Orthonormal Bases, Outer Products and Hilbert Spaces

Last modified: 2026-11-24 09:19:45 (PST)

Orthonormal bases and Gram-Schmidt

Source

Like Section in Inner Products and the Cauchy-Schwarz Inequality, this section is adapted from Zhou [1], used under the MIT License (see the license text in Section in Subspaces and Rank). The source leaves the proof that orthonormal vectors are linearly independent to class; it is written out here.

Theorem 0.1 (The dot product of a vector with a linear combination)**.** *For vectors* $\underset{\sim}{x}, \underset{\sim}{u}_1, \dots, \underset{\sim}{u}_k \in \mathbb{R}^p$ *and numbers* $c_1, \dots, c_k$*, with* $k \geq 1$*,*

$$\underset{\sim}{x} \cdot \left(\sum_{j=1}^{k} c_j \underset{\sim}{u}_j \right) = \sum_{j=1}^{k} c_j \, (\underset{\sim}{x} \cdot \underset{\sim}{u}_j).$$

Proof

Proof. Write u_{ji} for entry i of $\underset{\sim}{u}_j$. Then

$$
\begin{aligned}
\underset{\sim}{x} \cdot \left(\sum_{j=1}^{k} c_j \underset{\sim}{u}_j \right) &= \sum_{i=1}^{p} x_i \sum_{j=1}^{k} c_j u_{ji} && \text{(Definition 10 in Vectors, Definition 11 in Vectors)} \\
&= \sum_{i=1}^{p} \sum_{j=1}^{k} x_i c_j u_{ji} && \text{(distribute each } x_i \text{ over the inner sum)} \\
&= \sum_{j=1}^{k} \sum_{i=1}^{p} x_i c_j u_{ji} && \text{(swap the order of the finite sums)} \\
&= \sum_{j=1}^{k} \sum_{i=1}^{p} c_j x_i u_{ji} && \text{(commute the factors in each product)} \\
&= \sum_{j=1}^{k} c_j \sum_{i=1}^{p} x_i u_{ji} && \text{(factor } c_j \text{ out of the inner sum)} \\
&= \sum_{j=1}^{k} c_j \, (\underset{\sim}{x} \cdot \underset{\sim}{u}_j). && \text{(Definition 10 in Vectors)}
\end{aligned}
$$

□

Example 0.1 (Dotting with a combination of three vectors)**.** Let $\underset{\sim}{x} = (1, 2, 3)$ and take the combination $2\,(1, 0, 0) - (0, 1, 0) + 4\,(0, 0, 1) = (2, -1, 4)$. Directly,

$$
\begin{aligned}
\underset{\sim}{x} \cdot (2, -1, 4) &= 2 - 2 + 12 \\
&= 12.
\end{aligned}
$$

By Theorem 0.1, $2 \cdot 1 - 1 \cdot 2 + 4 \cdot 3 = 12$, the same number.

Theorem 0.2 (Orthonormal vectors are linearly independent)**.** *If $\underset{\sim}{q}_1, \ldots, \underset{\sim}{q}_k$ are orthonormal (Definition 28 in Vectors), then they are linearly independent (Definition 1 in Subspaces and Rank).*

Proof

Proof. Suppose $\sum_{j=1}^{k} c_j \underset{\sim}{q}_j = \underset{\sim}{0}$. Take any $i \in \{1, \ldots, k\}$. Then

$$
\begin{aligned}
0 &= \underset{\sim}{q}_i \cdot \underset{\sim}{0} && \text{(every term of the dot product is 0)} \\
&= \underset{\sim}{q}_i \cdot \left(\sum_{j=1}^{k} c_j \underset{\sim}{q}_j \right) && \text{(substitute the supposed equation)} \\
&= \sum_{j=1}^{k} c_j \left(\underset{\sim}{q}_i \cdot \underset{\sim}{q}_j \right) && \text{(Theorem 0.1)} \\
&= c_i. && \left(\underset{\sim}{q}_i \cdot \underset{\sim}{q}_j \text{ is 1 if } j = i \text{ and 0 otherwise} \right)
\end{aligned}
$$

So every $c_i = 0$. □

Example 0.2 (Three orthonormal vectors in $\mathbb{R}^3$)**.** Let

$$
\underset{\sim}{q}_1 = (0, 0, 1), \quad \underset{\sim}{q}_2 = \tfrac{1}{\sqrt{2}} (1, 1, 0), \quad \underset{\sim}{q}_3 = \tfrac{1}{\sqrt{2}} (1, -1, 0).
$$

Each has norm 1:

$$
\begin{aligned}
\left\| \underset{\sim}{q}_2 \right\|^2 &= \tfrac{1}{2} (1 + 1 + 0) \\
&= 1,
\end{aligned}
$$

and likewise for the others. They are mutually orthogonal: $\underset{\sim}{q}_1 \cdot \underset{\sim}{q}_2 = 0$ and $\underset{\sim}{q}_1 \cdot \underset{\sim}{q}_3 = 0$, because $\underset{\sim}{q}_1$ has its only nonzero entry where the others have $\underset{\sim}{0}$, and

$$
\begin{aligned}
\underset{\sim}{q}_2 \cdot \underset{\sim}{q}_3 &= \tfrac{1}{2} (1 - 1 + 0) \\
&= 0.
\end{aligned}
$$

So by Theorem 0.2 they are linearly independent, with no need to solve $c_1 \underset{\sim}{q}_1 + c_2 \underset{\sim}{q}_2 + c_3 \underset{\sim}{q}_3 = \underset{\sim}{0}$.
The converse fails: $(1, 0)$ and $(1, 1)$ are linearly independent, since $c_1(1, 0) + c_2(1, 1) = (c_1 + c_2, c_2)$ is $\underset{\sim}{0}$ only if $c_2 = 0$ and then $c_1 = 0$, but they are not orthonormal, since $(1, 0) \cdot (1, 1) = 1 \neq 0$.

Definition 0.1 (Orthonormal basis)**.** An **orthonormal basis** [2] of a subspace $\mathcal{S}$ of $\mathbb{R}^p$ is a basis of $\mathcal{S}$ (Definition 1 in Bases and Dimension) whose vectors are orthonormal (Definition 28 in Vectors).

Example 0.3 (Orthonormal bases of $\mathbb{R}^3$)**.**

- The indicator vectors $\underset{\sim}{e}_1, \underset{\sim}{e}_2, \underset{\sim}{e}_3$ are orthonormal
  (1) and a basis of $\mathbb{R}^3$ (Example 6 in Bases and Dimension).
- The vectors $\underset{\sim}{q}_1, \underset{\sim}{q}_2, \underset{\sim}{q}_3$ of Example 0.2 are orthonormal and linearly independent,

so they are a basis of their span (Definition 1 in Bases and Dimension), which therefore has dimension 3 (Definition 4 in Bases and Dimension) and is all of $\mathbb{R}^3$ (Theorem 6 in Bases and Dimension, Example 6 in Bases and Dimension). So they are an orthonormal basis of $\mathbb{R}^3$ too.

Example 0.4 (Bases that are not orthonormal, and orthonormal lists that are not bases)**.**

- $(1,0),(1,1)$ is a basis of $\mathbb{R}^2$ (Example 1 in Bases and Dimension). It is not an orthonormal basis, since $(1,0)\cdot(1,1)=1$.
- $(1,0,0),(0,1,0)$ is orthonormal but not a basis of $\mathbb{R}^3$: its span is the plane $z=0$ (Example 6 in Subspaces and Rank), which does not contain $(0,0,1)$.

