y}$. Then all the following hold:

(a) $\{\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{z}\}$ is an orthogonal set of vectors;

(b) $\mathbf{z} = \mathbf{0}$ iff $\mathbf{v} \in \text{Span}(\mathbf{u}_1, \dots, \mathbf{u}_k)$;

(c) $\text{Span}(\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{v}) = \text{Span}(\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{z})$.

Proof. First of all, Proposition 6.3.2 guarantees that $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ is a linearly independent set, and we deduce that $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ is an orthogonal basis of $\text{Span}(\mathbf{u}_1, \dots, \mathbf{u}_k)$.

Proof (continued). Let us first prove (a).

Proof (continued). Let us first prove (a). By hypothesis, vectors $\mathbf{u}_1, \ldots, \mathbf{u}_k$ are pairwise orthogonal.

Proof (continued). Let us first prove (a). By hypothesis, vectors $\mathbf{u}_1, \dots, \mathbf{u}_k$ are pairwise orthogonal. On the other hand, for each $j \in \{1, \dots, k\}$, we have the following:

$$\begin{aligned}
\langle \mathbf{z}, \mathbf{u}_j \rangle &= \langle \mathbf{v} - \sum_{i=1}^{k} \tfrac{\langle \mathbf{v}, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \mathbf{u}_i, \mathbf{u}_j \rangle \\
&\overset{(*)}{=} \langle \mathbf{v}, \mathbf{u}_j \rangle - \sum_{i=1}^{k} \tfrac{\langle \mathbf{v}, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \langle \mathbf{u}_i, \mathbf{u}_j \rangle \\
&\overset{(**)}{=} \langle \mathbf{v}, \mathbf{u}_j \rangle - \tfrac{\langle \mathbf{v}, \mathbf{u}_j \rangle}{\langle \mathbf{u}_j, \mathbf{u}_j \rangle} \langle \mathbf{u}_j, \mathbf{u}_j \rangle \\
&= \langle \mathbf{v}, \mathbf{u}_j \rangle - \langle \mathbf{v}, \mathbf{u}_j \rangle \quad = \quad 0,
\end{aligned}$$

where (*) follows from r.2 and r.3 (in the real case) or from c.2 and c.3 (in the complex case), and (**) follows from the fact that $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ is an orthogonal set.

Proof (continued). Let us first prove (a). By hypothesis, vectors $\mathbf{u}_1, \dots, \mathbf{u}_k$ are pairwise orthogonal. On the other hand, for each $j \in \{1, \dots, k\}$, we have the following:

$$\begin{aligned}
\langle \mathbf{z}, \mathbf{u}_j \rangle &= \langle \mathbf{v} - \sum_{i=1}^{k} \tfrac{\langle \mathbf{v}, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \mathbf{u}_i, \mathbf{u}_j \rangle \\
&\overset{(*)}{=} \langle \mathbf{v}, \mathbf{u}_j \rangle - \sum_{i=1}^{k} \tfrac{\langle \mathbf{v}, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \langle \mathbf{u}_i, \mathbf{u}_j \rangle \\
&\overset{(**)}{=} \langle \mathbf{v}, \mathbf{u}_j \rangle - \tfrac{\langle \mathbf{v}, \mathbf{u}_j \rangle}{\langle \mathbf{u}_j, \mathbf{u}_j \rangle} \langle \mathbf{u}_j, \mathbf{u}_j \rangle \\
&= \langle \mathbf{v}, \mathbf{u}_j \rangle - \langle \mathbf{v}, \mathbf{u}_j \rangle \;=\; 0,
\end{aligned}$$

where (*) follows from r.2 and r.3 (in the real case) or from c.2 and c.3 (in the complex case), and (**) follows from the fact that $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ is an orthogonal set. Thus, $\{\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{z}\}$ is an orthogonal set of vectors. This proves (a).

Proof (continued). Next, we prove (b).

Proof (continued). Next, we prove (b). Clearly, $\mathbf{z} = \mathbf{0}$ iff $\mathbf{v} = \sum_{i=1}^{k} \frac{\langle \mathbf{v}, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \mathbf{u}_i$.

Proof (continued). Next, we prove (b). Clearly, $\mathbf{z} = \mathbf{0}$ iff $\mathbf{v} = \sum_{i=1}^{k} \frac{\langle \mathbf{v}, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \mathbf{u}_i$.

So, we need to show that $\mathbf{v} = \sum_{i=1}^{k} \frac{\langle \mathbf{v}, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \mathbf{u}_i$ iff $\mathbf{v} \in \text{Span}(\mathbf{u}_1, \ldots, \mathbf{u}_k)$.

Proof (continued). Next, we prove (b). Clearly, $\mathbf{z} = \mathbf{0}$ iff $\mathbf{v} = \sum_{i=1}^{k} \frac{\langle \mathbf{v}, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \mathbf{u}_i$.

So, we need to show that $\mathbf{v} = \sum_{i=1}^{k} \frac{\langle \mathbf{v}, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \mathbf{u}_i$ iff $\mathbf{v} \in \text{Span}(\mathbf{u}_1, \ldots, \mathbf{u}_k)$.

If $\mathbf{v} = \sum_{i=1}^{k} \frac{\langle \mathbf{v}, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \mathbf{u}_i$, then $\mathbf{v}$ is a linear combination of the vectors $\mathbf{u}_1, \ldots, \mathbf{u}_k$, and consequently, $\mathbf{v} \in \text{Span}(\mathbf{u}_1, \ldots, \mathbf{u}_k)$.

Proof (continued). Next, we prove (b). Clearly, $\mathbf{z} = \mathbf{0}$ iff $\mathbf{v} = \sum_{i=1}^{k} \frac{\langle \mathbf{v}, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \mathbf{u}_i$.

So, we need to show that $\mathbf{v} = \sum_{i=1}^{k} \frac{\langle \mathbf{v}, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \mathbf{u}_i$ iff $\mathbf{v} \in \text{Span}(\mathbf{u}_1, \dots, \mathbf{u}_k)$.

If $\mathbf{v} = \sum_{i=1}^{k} \frac{\langle \mathbf{v}, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \mathbf{u}_i$, then $\mathbf{v}$ is a linear combination of the vectors $\mathbf{u}_1, \dots, \mathbf{u}_k$, and consequently, $\mathbf{v} \in \text{Span}(\mathbf{u}_1, \dots, \mathbf{u}_k)$.

On the other hand, if $\mathbf{v} \in \text{Span}(\mathbf{u}_1, \dots, \mathbf{u}_k)$, then Theorem 6.3.5 guarantees $\mathbf{v} = \sum_{i=1}^{k} \frac{\langle \mathbf{v}, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \mathbf{u}_i$ (because $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ is an orthogonal basis of $\text{Span}(\mathbf{u}_1, \dots, \mathbf{u}_k)$). This proves (b).

Proof (continued).

Proof (continued). Finally, we prove (c). Fix any vector $\mathbf{x} \in V$. WTS $\mathbf{x} \in \text{Span}(\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{v})$ iff $\mathbf{x} \in \text{Span}(\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{z})$. We prove both directions (they are very similar).

