LECTURE 17

Orthogonal decomposition

Splitting a vector into perpendicular components

Dr Ilia Pirashvili
Department of Mathematics, University of Galway
3 November 2026 · Winter 2026

Where we got to

LAST LECTURE

- $\mathbf{u} \cdot \mathbf{v} = \sum u_i v_i = \mathbf{u}^{\mathsf{T}}\mathbf{v}$, with $\|\mathbf{v}\| = \sqrt{\mathbf{v} \cdot \mathbf{v}}$ and $\mathbf{u} \cdot \mathbf{v} = \|\mathbf{u}\|\|\mathbf{v}\| \cos\theta$.
- $\mathbf{u} \perp \mathbf{v}$ means $\mathbf{u} \cdot \mathbf{v} = 0$; then Pythagoras holds.
- Every $\mathbf{v}$ splits as $\mathbf{p} + \mathbf{w}$ with $\mathbf{p}$ along $\mathbf{u}$ and $\mathbf{w} \perp \mathbf{u}$.
- $H^{\perp}$ is a subspace, and checking a spanning set of H is enough.

That splitting was for a single $\mathbf{u}$. Today we do it for a whole subspace, and find that it is the key to approximation.

Warm-up

WARM-UP

Let $\mathbf{u} = \begin{bmatrix} 3 \\ 0 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} 2 \\ 5 \end{bmatrix}$. Write $\mathbf{v} = \mathbf{p} + \mathbf{w}$ ▸ with $\mathbf{p}$ a multiple of $\mathbf{u}$ and $\mathbf{w} \perp \mathbf{u}$. Which of $\mathbf{p}, \mathbf{w}$ is longer?

Warm-up

WARM-UP

Let $\mathbf{u} = \begin{bmatrix} 3 \\ 0 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} 2 \\ 5 \end{bmatrix}$. Write $\mathbf{v} = \mathbf{p} + \mathbf{w}$ ▸ with $\mathbf{p}$ a multiple of $\mathbf{u}$ and $\mathbf{w} \perp \mathbf{u}$. Which of $\mathbf{p}, \mathbf{w}$ is longer?

SOLUTION

$\mathbf{v} \cdot \mathbf{u} = 6$ and $\mathbf{u} \cdot \mathbf{u} = 9$, so $\mathbf{p} = \frac{6}{9} \begin{bmatrix} 3 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$ and $\mathbf{w} = \begin{bmatrix} 0 \\ 5 \end{bmatrix}$. Indeed $\mathbf{w} \cdot \mathbf{u} = 0$.

$\|\mathbf{p}\| = 2$ and $\|\mathbf{w}\| = 5$, and Pythagoras checks out: $4 + 25 = 29 = \|\mathbf{v}\|^2$.

SECTION 1

Orthogonal sets

Orthogonal and orthonormal sets

DEFINITION 17.1 ▸

A set $\{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ of non-zero vectors is called **orthogonal** if $\mathbf{u}_i \cdot \mathbf{u}_j = 0$ whenever $i \neq j$. It is **orthonormal** if in addition $\|\mathbf{u}_i\| = 1$ for each i.

Orthogonal and orthonormal sets

DEFINITION 17.1 ▸

A set $\{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ of non-zero vectors is called **orthogonal** if $\mathbf{u}_i \cdot \mathbf{u}_j = 0$ whenever $i \neq j$. It is **orthonormal** if in addition $\|\mathbf{u}_i\| = 1$ for each i.

EXAMPLE 17.2 ▸

$\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ is orthonormal. $\left\{\begin{bmatrix}1\\1\end{bmatrix}, \begin{bmatrix}1\\-1\end{bmatrix}\right\}$ is orthogonal but not orthonormal; dividing each by $\sqrt{2}$ makes it an orthonormal basis.

Orthogonal and orthonormal sets

DEFINITION 17.1 ▸

A set $\{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ of non-zero vectors is called **orthogonal** if $\mathbf{u}_i \cdot \mathbf{u}_j = 0$ whenever $i \neq j$. It is **orthonormal** if in addition $\|\mathbf{u}_i\| = 1$ for each i.

EXAMPLE 17.2 ▸

$\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ is orthonormal. $\left\{\begin{bmatrix}1\\1\end{bmatrix}, \begin{bmatrix}1\\-1\end{bmatrix}\right\}$ is orthogonal but not orthonormal; dividing each by $\sqrt{2}$ makes it an orthonormal basis.

Every orthogonal set can be made orthonormal by normalising each vector, which changes no direction and no angle. It is also something done to individual vectors one at a time, not to the collection, making it fairly simple.

Orthogonal sets are independent

THEOREM 17.3

An orthogonal set of non-zero vectors is linearly independent.

PROOF.

Suppose $c_1\mathbf{u}_1 + \cdots + c_k\mathbf{u}_k = \mathbf{0}$. Take the inner product of both sides with $\mathbf{u}_j$. Every term with $i \neq j$ vanishes, leaving

$$c_j\,\mathbf{u}_j \cdot \mathbf{u}_j \;=\; c_j\|\mathbf{u}_j\|^2 \;=\; 0.$$

Since $\mathbf{u}_j \neq \mathbf{0}$ we get $c_j = 0$; and j was arbitrary. □

Orthogonal sets are independent

THEOREM 17.3

An orthogonal set of non-zero vectors is linearly independent.

PROOF.