Theorem 0.3 (Coordinates in an orthonormal basis are dot products)**.** *If $\underset{\sim}{q}_1,\dots,\underset{\sim}{q}_k$ is an orthonormal basis of a subspace $\mathcal{S}$ (Definition 0.1), then every $\underset{\sim}{x}\in\mathcal{S}$ satisfies*

$$\underset{\sim}{x}=\sum_{i=1}^{k}(\underset{\sim}{q}_i\cdot\underset{\sim}{x})\,\underset{\sim}{q}_i.$$

Proof

Proof. The basis spans $\mathcal{S}$, so $\underset{\sim}{x}=\sum_{j=1}^{k}c_j\underset{\sim}{q}_j$ for some numbers c_j. For each i,

$$\begin{aligned}\underset{\sim}{q}_i\cdot\underset{\sim}{x}&=\underset{\sim}{q}_i\cdot\left(\sum_{j=1}^{k}c_j\underset{\sim}{q}_j\right) && \text{(substitute)}\\ &=\sum_{j=1}^{k}c_j\,(\underset{\sim}{q}_i\cdot\underset{\sim}{q}_j) && \text{(Theorem 0.1)}\\ &=c_i, && \text{(Definition 28 in Vectors)}\end{aligned}$$

so each coefficient c_i is $\underset{\sim}{q}_i\cdot\underset{\sim}{x}$. □

Example 0.5 (Expanding a vector in an orthonormal basis)**.** With the orthonormal basis $\underset{\sim}{q}_1,\underset{\sim}{q}_2,\underset{\sim}{q}_3$ of $\mathbb{R}^3$ from Example 0.3 and $\underset{\sim}{x}=(1,2,3)$: $\underset{\sim}{q}_1\cdot\underset{\sim}{x}=3$,

$$\begin{aligned}\underset{\sim}{q}_2\cdot\underset{\sim}{x}&=\tfrac{1}{\sqrt{2}}\,(1+2)\\ &=\tfrac{3}{\sqrt{2}}\end{aligned}$$

and

$$\begin{aligned} \underset{\sim}{q}_3 \cdot \underset{\sim}{x} &= \tfrac{1}{\sqrt{2}}(1-2) \\ &= -\tfrac{1}{\sqrt{2}}. \end{aligned}$$

Then

$$\begin{aligned} \sum_{i=1}^{3} (\underset{\sim}{q}_i \cdot \underset{\sim}{x})\, \underset{\sim}{q}_i &= 3\,(0,0,1) + \tfrac{3}{\sqrt{2}} \cdot \tfrac{1}{\sqrt{2}}\,(1,1,0) - \tfrac{1}{\sqrt{2}} \cdot \tfrac{1}{\sqrt{2}}\,(1,-1,0) && \text{(substitute)} \\ &= (0,0,3) + \left(\tfrac{3}{2}, \tfrac{3}{2}, 0\right) + \left(-\tfrac{1}{2}, \tfrac{1}{2}, 0\right) && \text{(multiply out each term)} \\ &= (1,2,3), && \text{(add entrywise)} \end{aligned}$$

which is $\underset{\sim}{x}$, as Theorem 0.3 says.

Definition 0.2 (Gram-Schmidt process)**.** The **Gram-Schmidt process** [3] takes vectors $\underset{\sim}{a}_1, \ldots, \underset{\sim}{a}_k \in \mathbb{R}^p$ and, for $i = 1, 2, \ldots, k$ in turn:

1. **Orthogonalize:** $\underset{\sim}{q}'_i \stackrel{\text{def}}{=} \underset{\sim}{a}_i - \sum_{j=1}^{i-1} (\underset{\sim}{q}_j \cdot \underset{\sim}{a}_i)\, \underset{\sim}{q}_j$ (for $i = 1$ the sum is empty, so $\underset{\sim}{q}'_1 = \underset{\sim}{a}_1$);
2. **Test:** if $\underset{\sim}{q}'_i = \underset{\sim}{0}$, stop;
3. **Normalize:** $\underset{\sim}{q}_i \stackrel{\text{def}}{=} \underset{\sim}{q}'_i / \|\underset{\sim}{q}'_i\|$.

The output is the list $\underset{\sim}{q}_1, \underset{\sim}{q}_2, \ldots$ produced before the process stops, or all k of them if it never stops.

Example 0.6 (Two steps of Gram-Schmidt in $\mathbb{R}^3$)**.** Let $\underset{\sim}{a}_1 = (1,1,0)$ and $\underset{\sim}{a}_2 = (1,0,1)$.

- **Step 1.** $\underset{\sim}{q}'_1 = (1,1,0)$, with norm $\sqrt{2}$, so $\underset{\sim}{q}_1 = \frac{1}{\sqrt{2}}(1,1,0)$.
- **Step 2.**

  $$\begin{aligned} \underset{\sim}{q}_1 \cdot \underset{\sim}{a}_2 &= \tfrac{1}{\sqrt{2}}(1+0+0) \\ &= \tfrac{1}{\sqrt{2}}, \end{aligned}$$

  so

  $$\begin{aligned} \underset{\sim}{q}'_2 &= (1,0,1) - \tfrac{1}{\sqrt{2}} \cdot \tfrac{1}{\sqrt{2}}\,(1,1,0) && \text{(orthogonalize)} \\ &= (1,0,1) - \left(\tfrac{1}{2}, \tfrac{1}{2}, 0\right) && \text{(multiply out)} \\ &= \left(\tfrac{1}{2}, -\tfrac{1}{2}, 1\right), && \text{(subtract entrywise)} \end{aligned}$$

  with norm $\sqrt{\frac{1}{4} + \frac{1}{4} + 1} = \sqrt{\frac{3}{2}}$, so

$$\begin{aligned} \underset{\sim}{q}_2 &= \sqrt{\tfrac{2}{3}}\,(\tfrac{1}{2}, -\tfrac{1}{2}, 1) \\ &= \tfrac{1}{\sqrt{6}}\,(1, -1, 2). \end{aligned}$$

As a check,

$$\begin{aligned} \underset{\sim}{q}_1 \cdot \underset{\sim}{q}_2 &= \tfrac{1}{\sqrt{12}}\,(1 - 1 + 0) \\ &= 0. \end{aligned}$$

Example 0.7 (Gram-Schmidt stops on dependent vectors)**.** Let $\underset{\sim}{a}_1 = (1, 2)$ and $\underset{\sim}{a}_2 = (2, 4)$. Step 1 gives $\underset{\sim}{q}_1 = \frac{1}{\sqrt{5}}\,(1, 2)$. In step 2,

$$\begin{aligned} \underset{\sim}{q}_1 \cdot \underset{\sim}{a}_2 &= \tfrac{1}{\sqrt{5}}\,(2 + 8) \\ &= \tfrac{10}{\sqrt{5}}, \end{aligned}$$

and

$$\begin{aligned} \underset{\sim}{q}_2' &= (2, 4) - \tfrac{10}{\sqrt{5}} \cdot \tfrac{1}{\sqrt{5}}\,(1, 2) \\ &= (2, 4) - 2\,(1, 2) \\ &= \underset{\sim}{0}, \end{aligned}$$

so the process stops: $\underset{\sim}{a}_2$ is a multiple of $\underset{\sim}{a}_1$.

