Linear Algebra 2: Lecture 12

Irena Penev

Summer 2023

Recall that for a vector space V over $\mathbb{R}$ or $\mathbb{C}$, equipped with a scalar product $\langle \cdot, \cdot \rangle$, we say that vectors $\mathbf{x}$ and $\mathbf{y}$ are *orthogonal*, and we write $\mathbf{x} \perp \mathbf{y}$, provided that $\langle \mathbf{x}, \mathbf{y} \rangle = 0$. As we explained in Lecture Notes 11, the orthogonality relation is symmetric, i.e. $\mathbf{x} \perp \mathbf{y}$ if and only if $\mathbf{y} \perp \mathbf{x}$. Moreover, we saw that $\mathbf{0}$ is orthogonal to every vector in V.

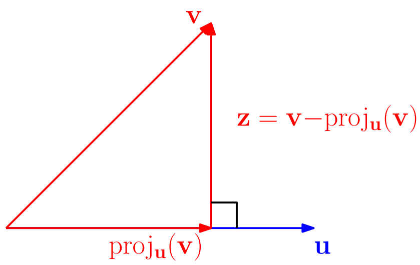
1 Vector projection

Suppose we are given a vector space V over $\mathbb{R}$ or $\mathbb{C}$, 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

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

As we can see, $\text{proj}_{\mathbf{u}}(\mathbf{v})$ is a scalar multiple of $\mathbf{u}$. Moreover, we have the following proposition.

Proposition 1.1. *Let V be a vector space over $\mathbb{R}$ or $\mathbb{C}$, 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 have that

$$\begin{aligned}
\langle \mathbf{z}, \mathbf{u} \rangle &= \langle \mathbf{v} - \mathrm{proj}_{\mathbf{u}}(\mathbf{v}), \mathbf{u} \rangle \\
&= \langle \mathbf{v} - \tfrac{\langle \mathbf{v}, \mathbf{u} \rangle}{\langle \mathbf{u}, \mathbf{u} \rangle} \mathbf{u}, \mathbf{u} \rangle \\
&= \langle \mathbf{v}, \mathbf{u} \rangle - \tfrac{\langle \mathbf{v}, \mathbf{u} \rangle}{\langle \mathbf{u}, \mathbf{u} \rangle} \langle \mathbf{u}, \mathbf{u} \rangle \\
&= \langle \mathbf{v}, \mathbf{u} \rangle - \langle \mathbf{v}, \mathbf{u} \rangle \\
&= 0,
\end{aligned}$$

and consequently, $\mathbf{z} \perp \mathbf{u}$, which is what we needed to show. □

2 Orthogonal and orthonormal bases. Gram-Schmidt orthogonalization

Suppose we are given a vector space V over $\mathbb{R}$ or $\mathbb{C}$, equipped with a scalar product $\langle \cdot, \cdot \rangle$ and the norm $||\cdot||$ induced by $\langle \cdot, \cdot \rangle$. An *orthogonal set of vectors* in V is a collection of pairwise orthogonal vectors in V. An *orthonormal set of vectors* is an orthogonal set of vectors that satisfies the additional property that all vectors in this set are of length 1 (i.e. every vector $\mathbf{v}$ in the set satisfies $||\mathbf{v}|| = 1$). An *orthogonal basis* (resp. *orthonormal basis*) for V is an orthogonal (resp. orthonormal) set in V that is also a basis of V.

If we have an orthogonal basis of a vector space over $\mathbb{R}$ or $\mathbb{C}$, 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 2.1). If our basis is orthonormal, then we get an even nicer formula for the coefficients (see Corollary 2.2).

Theorem 2.1. *Let V be a vector space over $\mathbb{R}$ or $\mathbb{C}$, equipped with a scalar product $\langle \cdot, \cdot \rangle$. Let $\{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ be an orthogonal basis of V. Then for all $\mathbf{v} \in V$, we have that*

$$\mathbf{v} = \sum_{i=1}^{n} \tfrac{\langle \mathbf{v}, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \mathbf{u}_i = \sum_{i=1}^{n} proj_{\mathbf{u}_i}(\mathbf{v}).$$

Proof. The second equality follows immediately from the definition of projection. Let us prove the first equality. Since $\mathbf{v} \in V$ and $\{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ is a basis of V, there exist scalars $\alpha_1, \dots, \alpha_n$ such that

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

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

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

where (*) 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), we see 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

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

and we are done. □

Corollary 2.2. *Let V be a vector space over $\mathbb{R}$ or $\mathbb{C}$, equipped with a scalar product $\langle \cdot, \cdot \rangle$ and the norm $|| \cdot ||$ induced by $\langle \cdot, \cdot \rangle$. Let $\{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ be an orthonormal basis of V. Then for all $\mathbf{v} \in V$, we have that*

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

Proof. Since $\{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ be an orthonormal basis of V, we know that $||\mathbf{u}_1|| = \dots = ||\mathbf{u}_n|| = 1$, and consequently (by the construction of $|| \cdot ||$), we have that $\langle \mathbf{u}_1, \mathbf{u}_1 \rangle = \dots = \langle \mathbf{u}_n, \mathbf{u}_n \rangle = 1$. The result now follows immediately from Theorem 2.2. □

We remark that the coefficients $\langle \mathbf{v}, \mathbf{u}_i \rangle$ from Corollary 2.2 are called the *Fourier coefficients.*

Our next goal is to describe the "Gram-Schmidt orthogonalization process," which gives a recipe for transforming an arbitrary basis of a vector space over $\mathbb{R}$ or $\mathbb{C}$, equipped with a scalar product and the norm induced by it, into an orthogonal (and even orthonormal) basis. But first, we need a couple of technical propositions.