Suppose $c_1\mathbf{u}_1 + \cdots + c_k\mathbf{u}_k = \mathbf{0}$. Take the inner product of both sides with $\mathbf{u}_j$. Every term with $i \neq j$ vanishes, leaving

$$c_j\,\mathbf{u}_j \cdot \mathbf{u}_j \;=\; c_j\|\mathbf{u}_j\|^2 \;=\; 0.$$

Since $\mathbf{u}_j \neq \mathbf{0}$ we get $c_j = 0$; and j was arbitrary. □

By the Basis Theorem, n non-zero orthogonal vectors in $\mathbb{R}^n$ are automatically a basis, called **orthogonal basis**.

Why an orthogonal basis is worth having

With an ordinary basis, finding coordinates means solving a system. With an orthogonal one, the coefficients are projects, and thus, we have a formula:

THEOREM 17.4

If $\{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ *is an orthogonal basis of* V*, then every* $\mathbf{v} \in V$ *is*

$$\mathbf{v} = \sum_{i=1}^{n} c_i \mathbf{u}_i \qquad \textit{with} \qquad c_i = \frac{\mathbf{v} \cdot \mathbf{u}_i}{\mathbf{u}_i \cdot \mathbf{u}_i}.$$

PROOF.

Write $\mathbf{v} = \sum c_i \mathbf{u}_i$ and take the inner product with $\mathbf{u}_j$. All cross terms vanish, leaving $\mathbf{v} \cdot \mathbf{u}_j = c_j \mathbf{u}_j \cdot \mathbf{u}_j$. □

One inner product per coefficient, instead of one row reduction for the lot.

SECTION 2

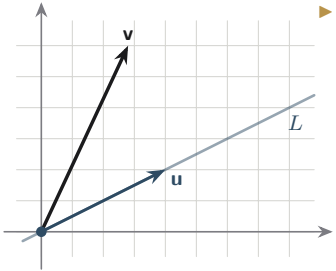
Projection onto a subspace

Projection onto a line, in detail

For a line $L = \text{Span}\{\mathbf{u}\}$, the projection of $\mathbf{v}$ is

$$\text{Proj}_L(\mathbf{v}) = \frac{\mathbf{v} \cdot \mathbf{u}}{\mathbf{u} \cdot \mathbf{u}} \mathbf{u},$$

the single point of L sitting at the foot of the perpendicular from $\mathbf{v}$.

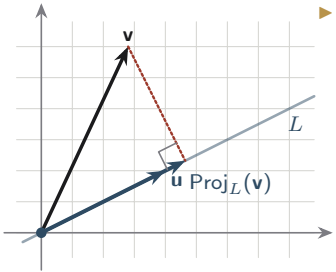


Projection onto a line, in detail

For a line $L = \text{Span}\{\mathbf{u}\}$, the projection of $\mathbf{v}$ is

$$\text{Proj}_L(\mathbf{v}) = \frac{\mathbf{v} \cdot \mathbf{u}}{\mathbf{u} \cdot \mathbf{u}} \mathbf{u},$$

the single point of L sitting at the foot of the perpendicular from $\mathbf{v}$.



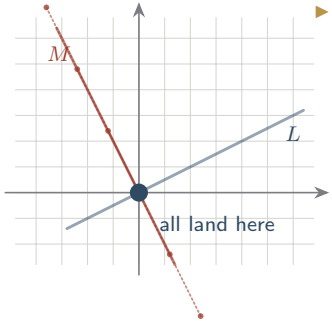
Note that Proj_L is itself a linear map $\mathbb{R}^n \to \mathbb{R}^n$ whose image is L. So it collapses n dimensions down to one, and its kernel is everything perpendicular to $\mathbf{u}$, an $n-1$ dimensional space.

One line projected onto another

Take $L = \text{Span}\{\begin{bmatrix}2\\1\end{bmatrix}\}$ and $M = \text{Span}\{\begin{bmatrix}-1\\2\end{bmatrix}\}$, which are perpendicular to each other.

Every point of M is $t\begin{bmatrix}-1\\2\end{bmatrix}$, and

$$\text{Proj}_L\left(t\begin{bmatrix}-1\\2\end{bmatrix}\right) = \frac{\mathbf{v}\cdot\mathbf{u}}{\mathbf{u}\cdot\mathbf{u}}\mathbf{u}$$



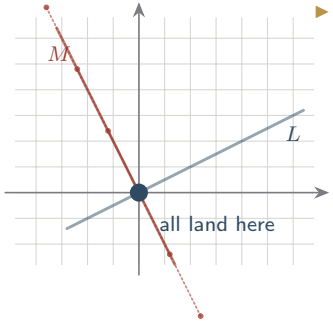
where $\mathbf{v} = t\begin{bmatrix}-1\\2\end{bmatrix}$ and $\mathbf{u} = \begin{bmatrix}2\\1\end{bmatrix}$.

One line projected onto another

Take $L = \text{Span}\{\begin{bmatrix}2\\1\end{bmatrix}\}$ and $M = \text{Span}\{\begin{bmatrix}-1\\2\end{bmatrix}\}$, which are perpendicular to each other.

Every point of M is $t\begin{bmatrix}-1\\2\end{bmatrix}$, and

$$\text{Proj}_L\left(t\begin{bmatrix}-1\\2\end{bmatrix}\right) = \frac{t(-2+2)}{5}\begin{bmatrix}2\\1\end{bmatrix}$$



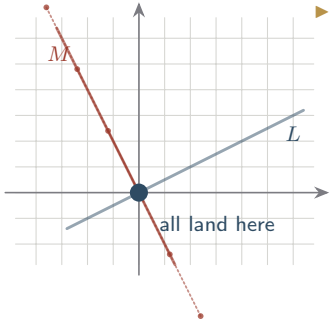
as $\mathbf{v}\cdot\mathbf{u} = \begin{bmatrix}-t\\2t\end{bmatrix}\cdot\begin{bmatrix}2\\1\end{bmatrix} = -2t+2t$ and $\mathbf{u}\cdot\mathbf{u} = \begin{bmatrix}2\\1\end{bmatrix}\cdot\begin{bmatrix}2\\1\end{bmatrix} = 4+1$.