Proof (continued). Finally, we prove (c). Fix any vector $\mathbf{x} \in V$. WTS $\mathbf{x} \in \text{Span}(\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{v})$ iff $\mathbf{x} \in \text{Span}(\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{z})$. We prove both directions (they are very similar).

Suppose first that $\mathbf{x} \in \text{Span}(\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{v})$.

Proof (continued). Finally, we prove (c). Fix any vector $\mathbf{x} \in V$. WTS $\mathbf{x} \in \text{Span}(\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{v})$ iff $\mathbf{x} \in \text{Span}(\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{z})$. We prove both directions (they are very similar).

Suppose first that $\mathbf{x} \in \text{Span}(\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{v})$. Then there exist scalars $\alpha_1, \dots, \alpha_k, \beta$ s.t. $\mathbf{x} = \alpha_1 \mathbf{u}_1 + \dots + \alpha_k \mathbf{u}_k + \beta \mathbf{v}$.

Proof (continued). Finally, we prove (c). Fix any vector $\mathbf{x} \in V$. WTS $\mathbf{x} \in \text{Span}(\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{v})$ iff $\mathbf{x} \in \text{Span}(\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{z})$. We prove both directions (they are very similar).

Suppose first that $\mathbf{x} \in \text{Span}(\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{v})$. Then there exist scalars $\alpha_1, \dots, \alpha_k, \beta$ s.t. $\mathbf{x} = \alpha_1\mathbf{u}_1 + \dots + \alpha_k\mathbf{u}_k + \beta\mathbf{v}$. But now

$$\begin{aligned}
\mathbf{x} &= \alpha_1\mathbf{u}_1 + \dots + \alpha_k\mathbf{u}_k + \beta\mathbf{v} \\
&= \Big(\sum_{i=1}^{k} \alpha_i\mathbf{u}_i\Big) + \beta(\mathbf{y} + \mathbf{z}) \\
&= \Big(\sum_{i=1}^{k} \alpha_i\mathbf{u}_i\Big) + \beta\Big(\big(\sum_{i=1}^{k} \tfrac{\langle \mathbf{v}, \mathbf{u}_i\rangle}{\langle \mathbf{u}_i, \mathbf{u}_i\rangle}\mathbf{u}_i\big) + \mathbf{z}\Big) \\
&= \Big(\sum_{i=1}^{k} \big(\alpha_i + \beta\tfrac{\langle \mathbf{v}, \mathbf{u}_i\rangle}{\langle \mathbf{u}_i, \mathbf{u}_i\rangle}\big)\mathbf{u}_i\Big) + \beta\mathbf{z},
\end{aligned}$$

and we deduce that $\mathbf{x} \in \text{Span}(\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{z})$.

Proof (continued). Suppose, conversely, that $\mathbf{x} \in \text{Span}(\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{z})$.

Proof (continued). Suppose, conversely, that $\mathbf{x} \in \text{Span}(\mathbf{u}_1, \ldots, \mathbf{u}_k, \mathbf{z})$. Then there exist scalars $\alpha_1, \ldots, \alpha_k, \beta$ s.t. $\mathbf{x} = \alpha_1 \mathbf{u}_1 + \cdots + \alpha_k \mathbf{u}_k + \beta \mathbf{z}$.

Proof (continued). Suppose, conversely, that $\mathbf{x} \in \text{Span}(\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{z})$. Then there exist scalars $\alpha_1, \dots, \alpha_k, \beta$ s.t. $\mathbf{x} = \alpha_1 \mathbf{u}_1 + \dots + \alpha_k \mathbf{u}_k + \beta \mathbf{z}$. But now

$$\begin{aligned}
\mathbf{x} &= \alpha_1 \mathbf{u}_1 + \dots + \alpha_k \mathbf{u}_k + \beta \mathbf{z} \\
&= \Big(\sum_{i=1}^{k} \alpha_i \mathbf{u}_i \Big) + \beta(\mathbf{v} - \mathbf{y}) \\
&= \Big(\sum_{i=1}^{k} \alpha_i \mathbf{u}_i \Big) + \beta\Big(\mathbf{v} - \Big(\sum_{i=1}^{k} \tfrac{\langle \mathbf{v}, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \mathbf{u}_i \Big)\Big) \\
&= \Big(\sum_{i=1}^{k} \big(\alpha_i - \beta \tfrac{\langle \mathbf{v}, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle}\big) \mathbf{u}_i \Big) + \beta \mathbf{v},
\end{aligned}$$

and we deduce that $\mathbf{x} \in \text{Span}(\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{v})$. This proves (c). □

Proposition 6.3.7

Let V be a real or complex vector space, equipped with a scalar product $\langle \cdot, \cdot \rangle$. Let $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ be an orthogonal set of non-zero vectors in V. Let $\mathbf{v} \in V$, and set $\mathbf{y} := \sum_{i=1}^{k} \text{proj}_{\mathbf{u}_i}(\mathbf{v}) = \sum_{i=1}^{k} \frac{\langle \mathbf{v}, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \mathbf{u}_i$ and $\mathbf{z} := \mathbf{v} - \mathbf{y}$. Then all the following hold:

- (a) $\{\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{z}\}$ is an orthogonal set of vectors;
- (b) $\mathbf{z} = \mathbf{0}$ iff $\mathbf{v} \in \text{Span}(\mathbf{u}_1, \dots, \mathbf{u}_k)$;
- (c) $\text{Span}(\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{v}) = \text{Span}(\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{z})$.

Proposition 6.3.7

Let V be a real or complex vector space, equipped with a scalar product $\langle\cdot,\cdot\rangle$. Let $\{\mathbf{u}_1,\dots,\mathbf{u}_k\}$ be an orthogonal set of non-zero vectors in V. Let $\mathbf{v}\in V$, and set $\mathbf{y} := \sum_{i=1}^{k} \text{proj}_{\mathbf{u}_i}(\mathbf{v}) = \sum_{i=1}^{k} \frac{\langle \mathbf{v},\mathbf{u}_i\rangle}{\langle \mathbf{u}_i,\mathbf{u}_i\rangle}\mathbf{u}_i$ and $\mathbf{z} := \mathbf{v}-\mathbf{y}$. Then all the following hold:

a. $\{\mathbf{u}_1,\dots,\mathbf{u}_k,\mathbf{z}\}$ is an orthogonal set of vectors;
b. $\mathbf{z}=\mathbf{0}$ iff $\mathbf{v}\in \text{Span}(\mathbf{u}_1,\dots,\mathbf{u}_k)$;
c. $\text{Span}(\mathbf{u}_1,\dots,\mathbf{u}_k,\mathbf{v}) = \text{Span}(\mathbf{u}_1,\dots,\mathbf{u}_k,\mathbf{z})$.

- Using Proposition 6.3.7, we can now prove the correctness of the Gram-Schmidt orthogonalization process (version 1).