Proposition 2.3. *Let V be a vector space over $\mathbb{R}$ or $\mathbb{C}$, 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. Then $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ is linearly independent.*

Proof. Fix scalars $\alpha_1, \dots, \alpha_k$ such that

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

We must show that $\alpha_1 = \cdots = \alpha_k = 0$. Fix any $i \in \{1, \dots, k\}$. Then

$$\langle \alpha_1 \mathbf{u}_1 + \cdots + \alpha_k \mathbf{u}_k, \mathbf{u}_i \rangle \quad = \quad \langle \mathbf{0}, \mathbf{u}_i \rangle.$$

Now, note that

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

where (*) follows from the fact that $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ is an orthogonal set. On the other hand, we know that $\langle \mathbf{0}, \mathbf{u}_i \rangle = 0$. So,

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

Since $\mathbf{u}_i \neq \mathbf{0}$, we see 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 2.4. *Let V be a vector space over $\mathbb{R}$ or $\mathbb{C}$, 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 $\mathbf{y} := \sum_{i=1}^{k} proj_{\mathbf{u}_i}(\mathbf{v}) = \sum_{i=1}^{k} \frac{\langle \mathbf{v}, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \mathbf{u}_i$ 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) $\mathbf{z} = \mathbf{0}$ if and only if $\mathbf{v} \in Span(\mathbf{u}_1, \dots, \mathbf{u}_k)$;

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

Proof. First of all, we remark that Proposition 2.3 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 $\text{Span}(\mathbf{u}_1, \dots, \mathbf{u}_k)$.

We first prove (a). Since $\{\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{z}\}$ is an orthogonal set of vectors, it suffices to show that $\langle \mathbf{z}, \mathbf{u}_j \rangle = 0$ for all $j \in \{1, \dots, k\}$. But indeed, for each

$j \in \{1, \dots, k\}$, we have that

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

where (*) follows from the fact that $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ is an orthogonal set. This proves (a).

Next, we prove (b). Clearly, $\mathbf{z} = \mathbf{0}$ if and only if $\mathbf{v} = \sum_{i=1}^{k} \frac{\langle \mathbf{v}, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \mathbf{u}_i$. 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 2.1 guarantees $\mathbf{v} = \sum_{i=1}^{k} \frac{\langle \mathbf{v}, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \mathbf{u}_i$.[1] This proves (b).

Finally, we prove (c). Fix any vector $\mathbf{x} \in V$. We must show that $\mathbf{x} \in \text{Span}(\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{v})$ if and only if $\mathbf{x} \in \text{Span}(\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{z})$. We prove both directions (as we shall see, 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 $\alpha_1, \dots, \alpha_k, \beta$ such that

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

But now

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

[1] This is because $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ is an orthogonal basis of $\text{Span}(\mathbf{u}_1, \dots, \mathbf{u}_k)$.

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

Suppose, conversely, that $\mathbf{x} \in \text{Span}(\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{z})$. Then there exist scalars $\alpha_1, \dots, \alpha_k, \beta$ such that

$$\mathbf{x} = \alpha_1 \mathbf{u}_1 + \cdots + \alpha_k \mathbf{u}_k + \beta \mathbf{z}.$$

But now

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

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

The Gram-Schmidt orthogonalization process (version 1). *Let V be a vector space over $\mathbb{R}$ or $\mathbb{C}$, equipped with a scalar product $\langle \cdot, \cdot \rangle$ and the norm $|| \cdot ||$ induced by $\langle \cdot, \cdot \rangle$. Let $\mathbf{v}_1, \dots, \mathbf{v}_k$ be a linearly independent set of vectors in V. For all $\ell \in \{1, \dots, k\}$, set*

$$\mathbf{u}_\ell = \mathbf{v}_\ell - \sum_{i=1}^{\ell-1} proj_{\mathbf{u}_i}(\mathbf{v}_\ell) = \mathbf{v}_\ell - \sum_{i=1}^{\ell-1} \tfrac{\langle \mathbf{v}_\ell, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \mathbf{u}_i.$$

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

Remark: Before proceeding with the proof, it may be helpful to note that the sequence $\mathbf{u}_1, \dots, \mathbf{u}_k$ is obtained (recursively) as follows:

- $\mathbf{u}_1 = \mathbf{v}_1$;
- $\mathbf{u}_2 = \mathbf{v}_2 - \text{proj}_{\mathbf{u}_1}(\mathbf{v}_2)$;
- $\mathbf{u}_3 = \mathbf{v}_3 - \Big(\text{proj}_{\mathbf{u}_1}(\mathbf{v}_3) + \text{proj}_{\mathbf{u}_2}(\mathbf{v}_3)\Big)$;

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

This describes precisely the sequence $\mathbf{u}_1, \dots, \mathbf{u}_k$ from the statement of the Gram-Schmidt orthogonalization process (version 1), only in a less compact (but perhaps more readable) form. We also note that it may be helpful to read Example 2.6 before reading the proof below, since Example 2.6 illustrates the Gram-Schmidt orthogonalization process on a concrete numerical example.

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 let $U_\ell := \text{Span}(\mathbf{v}_1, \dots, \mathbf{v}_\ell)$, and we prove 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 $U_k = \text{Span}(\mathbf{v}_1, \dots, \mathbf{v}_k)$.

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.

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. We must show that $\{\mathbf{u}_1, \dots, \mathbf{u}_\ell, \mathbf{u}_{\ell+1}\}$ is an orthogonal basis of $U_{\ell+1}$. Since $\{\mathbf{u}_1, \dots, \mathbf{u}_\ell\}$ and $\{\mathbf{v}_1, \dots, \mathbf{v}_\ell\}$ are two bases of U_ℓ,[2] it is clear that $\text{Span}(\mathbf{u}_1, \dots, \mathbf{u}_\ell, \mathbf{v}_{\ell+1}) = \text{Span}(\mathbf{v}_1, \dots, \mathbf{v}_\ell, \mathbf{v}_{\ell+1}) = U_{\ell+1}$.[3] On the other hand, by the construction of $\mathbf{u}_{\ell+1}$ and by Proposition 2.4(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 $\dim(U_{\ell+1}) = \ell + 1$,[4] 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 $U_{\ell+1}$.[5] By the induction hypothesis, vectors $\mathbf{u}_1, \dots, \mathbf{u}_\ell$ are pairwise orthogonal non-zero vectors,[6] and so by the costruction of $\mathbf{u}_{\ell+1}$ and by Proposition 2.4(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.