One line projected onto another

Take $L = \text{Span}\{\begin{bmatrix}2\\1\end{bmatrix}\}$ and $M = \text{Span}\{\begin{bmatrix}-1\\2\end{bmatrix}\}$, which are perpendicular to each other.

Every point of M is $t\begin{bmatrix}-1\\2\end{bmatrix}$, and

$$\text{Proj}_L\left(t\begin{bmatrix}-1\\2\end{bmatrix}\right) = \frac{t(-2+2)}{5}\begin{bmatrix}2\\1\end{bmatrix} = \mathbf{0}.$$



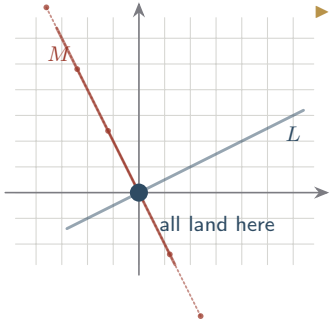
One line projected onto another

Take $L = \text{Span}\{\begin{bmatrix}2\\1\end{bmatrix}\}$ and $M = \text{Span}\{\begin{bmatrix}-1\\2\end{bmatrix}\}$, which are perpendicular to each other.

Every point of M is $t\begin{bmatrix}-1\\2\end{bmatrix}$, and

$$\text{Proj}_L\left(t\begin{bmatrix}-1\\2\end{bmatrix}\right) = \frac{t(-2+2)}{5}\begin{bmatrix}2\\1\end{bmatrix} = \mathbf{0}.$$

The whole of M lands on the single point $\mathbf{0}$.



For any line not perpendicular to L, the image under the projection is all of L.

LEMMA 17.5

The only vector perpendicular to itself is zero. Formally, $\mathbf{v} \in W \cap W^\perp \implies \mathbf{v} = \mathbf{0}$.

PROOF. A vector in both is orthogonal to itself, so $\mathbf{v} \cdot \mathbf{v} = \|\mathbf{v}\|^2 = 0$.

Projection onto a subspace

DEFINITION 17.6 ▸

Let $W \subseteq \mathbb{R}^n$ have orthogonal basis $\{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$. The **orthogonal projection** of $\mathbf{v}$ onto W is

$$\text{Proj}_W(\mathbf{v}) = \sum_{i=1}^{k} \frac{\mathbf{v} \cdot \mathbf{u}_i}{\mathbf{u}_i \cdot \mathbf{u}_i} \mathbf{u}_i.$$

Each term is the projection onto one basis vector (line), and then we simply add them.

Projection onto a subspace

DEFINITION 17.6 ▸

Let $W \subseteq \mathbb{R}^n$ have orthogonal basis $\{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$. The **orthogonal projection** of $\mathbf{v}$ onto W is

$$\text{Proj}_W(\mathbf{v}) = \sum_{i=1}^{k} \frac{\mathbf{v} \cdot \mathbf{u}_i}{\mathbf{u}_i \cdot \mathbf{u}_i} \mathbf{u}_i.$$

Each term is the projection onto one basis vector (line), and then we simply add them.

CAUTION

The basis must be orthogonal for this formula. With a general basis the terms interfere and the answer is wrong, see the next section (Gram–Schmidt).

The Orthogonal Decomposition Theorem

THEOREM 17.7

Let $W \subseteq \mathbb{R}^n$ be a subspace. Every $\mathbf{v} \in \mathbb{R}^n$ can be written in exactly one way as

$$\mathbf{v} = \hat{\mathbf{v}} + \mathbf{z}, \qquad \text{with} \quad \hat{\mathbf{v}} \in W, \quad \mathbf{z} \in W^{\perp},$$

and $\hat{\mathbf{v}} = \text{Proj}_W(\mathbf{v})$.

The Orthogonal Decomposition Theorem

THEOREM 17.7

Let $W \subseteq \mathbb{R}^n$ be a subspace. Every $\mathbf{v} \in \mathbb{R}^n$ can be written in exactly one way as

$$\mathbf{v} \;=\; \hat{\mathbf{v}} + \mathbf{z}, \qquad \textit{with} \quad \hat{\mathbf{v}} \in W, \quad \mathbf{z} \in W^{\perp},$$

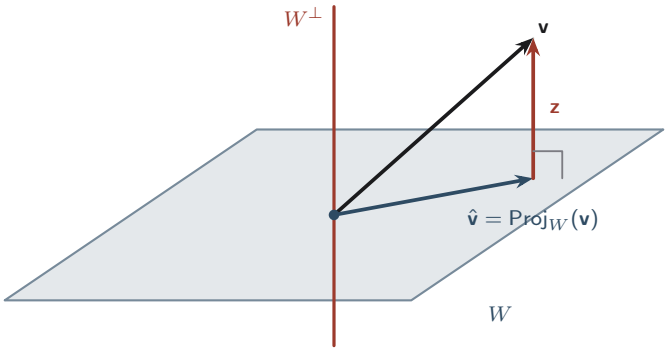
and $\hat{\mathbf{v}} = \mathrm{Proj}_W(\mathbf{v})$.