Theorem 0.4 (What Gram-Schmidt produces)**.** *Run the Gram-Schmidt process (Definition 0.2) on* $\underset{\sim}{a}_1, \dots, \underset{\sim}{a}_k \in \mathbb{R}^p$.

1. *If it has produced* $\underset{\sim}{q}_1, \dots, \underset{\sim}{q}_i$ *without stopping, then* $\underset{\sim}{q}_1, \dots, \underset{\sim}{q}_i$ *are orthonormal and* $\operatorname{span}\{\underset{\sim}{q}_1, \dots, \underset{\sim}{q}_i\} = \operatorname{span}\{\underset{\sim}{a}_1, \dots, \underset{\sim}{a}_i\}$.
2. *If it reaches step* i *(that is, it did not stop at steps* $1, \dots, i-1$*), it stops there exactly when* $\underset{\sim}{a}_i \in \operatorname{span}\{\underset{\sim}{a}_1, \dots, \underset{\sim}{a}_{i-1}\}$.
3. *It completes all* k *steps exactly when* $\underset{\sim}{a}_1, \dots, \underset{\sim}{a}_k$ *are linearly independent; then* $\underset{\sim}{q}_1, \dots, \underset{\sim}{q}_k$ *is an orthonormal basis of* $\operatorname{span}\{\underset{\sim}{a}_1, \dots, \underset{\sim}{a}_k\}$ *(Definition 0.1).*

Proof

Proof. **Parts 1 and 2, by induction on** i**.** Before step 1, the list of q's is empty, which is orthonormal, and both spans are $\{\underset{\sim}{0}\}$ (Definition 5 in Subspaces and Rank). Suppose part 1 holds after step $i - 1$, and write

$$\begin{aligned}\mathcal{S}_{i-1} &= \operatorname{span}\{\underset{\sim}{a}_1, \dots, \underset{\sim}{a}_{i-1}\} \\ &= \operatorname{span}\{\underset{\sim}{q}_1, \dots, \underset{\sim}{q}_{i-1}\}.\end{aligned}$$

The orthonormal vectors $\underset{\sim}{q}_1, \dots, \underset{\sim}{q}_{i-1}$ are linearly independent (Theorem 0.2) and span $\mathcal{S}_{i-1}$, so they are an orthonormal basis of it.

$\underset{\sim}{q}'_i$ is orthogonal to the earlier $\underset{\sim}{q}$'s. For each $l < i$,

$$\begin{aligned}\underset{\sim}{q}_l \cdot \underset{\sim}{q}'_i &= \underset{\sim}{q}_l \cdot \left(1\,\underset{\sim}{a}_i + \sum_{j=1}^{i-1}\left(-\underset{\sim}{q}_j \cdot \underset{\sim}{a}_i\right)\underset{\sim}{q}_j\right) && \text{(Definition 0.2, as a linear combination)} \\ &= 1\,(\underset{\sim}{q}_l \cdot \underset{\sim}{a}_i) + \sum_{j=1}^{i-1}\left(-\underset{\sim}{q}_j \cdot \underset{\sim}{a}_i\right)(\underset{\sim}{q}_l \cdot \underset{\sim}{q}_j) && \text{(Theorem 0.1)} \\ &= \underset{\sim}{q}_l \cdot \underset{\sim}{a}_i - \sum_{j=1}^{i-1}(\underset{\sim}{q}_j \cdot \underset{\sim}{a}_i)(\underset{\sim}{q}_l \cdot \underset{\sim}{q}_j) && (1\,z = z\text{, and pull the minus signs out of the sum}) \\ &= \underset{\sim}{q}_l \cdot \underset{\sim}{a}_i - \underset{\sim}{q}_l \cdot \underset{\sim}{a}_i && \text{(only the } j = l \text{ term survives, Definition 28 in Vectors)} \\ &= 0. && \text{(arithmetic)}\end{aligned}$$

Part 2 at step i. If $\underset{\sim}{q}'_i = \underset{\sim}{0}$, then $\underset{\sim}{a}_i = \sum_{j<i}(\underset{\sim}{q}_j \cdot \underset{\sim}{a}_i)\,\underset{\sim}{q}_j \in \mathcal{S}_{i-1}$. Conversely, if $\underset{\sim}{a}_i \in \mathcal{S}_{i-1}$, then $\underset{\sim}{a}_i = \sum_{j<i}(\underset{\sim}{q}_j \cdot \underset{\sim}{a}_i)\,\underset{\sim}{q}_j$ (Theorem 0.3), so $\underset{\sim}{q}'_i = \underset{\sim}{0}$.

Part 1 at step i, if the process does not stop. Then $\underset{\sim}{q}'_i \neq \underset{\sim}{0}$, so $\|\underset{\sim}{q}'_i\| > 0$, and $\underset{\sim}{q}_i = \underset{\sim}{q}'_i / \|\underset{\sim}{q}'_i\|$ has norm $\|\underset{\sim}{q}'_i\| / \|\underset{\sim}{q}'_i\| = 1$ (Theorem 1 in Inner Products and the Cauchy-Schwarz Inequality) and for $l < i$

$$\begin{aligned}\underset{\sim}{q}_l \cdot \underset{\sim}{q}_i &= \frac{1}{\|\underset{\sim}{q}'_i\|}\,(\underset{\sim}{q}_l \cdot \underset{\sim}{q}'_i) && \text{(Theorem 1 in Orthogonal Complements and the Fundamental Theorem, special case} \\ &= 0, && (\underset{\sim}{q}'_i \text{ is orthogonal to } \underset{\sim}{q}_l)\end{aligned}$$

and $\underset{\sim}{q}_i \cdot \underset{\sim}{q}_l = 0$ too (Theorem 1 in Vectors). So $\underset{\sim}{q}_1, \dots, \underset{\sim}{q}_i$ are orthonormal. For the spans:

- $\underset{\sim}{a}_i = \|\underset{\sim}{q}'_i\|\,\underset{\sim}{q}_i + \sum_{j<i}(\underset{\sim}{q}_j \cdot \underset{\sim}{a}_i)\,\underset{\sim}{q}_j$ is in $\operatorname{span}\{\underset{\sim}{q}_1, \dots, \underset{\sim}{q}_i\}$, and so is each of $\underset{\sim}{a}_1, \dots, \underset{\sim}{a}_{i-1}$, which lie in $\mathcal{S}_{i-1} = \operatorname{span}\{\underset{\sim}{q}_1, \dots, \underset{\sim}{q}_{i-1}\} \subseteq \operatorname{span}\{\underset{\sim}{q}_1, \dots, \underset{\sim}{q}_i\}$.
- $\underset{\sim}{q}_i = \left(\underset{\sim}{a}_i - \sum_{j<i}(\underset{\sim}{q}_j \cdot \underset{\sim}{a}_i)\,\underset{\sim}{q}_j\right) / \|\underset{\sim}{q}'_i\|$ is in $\operatorname{span}\{\underset{\sim}{a}_1, \dots, \underset{\sim}{a}_i\}$, because each $\underset{\sim}{q}_j$ with $j < i$ lies in $\mathcal{S}_{i-1}$; and so is each of $\underset{\sim}{q}_1, \dots, \underset{\sim}{q}_{i-1}$.