So far, we have shown that $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ is an orthogonal basis of $\text{Span}(\mathbf{v}_1, \dots, \mathbf{v}_k)$. It remains to show that $\{\frac{\mathbf{u}_1}{||\mathbf{u}_1||}, \dots, \frac{\mathbf{u}_k}{||\mathbf{u}_k||}\}$ is an orthonormal basis of $\text{Span}(\mathbf{v}_1, \dots, \mathbf{v}_k)$. Since $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ is a basis of $\text{Span}(\mathbf{v}_1, \dots, \mathbf{v}_k)$, we know that vectors $\mathbf{u}_1, \dots, \mathbf{u}_k$ are all non-zero, and consequently, $||\mathbf{u}_1||, \dots, ||\mathbf{u}_k||$ are all strictly positive. In particular, vectors $\frac{\mathbf{u}_1}{||\mathbf{u}_1||}, \dots, \frac{\mathbf{u}_k}{||\mathbf{u}_k||}$ are defined.

Since $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ is an orthogonal basis of $\text{Span}(\mathbf{v}_1, \dots, \mathbf{v}_k)$, and since the set $\{\frac{\mathbf{u}_1}{||\mathbf{u}_1||}, \dots, \frac{\mathbf{u}_k}{||\mathbf{u}_k||}\}$ is obtained simply by rescaling vectors in that basis by non-zero scalars, it is clear that $\{\frac{\mathbf{u}_1}{||\mathbf{u}_1||}, \dots, \frac{\mathbf{u}_k}{||\mathbf{u}_k||}\}$ is an orthogonal basis of $\text{Span}(\mathbf{v}_1, \dots, \mathbf{v}_k)$.[7] Moreover, for all $i \in \{1, \dots, k\}$, we have that

$$||\frac{\mathbf{u}_i}{||\mathbf{u}_i||}|| \quad = \quad |\frac{1}{||\mathbf{u}_i||}| \ \mathbf{u}_i \quad \overset{(*)}{=} \quad \frac{1}{||\mathbf{u}_i||} \ \mathbf{u}_i \quad = \quad 1,$$

where (*) follows from the fact that $||\mathbf{u}_i|| > 0$. So, $\{\frac{\mathbf{u}_1}{||\mathbf{u}_1||}, \dots, \frac{\mathbf{u}_k}{||\mathbf{u}_k||}\}$ is indeed an orthonormal basis of $\text{Span}(\mathbf{v}_1, \dots, \mathbf{v}_k)$. □

[2] The fact that $\{\mathbf{u}_1, \dots, \mathbf{u}_\ell\}$ is a basis of U_ℓ follows from the induction hypothesis. The fact that $\{\mathbf{v}_1, \dots, \mathbf{v}_\ell\}$ is a basis of U_ℓ follows from the fact that $\{\mathbf{v}_1, \dots, \mathbf{v}_\ell\}$ is linearly independent (because $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ is linearly independent) and $U_\ell = \text{Span}(\mathbf{v}_1, \dots, \mathbf{v}_\ell)$ (by construction).

[3] Details?

[4] This is because $\{\mathbf{v}_1, \dots, \mathbf{v}_{\ell+1}\}$ is a basis of $U_{\ell+1}$.

[5] We are using Proposition 1.11(b) from Lecture Notes 7.

[6] The fact that $\mathbf{u}_1, \dots, \mathbf{u}_\ell$ are all non-zero follows from the fact that $\{\mathbf{u}_1, \dots, \mathbf{u}_\ell\}$ is a basis of U_ℓ (by the induction hypothesis).

[7] Details?

Corollary 2.5. *Let V be a finite-dimensional vector space over $\mathbb{R}$ or $\mathbb{C}$, equipped with a scalar product $\langle\cdot,\cdot\rangle$ and the norm $||\cdot||$ induced by $\langle\cdot,\cdot\rangle$. 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;[8]

(c) U has an orthonormal basis;

(d) any orthonormal basis of U can be extended to an orthonormal basis of V.[9]

Proof. Since V is finite-dimensional, so is its subspace U.[10] Consider any basis $\{\mathbf{v}_1,\dots,\mathbf{v}_k\}$ of U. Then the Gram-Schmidt orthogonalization process applied to the vectors $\mathbf{v}_1,\dots,\mathbf{v}_k$ yields a sequence of vectors $\mathbf{u}_1,\dots,\mathbf{u}_k$ such that $\{\mathbf{u}_1,\dots,\mathbf{u}_k\}$ is an orthogonal and $\{\frac{\mathbf{u}_1}{||\mathbf{u}_1||},\dots,\frac{\mathbf{u}_k}{||\mathbf{u}_k||}\}$ an orthonormal basis of $U = \text{Span}(\mathbf{v}_1,\dots,\mathbf{v}_k)$. This proves (a) and (c).

For (b), consider any orthogonal basis $\{\mathbf{v}_1,\dots,\mathbf{v}_k\}$ of U, and extend it to a basis $\{\mathbf{v}_1,\dots,\mathbf{v}_k,\mathbf{v}_{k+1},\dots,\mathbf{v}_n\}$ of V. We then apply the Gram-Schmidt orthogonaliation process to the sequence $\mathbf{v}_1,\dots,\mathbf{v}_k,\mathbf{v}_{k+1},\dots,\mathbf{v}_n$, and we obtain some sequence $\mathbf{u}_1,\dots,\mathbf{u}_k,\mathbf{u}_{k+1},\dots,\mathbf{u}_n$ such that $\{\mathbf{u}_1,\dots,\mathbf{u}_k,\mathbf{u}_{k+1},\dots,\mathbf{u}_n\}$ 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).

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. But then $\{\frac{\mathbf{u}_1}{||\mathbf{u}_1||},\dots,\frac{\mathbf{u}_n}{||\mathbf{u}_n||}\}$ is an orthonormal basis of V.[11] 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 $\{\frac{\mathbf{u}_1}{||\mathbf{u}_1||},\dots,\frac{\mathbf{u}_n}{||\mathbf{u}_n||}\}$ of V in fact extends the orthonormal basis $\{\mathbf{u}_1,\dots,\mathbf{u}_k\}$ of U. This proves (d). □

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

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

[10] This follows from Theorem 1.12 from Lecture Notes 7.