PROOF.

Existence: Let $\hat{\mathbf{v}} = \mathrm{Proj}_W(\mathbf{v})$ and put $\mathbf{z} = \mathbf{v} - \hat{\mathbf{v}}$. For each orthogonal basis vector $\mathbf{u}_j$ of W, all terms of $\mathrm{Proj}_W(\mathbf{v})$ but one are orthogonal to $\mathbf{u}_j$, and that one contributes exactly $\mathbf{v} \cdot \mathbf{u}_j$; so $\mathbf{z} \cdot \mathbf{u}_j = (\mathbf{v} - \hat{\mathbf{v}}) \cdot \mathbf{u}_j = \mathbf{v} \cdot \mathbf{u}_j - \mathbf{v} \cdot \mathbf{u}_j = 0$, and $\mathbf{z} \perp W$.

Uniqueness: Assume $\mathbf{v} = \hat{\mathbf{v}}_1 + \mathbf{z}_1 = \hat{\mathbf{v}}_2 + \mathbf{z}_2$. Subtracting the two expressions of $\mathbf{v}$ from each other gives $\hat{\mathbf{v}}_1 - \hat{\mathbf{v}}_2 = \mathbf{z}_2 - \mathbf{z}_1$. By definition $\hat{\mathbf{v}}_1, \hat{\mathbf{v}}_2$ lie in W and $\mathbf{z}_2, \mathbf{z}_1$ lie in $W^{\perp}$. Hence, $\hat{\mathbf{v}}_1 - \hat{\mathbf{v}}_2 = \mathbf{z}_2 - \mathbf{z}_1$ lies in both W and $W^{\perp}$, and is thus orthogonal to itself and must be $\mathbf{0}$ (Lemma 17.5). This gives $\hat{\mathbf{v}}_1 = \hat{\mathbf{v}}_2$ and $\mathbf{z}_2 - \mathbf{z}_1$ as desired. □

The decomposition, seen in $\mathbb{R}^3$



The shadow $\hat{\mathbf{v}}$ lies in W; the leftover $\mathbf{z}$ is perpendicular to all of W. Every vector splits this way, and only this way.

A worked projection

Let $W = \text{Span}\left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \right\}$. $\mathbf{u}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \mathbf{u}_2 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$ are orthogonal as $-1 + 1 + 0 = 0$.

What is the projection of $\mathbf{v} = \begin{bmatrix} 2 \\ 4 \\ 3 \end{bmatrix}$ onto W ▸?.

A worked projection

Let $W = \text{Span}\left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \right\}$. $\mathbf{u}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \mathbf{u}_2 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$ are orthogonal as $-1 + 1 + 0 = 0$.

What is the projection of $\mathbf{v} = \begin{bmatrix} 2 \\ 4 \\ 3 \end{bmatrix}$ onto W ▸?. We have

$$\frac{\mathbf{v} \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1} = \frac{6}{2} = 3, \qquad \frac{\mathbf{v} \cdot \mathbf{u}_2}{\mathbf{u}_2 \cdot \mathbf{u}_2} = \frac{2}{2} = 1,$$

and thus

$$\text{Proj}_W(\mathbf{v}) = 3 \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + 1 \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \\ 0 \end{bmatrix}, \qquad \mathbf{z} = \mathbf{v} - \text{Proj}_W(\mathbf{v}) = \begin{bmatrix} 0 \\ 0 \\ 3 \end{bmatrix}$$

A worked projection

Let $W = \text{Span}\left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \right\}$. $\mathbf{u}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \mathbf{u}_2 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$ are orthogonal as $-1 + 1 + 0 = 0$.

What is the projection of $\mathbf{v} = \begin{bmatrix} 2 \\ 4 \\ 3 \end{bmatrix}$ onto W ▸?. We have

$$\frac{\mathbf{v} \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1} = \frac{6}{2} = 3, \qquad \frac{\mathbf{v} \cdot \mathbf{u}_2}{\mathbf{u}_2 \cdot \mathbf{u}_2} = \frac{2}{2} = 1,$$

and thus

$$\text{Proj}_W(\mathbf{v}) = 3 \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + 1 \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \\ 0 \end{bmatrix}, \qquad \mathbf{z} = \mathbf{v} - \text{Proj}_W(\mathbf{v}) = \begin{bmatrix} 0 \\ 0 \\ 3 \end{bmatrix}$$

Here, $\mathbf{z}$ is *visibly* orthogonal to both $\mathbf{u}_1$ and $\mathbf{u}_2$, as it must be.

Here W was the plane $z = 0$, so the answer was predictable. In general, special cases are good ways to build intution and test formulas.

SECTION 3

Gram–Schmidt

The problem ▸

Everything above needed an *orthogonal* basis. Most bases are not orthogonal.

GRAM–SCHMIDT

Given any basis $\mathbf{x}_1, \ldots, \mathbf{x}_k$ of W, build an orthogonal basis one vector at a time. Keep the first vector as it is. Subtract from each subsequent vector its projection onto every direction already built. What remains has no component along any previous direction, and is thus orthogonal to all of them. It is nonzero as otherwise the original vector would lie in the span of the earlier ones, contradicting independence.

Repeating this gives an orthogonal basis of W.

The problem ▸

Everything above needed an *orthogonal* basis. Most bases are not orthogonal.