A span is a subspace (Theorem 2 in Subspaces and Rank), so it contains every linear combination of vectors in it; each span therefore contains the other, and they are equal.

Part 3. If the process completes, then $\underset{\sim}{a}_1 \notin$ span of the empty list, and for each $i \geq 2$, $\underset{\sim}{a}_i \notin \operatorname{span}\{\underset{\sim}{a}_1, \dots, \underset{\sim}{a}_{i-1}\}$ (part 2). Applying Theorem 3 in Bases and Dimension k times, starting from the empty list, shows that $\underset{\sim}{a}_1, \dots, \underset{\sim}{a}_k$ are linearly independent. Conversely, if

they are linearly independent, no $\underset{\sim}{a}_i$ is a linear combination $\sum_{j<i} c_j \underset{\sim}{a}_j$ of the earlier ones, because then $\sum_{j<i} c_j \underset{\sim}{a}_j - \underset{\sim}{a}_i = \underset{\sim}{0}$ would be a combination with a nonzero coefficient; so by part 2 the process never stops. When it completes, $\underset{\sim}{q}_1, \dots, \underset{\sim}{q}_k$ are orthonormal, hence linearly independent (Theorem 0.2), and span $\operatorname{span}\{\underset{\sim}{a}_1, \dots, \underset{\sim}{a}_k\}$ (part 1), so they are an orthonormal basis of it. □

Example 0.8 (Finishing an orthonormal basis of $\mathbb{R}^3$)**.** Continue Example 0.6 with $\underset{\sim}{a}_3 = (0, 1, 1)$. $\underset{\sim}{q}_1 \cdot \underset{\sim}{a}_3 = \frac{1}{\sqrt{2}}$ and

$$\begin{aligned} \underset{\sim}{q}_2 \cdot \underset{\sim}{a}_3 &= \tfrac{1}{\sqrt{6}}\,(0 - 1 + 2) \\ &= \tfrac{1}{\sqrt{6}}, \end{aligned}$$

so

$$\begin{aligned} \underset{\sim}{q}_3' &= (0,1,1) - \tfrac{1}{\sqrt{2}} \cdot \tfrac{1}{\sqrt{2}}\,(1,1,0) - \tfrac{1}{\sqrt{6}} \cdot \tfrac{1}{\sqrt{6}}\,(1,-1,2) && \text{(orthogonalize)} \\ &= (0,1,1) - \tfrac{1}{2}\,(1,1,0) - \tfrac{1}{6}\,(1,-1,2) && \text{(multiply the scalars)} \\ &= \left(0 - \tfrac{1}{2} - \tfrac{1}{6},\ 1 - \tfrac{1}{2} + \tfrac{1}{6},\ 1 - 0 - \tfrac{2}{6}\right) && \text{(subtract entrywise)} \\ &= \left(-\tfrac{2}{3}, \tfrac{2}{3}, \tfrac{2}{3}\right), && \text{(arithmetic)} \end{aligned}$$

with norm $\sqrt{3 \cdot \frac{4}{9}} = \frac{2}{\sqrt{3}}$, so $\underset{\sim}{q}_3 = \frac{1}{\sqrt{3}}\,(-1, 1, 1)$. The process did not stop, so by Theorem 0.4 $\underset{\sim}{a}_1, \underset{\sim}{a}_2, \underset{\sim}{a}_3$ are linearly independent and $\underset{\sim}{q}_1, \underset{\sim}{q}_2, \underset{\sim}{q}_3$ is an orthonormal basis of their span. That span therefore has dimension 3 (Definition 4 in Bases and Dimension), so it is $\mathbb{R}^3$ (Theorem 6 in Bases and Dimension, Example 6 in Bases and Dimension). As a check,

$$\begin{aligned} \underset{\sim}{q}_1 \cdot \underset{\sim}{q}_3 &= \tfrac{1}{\sqrt{6}}\,(-1 + 1 + 0) \\ &= 0 \end{aligned}$$

and

$$\begin{aligned} \underset{\sim}{q}_2 \cdot \underset{\sim}{q}_3 &= \tfrac{1}{\sqrt{18}}\,(-1 - 1 + 2) \\ &= 0. \end{aligned}$$

Corollary 0.1 (Every subspace has an orthonormal basis)**.** *Let $\mathcal{S}$ be a subspace of $\mathbb{R}^p$. Every orthonormal list of vectors in $\mathcal{S}$ extends to an orthonormal basis of $\mathcal{S}$ (Definition 0.1). Starting from the empty list shows that $\mathcal{S}$ has an orthonormal basis.*

Proof

Proof. Let $\underset{\sim}{u}_1, \dots, \underset{\sim}{u}_r \in \mathcal{S}$ be orthonormal. They are linearly independent (Theorem 0.2), so they extend to a basis $\underset{\sim}{u}_1, \dots, \underset{\sim}{u}_r, \underset{\sim}{a}_{r+1}, \dots, \underset{\sim}{a}_d$ of $\mathcal{S}$ (Theorem 4 in Bases and Dimension). Run the Gram-Schmidt process on this basis. It is linearly independent, so the process completes, and the output is an orthonormal basis of its span, which is $\mathcal{S}$ (Theorem 0.4, part 3).

The first r outputs are $\underset{\sim}{u}_1, \dots, \underset{\sim}{u}_r$ themselves, by induction on $i \leq r$. Suppose the outputs before step i are $\underset{\sim}{u}_1, \dots, \underset{\sim}{u}_{i-1}$ (for $i = 1$ there are none, and the sum below is empty). Then

$$\begin{aligned} \underset{\sim}{q}_i' &= \underset{\sim}{u}_i - \sum_{j<i} (\underset{\sim}{u}_j \cdot \underset{\sim}{u}_i)\, \underset{\sim}{u}_j \\ &= \underset{\sim}{u}_i, \end{aligned}$$

because each $\underset{\sim}{u}_j \cdot \underset{\sim}{u}_i = 0$; $\underset{\sim}{q}_i' = \underset{\sim}{u}_i \neq \underset{\sim}{0}$, because $\|\underset{\sim}{u}_i\| = 1$, so the process does not stop; and

$$\begin{aligned} \underset{\sim}{q}_i &= \underset{\sim}{u}_i / \|\underset{\sim}{u}_i\| \\ &= \underset{\sim}{u}_i. \end{aligned}$$

So the orthonormal basis contains the starting list. □

Example 0.9 (Extending one unit vector to an orthonormal basis of $\mathbb{R}^3$)**.** Start from the unit vector $\underset{\sim}{u}_1 = \frac{1}{\sqrt{2}}(1,1,0)$. Extend it to a basis of $\mathbb{R}^3$ as in Example 8 in Bases and Dimension, starting from the empty list: $\underset{\sim}{u}_1$ is not in the span of the empty list, $\{\underset{\sim}{0}\}$, because $\|\underset{\sim}{u}_1\| = 1$; $(1,0,0)$ is not in $\operatorname{span}\{\underset{\sim}{u}_1\}$, because every multiple of $\underset{\sim}{u}_1$ has equal first and second entries, and $(0,0,1)$ is not in $\operatorname{span}\{\underset{\sim}{u}_1, (1,0,0)\}$, because every combination of those two vectors has third entry 0. So $\underset{\sim}{u}_1, (1,0,0), (0,0,1)$ are linearly independent (Theorem 3 in Bases and Dimension), and three linearly independent vectors in $\mathbb{R}^3$ are a basis of it (they are a basis of their span, which has dimension 3 (Definition 4 in Bases and Dimension) and so is $\mathbb{R}^3$ (Theorem 6 in Bases and Dimension, Example 6 in Bases and Dimension)). Gram-Schmidt on $\underset{\sim}{u}_1, (1,0,0), (0,0,1)$:

1. $\underset{\sim}{q}_1 = \underset{\sim}{u}_1$.
2. $\underset{\sim}{q}_1 \cdot (1,0,0) = \frac{1}{\sqrt{2}}$, so

$$\begin{aligned} \underset{\sim}{q}_2' &= (1,0,0) - \tfrac{1}{2}(1,1,0) \\ &= (\tfrac{1}{2}, -\tfrac{1}{2}, 0), \end{aligned}$$

with norm $\frac{1}{\sqrt{2}}$, and $\underset{\sim}{q}_2 = \frac{1}{\sqrt{2}}(1,-1,0)$.

3. $(0,0,1)$ is orthogonal to $\underset{\sim}{q}_1$ and $\underset{\sim}{q}_2$, so $\underset{\sim}{q}'_3 = (0,0,1)$, which already has norm 1, and $\underset{\sim}{q}_3 = (0,0,1)$.

The result is the orthonormal basis of Example 0.3, in a different order.

Outer product

Definition 0.1 (Outer product)**.** The **outer product** [4] of a vector $\underset{\sim}{u}$ of length m and a vector $\underset{\sim}{v}$ of length n is the $m \times n$ matrix product (Definition 1 in Matrix Multiplication) $\underset{\sim}{u}\,\underset{\sim}{v}^\top$ of the column vector $\underset{\sim}{u}$ with the row vector $\underset{\sim}{v}^\top$. Its entries are

$$\left(\underset{m\times 1}{\underset{\sim}{u}}\ \underset{1\times n}{\underset{\sim}{v}^\top}\right)_{ij} = u_i v_j \qquad \text{(definition of matrix multiplication; the sum has one term)}$$

Example 0.1 (An outer product)**.** For $\underset{\sim}{u} = (1,2,3)$ and $\underset{\sim}{v} = (4,5)$:

$$\begin{aligned}
\underset{\sim}{u}\,\underset{\sim}{v}^\top &= \begin{bmatrix}1\\2\\3\end{bmatrix}\begin{bmatrix}4 & 5\end{bmatrix} && \text{(definition of the outer product)}\\
&= \begin{bmatrix}1\cdot 4 & 1\cdot 5\\ 2\cdot 4 & 2\cdot 5\\ 3\cdot 4 & 3\cdot 5\end{bmatrix} && \text{(entry } (i,j) \text{ is } u_i v_j)\\
&= \begin{bmatrix}4 & 5\\ 8 & 10\\ 12 & 15\end{bmatrix} && \text{(multiply)}
\end{aligned}$$

ℹ Remark (Outer product and dot product)

Remark 0.1 (Outer product and dot product). The dot product $\underset{\sim}{u}^\top\underset{\sim}{v}$ (Example 3 in Matrix Multiplication) needs two vectors of the same length and gives a number. The outer product $\underset{\sim}{u}\,\underset{\sim}{v}^\top$ takes vectors of any two lengths and gives a matrix [5, Ch. 1], p. 11. For example, $\underset{\sim}{u} = (1,2,3)$ and $\underset{\sim}{v} = (4,5)$ in Example 0.1 have different lengths, so they have no dot product, but their outer product is a 3×2 matrix. For $\underset{\sim}{a} = (1,2)$ and $\underset{\sim}{b} = (3,4)$, which have the same length, both products exist:

$$\begin{aligned} \underset{\sim}{a}^\top \underset{\sim}{b} &= 1 \cdot 3 + 2 \cdot 4 \\ &= 11 \\ \underset{\sim}{a}\,\underset{\sim}{b}^\top &= \begin{bmatrix} 1 \cdot 3 & 1 \cdot 4 \\ 2 \cdot 3 & 2 \cdot 4 \end{bmatrix} \\ &= \begin{bmatrix} 3 & 4 \\ 6 & 8 \end{bmatrix} \end{aligned}$$

Definition 0.2 (Self outer product)**.** The **self outer product** of a vector $\underset{\sim}{x}$ with n entries is the outer product of $\underset{\sim}{x}$ with itself, the $n \times n$ matrix

$$\underset{\sim}{x}\underset{\sim}{x}^\top \stackrel{\text{def}}{=} \underset{\sim}{x}\underset{\sim}{x}^\top$$

Example 0.2 (A self outer product)**.** For $\underset{\sim}{x} = (1, 2)$, the self outer product is

$$\begin{aligned} \underset{\sim}{x}\underset{\sim}{x}^\top &= \begin{bmatrix} 1 \\ 2 \end{bmatrix} \begin{bmatrix} 1 & 2 \end{bmatrix} && \text{(definition of the self outer product)} \\ &= \begin{bmatrix} 1 \cdot 1 & 1 \cdot 2 \\ 2 \cdot 1 & 2 \cdot 2 \end{bmatrix} && \text{(multiply each entry of the column by each entry of the row)} \\ &= \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} && \text{(multiply)} \end{aligned}$$

Remark (Self outer products in statistics)

Remark 0.2 (Self outer products in statistics). Statistics meets the self outer product mostly through sums of them. The Gram matrix $\mathbf{X}^\top\mathbf{X}$ of a data matrix $\mathbf{X}$ is the sum of the self outer products of its rows, as the Gram matrix example shows. For observation vectors $\underset{\sim}{x}_1, \ldots, \underset{\sim}{x}_n$, the sum $\sum_{i=1}^{n} (\underset{\sim}{x}_i - \bar{\underset{\sim}{x}})(\underset{\sim}{x}_i - \bar{\underset{\sim}{x}})^\top$ of the self outer products of their deviations from the vector $\bar{\underset{\sim}{x}}$ of entrywise means is what other sources call the *scatter matrix* [6]. Each term of either sum is one observation's contribution to the total.

Theorem 0.1 (An outer product has rank one)**.** *If $\underset{\sim}{u}$ is a nonzero vector of length m and $\underset{\sim}{v}$ is a nonzero vector of length n, then* $\operatorname{rank}(\underset{\sim}{u}\,\underset{\sim}{v}^\top) = 1$ *(Definition 2 in Subspaces and Rank).*