[11] This follows from the proof of correctness of the Gram-Schmidt orthogonalization process (version 1).

Example 2.6. *Consider the vectors following vectors*

$$\mathbf{v}_1 = \begin{bmatrix} 3 \\ 1 \\ -1 \\ 3 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} -5 \\ 1 \\ 5 \\ -7 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} 1 \\ 1 \\ -2 \\ 8 \end{bmatrix}.$$

in $\mathbb{R}^4$*, and set* $U := Span(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$*. Using the Gram-Schmidt orthogonalization process (version 1):*

(a) find an orthogonal basis of U *(with respect to the standard scalar product* $\cdot$ *in* $\mathbb{R}^4$*).*

(b) find an orthonormal basis of U *(with respect to the standard scalar product* $\cdot$ *in* $\mathbb{R}^4$ *and the norm* $||\cdot||$ *induced by it).*

Solution. (a) We set $\mathbf{u}_1 := \mathbf{v}_1 = \begin{bmatrix} 3 & 1 & -1 & 3 \end{bmatrix}^T$. Next, we compute

$$\begin{aligned}
\mathbf{u}_2 &:= \mathbf{v}_2 - \tfrac{\mathbf{v}_2 \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1} \mathbf{u}_1 \\
&= \begin{bmatrix} -5 \\ 1 \\ 5 \\ -7 \end{bmatrix} - \frac{\begin{bmatrix} -5 \\ 1 \\ 5 \\ -7 \end{bmatrix} \cdot \begin{bmatrix} 3 \\ 1 \\ -1 \\ 3 \end{bmatrix}}{\begin{bmatrix} 3 \\ 1 \\ -1 \\ 3 \end{bmatrix} \cdot \begin{bmatrix} 3 \\ 1 \\ -1 \\ 3 \end{bmatrix}} \begin{bmatrix} 3 \\ 1 \\ -1 \\ 3 \end{bmatrix} \\
&= \begin{bmatrix} -5 \\ 1 \\ 5 \\ -7 \end{bmatrix} - \tfrac{(-5)\cdot 3 + 1\cdot 1 + 5\cdot(-1) + (-7)\cdot 3}{3\cdot 3 + 1\cdot 1 + (-1)\cdot(-1) + 3\cdot 3} \begin{bmatrix} 3 \\ 1 \\ -1 \\ 3 \end{bmatrix} \\
&= \begin{bmatrix} 1 \\ 3 \\ 3 \\ -1 \end{bmatrix}.
\end{aligned}$$

Finally, we compute

$$\begin{aligned}
\mathbf{u}_3 &:= \mathbf{v}_3 - \left(\tfrac{\mathbf{v}_3\cdot\mathbf{u}_1}{\mathbf{u}_1\cdot\mathbf{u}_1}\mathbf{u}_1 + \tfrac{\mathbf{v}_3\cdot\mathbf{u}_2}{\mathbf{u}_2\cdot\mathbf{u}_2}\mathbf{u}_2\right) \\
&= \begin{bmatrix} 1 \\ 1 \\ -2 \\ 8 \end{bmatrix} - \Bigg(\frac{\begin{bmatrix} 1 \\ 1 \\ -2 \\ 8 \end{bmatrix} \cdot \begin{bmatrix} 3 \\ 1 \\ -1 \\ 3 \end{bmatrix}}{\begin{bmatrix} 3 \\ 1 \\ -1 \\ 3 \end{bmatrix} \cdot \begin{bmatrix} 3 \\ 1 \\ -1 \\ 3 \end{bmatrix}} \begin{bmatrix} 3 \\ 1 \\ -1 \\ 3 \end{bmatrix} + \frac{\begin{bmatrix} 1 \\ 1 \\ -2 \\ 8 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 3 \\ 3 \\ -1 \end{bmatrix}}{\begin{bmatrix} 1 \\ 3 \\ 3 \\ -1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 3 \\ 3 \\ -1 \end{bmatrix}} \begin{bmatrix} 1 \\ 3 \\ 3 \\ -1 \end{bmatrix} \Bigg) \\
&= \begin{bmatrix} 1 \\ 1 \\ -2 \\ 8 \end{bmatrix} - \left(\tfrac{1\cdot3+1\cdot1+(-2)\cdot(-1)+8\cdot3}{3\cdot3+1\cdot1+(-1)\cdot(-1)+3\cdot3} \begin{bmatrix} 3 \\ 1 \\ -1 \\ 3 \end{bmatrix} + \tfrac{1\cdot1+1\cdot3+(-2)\cdot3+8\cdot(-1)}{1\cdot1+3\cdot3+3\cdot3+(-1)\cdot(-1)} \begin{bmatrix} 1 \\ 3 \\ 3 \\ -1 \end{bmatrix} \right) \\
&= \begin{bmatrix} -3 \\ 1 \\ 1 \\ 3 \end{bmatrix}.
\end{aligned}$$

So,

$$\mathcal{B} := \{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\} = \left\{ \begin{bmatrix} 3 \\ 1 \\ -1 \\ 3 \end{bmatrix}, \begin{bmatrix} 1 \\ 3 \\ 3 \\ -1 \end{bmatrix}, \begin{bmatrix} -3 \\ 1 \\ 1 \\ 3 \end{bmatrix} \right\}$$

is an orthogonal basis of U.