> GRAM–SCHMIDT
>
> Given any basis $\mathbf{x}_1, \ldots, \mathbf{x}_k$ of W, build an orthogonal basis one vector at a time. Keep the first vector as it is. Subtract from each subsequent vector its projection onto every direction already built. What remains has no component along any previous direction, and is thus orthogonal to all of them. It is nonzero as otherwise the original vector would lie in the span of the earlier ones, contradicting independence.
>
> Repeating this gives an orthogonal basis of W.

This is just the single-vector splitting of Lecture 16, applied repeatedly.

Normalising each entry gives an **orthonormal basis**. More on it later.

The process, step by step

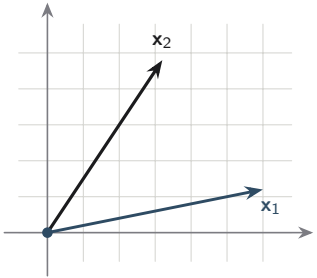
$$
\begin{aligned}
\mathbf{u}_1 &= \mathbf{x}_1 \\
\mathbf{u}_2 &= \mathbf{x}_2 - \frac{\mathbf{x}_2 \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1}\mathbf{u}_1 \\
\mathbf{u}_3 &= \mathbf{x}_3 - \frac{\mathbf{x}_3 \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1}\mathbf{u}_1 - \frac{\mathbf{x}_3 \cdot \mathbf{u}_2}{\mathbf{u}_2 \cdot \mathbf{u}_2}\mathbf{u}_2 \\
&\vdots \\
\mathbf{u}_k &= \mathbf{x}_k - \mathrm{Proj}_{\mathrm{Span}\{\mathbf{u}_1,\ldots,\mathbf{u}_{k-1}\}}(\mathbf{x}_k)
\end{aligned}
$$

The process, step by step

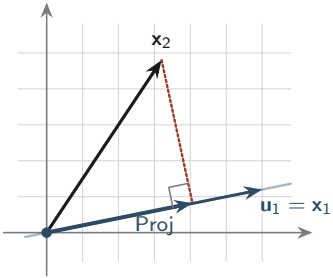
$$
\begin{aligned}
\mathbf{u}_1 &= \mathbf{x}_1 \\
\mathbf{u}_2 &= \mathbf{x}_2 - \frac{\mathbf{x}_2 \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1}\mathbf{u}_1 \\
\mathbf{u}_3 &= \mathbf{x}_3 - \frac{\mathbf{x}_3 \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1}\mathbf{u}_1 - \frac{\mathbf{x}_3 \cdot \mathbf{u}_2}{\mathbf{u}_2 \cdot \mathbf{u}_2}\mathbf{u}_2 \\
&\vdots \\
\mathbf{u}_k &= \mathbf{x}_k - \mathrm{Proj}_{\mathrm{Span}\{\mathbf{u}_1,\ldots,\mathbf{u}_{k-1}\}}(\mathbf{x}_k)
\end{aligned}
$$

At every stage $\mathrm{Span}\{\mathbf{u}_1, \ldots, \mathbf{u}_j\} = \mathrm{Span}\{\mathbf{x}_1, \ldots, \mathbf{x}_j\}$, since each $\mathbf{u}_j$ differs from $\mathbf{x}_j$ by a linear combination of elements of the current span. So nothing is lost, and the result is an orthogonal basis of W.

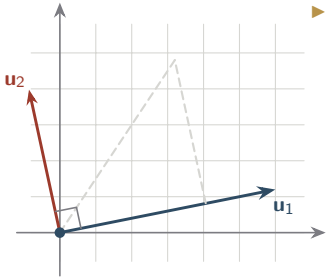
Step 2, drawn



a basis,
not orthogonal



project $\mathbf{x}_2$
onto $\mathbf{u}_1$



subtract: what is left
is orthogonal

A worked example (1/2)

Apply Gram–Schmidt ▸ to $\mathbf{x}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$, $\mathbf{x}_2 = \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}$, $\mathbf{x}_3 = \begin{bmatrix} 0 \\ 0 \\ 3 \end{bmatrix}$.

A worked example (1/2)

Apply Gram–Schmidt ▸ to $\mathbf{x}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$, $\mathbf{x}_2 = \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}$, $\mathbf{x}_3 = \begin{bmatrix} 0 \\ 0 \\ 3 \end{bmatrix}$.

Step 1. $\mathbf{u}_1 = \mathbf{x}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$, with $\mathbf{u}_1 \cdot \mathbf{u}_1 = 2$.

A worked example (1/2)

Apply Gram–Schmidt ▸ to $\mathbf{x}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$, $\mathbf{x}_2 = \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}$, $\mathbf{x}_3 = \begin{bmatrix} 0 \\ 0 \\ 3 \end{bmatrix}$.

Step 1. $\mathbf{u}_1 = \mathbf{x}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$, with $\mathbf{u}_1 \cdot \mathbf{u}_1 = 2$.

Step 2. $\mathbf{x}_2 \cdot \mathbf{u}_1 = 2 + 0 + 0 = 2$, so

$$\mathbf{u}_2 \;=\; \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix} - \frac{2}{2} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \;=\; \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}.$$

A worked example (1/2)

Apply Gram–Schmidt ▸ to $\mathbf{x}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$, $\mathbf{x}_2 = \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}$, $\mathbf{x}_3 = \begin{bmatrix} 0 \\ 0 \\ 3 \end{bmatrix}$.

Step 1. $\mathbf{u}_1 = \mathbf{x}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$, with $\mathbf{u}_1 \cdot \mathbf{u}_1 = 2$.