Proof

Proof. Column j of $\underset{\sim}{u}\,\underset{\sim}{v}^\top$ has entries $u_1 v_j, \ldots, u_m v_j$, so it is the vector $v_j \underset{\sim}{u}$.
The rank is at least one. Because $\underset{\sim}{v} \neq \underset{\sim}{0}$, some entry v_j is nonzero. Then column j, $v_j \underset{\sim}{u}$, is a nonzero vector, and a single nonzero vector is linearly independent (Definition 1 in Subspaces and Rank): $c\,(v_j \underset{\sim}{u}) = \underset{\sim}{0}$ forces $c = 0$.
The rank is at most one. Take any two columns $j \neq k$. If

$$\begin{aligned} v_j &= v_k \\ &= 0, \end{aligned}$$

both columns are $\underset{\sim}{0}$, and $1 \cdot (v_j \underset{\sim}{u}) + 1 \cdot (v_k \underset{\sim}{u}) = \underset{\sim}{0}$. Otherwise, use the coefficients $c_j = v_k$ and $c_k = -v_j$, which are not both zero:

$$\begin{aligned} v_k\,(v_j \underset{\sim}{u}) - v_j\,(v_k \underset{\sim}{u}) &= (v_k v_j - v_j v_k)\,\underset{\sim}{u} && \text{(collect the scalar multiples of } \underset{\sim}{u}) \\ &= 0 \cdot \underset{\sim}{u} && \text{(multiplication of numbers is commutative)} \\ &= \underset{\sim}{0} && \text{(a zero multiple of a vector is } \underset{\sim}{0}) \end{aligned}$$

Either way, every pair of columns has a combination with coefficients that are not all zero and that equals $\underset{\sim}{0}$. Now take any set of two or more columns. It contains a pair $j \neq k$. Use that pair's coefficients for columns j and k, and the coefficient 0 for every other column in the set. This combination equals $\underset{\sim}{0}$, and its coefficients are not all zero, so the set is not linearly independent (Definition 1 in Subspaces and Rank). The largest linearly independent set of columns therefore has one column. □

Example 0.3 (The rank of an outer product)**.** In Example 0.1, the second column $(5, 10, 15)$ of $\underset{\sim}{u}\,\underset{\sim}{v}^\top$ is $\frac{5}{4}$ times the first column $(4, 8, 12)$, because the columns are $v_1 \underset{\sim}{u} = 4\underset{\sim}{u}$ and $v_2 \underset{\sim}{u} = 5\underset{\sim}{u}$. So the two columns are not linearly independent, and the 3×2 matrix has rank 1.

Corollary 0.1 (A self outer product has rank one)**.** *For a nonzero vector* $\underset{\sim}{x}$*, the self outer product* $\underset{\sim}{x}\underset{\sim}{x}^\top$ *has rank* 1*:*

$$\operatorname{rank}(\underset{\sim}{x}\underset{\sim}{x}^\top) = 1$$

Proof

Proof. The self outer product $\underset{\sim}{x}\underset{\sim}{x}^\top$ is the outer product $\underset{\sim}{x}\underset{\sim}{x}^\top$. Both of its factors are the nonzero vector $\underset{\sim}{x}$, so Theorem 0.1, with $\underset{\sim}{x}$ in place of both $\underset{\sim}{u}$ and $\underset{\sim}{v}$, gives rank 1. □

Example 0.4 (The rank of a self outer product)**.** For $\underset{\sim}{x} = (1, 2)$, the example of a self outer product found the matrix $\underset{\sim}{\underset{\sim}{x}}\underset{\sim}{x}^\top = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$. Its second column $(2, 4)$ is 2 times its first column $(1, 2)$, so the two columns are not linearly independent, and the matrix has rank 1, as Corollary 0.1 says.

Theorem 0.2 (The transpose of an outer product)**.** *For a vector $\underset{\sim}{u}$ with m entries and a vector $\underset{\sim}{v}$ with n entries, the transpose of the outer product $\underset{\sim}{\underset{\sim}{u}}\underset{\sim}{v}^\top$ equals the outer product $\underset{\sim}{\underset{\sim}{v}}\underset{\sim}{u}^\top$:*

$$\underbrace{\left(\underset{\sim}{\underset{\sim}{u}}\underset{\sim}{v}^\top\right)^\top}_{n \times m} = \underset{n \times m}{\underset{\sim}{\underset{\sim}{v}}\underset{\sim}{u}^\top}$$

Proof

Proof. Entry (i, j) of each side, for i from 1 to n and j from 1 to m:

$$\begin{aligned} \left[\left(\underset{\sim}{\underset{\sim}{u}}\underset{\sim}{v}^\top\right)^\top\right]_{ij} &= \left(\underset{\sim}{\underset{\sim}{u}}\underset{\sim}{v}^\top\right)_{ji} && \text{(definition of the transpose)} \\ &= u_j v_i && \text{(entries of an outer product)} \\ &= v_i u_j && \text{(multiplication of numbers is commutative)} \\ &= \left(\underset{\sim}{\underset{\sim}{v}}\underset{\sim}{u}^\top\right)_{ij} && \text{(entries of an outer product)} \end{aligned}$$

□

Example 0.5 (Transposing an outer product)**.** For $\underset{\sim}{u} = (1, 2, 3)$ and $\underset{\sim}{v} = (4, 5)$, the example of an outer product found the 3×2 matrix $\underset{\sim}{\underset{\sim}{u}}\underset{\sim}{v}^\top$. Swapping the two vectors gives its transpose:

$$\begin{aligned} \underset{\sim}{\underset{\sim}{v}}\underset{\sim}{u}^\top &= \begin{bmatrix} 4 \\ 5 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \end{bmatrix} && \text{(definition of the outer product)} \\ &= \begin{bmatrix} 4 \cdot 1 & 4 \cdot 2 & 4 \cdot 3 \\ 5 \cdot 1 & 5 \cdot 2 & 5 \cdot 3 \end{bmatrix} && \text{(entry } (i, j) \text{ is } v_i u_j) \\ &= \begin{bmatrix} 4 & 8 & 12 \\ 5 & 10 & 15 \end{bmatrix} && \text{(multiply)} \end{aligned}$$

Each row of this matrix is a column of $\underset{\sim}{\underset{\sim}{u}}\underset{\sim}{v}^\top$, as Theorem 0.2 says.

Theorem 0.3 (An outer product times a vector)**.** *For a vector $\underset{\sim}{u}$ with m entries and two vectors $\underset{\sim}{v}$ and $\underset{\sim}{w}$ with n entries, the product of the outer product $\underset{\sim}{\underset{\sim}{u}}\underset{\sim}{v}^\top$ with $\underset{\sim}{w}$ equals the*

scalar multiple of $\underset{\sim}{u}$ *by the dot product* $\underset{\sim}{v}^\top \underset{\sim}{w}$*:*

$$\underbrace{(\underset{\sim}{u}\underset{\sim}{v}^\top)\underset{\sim}{w}}_{m\times 1} = (\underset{\sim}{v}^\top \underset{\sim}{w})\ \underset{m\times 1}{\underset{\sim}{u}}$$

Proof

Proof. Entry i of each side, for i from 1 to m:

$$\begin{aligned}
[(\underset{\sim}{u}\underset{\sim}{v}^\top)\underset{\sim}{w}]_i &= \sum_{j=1}^{n} (\underset{\sim}{u}\underset{\sim}{v}^\top)_{ij} w_j && \text{(definition of matrix-vector multiplication)} \\
&= \sum_{j=1}^{n} u_i v_j w_j && \text{(entries of an outer product)} \\
&= u_i \sum_{j=1}^{n} v_j w_j && \text{(factor } u_i \text{ out of the sum)} \\
&= u_i (\underset{\sim}{v}^\top \underset{\sim}{w}) && \text{(definition of the dot product)} \\
&= (\underset{\sim}{v}^\top \underset{\sim}{w}) u_i && \text{(multiplication of numbers is commutative)} \\
&= [(\underset{\sim}{v}^\top \underset{\sim}{w})\underset{\sim}{u}]_i && \text{(definition of scalar multiplication)}
\end{aligned}$$