(b) To obtain an orthonormal basis of U, we "normalize" the vectors from the orthogonal basis $\mathcal{B}$ of U from part (a), that is, we rescale them so that their length becomes 1. We compute:

$$\begin{aligned}
||\mathbf{u}_1|| &= \sqrt{\mathbf{u}_1 \cdot \mathbf{u}_1} \\
&= \sqrt{\begin{bmatrix} 3 \\ 1 \\ -1 \\ 3 \end{bmatrix} \cdot \begin{bmatrix} 3 \\ 1 \\ -1 \\ 3 \end{bmatrix}} \\
&= \sqrt{3\cdot3 + 1\cdot1 + (-1)\cdot(-1) + 3\cdot3} \\
&= 2\sqrt{5}
\end{aligned}$$

$$
\begin{aligned}
||\mathbf{u}_2|| &= \sqrt{\mathbf{u}_2 \cdot \mathbf{u}_2} \\
&= \sqrt{\begin{bmatrix} 1 \\ 3 \\ 3 \\ -1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 3 \\ 3 \\ -1 \end{bmatrix}} \\
&= \sqrt{1 \cdot 1 + 3 \cdot 3 + 3 \cdot 3 + (-1) \cdot (-1)} \\
&= 2\sqrt{5}
\end{aligned}
$$

$$
\begin{aligned}
||\mathbf{u}_3|| &= \sqrt{\mathbf{u}_3 \cdot \mathbf{u}_3} \\
&= \sqrt{\begin{bmatrix} -3 \\ 1 \\ 1 \\ 3 \end{bmatrix} \cdot \begin{bmatrix} -3 \\ 1 \\ 1 \\ 3 \end{bmatrix}} \\
&= \sqrt{(-3) \cdot (-3) + 1 \cdot 1 + 1 \cdot 1 + 3 \cdot 3} \\
&= 2\sqrt{5}.
\end{aligned}
$$

(**Note:** All three vectors in $\mathcal{B}$ have length $2\sqrt{5}$. This is just a coincidence. Do not expect this to happen in general.)

We now see that

$$
\mathcal{C} := \{\tfrac{\mathbf{u}_1}{||\mathbf{u}_1||}, \tfrac{\mathbf{u}_2}{||\mathbf{u}_2||}, \tfrac{\mathbf{u}_3}{||\mathbf{u}_3||}\} = \left\{\tfrac{1}{2\sqrt{5}} \begin{bmatrix} 3 \\ 1 \\ -1 \\ 3 \end{bmatrix}, \tfrac{1}{2\sqrt{5}} \begin{bmatrix} 1 \\ 3 \\ 3 \\ -1 \end{bmatrix}, \tfrac{1}{2\sqrt{5}} \begin{bmatrix} -3 \\ 1 \\ 1 \\ 3 \end{bmatrix}\right\}
$$

is an orthonormal basis of U. □

The Gram-Schmidt orthogonalization process (version 2). *Let V be a vector space over $\mathbb{R}$ or $\mathbb{C}$, equipped with a scalar product $\langle \cdot, \cdot \rangle$ and the norm $|| \cdot ||$ induced by $\langle \cdot, \cdot \rangle$. Let $\mathbf{v}_1, \dots, \mathbf{v}_k$ be a linearly independent set of vectors in V. For all $\ell \in \{1, \dots, k\}$, set*

$$
\begin{aligned}
\mathbf{u}_\ell &= \mathbf{v}_\ell - \sum_{i=1}^{\ell-1} proj_{\mathbf{u}_i}(\mathbf{v}_\ell) = \mathbf{v}_\ell - \sum_{i=1}^{\ell-1} \langle \mathbf{v}_\ell, \mathbf{z}_i \rangle \, \mathbf{z}_i; \\
\mathbf{z}_\ell &= \tfrac{\mathbf{u}_\ell}{||\mathbf{u}_\ell||}.
\end{aligned}
$$

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

The proof of correctness of the Gram-Schmidt orthogonalization process (version 2) is similar to that of version 1, and we omit it. Let us, however, explain the main difference. The Gram-Schmidt orthogonalization process (version 2) recursively creates two sequences of vectors, namely, $\mathbf{u}_1, \dots, \mathbf{u}_k$ and $\mathbf{z}_1, \dots, \mathbf{z}_k$, as follows:

- $\mathbf{u}_1 = \mathbf{v}_1$;
- $\mathbf{z}_1 = \frac{\mathbf{u}_1}{||\mathbf{u}_1||}$;
- $\mathbf{u}_2 = \mathbf{v}_2 - \text{proj}_{\mathbf{z}_1}(\mathbf{v}_2)$;
- $\mathbf{z}_2 = \frac{\mathbf{u}_2}{||\mathbf{u}_2||}$;
- $\mathbf{u}_3 = \mathbf{v}_3 - \Big(\text{proj}_{\mathbf{z}_1}(\mathbf{v}_3) + \text{proj}_{\mathbf{z}_2}(\mathbf{v}_3)\Big)$;
- $\mathbf{z}_3 = \frac{\mathbf{u}_3}{||\mathbf{u}_3||}$;

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

So, at each step, we obtain a vector $\mathbf{u}_\ell$ that is orthogonal to the previously constructed vectors $\mathbf{z}_1, \dots, \mathbf{z}_{\ell-1}$, and then we "normalize" $\mathbf{u}_\ell$, i.e. we rescale it so that we obtain a vector $\mathbf{z}_\ell$ that points in the same direction as $\mathbf{u}_\ell$, but has length 1. (In version 1, we skip this normalization process during our recursive construction.) Note that, for each $\ell \in \{1, \dots, k\}$ and $i \in \{1, \dots, \ell - 1\}$, we have that $\text{proj}_{\mathbf{z}_i}(\mathbf{v}_\ell) = \langle \mathbf{v}_\ell, \mathbf{z}_i \rangle\, \mathbf{z}_i$. This is because $||\mathbf{z}_i|| = 1$, and so $\langle \mathbf{z}_i, \mathbf{z}_i \rangle = 1$, and therefore, $\text{proj}_{\mathbf{z}_i}(\mathbf{v}_\ell) = \frac{\langle \mathbf{v}_\ell, \mathbf{z}_i \rangle}{\langle \mathbf{z}_i, \mathbf{z}_i \rangle}\mathbf{z}_i = \langle \mathbf{v}_\ell, \mathbf{z}_i \rangle\, \mathbf{z}_i$.

We now return to Example 1, and we compute an orthonormal basis using the Gram-Schmidt process (version 2).