Gram-Schmidt orthogonalization process (version 1)

Let V be a real or complex vector space, equipped with a scalar product $\langle\cdot,\cdot\rangle$ and the induced norm $||\cdot||$, and let $\mathbf{v}_1,\dots,\mathbf{v}_k$ be linearly independent vectors in V. For all $\ell \in \{1,\dots,k\}$, set

$$\mathbf{u}_\ell \;:=\; \mathbf{v}_\ell - \sum_{i=1}^{\ell-1} \text{proj}_{\mathbf{u}_i}(\mathbf{v}_\ell) \;=\; \mathbf{v}_\ell - \sum_{i=1}^{\ell-1} \frac{\langle \mathbf{v}_\ell, \mathbf{u}_i\rangle}{\langle \mathbf{u}_i, \mathbf{u}_i\rangle}\mathbf{u}_i.$$

Then $\{\mathbf{u}_1,\dots,\mathbf{u}_k\}$ is an orthogonal basis of $\text{Span}(\mathbf{v}_1,\dots,\mathbf{v}_k)$, and $\left\{\frac{\mathbf{u}_1}{||\mathbf{u}_1||},\dots,\frac{\mathbf{u}_k}{||\mathbf{u}_k||}\right\}$ is an orthonormal basis of $\text{Span}(\mathbf{v}_1,\dots,\mathbf{v}_k)$.

Proof.

Gram-Schmidt orthogonalization process (version 1)

Let V be a real or complex vector space, equipped with a scalar product $\langle \cdot, \cdot \rangle$ and the induced norm $|| \cdot ||$, and let $\mathbf{v}_1, \dots, \mathbf{v}_k$ be linearly independent vectors in V. For all $\ell \in \{1, \dots, k\}$, set

$$\mathbf{u}_\ell \;:=\; \mathbf{v}_\ell - \sum_{i=1}^{\ell-1} \text{proj}_{\mathbf{u}_i}(\mathbf{v}_\ell) \;=\; \mathbf{v}_\ell - \sum_{i=1}^{\ell-1} \frac{\langle \mathbf{v}_\ell, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \mathbf{u}_i.$$

Then $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ is an orthogonal basis of $\text{Span}(\mathbf{v}_1, \dots, \mathbf{v}_k)$, and $\left\{ \frac{\mathbf{u}_1}{||\mathbf{u}_1||}, \dots, \frac{\mathbf{u}_k}{||\mathbf{u}_k||} \right\}$ is an orthonormal basis of $\text{Span}(\mathbf{v}_1, \dots, \mathbf{v}_k)$.

Proof. We first prove that $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ is an orthogonal basis of $\text{Span}(\mathbf{v}_1, \dots, \mathbf{v}_k)$.

Gram-Schmidt orthogonalization process (version 1)

Let V be a real or complex vector space, equipped with a scalar product $\langle \cdot, \cdot \rangle$ and the induced norm $|| \cdot ||$, and let $\mathbf{v}_1, \dots, \mathbf{v}_k$ be linearly independent vectors in V. For all $\ell \in \{1, \dots, k\}$, set

$$\mathbf{u}_\ell \quad := \quad \mathbf{v}_\ell - \sum_{i=1}^{\ell-1} \text{proj}_{\mathbf{u}_i}(\mathbf{v}_\ell) \quad = \quad \mathbf{v}_\ell - \sum_{i=1}^{\ell-1} \frac{\langle \mathbf{v}_\ell, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \mathbf{u}_i.$$

Then $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ is an orthogonal basis of $\text{Span}(\mathbf{v}_1, \dots, \mathbf{v}_k)$, and $\left\{ \frac{\mathbf{u}_1}{||\mathbf{u}_1||}, \dots, \frac{\mathbf{u}_k}{||\mathbf{u}_k||} \right\}$ is an orthonormal basis of $\text{Span}(\mathbf{v}_1, \dots, \mathbf{v}_k)$.

Proof. We first prove that $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ is an orthogonal basis of $\text{Span}(\mathbf{v}_1, \dots, \mathbf{v}_k)$. For each $\ell \in \{1, \dots, k\}$, we set $U_\ell := \text{Span}(\mathbf{v}_1, \dots, \mathbf{v}_\ell)$, and we prove (inductively) that $\{\mathbf{u}_1, \dots, \mathbf{u}_\ell\}$ is an orthogonal basis of U_ℓ.

Gram-Schmidt orthogonalization process (version 1)

Let V be a real or complex vector space, equipped with a scalar product $\langle \cdot, \cdot \rangle$ and the induced norm $|| \cdot ||$, and let $\mathbf{v}_1, \dots, \mathbf{v}_k$ be linearly independent vectors in V. For all $\ell \in \{1, \dots, k\}$, set

$$\mathbf{u}_\ell \quad := \quad \mathbf{v}_\ell - \sum_{i=1}^{\ell-1} \text{proj}_{\mathbf{u}_i}(\mathbf{v}_\ell) \quad = \quad \mathbf{v}_\ell - \sum_{i=1}^{\ell-1} \frac{\langle \mathbf{v}_\ell, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \mathbf{u}_i.$$

Then $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ is an orthogonal basis of $\text{Span}(\mathbf{v}_1, \dots, \mathbf{v}_k)$, and $\left\{ \frac{\mathbf{u}_1}{||\mathbf{u}_1||}, \dots, \frac{\mathbf{u}_k}{||\mathbf{u}_k||} \right\}$ is an orthonormal basis of $\text{Span}(\mathbf{v}_1, \dots, \mathbf{v}_k)$.

Proof. We first prove that $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ is an orthogonal basis of $\text{Span}(\mathbf{v}_1, \dots, \mathbf{v}_k)$. For each $\ell \in \{1, \dots, k\}$, we set $U_\ell := \text{Span}(\mathbf{v}_1, \dots, \mathbf{v}_\ell)$, and we prove (inductively) that $\{\mathbf{u}_1, \dots, \mathbf{u}_\ell\}$ is an orthogonal basis of U_ℓ. Obviously, this is enough, because for $k = \ell$, we get that $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ is an orthogonal basis of $U_k = \text{Span}(\mathbf{v}_1, \dots, \mathbf{v}_k)$.

Proof (continued). Reminder: WTS $\forall \ell \in \{1, \ldots, k\}$: $\{\mathbf{u}_1, \ldots, \mathbf{u}_\ell\}$ is an orthogonal basis of $U_\ell := \text{Span}(\mathbf{v}_1, \ldots, \mathbf{v}_\ell)$.

Proof (continued). Reminder: WTS $\forall \ell \in \{1, \ldots, k\}$: $\{\mathbf{u}_1, \ldots, \mathbf{u}_\ell\}$ is an orthogonal basis of $U_\ell := \text{Span}(\mathbf{v}_1, \ldots, \mathbf{v}_\ell)$.

Since $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ is linearly independent, we see that $\mathbf{v}_1, \ldots, \mathbf{v}_k$ are all non-zero, and in particular, $\{\mathbf{v}_1\}$ is linearly independent.

