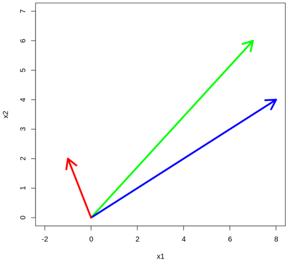
Annotated slides from Tuesday

Week 10: Orthogonal Everything

Dr Niall Madden

8 and 11 November, 2022



R code

```
v <- c(7,6)
w <- c(8,4)
z <- c(-1,2)
plot(NULL, xlim=c(-2,8), ylim=c(0,7)),
   xlab="x1", ylab="x2")
arrows(0,0, v[1], v[2],lwd=4,col="green")
arrows(0,0, w[1], w[2],lwd=4,col="blue")
arrows(0,0, z[1], z[2],lwd=4,col="red")
```

These slides are adapted (slightly) from ones by Tobias Rossmann.

Outline

For more details,

- Section 6.1 (Inner Product, Length and Orthogonality) of the Lay et al text-book https://nuigalway-primo.hosted.exlibrisgroup.com/permalink/f/1pmb9lf/353GAL_ALMA_DS5192067630003626
- Chapters 6 and 9 of *Linear Algebra for Data Science* https://shainarace.github.io/LinearAlgebra/norms.html and https://shainarace.github.io/LinearAlgebra/orthog.html

Announcements, etc

Assignment 5

Assignment 5 opens today. Deadline is 5pm, Monday, 21st of November.

Communication Skills : Next steps...

MA313
Week 10: Orthogonal Everything

Start of ...

PART 1: Announcements and Preview of Week 10

The big ideas from this week will be **Orthogonality**.

- How to find the orthogonal projector of a vector onto a subspace
- What it means if the columns of a matrix are orthogonal to each other.

These are the essential ideas from recent lectures that you need for this week.

- A **BASIS** of a vector space V is the **smallest** sequence $(v_1, v_2, \dots, v_r)$ of vectors which which spans V.
- The **DIMENSION** of V is the number of vectors in any basis for V.
- The vector space $\mathbb{R}^n$ has dimension n. Any sequence of n linearly independent vectors is a basis for $\mathbb{R}^n$.
- The **INNER PRODUCT** of vectors u and v in $\mathbb{R}^n$ is the real number given by

$$u \cdot v = \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix} \cdot \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} = u_1 v_1 + \cdots + u_n v_n.$$

- $u \cdot v = u^\top v$.
- The **LENGTH** of a vector $v \in \mathbb{R}^n$ is $\|v\| := \sqrt{v \cdot v} = \sqrt{v_1^2 + \cdots + v_n^2}$.
- If $u, v \in \mathbb{R}^n$ are both be non-zero, then the **angle** $\measuredangle(u, v) \in [0, \pi]$ between u and v is defined by $\cos(\measuredangle(u, v)) = \dfrac{u \cdot v}{\|u\|\|v\|}$.
- We say $v \in \mathbb{R}^n$ is a **unit vector** if $\|v\| = 1$.
  The unit vector in the same direction as v is $v/\|v\|$.
- $u, v \in \mathbb{R}^n$ are **orthogonal** if $u \cdot v = 0$. We may write this as $u \perp v$.
- **Pythagorean Theorem:** If $u \perp v$, then $\|u + v\|^2 = \|u\|^2 + \|v\|^2$.

- If u and v are non-zero vectors in $\mathbb{R}^n$, then

  $$w = u - \frac{u \cdot v}{v \cdot v} v$$

  is orthogonal to v.

- **The Cauchy-Schwarz inequality:** $|u \cdot v| \leqslant \|u\|\|v\|$. And $|u \cdot v| = \|u\|\|v\|$, if and only if u and v are linearly dependent.

- The Cauchy-Schwarz inequality implies that, if $u \neq 0 \neq v$, then

  $$-1 \leqslant \frac{u \cdot v}{\|u\|\|v\|} \leqslant 1.$$

  Therefore, the definition of the angle between u and v via

  $$\cos(\measuredangle(u, v)) = \frac{u \cdot v}{\|u\|\|v\|}$$

  makes sense.

The Triangle inequality

$\|u + v\| \leqslant \|u\| + \|v\|$ for all $u, v \in \mathbb{R}^n$.

$u \cdot v = v \cdot u$

Proof

$$\|u+v\|^2 = (u+v)\cdot(u+v)$$
$$= u\cdot u + v\cdot u + u\cdot v + v\cdot v$$
$$= \|u\|^2 + 2\, u\cdot v + \|v\|^2$$
$$\leq \|u\|^2 + 2\,|u\cdot v| + \|v\|^2$$
$$\leq \|u\|^2 + 2\|u\|\,\|v\| + \|v\|^2 \quad \text{(C-S inequality)}$$
$$= (\|u\| + \|v\|)^2$$

So $\|u+v\| \leq \|u\| + \|v\|$.

MA313
Week 10: Orthogonal Everything

Start of ...

PART 2: Orthogonal Projections

Definition (ORTHOGONAL to a subspace)

Let W be a subspace of $\mathbb{R}^n$. We say that a vector $z \in \mathbb{R}^n$ is **orthogonal** to W if $z \perp w$ for all $w \in W$.