□

Example 0.6 (Multiplying an outer product by a vector)**.** Take $\underset{\sim}{u} = (1, 2, 3)$ and $\underset{\sim}{v} = (4, 5)$ from the example of an outer product, and $\underset{\sim}{w} = (1, -1)$. Multiplying the outer product by $\underset{\sim}{w}$ directly gives

$$\begin{aligned}
(\underset{\sim}{u}\underset{\sim}{v}^\top)\underset{\sim}{w} &= \begin{bmatrix} 4 & 5 \\ 8 & 10 \\ 12 & 15 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} && \text{(the outer product found in that example)} \\
&= \begin{bmatrix} 4-5 \\ 8-10 \\ 12-15 \end{bmatrix} && \text{(definition of matrix-vector multiplication)} \\
&= \begin{bmatrix} -1 \\ -2 \\ -3 \end{bmatrix} && \text{(subtract)}
\end{aligned}$$

The dot product of $\underset{\sim}{v}$ and $\underset{\sim}{w}$ is

$$\begin{aligned}
\underset{\sim}{v}^\top \underset{\sim}{w} &= 4 \cdot 1 + 5 \cdot (-1) && \text{(definition of the dot product)} \\
&= -1 && \text{(multiply and add)}
\end{aligned}$$

and $(-1)\,\underset{\sim}{u} = (-1, -2, -3)$ is the same vector, as Theorem 0.3 says.

Kronecker product

Definition 0.1 (Kronecker product)**.** The **Kronecker product** of an $m \times n$ matrix $\mathbf{A}$ and a $p \times q$ matrix $\mathbf{B}$ is the $mp \times nq$ block matrix whose block in row i and column j is $a_{ij}\mathbf{B}$, where a_{ij} is the (i, j) entry of $\mathbf{A}$:

$$\mathbf{A} \otimes \mathbf{B} = \begin{bmatrix} a_{11}\mathbf{B} & \cdots & a_{1n}\mathbf{B} \\ \vdots & \ddots & \vdots \\ a_{m1}\mathbf{B} & \cdots & a_{mn}\mathbf{B} \end{bmatrix}.$$

Example 0.1 (A Kronecker product)**.** For $\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ and $\mathbf{B} = \begin{bmatrix} 0 & 1 \end{bmatrix}$:

$$\begin{aligned} \mathbf{A} \otimes \mathbf{B} &= \begin{bmatrix} 1 \cdot \mathbf{B} & 2 \cdot \mathbf{B} \\ 3 \cdot \mathbf{B} & 4 \cdot \mathbf{B} \end{bmatrix} && \text{(definition of the Kronecker product)} \\ &= \begin{bmatrix} 0 & 1 & 0 & 2 \\ 0 & 3 & 0 & 4 \end{bmatrix} && \text{(multiply each entry of } \mathbf{A} \text{ by } \mathbf{B}\text{)} \end{aligned}$$

The result is 2×4: $\mathbf{A}$ is 2×2 and $\mathbf{B}$ is 1×2, so $\mathbf{A} \otimes \mathbf{B}$ is $(2 \cdot 1) \times (2 \cdot 2)$.

Theorem 0.1 (An outer product is a Kronecker product)**.** *For a vector $\underset{\sim}{u}$ with m entries and a vector $\underset{\sim}{v}$ with n entries, the Kronecker product of the column vector $\underset{\sim}{u}$ with the transpose $\underset{\sim}{v}^\top$ of $\underset{\sim}{v}$ equals the outer product $\underset{\sim\sim}{uv}^\top$:*

$$\underset{m\times 1}{\underset{\sim}{u}} \otimes \underset{1\times n}{\underset{\sim}{v}^\top} = \underset{m\times n}{\underset{\sim\sim}{uv}^\top}$$

Proof

Proof. The Kronecker product $\underset{\sim}{u} \otimes \underset{\sim}{v}^\top$ stacks $u_1\underset{\sim}{v}^\top, \ldots, u_m\underset{\sim}{v}^\top$, so its ith row is $u_i\underset{\sim}{v}^\top$. So entry (i, j) of each side, for i from 1 to m and j from 1 to n, is

$$\begin{aligned} (\underset{\sim}{u} \otimes \underset{\sim}{v}^\top)_{ij} &= (u_i\underset{\sim}{v}^\top)_j && \text{(definition of the Kronecker product)} \\ &= u_i v_j && \text{(entry } j \text{ of the scalar multiple } u_i\underset{\sim}{v}^\top\text{)} \\ &= (\underset{\sim\sim}{uv}^\top)_{ij} && \text{(entries of an outer product)} \end{aligned}$$

□

Example 0.2 (An outer product as a Kronecker product)**.** For $\underset{\sim}{u} = (1, 2, 3)$ and $\underset{\sim}{v} = (4, 5)$, the Kronecker product of $\underset{\sim}{u}$ with $\underset{\sim}{v}^\top$ is

$$\underset{\sim}{u} \otimes \underset{\sim}{v}^\top = \begin{bmatrix} 1 \cdot \begin{bmatrix} 4 & 5 \end{bmatrix} \\ 2 \cdot \begin{bmatrix} 4 & 5 \end{bmatrix} \\ 3 \cdot \begin{bmatrix} 4 & 5 \end{bmatrix} \end{bmatrix} \quad \text{(definition of the Kronecker product)}$$

$$= \begin{bmatrix} 4 & 5 \\ 8 & 10 \\ 12 & 15 \end{bmatrix} \quad \text{(multiply)}$$

This matrix is the outer product $\underset{\sim}{u}\underset{\sim}{v}^\top$ that the example of an outer product found, as Theorem 0.1 says.

Definition 0.2 (Vectorization)**.** The **vectorization** [7] of an $m \times n$ matrix $\mathbf{A}$ is the column vector with mn entries that stacks the columns $\mathbf{A}_{:,1}, \ldots, \mathbf{A}_{:,n}$ of $\mathbf{A}$ in order:

$$\operatorname{vec}(\mathbf{A}) \stackrel{\text{def}}{=} \begin{bmatrix} \mathbf{A}_{:,1} \\ \vdots \\ \mathbf{A}_{:,n} \end{bmatrix}$$

Example 0.3 (Vectorizing a 2×2 matrix)**.** For $\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$, the columns are $(1, 3)$ and $(2, 4)$, so the vectorization stacks them:

$$\operatorname{vec}(\mathbf{A}) = \begin{bmatrix} 1 \\ 3 \\ 2 \\ 4 \end{bmatrix}$$

Theorem 0.2 (The vectorization of an outer product is a Kronecker product)**.** *For a vector $\underset{\sim}{u}$ with m entries and a vector $\underset{\sim}{v}$ with n entries, the vectorization of the outer product $\underset{\sim}{u}\underset{\sim}{v}^\top$ equals the Kronecker product $\underset{\sim}{v} \otimes \underset{\sim}{u}$:*

$$\underbrace{\operatorname{vec}(\underset{\sim}{u}\underset{\sim}{v}^\top)}_{mn \times 1} = \underset{n \times 1}{\underset{\sim}{v}} \otimes \underset{m \times 1}{\underset{\sim}{u}}$$