Example 2.7. *Consider the vectors following vectors*

$$\mathbf{v}_1 = \begin{bmatrix} 3 \\ 1 \\ -1 \\ 3 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} -5 \\ 1 \\ 5 \\ -7 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} 1 \\ 1 \\ -2 \\ 8 \end{bmatrix}.$$

in $\mathbb{R}^4$*, and set* $U := \text{Span}(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$*. Using the Gram-Schmidt orthogonalization process (version 2), find an orthonormal basis of* U *(with respect to the standard scalar product* $\cdot$ *in* $\mathbb{R}^4$ *and the norm* $||\cdot||$ *induced by it).*

Proof. We set $\mathbf{u}_1 := \mathbf{v}_1 = \begin{bmatrix} 3 & 1 & -1 & 3 \end{bmatrix}^T$, we compute $||\mathbf{u}_1|| = \sqrt{\mathbf{u}_1 \cdot \mathbf{u}_1} = 2\sqrt{5}$, and we set

$$\mathbf{z}_1 \quad := \quad \frac{\mathbf{u}_1}{||\mathbf{u}_1||} \quad = \quad \frac{1}{2\sqrt{5}} \begin{bmatrix} 3 \\ 1 \\ -1 \\ 3 \end{bmatrix}.$$

Next, we set

$$\begin{aligned}
\mathbf{u}_2 \quad &:= \quad \mathbf{v}_2 - \langle \mathbf{v}_2, \mathbf{z}_1 \rangle\, \mathbf{z}_1 \\
&= \begin{bmatrix} -5 \\ 1 \\ 5 \\ -7 \end{bmatrix} - \left(\begin{bmatrix} -5 \\ 1 \\ 5 \\ -7 \end{bmatrix} \cdot (\tfrac{1}{2\sqrt{5}} \begin{bmatrix} 3 \\ 1 \\ -1 \\ 3 \end{bmatrix}) \right) \left(\tfrac{1}{2\sqrt{5}} \begin{bmatrix} 3 \\ 1 \\ -1 \\ 3 \end{bmatrix} \right) \\
&= \begin{bmatrix} 1 \\ 3 \\ 3 \\ -1 \end{bmatrix}.
\end{aligned}$$

We compute $||\mathbf{u}_2|| = \sqrt{\mathbf{u}_2 \cdot \mathbf{u}_2} = 2\sqrt{5}$, and we set

$$\mathbf{z}_2 \quad := \quad \frac{\mathbf{u}_2}{||\mathbf{u}_2||} \quad = \quad \frac{1}{2\sqrt{5}} \begin{bmatrix} 1 \\ 3 \\ 3 \\ -1 \end{bmatrix}.$$

Next, we set

$$\begin{aligned}
\mathbf{u}_3 \quad &:= \quad \mathbf{v}_3 - \left(\langle \mathbf{v}_3, \mathbf{z}_1 \rangle\, \mathbf{z}_1 + \langle \mathbf{v}_3, \mathbf{z}_2 \rangle\, \mathbf{z}_2 \right) \\
&= \begin{bmatrix} 1 \\ 1 \\ -2 \\ 8 \end{bmatrix} - \Bigg(\left(\begin{bmatrix} 1 \\ 1 \\ -2 \\ 8 \end{bmatrix} \cdot (\tfrac{1}{2\sqrt{5}} \begin{bmatrix} 3 \\ 1 \\ -1 \\ 3 \end{bmatrix}) \right) (\tfrac{1}{2\sqrt{5}} \begin{bmatrix} 3 \\ 1 \\ -1 \\ 3 \end{bmatrix}) + \\
&\qquad + \left(\begin{bmatrix} 1 \\ 1 \\ -2 \\ 8 \end{bmatrix} \cdot (\tfrac{1}{2\sqrt{5}} \begin{bmatrix} 1 \\ 3 \\ 3 \\ -1 \end{bmatrix}) \right) (\tfrac{1}{2\sqrt{5}} \begin{bmatrix} 1 \\ 3 \\ 3 \\ -1 \end{bmatrix}) \Bigg) \\
&= \begin{bmatrix} -3 \\ 1 \\ 1 \\ 3 \end{bmatrix}.
\end{aligned}$$

We compute $||\mathbf{u}_3|| = \sqrt{\mathbf{u}_3 \cdot \mathbf{u}_3} = 2\sqrt{5}$, and we set

$$\mathbf{z}_3 \quad := \quad \tfrac{\mathbf{u}_3}{||\mathbf{u}_3||} \quad = \quad \tfrac{1}{2\sqrt{5}} \begin{bmatrix} -3 \\ 1 \\ 1 \\ 3 \end{bmatrix}.$$

We now have that

$$\mathcal{C} \quad := \quad \{\mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_3\} \quad = \quad \left\{ \tfrac{1}{2\sqrt{5}} \begin{bmatrix} 3 \\ 1 \\ -1 \\ 3 \end{bmatrix}, \tfrac{1}{2\sqrt{5}} \begin{bmatrix} 1 \\ 3 \\ 3 \\ -1 \end{bmatrix}, \tfrac{1}{2\sqrt{5}} \begin{bmatrix} -3 \\ 1 \\ 1 \\ 3 \end{bmatrix} \right\}$$

is an orthonormal basis of U. □

3 Orthogonal complement

Suppose we are given a vector space V over $\mathbb{R}$ or $\mathbb{C}$, equipped with a scalar product $\langle \cdot, \cdot \rangle$. For a vector $\mathbf{v} \in V$ and a set of vectors $A \subseteq V$,[12] we say that $\mathbf{v}$ is *orthogonal to* A, and we write $\mathbf{v} \perp A$, provided that $\mathbf{v}$ is orthogonal to all vectors in A.[13] For a set $A \subseteq V$, the *orthogonal complement* of A, denoted by $A^\perp$, is the set of all vectors in V that are orthogonal to A, i.e.