Proof (continued). Reminder: WTS $\forall \ell \in \{1, \dots, k\}$: $\{\mathbf{u}_1, \dots, \mathbf{u}_\ell\}$ is an orthogonal basis of $U_\ell := \text{Span}(\mathbf{v}_1, \dots, \mathbf{v}_\ell)$.

Since $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ is linearly independent, we see that $\mathbf{v}_1, \dots, \mathbf{v}_k$ are all non-zero, and in particular, $\{\mathbf{v}_1\}$ is linearly independent. Since $U_1 = \text{Span}(\mathbf{v}_1)$ and $\mathbf{u}_1 = \mathbf{v}_1$, we deduce that $\{\mathbf{u}_1\}$ is a basis of U_1, and this basis is obviously orthogonal (since it contains only one vector).

Now, fix $\ell \in \{1, \dots, k-1\}$, and assume inductively that $\{\mathbf{u}_1, \dots, \mathbf{u}_\ell\}$ is an orthogonal basis of U_ℓ.

Proof (continued). Reminder: WTS $\forall \ell \in \{1, \dots, k\}$: $\{\mathbf{u}_1, \dots, \mathbf{u}_\ell\}$ is an orthogonal basis of $U_\ell := \text{Span}(\mathbf{v}_1, \dots, \mathbf{v}_\ell)$.

Since $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ is linearly independent, we see that $\mathbf{v}_1, \dots, \mathbf{v}_k$ are all non-zero, and in particular, $\{\mathbf{v}_1\}$ is linearly independent. Since $U_1 = \text{Span}(\mathbf{v}_1)$ and $\mathbf{u}_1 = \mathbf{v}_1$, we deduce that $\{\mathbf{u}_1\}$ is a basis of U_1, and this basis is obviously orthogonal (since it contains only one vector).

Now, fix $\ell \in \{1, \dots, k-1\}$, and assume inductively that $\{\mathbf{u}_1, \dots, \mathbf{u}_\ell\}$ is an orthogonal basis of U_ℓ. WTS $\{\mathbf{u}_1, \dots, \mathbf{u}_\ell, \mathbf{u}_{\ell+1}\}$ is an orthogonal basis of $U_{\ell+1}$.

Proof (continued). Reminder: WTS $\forall \ell \in \{1, \dots, k\}$: $\{\mathbf{u}_1, \dots, \mathbf{u}_\ell\}$ is an orthogonal basis of $U_\ell := \text{Span}(\mathbf{v}_1, \dots, \mathbf{v}_\ell)$.

Since $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ is linearly independent, we see that $\mathbf{v}_1, \dots, \mathbf{v}_k$ are all non-zero, and in particular, $\{\mathbf{v}_1\}$ is linearly independent. Since $U_1 = \text{Span}(\mathbf{v}_1)$ and $\mathbf{u}_1 = \mathbf{v}_1$, we deduce that $\{\mathbf{u}_1\}$ is a basis of U_1, and this basis is obviously orthogonal (since it contains only one vector).

Now, fix $\ell \in \{1, \dots, k-1\}$, and assume inductively that $\{\mathbf{u}_1, \dots, \mathbf{u}_\ell\}$ is an orthogonal basis of U_ℓ. WTS $\{\mathbf{u}_1, \dots, \mathbf{u}_\ell, \mathbf{u}_{\ell+1}\}$ is an orthogonal basis of $U_{\ell+1}$.

We first prove that $\{\mathbf{u}_1, \dots, \mathbf{u}_\ell, \mathbf{u}_{\ell+1}\}$ is a basis of $U_{\ell+1}$, and then that it is an orthogonal set of vectors.

Proof (continued). Reminder: WTS $\forall \ell \in \{1, \ldots, k\}$: $\{\mathbf{u}_1, \ldots, \mathbf{u}_\ell\}$ is an orthogonal basis of $U_\ell := \text{Span}(\mathbf{v}_1, \ldots, \mathbf{v}_\ell)$.

Since $\{\mathbf{u}_1, \ldots, \mathbf{u}_\ell\}$ and $\{\mathbf{v}_1, \ldots, \mathbf{v}_\ell\}$ are two bases of U_ℓ, it is clear that $\text{Span}(\mathbf{u}_1, \ldots, \mathbf{u}_\ell, \mathbf{v}_{\ell+1}) = \text{Span}(\mathbf{v}_1, \ldots, \mathbf{v}_\ell, \mathbf{v}_{\ell+1}) = U_{\ell+1}$.

- Details?

Proof (continued). Reminder: WTS $\forall \ell \in \{1, \ldots, k\}$: $\{\mathbf{u}_1, \ldots, \mathbf{u}_\ell\}$ is an orthogonal basis of $U_\ell := \text{Span}(\mathbf{v}_1, \ldots, \mathbf{v}_\ell)$.

Since $\{\mathbf{u}_1, \ldots, \mathbf{u}_\ell\}$ and $\{\mathbf{v}_1, \ldots, \mathbf{v}_\ell\}$ are two bases of U_ℓ, it is clear that $\text{Span}(\mathbf{u}_1, \ldots, \mathbf{u}_\ell, \mathbf{v}_{\ell+1}) = \text{Span}(\mathbf{v}_1, \ldots, \mathbf{v}_\ell, \mathbf{v}_{\ell+1}) = U_{\ell+1}$.

- Details?

On the other hand, by the construction of $\mathbf{u}_{\ell+1}$ and by Proposition 6.3.7(c), we have that

$$\text{Span}(\mathbf{u}_1, \ldots, \mathbf{u}_\ell, \mathbf{v}_{\ell+1}) = \text{Span}(\mathbf{u}_1, \ldots, \mathbf{u}_\ell, \mathbf{u}_{\ell+1}).$$

Proof (continued). Reminder: WTS $\forall \ell \in \{1, \dots, k\}$: $\{\mathbf{u}_1, \dots, \mathbf{u}_\ell\}$ is an orthogonal basis of $U_\ell := \text{Span}(\mathbf{v}_1, \dots, \mathbf{v}_\ell)$.

Since $\{\mathbf{u}_1, \dots, \mathbf{u}_\ell\}$ and $\{\mathbf{v}_1, \dots, \mathbf{v}_\ell\}$ are two bases of U_ℓ, it is clear that $\text{Span}(\mathbf{u}_1, \dots, \mathbf{u}_\ell, \mathbf{v}_{\ell+1}) = \text{Span}(\mathbf{v}_1, \dots, \mathbf{v}_\ell, \mathbf{v}_{\ell+1}) = U_{\ell+1}$.

- Details?

On the other hand, by the construction of $\mathbf{u}_{\ell+1}$ and by Proposition 6.3.7(c), we have that

$$\text{Span}(\mathbf{u}_1, \dots, \mathbf{u}_\ell, \mathbf{v}_{\ell+1}) = \text{Span}(\mathbf{u}_1, \dots, \mathbf{u}_\ell, \mathbf{u}_{\ell+1}).$$

So, $\text{Span}(\mathbf{u}_1, \dots, \mathbf{u}_\ell, \mathbf{u}_{\ell+1}) = U_{\ell+1}$.