Step 2. $\mathbf{x}_2 \cdot \mathbf{u}_1 = 2 + 0 + 0 = 2$, so

$$\mathbf{u}_2 \;=\; \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix} - \frac{2}{2} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \;=\; \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}.$$

Check: $\mathbf{u}_2 \cdot \mathbf{u}_1 = 1 - 1 + 0 = 0$.

It is not a bad idea to check the first few, as its quick and one can catch a slip early and avoid followup mistakes.

A worked example (2/2)

Step 3. ▸ $\mathbf{x}_3 \cdot \mathbf{u}_1 = 0$ and $\mathbf{x}_3 \cdot \mathbf{u}_2 = 3$, with $\mathbf{u}_2 \cdot \mathbf{u}_2 = 3$. So

$$\mathbf{u}_3 = \begin{bmatrix} 0 \\ 0 \\ 3 \end{bmatrix} - \frac{0}{2} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} - \frac{3}{3} \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \\ 2 \end{bmatrix}.$$

A worked example (2/2)

Step 3. ▸ $\mathbf{x}_3 \cdot \mathbf{u}_1 = 0$ and $\mathbf{x}_3 \cdot \mathbf{u}_2 = 3$, with $\mathbf{u}_2 \cdot \mathbf{u}_2 = 3$. So

$$\mathbf{u}_3 = \begin{bmatrix} 0 \\ 0 \\ 3 \end{bmatrix} - \frac{0}{2} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} - \frac{3}{3} \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \\ 2 \end{bmatrix}.$$

Checks: $\mathbf{u}_3 \cdot \mathbf{u}_1 = -1 + 1 + 0 = 0$ and $\mathbf{u}_3 \cdot \mathbf{u}_2 = -1 - 1 + 2 = 0$. **Yes.**

A worked example (2/2)

Step 3. ▸ $\mathbf{x}_3 \cdot \mathbf{u}_1 = 0$ and $\mathbf{x}_3 \cdot \mathbf{u}_2 = 3$, with $\mathbf{u}_2 \cdot \mathbf{u}_2 = 3$. So

$$\mathbf{u}_3 = \begin{bmatrix} 0 \\ 0 \\ 3 \end{bmatrix} - \frac{0}{2} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} - \frac{3}{3} \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \\ 2 \end{bmatrix}.$$

Checks: $\mathbf{u}_3 \cdot \mathbf{u}_1 = -1 + 1 + 0 = 0$ and $\mathbf{u}_3 \cdot \mathbf{u}_2 = -1 - 1 + 2 = 0$. **Yes.**

Normalising gives an *orthonormal* basis:

$$\frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \qquad \frac{1}{\sqrt{3}} \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}, \qquad \frac{1}{\sqrt{6}} \begin{bmatrix} -1 \\ 1 \\ 2 \end{bmatrix}.$$

Consequences

PROPOSITION 17.8

For a subspace $W \subseteq \mathbb{R}^n$:

1. $W \cap W^\perp = \{\mathbf{0}\}$;
2. $\dim W + \dim W^\perp = n$;
3. $(W^\perp)^\perp = W$.

Consequences

PROPOSITION 17.8

For a subspace $W \subseteq \mathbb{R}^n$:

1. $W \cap W^\perp = \{\mathbf{0}\}$;
2. $\dim W + \dim W^\perp = n$;
3. $(W^\perp)^\perp = W$.

PROOF.

(1) Lemma 17.5. (2) By Gram-Schmidt, W and $W^\perp$ admit an orthogonal basis. Putting these two basis together they remain orthogonal, as every element of $W^\perp$ is orthogonal to all of W. Moreover, this set is linearly independent by Theorem 17.3 and must by dimensionality span all of W. (3) $W \subseteq (W^\perp)^\perp$ is immediate, and the dimensions agree by (2). □

With $(\text{Row } A)^\perp = \text{Ker } A$, (2) gives $\dim \text{Row } A + \dim \text{Ker } A = n$: rank–nullity, from the geometric side.

SECTION 4

Orthogonal matrices

Orthogonal matrices have orthonormal columns

THEOREM 17.9

For a square $n \times n$ matrix A, the following are equivalent:

1. *The columns of A form an orthonormal basis of $\mathbb{R}^n$;*
2. *$A^{\mathsf{T}}A = I_n = AA^{\mathsf{T}}$, that is A is invertible with $A^{-1} = A^{\mathsf{T}}$;*
3. *$A\mathbf{x} \cdot A\mathbf{y} = \mathbf{x} \cdot \mathbf{y}$ for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$;*
4. *$\|A\mathbf{x}\| = \|\mathbf{x}\|$ for all $\mathbf{x} \in \mathbb{R}^n$.*

Such an A is called an ***orthogonal matrix****.*

Orthogonal matrices have orthonormal columns

THEOREM 17.9

For a square $n \times n$ matrix A, the following are equivalent:

1. *The columns of A form an orthonormal basis of $\mathbb{R}^n$;*
2. *$A^{\mathsf{T}}A = I_n = AA^{\mathsf{T}}$, that is A is invertible with $A^{-1} = A^{\mathsf{T}}$;*
3. *$A\mathbf{x} \cdot A\mathbf{y} = \mathbf{x} \cdot \mathbf{y}$ for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$;*
4. *$\|A\mathbf{x}\| = \|\mathbf{x}\|$ for all $\mathbf{x} \in \mathbb{R}^n$.*

Such an A is called an ***orthogonal matrix****.*

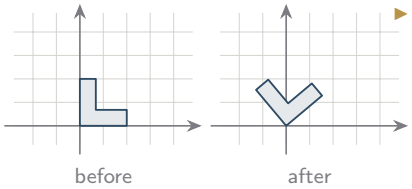
CAUTION

It is easy to think an orthogonal matrix would have orthogonal columns. But that would *not* preserve orthogonality. Instead, for A to map an orthogonal basis to an orthogonal one, the columns of A *must* be orthonorml.