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

Proposition 3.1. *Let V be a vector space over $\mathbb{R}$ or $\mathbb{C}$, equipped with a scalar product $\langle \cdot, \cdot \rangle$. Let $A, B \subseteq V$. Then*

(a) $A^\perp$ is a subspace of V;[14]

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

Proof. (a) We use Theorem 2.7 from Lecture Notes 6.

First, we know that for all $\mathbf{v} \in V$, $\langle \mathbf{0}, \mathbf{v} \rangle = 0$, i.e. $\mathbf{0} \perp \mathbf{v}$. So, $\mathbf{0} \in A^\perp$.

Next, suppose $\mathbf{x}_1, \mathbf{x}_2 \in A^\perp$. Then for all $\mathbf{a} \in A$, we have that $\langle \mathbf{x}_1, \mathbf{a} \rangle = \langle \mathbf{x}_2, \mathbf{a} \rangle = 0$, and consequently,

$$\langle \mathbf{x}_1 + \mathbf{x}_2, \mathbf{a} \rangle \quad = \quad \langle \mathbf{x}_1, \mathbf{a} \rangle + \langle \mathbf{x}_2, \mathbf{a} \rangle \quad = \quad 0 + 0 \quad = 0,$$

[12] A may, but need not be, a subspace of V.

[13] By definition, this means that for all $\mathbf{a} \in A$, we have that $\langle \mathbf{v}, \mathbf{a} \rangle = 0$.

[14] 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.

i.e. $(\mathbf{x}_1 + \mathbf{x}_2) \perp \mathbf{a}$. So, $\mathbf{x}_1 + \mathbf{x}_2 \in A^\perp$.

Finally, suppose that $\mathbf{x} \in A^\perp$ and that α is a scalar. Then for all $\mathbf{a} \in A$, we have that $\langle \mathbf{x}, \mathbf{a} \rangle = 0$, and consequently,

$$\langle \alpha\mathbf{x}, \mathbf{a} \rangle \quad = \quad \alpha\langle \mathbf{x}, \mathbf{a} \rangle \quad = \quad \alpha 0 \quad = \quad 0,$$

i.e. $\alpha\mathbf{x} \perp \mathbf{a}$. So, $\alpha\mathbf{x} \in A^\perp$.

By Theorem 2.7 from Lecture Notes 6, it now follows that $A^\perp$ is a subspace of V.

(b) Suppose that $A \subseteq B$. We must show that $A^\perp \supseteq B^\perp$. Fix $\mathbf{x} \in B^\perp$; we must show that $\mathbf{x} \in A^\perp$. Fix $\mathbf{a} \in A$. Then $\mathbf{a} \in B$ (because $A \subseteq B$), and so $\mathbf{x} \perp \mathbf{a}$ (because $\mathbf{x} \in B^\perp$). So, $\mathbf{x} \in A^\perp$. This proves that $A^\perp \supseteq B^\perp$. □

Proposition 3.2. *Let V be a vector space over $\mathbb{R}$ or $\mathbb{C}$, equipped with a scalar product $\langle \cdot, \cdot \rangle$. Let $\mathbf{u}_1, \dots, \mathbf{u}_k \in V$. Then $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}^\perp = Span(\mathbf{u}_1, \dots, \mathbf{u}_k)^\perp$.*

Proof. Since $\{\mathbf{u}_1, \dots, \mathbf{u}_k\} \subseteq \text{Span}(\mathbf{u}_1, \dots, \mathbf{u}_k)$, Proposition 3.1(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 $\mathbf{x} \in \{\mathbf{u}_1, \dots, \mathbf{u}_k\}^\perp$. We must show that $\mathbf{x} \in \text{Span}(\mathbf{u}_1, \dots, \mathbf{u}_k)^\perp$. Fix $\mathbf{u} \in \text{Span}(\mathbf{u}_1, \dots, \mathbf{u}_k)$. Then there exist scalars $\alpha_1, \dots, \alpha_k$ such that $\mathbf{u} = \alpha_1\mathbf{u}_1 + \dots + \alpha_k\mathbf{u}_k$. But now

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

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, $\mathbf{x} \in \text{Span}(\mathbf{u}_1, \dots, \mathbf{u}_k)^\perp$. □

Recall 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.[15]

Theorem 3.3. *Let V be a finite-dimensional vector space over $\mathbb{R}$ or $\mathbb{C}$, equipped with a scalar product $\langle \cdot, \cdot \rangle$ and the norm $|| \cdot ||$ induced by $\langle \cdot, \cdot \rangle$. Let U be a subspace of V.*[16] *Then all the following hold:*

[15] This was Problem 3 from HW#7 from Linear Algebra 1 (winter 2022).

[16] By Theorem 1.12 from Lecture Notes 7, the fact that V is finite-dimensional implies that U is also finite-dimensional.

(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*,*[17] *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*,*[18] *then* $\{\mathbf{u}_{k+1}, \ldots, \mathbf{u}_n\}$ *is an orthonormal basis of* $U^\perp$*;*

(c) $dim(U) + dim(U^\perp) = dim(V)$*;*

(d) $V = U + U^\perp$*;*

(e) $(U^\perp)^\perp = U$*;*

(f) $U \cap U^\perp = \{\mathbf{0}\}$.

Proof. We first prove (a). Assume that $\{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ is an orthogonal basis of U, and that $\{\mathbf{u}_1, \ldots, \mathbf{u}_k, \mathbf{u}_{k+1}, \ldots, \mathbf{u}_n\}$ is an extension of that basis to an orthogonal basis of V. (This implies that $\{\mathbf{u}_1, \ldots, \mathbf{u}_k, \mathbf{u}_{k+1}, \ldots, \mathbf{u}_n\}$ is linearly independent, and consequently, $\mathbf{u}_1, \ldots, \mathbf{u}_k, \mathbf{u}_{k+1}, \ldots, \mathbf{u}_n$ are all non-zero.) We must show that $\{\mathbf{u}_{k+1}, \ldots, \mathbf{u}_n\}$ is an orthogonal basis of $U^\perp$. Clearly, $\{\mathbf{u}_{k+1}, \ldots, \mathbf{u}_n\}$ is an orthogonal set of vectors, and so it suffices to show that $\{\mathbf{u}_{k+1}, \ldots, \mathbf{u}_n\}$ is in fact a basis of $U^\perp$. By Proposition 2.3, $\{\mathbf{u}_{k+1}, \ldots, \mathbf{u}_n\}$ is linearly independent, and so we need only show that $\text{Span}(\{\mathbf{u}_{k+1}, \ldots, \mathbf{u}_n\}) = U^\perp$.