Proof (continued). Reminder: WTS $\forall \ell \in \{1, \dots, k\}$: $\{\mathbf{u}_1, \dots, \mathbf{u}_\ell\}$ is an orthogonal basis of $U_\ell := \text{Span}(\mathbf{v}_1, \dots, \mathbf{v}_\ell)$.

Since $\{\mathbf{u}_1, \dots, \mathbf{u}_\ell\}$ and $\{\mathbf{v}_1, \dots, \mathbf{v}_\ell\}$ are two bases of U_ℓ, it is clear that $\text{Span}(\mathbf{u}_1, \dots, \mathbf{u}_\ell, \mathbf{v}_{\ell+1}) = \text{Span}(\mathbf{v}_1, \dots, \mathbf{v}_\ell, \mathbf{v}_{\ell+1}) = U_{\ell+1}$.

- Details?

On the other hand, by the construction of $\mathbf{u}_{\ell+1}$ and by Proposition 6.3.7(c), we have that

$$\text{Span}(\mathbf{u}_1, \dots, \mathbf{u}_\ell, \mathbf{v}_{\ell+1}) = \text{Span}(\mathbf{u}_1, \dots, \mathbf{u}_\ell, \mathbf{u}_{\ell+1}).$$

So, $\text{Span}(\mathbf{u}_1, \dots, \mathbf{u}_\ell, \mathbf{u}_{\ell+1}) = U_{\ell+1}$.

Since $\dim(U_{\ell+1}) = \ell + 1$ (because $\{\mathbf{v}_1, \dots, \mathbf{v}_\ell, \mathbf{v}_{\ell+1}\}$ is a basis of $U_{\ell+1}$), the fact that $\{\mathbf{u}_1, \dots, \mathbf{u}_\ell, \mathbf{u}_{\ell+1}\}$ spans $U_{\ell+1}$ implies that $\{\mathbf{u}_1, \dots, \mathbf{u}_\ell, \mathbf{u}_{\ell+1}\}$ is in fact a basis of $U_{\ell+1}$ (this follows from Corollary 3.2.20).

Proof (continued). Reminder: WTS $\forall \ell \in \{1, \ldots, k\}$: $\{\mathbf{u}_1, \ldots, \mathbf{u}_\ell\}$ is an orthogonal basis of $U_\ell := \text{Span}(\mathbf{v}_1, \ldots, \mathbf{v}_\ell)$.

So far, we have shown that $\{\mathbf{u}_1, \ldots, \mathbf{u}_\ell, \mathbf{u}_{\ell+1}\}$ is a basis of $U_{\ell+1}$.

Proof (continued). Reminder: WTS $\forall \ell \in \{1, \dots, k\}$: $\{\mathbf{u}_1, \dots, \mathbf{u}_\ell\}$ is an orthogonal basis of $U_\ell := \text{Span}(\mathbf{v}_1, \dots, \mathbf{v}_\ell)$.

So far, we have shown that $\{\mathbf{u}_1, \dots, \mathbf{u}_\ell, \mathbf{u}_{\ell+1}\}$ is a basis of $U_{\ell+1}$. It remains to show that $\{\mathbf{u}_1, \dots, \mathbf{u}_\ell, \mathbf{u}_{\ell+1}\}$ is an orthogonal set.

By the induction hypothesis, vectors $\mathbf{u}_1, \dots, \mathbf{u}_\ell$ are pairwise orthogonal non-zero vectors, and so by the construction of $\mathbf{u}_{\ell+1}$ and by Proposition 6.3.7(a), we have that $\mathbf{u}_1, \dots, \mathbf{u}_\ell, \mathbf{u}_{\ell+1}$ are pairwise orthogonal.

Proof (continued). Reminder: WTS $\forall \ell \in \{1, \dots, k\}$: $\{\mathbf{u}_1, \dots, \mathbf{u}_\ell\}$ is an orthogonal basis of $U_\ell := \text{Span}(\mathbf{v}_1, \dots, \mathbf{v}_\ell)$.

So far, we have shown that $\{\mathbf{u}_1, \dots, \mathbf{u}_\ell, \mathbf{u}_{\ell+1}\}$ is a basis of $U_{\ell+1}$. It remains to show that $\{\mathbf{u}_1, \dots, \mathbf{u}_\ell, \mathbf{u}_{\ell+1}\}$ is an orthogonal set.

By the induction hypothesis, vectors $\mathbf{u}_1, \dots, \mathbf{u}_\ell$ are pairwise orthogonal non-zero vectors, and so by the construction of $\mathbf{u}_{\ell+1}$ and by Proposition 6.3.7(a), we have that $\mathbf{u}_1, \dots, \mathbf{u}_\ell, \mathbf{u}_{\ell+1}$ are pairwise orthogonal.

Proof (continued). Reminder: WTS $\forall \ell \in \{1, \dots, k\}$: $\{\mathbf{u}_1, \dots, \mathbf{u}_\ell\}$ is an orthogonal basis of $U_\ell := \text{Span}(\mathbf{v}_1, \dots, \mathbf{v}_\ell)$.

So far, we have shown that $\{\mathbf{u}_1, \dots, \mathbf{u}_\ell, \mathbf{u}_{\ell+1}\}$ is a basis of $U_{\ell+1}$. It remains to show that $\{\mathbf{u}_1, \dots, \mathbf{u}_\ell, \mathbf{u}_{\ell+1}\}$ is an orthogonal set.

By the induction hypothesis, vectors $\mathbf{u}_1, \dots, \mathbf{u}_\ell$ are pairwise orthogonal non-zero vectors, and so by the construction of $\mathbf{u}_{\ell+1}$ and by Proposition 6.3.7(a), we have that $\mathbf{u}_1, \dots, \mathbf{u}_\ell, \mathbf{u}_{\ell+1}$ are pairwise orthogonal.

Proof (continued). Reminder: WTS $\forall \ell \in \{1, \dots, k\}$: $\{\mathbf{u}_1, \dots, \mathbf{u}_\ell\}$ is an orthogonal basis of $U_\ell := \text{Span}(\mathbf{v}_1, \dots, \mathbf{v}_\ell)$.

So far, we have shown that $\{\mathbf{u}_1, \dots, \mathbf{u}_\ell, \mathbf{u}_{\ell+1}\}$ is a basis of $U_{\ell+1}$. It remains to show that $\{\mathbf{u}_1, \dots, \mathbf{u}_\ell, \mathbf{u}_{\ell+1}\}$ is an orthogonal set.

By the induction hypothesis, vectors $\mathbf{u}_1, \dots, \mathbf{u}_\ell$ are pairwise orthogonal non-zero vectors, and so by the construction of $\mathbf{u}_{\ell+1}$ and by Proposition 6.3.7(a), we have that $\mathbf{u}_1, \dots, \mathbf{u}_\ell, \mathbf{u}_{\ell+1}$ are pairwise orthogonal.

Proof (continued). Reminder: WTS $\forall \ell \in \{1, \dots, k\}$: $\{\mathbf{u}_1, \dots, \mathbf{u}_\ell\}$ is an orthogonal basis of $U_\ell := \text{Span}(\mathbf{v}_1, \dots, \mathbf{v}_\ell)$.