What such a matrix does

We will not give a proof of the theorem, but the equivalence of (1) and (2) is easy. The (i, j)'th entry of $A^{\mathsf{T}}A$ is the inner product of column i with column j, so $A^{\mathsf{T}}A = I$ says exactly that the columns are orthonormal.

Condition (4) says that an orthogonal matrix does not change a vector's length; (3) says that it preserves angles as well. Hence, these are *rigid motions* of $\mathbb{R}^n$ fixing the origin, being a rotation or reflection; or a composite of these.



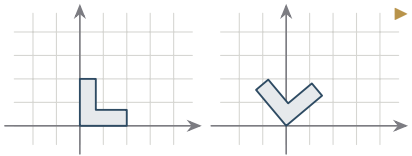
The polygon turns, but never stretches or bends.

What such a matrix does

We will not give a proof of the theorem, but the equivalence of (1) and (2) is easy. The (i, j)'th entry of $A^{\mathsf{T}}A$ is the inner product of column i with column j, so $A^{\mathsf{T}}A = I$ says exactly that the columns are orthonormal.

Condition (4) says that an orthogonal matrix does not change a vector's length; (3) says that it preserves angles as well. Hence, these are *rigid motions* of $\mathbb{R}^n$ fixing the origin, being a rotation or reflection; or a composite of these.

Consequently $\operatorname{Det} A = \pm 1$, from $\operatorname{Det}(A^{\mathsf{T}}A) = (\operatorname{Det} A)^2 = 1$. The sign distinguishes rotations $(+1)$ from reflections (-1).



The polygon turns, but never stretches or bends.

Examples

EXAMPLE 17.10 · Rotation

$A = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$. Each column has length $\cos^2\theta + \sin^2\theta = 1$, and their inner product is $-\cos\theta\sin\theta + \sin\theta\cos\theta = 0$. Here $\operatorname{Det} A = 1$.

Examples

EXAMPLE 17.10 · Rotation

$A = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$. Each column has length $\cos^2\theta + \sin^2\theta = 1$, and their inner product is $-\cos\theta\sin\theta + \sin\theta\cos\theta = 0$. Here $\operatorname{Det} A = 1$.

EXAMPLE 17.11 ▸ · Reflection

$A = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$ has orthonormal columns and $\operatorname{Det} A = -1$.

Examples

EXAMPLE 17.10 · Rotation

$A = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$. Each column has length $\cos^2\theta + \sin^2\theta = 1$, and their inner product is $-\cos\theta\sin\theta + \sin\theta\cos\theta = 0$. Here $\operatorname{Det} A = 1$.

EXAMPLE 17.11 ▸ · Reflection

$A = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$ has orthonormal columns and $\operatorname{Det} A = -1$.

EXAMPLE 17.12 · From Gram–Schmidt

The three normalised vectors of the last section ▸ are the columns of an orthogonal matrix. Every orthonormal basis gives one.

Examples

EXAMPLE 17.10 · Rotation

$A = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$. Each column has length $\cos^2\theta + \sin^2\theta = 1$, and their inner product is $-\cos\theta\sin\theta + \sin\theta\cos\theta = 0$. Here $\operatorname{Det} A = 1$.

EXAMPLE 17.11 ▸ · Reflection

$A = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$ has orthonormal columns and $\operatorname{Det} A = -1$.

EXAMPLE 17.12 · From Gram–Schmidt

The three normalised vectors of the last section ▸ are the columns of an orthogonal matrix. Every orthonormal basis gives one.

And for such an A, the projection formulas simplify sharply, since every denominator $\mathbf{u}_i \cdot \mathbf{u}_i$ becomes 1.

Summary of Lecture 17

ORTHOGONAL SETS	Non-zero and pairwise orthogonal $\Rightarrow$ independent; n of them in $\mathbb{R}^n$ must thus be an orthogonal basis.
COORDINATES	With an orthogonal basis $\mathcal{C} := \{\mathbf{u}_1, \ldots, u_n\}$, $[\mathbf{v}]_\mathcal{C}$ is found by $c_i = \mathbf{v} \cdot \mathbf{u}_i / \mathbf{u}_i \cdot \mathbf{u}_i$. We need solve no system of equations.
PROJECTION	$\mathrm{Proj}_W(\mathbf{v}) = \sum \frac{\mathbf{v}\cdot\mathbf{u}_i}{\mathbf{u}_i\cdot\mathbf{u}_i}\mathbf{u}_i$, for an orthogonal basis only.
DECOMPOSITION	Every vector $\mathbf{v}$ decomposes uniquely as $\mathbf{v} = \hat{\mathbf{v}} + \mathbf{z}$ with $\hat{\mathbf{v}} = \mathrm{Proj}_W(\mathbf{v}) \in W$, $\mathbf{z} \in W^\perp$.
HENCE	$W \cap W^\perp = \{\mathbf{0}\}$, $\dim W + \dim W^\perp = n$, $(W^\perp)^\perp = W$.
GRAM–SCHMIDT	Subtract from each $\mathbf{x}_j$ its projection onto what is built so far.
ORTHOGONAL MATRIX	Orthonormal columns $\iff A^{-1} = A^\mathsf{T} \iff$ lengths preserved; then $\mathrm{Det}\, A = \pm 1$.