Example

$$u = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} \text{ is orthogonal to the space } V = \mathrm{Span}\left\{ \overset{V_1}{\begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}}, \overset{V_2}{\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}} \right\}.$$

Check $u \cdot V_1 = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} = (1)(1) + (1)(1) + (-1)(2) = 0.$

$u \cdot V_2 = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} \cdot \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} = (1)(-1) + (+1)(1) + (-1)(0) = 0$

✓

Part 2: Orthogonal Projections

Definition (ORTHOGONAL to a subspace)

Let W be a subspace of $\mathbb{R}^n$. We say that a vector $z \in \mathbb{R}^n$ is **orthogonal** to W if $z \perp w$ for all $w \in W$.

Example

$$u = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} \text{ is orthogonal to the space } V = \text{Span}\left\{ \underbrace{\begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}}_{v_1}, \underbrace{\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}}_{v_2} \right\}.$$

More generally, any vector in V is of the form $c_1 v_1 + c_2 v_2$. But

$$u \cdot (c_1 v_1 + c_2 v_2) = u \cdot (c_1 v_1) + u \cdot (c_2 v_2)$$

$$= c_1 (u \cdot v_1) + c_2 (u \cdot v_2) = c_1 (0) + c_2 (0) = 0.$$

 $a^T b = a \cdot b.$

Orthogonal spaces

Two vector spaces, V and W are **orthogonal**, if, for every $v \in V$ is orthogonal to every $w \in W$, then $v \cdot w = 0$

Example

For any matrix A, its left null space is orthogonal to its column space.

If x is in the left null space of A, then $x^T A = 0$. (Or, $A^T x = 0$). If y is in the column space of A, the $Ab = y$ for some vector b. Then $x \cdot y = x^T y = x^T (Ab) =$

$$= (x^T A) b = (0)^T b = 0 \cdot b = 0.$$

Definition (ORTHOGONAL COMPLEMENT)

The **orthogonal complement** of a vector space W, denoted $W^\perp$, is the set of vectors that are orthogonal to W. That is,

$$W^\perp = \{z \in \mathbb{R}^n : z \perp w \text{ for all } w \in W\}.$$

Example. Let $W = \text{Span}\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\}$.

Then $W^\perp = \text{Span}\left\{ \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$

Check: $\left(a\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + b\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right) \cdot \left(c\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right)$

$= \begin{bmatrix} a \\ b \\ 0 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 0 \\ c \end{bmatrix} = a(0) + (b) + 0(c) = 0$ ✓

Part 2: Orthogonal Projections

Example

Let $W = \text{span}\left\{\begin{bmatrix}2\\-1\end{bmatrix}\right\} = \left\{\begin{bmatrix}2t\\-t\end{bmatrix} : t \in \mathbb{R}\right\}$.

Give a basis for $W^\perp$.

Let $v = \begin{bmatrix}v_1\\v_2\end{bmatrix} \in W^\perp$.

If $v \cdot \begin{bmatrix}2t\\-t\end{bmatrix} = 0$ then $2(v_1)t - (v_2)t = 0$.

$\Rightarrow t(2v_1 - v_2) = 0$ for $\underline{\text{all}}$ t.

If $t \neq 0$ then $2v_1 = v_2$. So $v = \begin{bmatrix}v_1\\2v_1\end{bmatrix}$

That is $W^\perp = \text{Span}\left\{\begin{bmatrix}1\\2\end{bmatrix}\right\}$.

Theorem: Unique representation/Orthogonal decomposition

Let W be a subspace of $\mathbb{R}^n$. Then:

- $W^\perp$ is a subspace of $\mathbb{R}^n$. This is true (i). $0 \in W^\perp$ since $0 \perp w$ $\forall w \in W$.
- If $W = \operatorname{span}\{w_1, \ldots, w_r\}$, then $W^\perp = \{z \in \mathbb{R}^n : z \perp w_1, \ldots, z \perp w_r\}$.
- Every vector $v \in \mathbb{R}^n$ has a **unique representation**

$$v = \hat{v} + z \qquad \text{for } \hat{v} \in W, \qquad \text{and } z \in W^\perp.$$

- The function $\operatorname{proj}_W \colon \mathbb{R}^n \to W, \quad v \mapsto \hat{v}$ is a linear transformation, called the **orthogonal projection** of $\mathbb{R}^n$ onto W.
- $W \cap W^\perp = \{0\}$.
- $\dim W^\perp = n - \dim W$.

Theorem: Unique representation/Orthogonal decomposition

Let W be a subspace of $\mathbb{R}^n$. Then:

If $w_1, w_2 \in W$, $z \in W^\perp$

Then $z \cdot (w_1 + w_2) =$

$z \cdot w_1 + z \cdot w_2 = 0 + 0$

$= 0.$

- $W^\perp$ is a subspace of $\mathbb{R}^n$.
- If $W = \operatorname{span}\{w_1, \ldots, w_r\}$, then $W^\perp = \{z \in \mathbb{R}^n : z \perp w_1, \ldots, z \perp w_r\}$.
- Every vector $v \in \mathbb{R}^n$ has a **unique representation**

$$v = \hat{v} + z \qquad \text{for } \hat{v} \in W, \qquad \text{and } z \in W^\perp.$$

- The function $\operatorname{proj}_W : \mathbb{R}^n \to W, \quad v \mapsto \hat{v}$ is a linear transformation, called the **orthogonal projection** of $\mathbb{R}^n$ onto W.
- $W \cap W^\perp = \{0\}$.
- $\dim W^\perp = n - \dim W$.

Finished here Tuesday.