So far, we have shown that $\{\mathbf{u}_1, \dots, \mathbf{u}_\ell, \mathbf{u}_{\ell+1}\}$ is a basis of $U_{\ell+1}$. It remains to show that $\{\mathbf{u}_1, \dots, \mathbf{u}_\ell, \mathbf{u}_{\ell+1}\}$ is an orthogonal set.

By the induction hypothesis, vectors $\mathbf{u}_1, \dots, \mathbf{u}_\ell$ are pairwise orthogonal non-zero vectors, and so by the construction of $\mathbf{u}_{\ell+1}$ and by Proposition 6.3.7(a), we have that $\mathbf{u}_1, \dots, \mathbf{u}_\ell, \mathbf{u}_{\ell+1}$ are pairwise orthogonal.

So, $\{\mathbf{u}_1, \dots, \mathbf{u}_\ell, \mathbf{u}_{\ell+1}\}$ is an orthogonal basis of $U_{\ell+1}$. This completes the induction.

Gram-Schmidt orthogonalization process (version 1)

Let V be a real or complex vector space, equipped with a scalar product $\langle \cdot, \cdot \rangle$ and the induced norm $|| \cdot ||$, and let $\mathbf{v}_1, \dots, \mathbf{v}_k$ be linearly independent vectors in V. For all $\ell \in \{1, \dots, k\}$, set

$$\mathbf{u}_\ell \;:=\; \mathbf{v}_\ell - \sum_{i=1}^{\ell-1} \text{proj}_{\mathbf{u}_i}(\mathbf{v}_\ell) \;=\; \mathbf{v}_\ell - \sum_{i=1}^{\ell-1} \frac{\langle \mathbf{v}_\ell, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \mathbf{u}_i.$$

Then $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ is an orthogonal basis of $\text{Span}(\mathbf{v}_1, \dots, \mathbf{v}_k)$, and $\left\{ \frac{\mathbf{u}_1}{||\mathbf{u}_1||}, \dots, \frac{\mathbf{u}_k}{||\mathbf{u}_k||} \right\}$ is an orthonormal basis of $\text{Span}(\mathbf{v}_1, \dots, \mathbf{v}_k)$.

Proof (continued). We have now shown that $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ is an orthogonal basis of $\text{Span}(\mathbf{v}_1, \dots, \mathbf{v}_k)$.

Gram-Schmidt orthogonalization process (version 1)

Let V be a real or complex vector space, equipped with a scalar product $\langle \cdot, \cdot \rangle$ and the induced norm $|| \cdot ||$, and let $\mathbf{v}_1, \dots, \mathbf{v}_k$ be linearly independent vectors in V. For all $\ell \in \{1, \dots, k\}$, set

$$\mathbf{u}_\ell \;:=\; \mathbf{v}_\ell - \sum_{i=1}^{\ell-1} \text{proj}_{\mathbf{u}_i}(\mathbf{v}_\ell) \;=\; \mathbf{v}_\ell - \sum_{i=1}^{\ell-1} \frac{\langle \mathbf{v}_\ell, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \mathbf{u}_i.$$

Then $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ is an orthogonal basis of $\text{Span}(\mathbf{v}_1, \dots, \mathbf{v}_k)$, and $\left\{ \frac{\mathbf{u}_1}{||\mathbf{u}_1||}, \dots, \frac{\mathbf{u}_k}{||\mathbf{u}_k||} \right\}$ is an orthonormal basis of $\text{Span}(\mathbf{v}_1, \dots, \mathbf{v}_k)$.

Proof (continued). We have now shown that $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ is an orthogonal basis of $\text{Span}(\mathbf{v}_1, \dots, \mathbf{v}_k)$.

By Proposition 6.3.3(b), this implies that $\left\{ \frac{\mathbf{u}_1}{||\mathbf{u}_1||}, \dots, \frac{\mathbf{u}_k}{||\mathbf{u}_k||} \right\}$ is an orthonormal basis of $\text{Span}(\mathbf{v}_1, \dots, \mathbf{v}_k)$. This completes the argument. $\square$

Gram-Schmidt orthogonalization process (version 2)

Let V be a real or complex vector space, equipped with a scalar product $\langle \cdot, \cdot \rangle$ and the induced norm $|| \cdot ||$, and let $\mathbf{v}_1, \dots, \mathbf{v}_k$ be linearly independent vectors in V. For all $\ell \in \{1, \dots, k\}$, set

$$\mathbf{u}_\ell = \mathbf{v}_\ell - \sum_{i=1}^{\ell-1} \text{proj}_{\mathbf{z}_i}(\mathbf{v}_\ell) = \mathbf{v}_\ell - \sum_{i=1}^{\ell-1} \langle \mathbf{v}_\ell, \mathbf{z}_i \rangle \, \mathbf{z}_i;$$

$$\mathbf{z}_\ell = \frac{\mathbf{u}_\ell}{||\mathbf{u}_\ell||}.$$

Then $\{\mathbf{z}_1, \dots, \mathbf{z}_k\}$ is an orthonormal basis of $\text{Span}(\mathbf{v}_1, \dots, \mathbf{v}_k)$.

- The Gram-Schmidt orthogonalization process (version 2) recursively constructs two sequences of vectors, namely, $\mathbf{u}_1, \ldots, \mathbf{u}_k$ and $\mathbf{z}_1, \ldots, \mathbf{z}_k$, as follows:
  - $\mathbf{u}_1 = \mathbf{v}_1$;
  - $\mathbf{z}_1 = \frac{\mathbf{u}_1}{||\mathbf{u}_1||}$;
  - $\mathbf{u}_2 = \mathbf{v}_2 - \text{proj}_{\mathbf{z}_1}(\mathbf{v}_2)$;
  - $\mathbf{z}_2 = \frac{\mathbf{u}_2}{||\mathbf{u}_2||}$;
  - $\mathbf{u}_3 = \mathbf{v}_3 - \left(\text{proj}_{\mathbf{z}_1}(\mathbf{v}_3) + \text{proj}_{\mathbf{z}_2}(\mathbf{v}_3)\right)$;
  - $\mathbf{z}_3 = \frac{\mathbf{u}_3}{||\mathbf{u}_3||}$;

    $\vdots$
  - $\mathbf{u}_k = \mathbf{v}_k - \left(\text{proj}_{\mathbf{z}_1}(\mathbf{v}_k) + \text{proj}_{\mathbf{z}_2}(\mathbf{v}_k) + \cdots + \text{proj}_{\mathbf{z}_{k-1}}(\mathbf{v}_k)\right)$;
  - $\mathbf{z}_k = \frac{\mathbf{u}_k}{||\mathbf{u}_k||}$.