Let us first prove that $\text{Span}(\{\mathbf{u}_{k+1}, \ldots, \mathbf{u}_n\}) \supseteq U^\perp$. Fix $\mathbf{x} \in U^\perp$. Then $\mathbf{x} \in V$, and so since $\{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ is an orthogonal basis of V, Theorem 2.1 guarantees that

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

Since $\mathbf{x} \in U^\perp$, and since $\mathbf{u}_1, \ldots, \mathbf{u}_k \in U$, we know that $\langle \mathbf{x}, \mathbf{u}_i \rangle = 0$ for all $i \in \{1, \ldots, k\}$. Consequently,

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

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

For the reverse inclusion, we fix an arbitrary $\mathbf{x} \in \text{Span}(\mathbf{u}_{k+1}, \ldots, \mathbf{u}_n)$, and we show that $\mathbf{x} \in U^\perp$. Fix scalars $\alpha_{k+1}, \ldots, \alpha_n$ such that

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

[17] The existence of $\{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ and $\{\mathbf{u}_1, \ldots, \mathbf{u}_k, \mathbf{u}_{k+1}, \mathbf{u}_n\}$ follows from Corollary 2.5(a,b).

[18] The existence of $\{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ and $\{\mathbf{u}_1, \ldots, \mathbf{u}_k, \mathbf{u}_{k+1}, \mathbf{u}_n\}$ follows from Corollary 2.5(c,d).

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

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

Now, if V is a vector space over $\mathbb{R}$, then we have that

$$\begin{aligned}\langle \mathbf{x}, \mathbf{u} \rangle &= \langle \alpha_{k+1}\mathbf{u}_{k+1} + \dots + \alpha_n \mathbf{u}_n, \alpha_1 \mathbf{u}_1 + \dots + \alpha_k \mathbf{u}_k \rangle \\ &= \sum_{i=k+1}^{n} \sum_{j=1}^{k} \alpha_i \alpha_j \langle \mathbf{u}_i, \mathbf{u}_j \rangle \\ &\overset{(*)}{=} 0,\end{aligned}$$

where (*) follows from the fact that $\{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ is an orthogonal set. Similarly, if V is a vector space over $\mathbb{C}$, then we have that

$$\begin{aligned}\langle \mathbf{x}, \mathbf{u} \rangle &= \langle \alpha_{k+1}\mathbf{u}_{k+1} + \dots + \alpha_n \mathbf{u}_n, \alpha_1 \mathbf{u}_1 + \dots + \alpha_k \mathbf{u}_k \rangle \\ &= \sum_{i=k+1}^{n} \sum_{j=1}^{k} \alpha_i \overline{\alpha_j} \langle \mathbf{u}_i, \mathbf{u}_j \rangle \\ &\overset{(*)}{=} 0,\end{aligned}$$

where (*) follows from the fact that $\{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ is an orthogonal set. In either case, we get that $\mathbf{x} \perp \mathbf{u}$, and consequently, $\mathbf{x} \in U^{\perp}$. It follows that $\text{Span}(\{\mathbf{u}_{k+1}, \dots, \mathbf{u}_n\}) \subseteq U^{\perp}$. This proves (a). The proof of (b) is analogous.

Next, we prove (f). By hypothesis, U is a subspace of V, and by Proposition 3.1(a), $U^{\perp}$ is also a subspace of V. So, both U and $U^{\perp}$ contain $\mathbf{0}$, i.e. $\{\mathbf{0}\} \subseteq U \cap U^{\perp}$. 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 (f).

It remains to prove (c), (d), and (e). First, since V is finite-dimensional, so is U. So, by Corollary 2.5(a), U has an orthogonal basis $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$. By Corollary 2.5(b), $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ 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 $U^{\perp}$. With this set-up, we prove (c), (d) and (e).

We first prove (c). Since $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ is a basis of U, we have that $\dim(U) = k$. Since $\{\mathbf{u}_{k+1}, \dots, \mathbf{u}_n\}$ is a basis of $U^{\perp}$, we have that $\dim(U^{\perp}) = n - k$. Finally, since $\{\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{u}_{k+1}, \dots, \mathbf{u}_n\}$ is a basis of V, we have that $\dim(V) = n$. So, $\dim(U) + \dim(U^{\perp}) = k + (n-k) = n = \dim(V)$. This proves (c).

For (d), it is clear that $U + U^{\perp} \subseteq V$, and we need only show that $V \subseteq U + U^{\perp}$. Fix $\mathbf{v} \in V$. 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 $\mathbf{v} = \alpha_1 \mathbf{u}_1 + \cdots + \alpha_n \mathbf{u}_n$. Set $\mathbf{v}_1 := \alpha_1 \mathbf{u}_1 + \cdots + \alpha_k \mathbf{u}_k$ and $\mathbf{v}_2 := \alpha_{k+1} \mathbf{u}_{k+1} + \cdots + \alpha_n \mathbf{u}_n$. 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).

It remains to prove (e). We know that $\{\mathbf{u}_{k+1}, \dots, \mathbf{u}_n\}$ is an orthogonal basis of $U^\perp$, and that $\{\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{u}_{k+1}, \dots, \mathbf{u}_n\}$ is an orthogonal basis of V extending $\{\mathbf{u}_{k+1}, \dots, \mathbf{u}_n\}$. So, by (a), $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ is an orthogonal basis of $(U^\perp)^\perp$. But by construction, $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ is an orthogonal basis of U. So, $U = \text{Span}(\mathbf{u}_1, \dots, \mathbf{u}_k) = (U^\perp)^\perp$. This proves (e). $\square$