Proof

Proof. The jth column of $\underset{\sim}{u}\underset{\sim}{v}^\top$ has entries $u_1 v_j, \ldots, u_m v_j$, so it is the vector $v_j \underset{\sim}{u}$. Then

$$\operatorname{vec}(\underset{\sim}{u}\underset{\sim}{v}^{\top}) = \begin{bmatrix} v_1 \underset{\sim}{u} \\ \vdots \\ v_n \underset{\sim}{u} \end{bmatrix} \quad \text{(definition of vectorization)}$$

$$= \underset{\sim}{v} \otimes \underset{\sim}{u} \quad \text{(definition of the Kronecker product, with one column of blocks)}$$

□

Example 0.4 (Vectorizing an outer product)**.** For $\underset{\sim}{u} = (1, 2, 3)$ and $\underset{\sim}{v} = (4, 5)$, the example of an outer product found the outer product $\underset{\sim}{u}\underset{\sim}{v}^{\top}$, whose columns are $(4, 8, 12)$ and $(5, 10, 15)$. Its vectorization and the Kronecker product $\underset{\sim}{v} \otimes \underset{\sim}{u}$ agree:

$$\operatorname{vec}(\underset{\sim}{u}\underset{\sim}{v}^{\top}) = \begin{bmatrix} 4 \\ 8 \\ 12 \\ 5 \\ 10 \\ 15 \end{bmatrix} \quad \text{(stack the columns)}$$

$$= \begin{bmatrix} 4 \cdot \underset{\sim}{u} \\ 5 \cdot \underset{\sim}{u} \end{bmatrix} \quad \text{(each column is a multiple of } \underset{\sim}{u}\text{)}$$

$$= \underset{\sim}{v} \otimes \underset{\sim}{u} \quad \text{(definition of the Kronecker product)}$$

as Theorem 0.2 says.

Hilbert spaces

Definition 0.1 (Cauchy sequence)**.** A sequence $(\underset{\sim}{x}_n)$ of vectors in a real inner product space V is a **Cauchy sequence** [8] if for every $\varepsilon > 0$ there is a natural number N such that

$$\|\underset{\sim}{x}_m - \underset{\sim}{x}_n\| < \varepsilon \quad \text{for all } m, n \geq N,$$

where $\|\underset{\sim}{x}\| = \sqrt{\langle \underset{\sim}{x}, \underset{\sim}{x} \rangle}$ for every $\underset{\sim}{x} \in V$.

Example 0.1 (A Cauchy sequence of real numbers)**.** The real numbers $\mathbb{R}$ with the inner product $\langle x, y \rangle = x\,y$ form an inner product space, and $\|x\| = |x|$ in it. The sequence $x_n = \frac{1}{n}$ is a Cauchy sequence. Given $\varepsilon > 0$, take N to be any natural number larger than $\frac{1}{\varepsilon}$. For all $m, n \geq N$, both $\frac{1}{m}$ and $\frac{1}{n}$ lie in $(0, \frac{1}{N}]$, so $\left|\frac{1}{m} - \frac{1}{n}\right| < \frac{1}{N} < \varepsilon$. For example, with

$\varepsilon = 0.01$, take $N = 101$: all $m, n \geq 101$ have $\left|\frac{1}{m} - \frac{1}{n}\right| < \frac{1}{101} < 0.01$.

Definition 0.2 (Complete inner product space)**.** A real inner product space V is **complete** [9] if for every Cauchy sequence $(\underset{\sim}{x}_n)$ in V there is a vector $\underset{\sim}{y} \in V$ such that for every $\varepsilon > 0$ there is a natural number N with

$$\|\underset{\sim}{x}_n - \underset{\sim}{y}\| < \varepsilon \quad \text{for every } n \geq N.$$

Example 0.2 (Polynomials are not complete)**.** Let V be the set of polynomials on $[0, 1]$, with the inner product $\langle f, g\rangle = \int_0^1 f(t)\, g(t)\, dt$. The partial sums $p_n(t) = \sum_{k=0}^{n} \frac{t^k}{k!}$ are a Cauchy sequence in V, because they converge uniformly on $[0, 1]$ to e^t. The only candidate for a limit is e^t, which is not a polynomial, so no polynomial $\underset{\sim}{y} \in V$ satisfies the condition of Definition 0.2, and V is not complete.

Definition 0.3 (Hilbert space)**.** A **Hilbert space** [10] is a complete real inner product space [5, Sec. 16.8], Definition 16.8, p. 547.

Example 0.3 (Real coordinate space is a Hilbert space)**.** The Euclidean space $\mathbb{R}^2$ with the dot product (Example 2 in Inner Products and the Cauchy-Schwarz Inequality) is a Hilbert space. Let $(\underset{\sim}{x}_n)$ with $\underset{\sim}{x}_n = (x_{n,1}, x_{n,2})$ be a Cauchy sequence in it. Each coordinate sequence is a Cauchy sequence of real numbers, because $|x_{m,j} - x_{n,j}| \leq \|\underset{\sim}{x}_m - \underset{\sim}{x}_n\|$ for $j = 1, 2$. Each coordinate sequence then converges to a real number y_j, because every Cauchy sequence of real numbers converges [9], and $y = (y_1, y_2)$ satisfies the condition of Definition 0.2. For example, $\underset{\sim}{x}_n = (\frac{1}{n}, 1 + \frac{1}{n})$ is a Cauchy sequence with limit $\underset{\sim}{y} = (0, 1)$.

[1] H. Zhou, "Vectors." Lecture notes for Biostat 216, Mathematical Methods for Biostatistics, University of California, Los Angeles, 2024. Accessed: Oct. 05, 2026. [Online]. Available: https://ucla-biostat-216.github.io/2024fall/slides/02-vector/02-vector.html

[2] Wikipedia contributors, "Orthonormal basis — Wikipedia, the free encyclopedia." 2026. Available: https://en.wikipedia.org/wiki/Orthonormal_basis

[3] Wikipedia contributors, "Gram–schmidt process — Wikipedia, the free encyclopedia." 2026. Available: https://en.wikipedia.org/wiki/Gram%E2%80%93Schmidt_process

[4] Wikipedia contributors, "Outer product — Wikipedia, the free encyclopedia." 2026. Available: https://en.wikipedia.org/wiki/Outer_product

[5] S. Banerjee and A. Roy, *Linear algebra and matrix analysis for statistics*, vol. 181. Crc Press Boca Raton, 2014. Available: https://www.routledge.com/Linear-Algebra-and-Matrix-Analysis-for-Statistics/Banerjee-Roy/p/book/9781420095388

[6] Wikipedia contributors, "Scatter matrix — Wikipedia, the free encyclopedia." 2026. Available: https://en.wikipedia.org/wiki/Scatter_matrix

[7] Wikipedia contributors, "Vectorization (mathematics) — Wikipedia, the free encyclopedia." 2026. Available: https://en.wikipedia.org/wiki/Vectorization_(mathematics)

[8] Wikipedia contributors, "Cauchy sequence — Wikipedia, the free encyclopedia." 2026. Available: https://en.wikipedia.org/wiki/Cauchy_sequence

[9] Wikipedia contributors, "Complete metric space — Wikipedia, the free encyclopedia." 2026. Available: https://en.wikipedia.org/wiki/Complete_metric_space

[10] Wikipedia contributors, "Hilbert space — Wikipedia, the free encyclopedia." 2026. Available: https://en.wikipedia.org/wiki/Hilbert_space