- The Gram-Schmidt orthogonalization process (version 2) recursively constructs two sequences of vectors, namely, $\mathbf{u}_1, \ldots, \mathbf{u}_k$ and $\mathbf{z}_1, \ldots, \mathbf{z}_k$, as follows:
  - $\mathbf{u}_1 = \mathbf{v}_1$;
  - $\mathbf{z}_1 = \frac{\mathbf{u}_1}{||\mathbf{u}_1||}$;
  - $\mathbf{u}_2 = \mathbf{v}_2 - \text{proj}_{\mathbf{z}_1}(\mathbf{v}_2)$;
  - $\mathbf{z}_2 = \frac{\mathbf{u}_2}{||\mathbf{u}_2||}$;
  - $\mathbf{u}_3 = \mathbf{v}_3 - \Big(\text{proj}_{\mathbf{z}_1}(\mathbf{v}_3) + \text{proj}_{\mathbf{z}_2}(\mathbf{v}_3)\Big)$;
  - $\mathbf{z}_3 = \frac{\mathbf{u}_3}{||\mathbf{u}_3||}$;

    $\vdots$
  - $\mathbf{u}_k = \mathbf{v}_k - \Big(\text{proj}_{\mathbf{z}_1}(\mathbf{v}_k) + \text{proj}_{\mathbf{z}_2}(\mathbf{v}_k) + \cdots + \text{proj}_{\mathbf{z}_{k-1}}(\mathbf{v}_k)\Big)$;
  - $\mathbf{z}_k = \frac{\mathbf{u}_k}{||\mathbf{u}_k||}$.
- So, at each step, we obtain a vector $\mathbf{u}_\ell$ that is orthogonal to the previously constructed vectors $\mathbf{z}_1, \ldots, \mathbf{z}_{\ell-1}$, and then we normalize $\mathbf{u}_\ell$ to obtain the unit vector $\mathbf{z}_\ell$ that points in the same direction as $\mathbf{u}_\ell$.

Gram-Schmidt orthogonalization process (version 2)

Let V be a real or complex vector space, equipped with a scalar product $\langle \cdot, \cdot \rangle$ and the induced norm $|| \cdot ||$, and let $\mathbf{v}_1, \dots, \mathbf{v}_k$ be linearly independent vectors in V. For all $\ell \in \{1, \dots, k\}$, set

$$\mathbf{u}_\ell = \mathbf{v}_\ell - \sum_{i=1}^{\ell-1} \text{proj}_{\mathbf{z}_i}(\mathbf{v}_\ell) = \mathbf{v}_\ell - \sum_{i=1}^{\ell-1} \langle \mathbf{v}_\ell, \mathbf{z}_i \rangle \, \mathbf{z}_i;$$

$$\mathbf{z}_\ell = \frac{\mathbf{u}_\ell}{||\mathbf{u}_\ell||}.$$

Then $\{\mathbf{z}_1, \dots, \mathbf{z}_k\}$ is an orthonormal basis of $\text{Span}(\mathbf{v}_1, \dots, \mathbf{v}_k)$.

- The proof of correctness is similar to that of version 1.
- A numerical example is given in the Lecture Notes.

Corollary 6.3.11

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 all the following hold:

(a) U has an orthogonal basis;

(b) any orthogonal basis of U can be extended to an orthogonal basis of V;[a]

(c) U has an orthonormal basis;

(d) any orthonormal basis of U can be extended to an orthonormal basis of V.[b]

---

[a] This means that for any orthogonal basis $\mathcal{B}$ of U, there exists an orthogonal basis $\mathcal{C}$ of V s.t. $\mathcal{B} \subseteq \mathcal{C}$.

[b] This means that for any orthonormal basis $\mathcal{B}$ of U, there exists an orthonormal basis $\mathcal{C}$ of V s.t. $\mathcal{B} \subseteq \mathcal{C}$.

Proof. We first prove (a) and (c).

Proof. We first prove (a) and (c). Since V is finite-dimensional, Theorem 3.2.21 guarantees that the subspace U of V is also finite-dimensional.

Proof. We first prove (a) and (c). Since V is finite-dimensional, Theorem 3.2.21 guarantees that the subspace U of V is also finite-dimensional. Consider any basis $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ of U.

Proof. We first prove (a) and (c). Since V is finite-dimensional, Theorem 3.2.21 guarantees that the subspace U of V is also finite-dimensional. Consider any basis $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ of U. Then the Gram-Schmidt orthogonalization process (version 1) applied to the vectors $\mathbf{v}_1, \dots, \mathbf{v}_k$ yields a sequence of vectors $\mathbf{u}_1, \dots, \mathbf{u}_k$ s.t. $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ is an orthogonal and $\left\{\frac{\mathbf{u}_1}{||\mathbf{u}_1||}, \dots, \frac{\mathbf{u}_k}{||\mathbf{u}_k||}\right\}$ an orthonormal basis of $U = \text{Span}(\mathbf{v}_1, \dots, \mathbf{v}_k)$.

Proof. We first prove (a) and (c). Since V is finite-dimensional, Theorem 3.2.21 guarantees that the subspace U of V is also finite-dimensional. Consider any basis $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ of U. Then the Gram-Schmidt orthogonalization process (version 1) applied to the vectors $\mathbf{v}_1, \dots, \mathbf{v}_k$ yields a sequence of vectors $\mathbf{u}_1, \dots, \mathbf{u}_k$ s.t. $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ is an orthogonal and $\left\{\frac{\mathbf{u}_1}{||\mathbf{u}_1||}, \dots, \frac{\mathbf{u}_k}{||\mathbf{u}_k||}\right\}$ an orthonormal basis of $U = \text{Span}(\mathbf{v}_1, \dots, \mathbf{v}_k)$. This proves (a) and (c).

Proof (continued). For (b), consider any orthogonal basis $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ of U, and using Theorem 3.2.19, extend it to a basis $\{\mathbf{v}_1, \dots, \mathbf{v}_k, \mathbf{v}_{k+1}, \dots, \mathbf{v}_n\}$ of V.

Proof (continued). For (b), consider any orthogonal basis $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ of U, and using Theorem 3.2.19, extend it to a basis $\{\mathbf{v}_1, \dots, \mathbf{v}_k, \mathbf{v}_{k+1}, \dots, \mathbf{v}_n\}$ of V.

We apply the Gram-Schmidt orthogonalization process (version 1) to the sequence $\mathbf{v}_1, \dots, \mathbf{v}_k, \mathbf{v}_{k+1}, \dots, \mathbf{v}_n$, and we obtain a sequence $\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{u}_{k+1}, \dots, \mathbf{u}_n$ s.t. $\{\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{u}_{k+1}, \dots, \mathbf{u}_n\}$ is an orthogonal basis of V.

Proof (continued). For (b), consider any orthogonal basis $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ of U, and using Theorem 3.2.19, extend it to a basis $\{\mathbf{v}_1, \dots, \mathbf{v}_k, \mathbf{v}_{k+1}, \dots, \mathbf{v}_n\}$ of V.