Before next time

AT HOME

1. Let $W = \text{Span}\left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\}$. Verify the basis is orthogonal ▸, then find $\text{Proj}_W \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix}$ ▸ and the component $\mathbf{z} \in W^\perp$. Check Pythagoras.
2. Apply Gram–Schmidt ▸ to $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$, $\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$, $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, then normalise.
3. Show that if A and B are orthogonal $n \times n$ matrices then so is AB. Is $A + B$ orthogonal?

END OF LECTURE 17

ilia.pirashvili@universityofgalway.ac.ie

Solution: question 1 ▸

SOLUTION

$\begin{bmatrix}1\\0\\1\end{bmatrix} \cdot \begin{bmatrix}0\\1\\0\end{bmatrix} = 0$ ▸, so the basis is orthogonal, with $\mathbf{u}_1 \cdot \mathbf{u}_1 = 2$ and $\mathbf{u}_2 \cdot \mathbf{u}_2 = 1$.

$\mathbf{v} \cdot \mathbf{u}_1 = 3 + 0 + 1 = 4$ and $\mathbf{v} \cdot \mathbf{u}_2 = 2$, so

$$\text{Proj}_W(\mathbf{v}) = \frac{4}{2}\begin{bmatrix}1\\0\\1\end{bmatrix} + \frac{2}{1}\begin{bmatrix}0\\1\\0\end{bmatrix} = \begin{bmatrix}2\\2\\2\end{bmatrix}, \qquad \mathbf{z} = \begin{bmatrix}1\\0\\-1\end{bmatrix}.$$

Solution: question 1 ▸

SOLUTION

$\begin{bmatrix}1\\0\\1\end{bmatrix} \cdot \begin{bmatrix}0\\1\\0\end{bmatrix} = 0$ ▸, so the basis is orthogonal, with $\mathbf{u}_1 \cdot \mathbf{u}_1 = 2$ and $\mathbf{u}_2 \cdot \mathbf{u}_2 = 1$.

$\mathbf{v} \cdot \mathbf{u}_1 = 3 + 0 + 1 = 4$ and $\mathbf{v} \cdot \mathbf{u}_2 = 2$, so

$$\text{Proj}_W(\mathbf{v}) = \frac{4}{2}\begin{bmatrix}1\\0\\1\end{bmatrix} + \frac{2}{1}\begin{bmatrix}0\\1\\0\end{bmatrix} = \begin{bmatrix}2\\2\\2\end{bmatrix}, \qquad \mathbf{z} = \begin{bmatrix}1\\0\\-1\end{bmatrix}.$$

$\mathbf{z} \cdot \mathbf{u}_1 = 1 - 1 = 0$ and $\mathbf{z} \cdot \mathbf{u}_2 = 0$. Pythagoras: $12 + 2 = 14 = 9 + 4 + 1$. **Yes.**

Solution: question 2 ▸

SOLUTION

$\mathbf{u}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ with $\mathbf{u}_1 \cdot \mathbf{u}_1 = 3$, and $\mathbf{x}_2 \cdot \mathbf{u}_1 = 2$, so $\mathbf{u}_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} - \frac{2}{3}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \frac{1}{3}\begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix}$.

Rescaling (which changes nothing), take $\mathbf{u}_2 = \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix}$, with $\mathbf{u}_2 \cdot \mathbf{u}_2 = 6$.

Solution: question 2 ▸

SOLUTION

$\mathbf{u}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ with $\mathbf{u}_1 \cdot \mathbf{u}_1 = 3$, and $\mathbf{x}_2 \cdot \mathbf{u}_1 = 2$, so $\mathbf{u}_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} - \frac{2}{3}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \frac{1}{3}\begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix}$.
Rescaling (which changes nothing), take $\mathbf{u}_2 = \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix}$, with $\mathbf{u}_2 \cdot \mathbf{u}_2 = 6$.

Then $\mathbf{x}_3 \cdot \mathbf{u}_1 = \mathbf{x}_3 \cdot \mathbf{u}_2 = 1$, giving $\mathbf{u}_3 = \frac{1}{2}\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$; take $\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$.

Normalised: $\frac{1}{\sqrt{3}}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$, $\frac{1}{\sqrt{6}}\begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix}$, $\frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$.

Solution: question 3

AB: YES

$$(AB)^{\mathsf{T}}(AB) = B^{\mathsf{T}}A^{\mathsf{T}}AB = B^{\mathsf{T}}IB = B^{\mathsf{T}}B = I,$$

using $(AB)^{\mathsf{T}} = B^{\mathsf{T}}A^{\mathsf{T}}$ from lecture 1.

Solution: question 3

AB: YES

$$(AB)^{\mathsf{T}}(AB) = B^{\mathsf{T}}A^{\mathsf{T}}AB = B^{\mathsf{T}}IB = B^{\mathsf{T}}B = I,$$

using $(AB)^{\mathsf{T}} = B^{\mathsf{T}}A^{\mathsf{T}}$ from lecture 1.

$A + B$: NO

Take $A = I$ and $B = I$. Then $A + B = 2I$, and $\|2I\mathbf{x}\| = 2\|\mathbf{x}\| \neq \|\mathbf{x}\|$ unless $\mathbf{x} = \mathbf{0}$.

Orthogonal matrices are closed under products and inverses, but not sums. They form a group, not a subspace.