We apply the Gram-Schmidt orthogonalization process (version 1) to the sequence $\mathbf{v}_1, \dots, \mathbf{v}_k, \mathbf{v}_{k+1}, \dots, \mathbf{v}_n$, and we obtain a sequence $\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{u}_{k+1}, \dots, \mathbf{u}_n$ s.t. $\{\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{u}_{k+1}, \dots, \mathbf{u}_n\}$ is an orthogonal basis of V.

However, since $\mathbf{v}_1, \dots, \mathbf{v}_k$ were pairwise orthogonal to begin with, we see from the description of the Gram-Schmidt orthogonalization process that $\mathbf{u}_1 = \mathbf{v}_1, \dots, \mathbf{u}_k = \mathbf{v}_k$.

Proof (continued). For (b), consider any orthogonal basis $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ of U, and using Theorem 3.2.19, extend it to a basis $\{\mathbf{v}_1, \dots, \mathbf{v}_k, \mathbf{v}_{k+1}, \dots, \mathbf{v}_n\}$ of V.

We apply the Gram-Schmidt orthogonalization process (version 1) to the sequence $\mathbf{v}_1, \dots, \mathbf{v}_k, \mathbf{v}_{k+1}, \dots, \mathbf{v}_n$, and we obtain a sequence $\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{u}_{k+1}, \dots, \mathbf{u}_n$ s.t. $\{\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{u}_{k+1}, \dots, \mathbf{u}_n\}$ is an orthogonal basis of V.

However, since $\mathbf{v}_1, \dots, \mathbf{v}_k$ were pairwise orthogonal to begin with, we see from the description of the Gram-Schmidt orthogonalization process that $\mathbf{u}_1 = \mathbf{v}_1, \dots, \mathbf{u}_k = \mathbf{v}_k$.

So, the orthogonal basis $\{\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{u}_{k+1}, \dots, \mathbf{u}_n\}$ of V extends the orthogonal basis $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ of U. This proves (b).

Proof (continued). For (d), consider any orthonormal basis $\{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ of U.

Proof (continued). For (d), consider any orthonormal basis $\{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ of U. In particular, the basis $\{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ of U is orthogonal, and so by (b), it can be extended to an orthogonal basis $\{\mathbf{u}_1, \ldots, \mathbf{u}_k, \mathbf{u}_{k+1}, \ldots, \mathbf{u}_n\}$ of V.

Proof (continued). For (d), consider any orthonormal basis $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ of U. In particular, the basis $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ of U is orthogonal, and so by (b), it can be extended to an orthogonal basis $\{\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{u}_{k+1}, \dots, \mathbf{u}_n\}$ of V.

Then by Proposition 6.3.3(c),

$$\left\{ \frac{\mathbf{u}_1}{||\mathbf{u}_1||}, \dots, \frac{\mathbf{u}_k}{||\mathbf{u}_k||}, \frac{\mathbf{u}_{k+1}}{||\mathbf{u}_{k+1}||}, \dots, \frac{\mathbf{u}_n}{||\mathbf{u}_n||} \right\}$$

is an orthonormal basis of V.

Proof (continued). For (d), consider any orthonormal basis $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ of U. In particular, the basis $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ of U is orthogonal, and so by (b), it can be extended to an orthogonal basis $\{\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{u}_{k+1}, \dots, \mathbf{u}_n\}$ of V.

Then by Proposition 6.3.3(c),

$$\left\{ \frac{\mathbf{u}_1}{||\mathbf{u}_1||}, \dots, \frac{\mathbf{u}_k}{||\mathbf{u}_k||}, \frac{\mathbf{u}_{k+1}}{||\mathbf{u}_{k+1}||}, \dots, \frac{\mathbf{u}_n}{||\mathbf{u}_n||} \right\}$$

is an orthonormal basis of V.

But since the basis $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ of U is orthonormal, we know that $||\mathbf{u}_1|| = \dots = ||\mathbf{u}_k|| = 1$,

Proof (continued). For (d), consider any orthonormal basis $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ of U. In particular, the basis $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ of U is orthogonal, and so by (b), it can be extended to an orthogonal basis $\{\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{u}_{k+1}, \dots, \mathbf{u}_n\}$ of V.

Then by Proposition 6.3.3(c),

$$\left\{ \frac{\mathbf{u}_1}{||\mathbf{u}_1||}, \dots, \frac{\mathbf{u}_k}{||\mathbf{u}_k||}, \frac{\mathbf{u}_{k+1}}{||\mathbf{u}_{k+1}||}, \dots, \frac{\mathbf{u}_n}{||\mathbf{u}_n||} \right\}$$

is an orthonormal basis of V.

But since the basis $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ of U is orthonormal, we know that $||\mathbf{u}_1|| = \cdots = ||\mathbf{u}_k|| = 1$, and it follows that $\frac{\mathbf{u}_1}{||\mathbf{u}_1||} = \mathbf{u}_1, \dots, \frac{\mathbf{u}_k}{||\mathbf{u}_k||} = \mathbf{u}_k$.

Proof (continued). For (d), consider any orthonormal basis $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ of U. In particular, the basis $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ of U is orthogonal, and so by (b), it can be extended to an orthogonal basis $\{\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{u}_{k+1}, \dots, \mathbf{u}_n\}$ of V.

Then by Proposition 6.3.3(c),

$$\left\{\frac{\mathbf{u}_1}{||\mathbf{u}_1||}, \dots, \frac{\mathbf{u}_k}{||\mathbf{u}_k||}, \frac{\mathbf{u}_{k+1}}{||\mathbf{u}_{k+1}||}, \dots, \frac{\mathbf{u}_n}{||\mathbf{u}_n||}\right\}$$

is an orthonormal basis of V.

But since the basis $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ of U is orthonormal, we know that $||\mathbf{u}_1|| = \dots = ||\mathbf{u}_k|| = 1$, and it follows that $\frac{\mathbf{u}_1}{||\mathbf{u}_1||} = \mathbf{u}_1, \dots, \frac{\mathbf{u}_k}{||\mathbf{u}_k||} = \mathbf{u}_k$.

So, our orthonormal basis $\left\{\frac{\mathbf{u}_1}{||\mathbf{u}_1||}, \dots, \frac{\mathbf{u}_k}{||\mathbf{u}_k||}, \frac{\mathbf{u}_{k+1}}{||\mathbf{u}_{k+1}||}, \dots, \frac{\mathbf{u}_n}{||\mathbf{u}_n||}\right\}$ of V in fact extends the orthonormal basis $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ of U. This proves (d). □

Corollary 6.3.11

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 all the following hold:

a. U has an orthogonal basis;
b. any orthogonal basis of U can be extended to an orthogonal basis of V;[a]
c. U has an orthonormal basis;
d. any orthonormal basis of U can be extended to an orthonormal basis of V.[b]

---

[a] This means that for any orthogonal basis $\mathcal{B}$ of U, there exists an orthogonal basis $\mathcal{C}$ of V s.t. $\mathcal{B} \subseteq \mathcal{C}$.

[b] This means that for any orthonormal basis $\mathcal{B}$ of U, there exists an orthonormal basis $\mathcal{C}$ of V s.t. $\mathcal{B} \subseteq \mathcal{C}$.