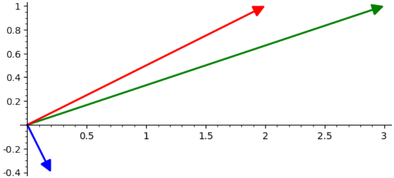
MA313 : Linear Algebra I

Week 9: (DRAFT) Inner Products and Orthogonality

Dr Niall Madden

1st and 4th of November, 2022



Sage code

```
u = vector([3, 1]);  v = vector([2, 1])
w = u-v*u.dot_product(v)/v.dot_product(v)
plot(u,color='green')+plot(v, color='red')+plot(w,color='blue')
```

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 4

Assignment 4 was posted last week. Deadline is 5pm, Monday, 7th November.

Upload you solutions, in PDF, to blackboard. If you prefer, you can give them to me in class on Tuesday, 8th Nov.

Communication Skills : Progress Report

Thanks for the progress reports. Jim and I will give feedback, and update you on the next steps next week.

Preview

The big ideas from this week will be

- dot products, and the angles between vectors
- the special case of when vectors are "perpendicular" (we say "orthogonal" in the general case).

To round off the previous section, I've posted two (old) videos to Week 9, om

- "Row Rank = Column Rank"
- Matrices of Linear Transformations.

MA313
Week 9: (DRAFT) Inner Products and Orthogonality

Start of ...

PART 1: Linear Transformations (Recorded)

This was section is "left over" from last week. I won't cover it in a "live" class, but have posted a video about it.

Part 1: Linear Transformations

Summary

Using **bases** and *coordinate vectors*, we essentially reduced the study of finitely generated (= finite-dimensional) vector spaces to that of $\mathbb{R}^n$.

Question

Can we similarly reduce the study of linear transformations (between finitely generated vector spaces) to that of matrices?

Part 1: Linear Transformations

From linear transformations to matrices

- Let V and W be vector spaces with bases $\mathcal{B} = (b_1, \ldots, b_n)$ and $\mathcal{C} = (c_1, \ldots, c_m)$, respectively.
- Let $T\colon V \to W$ be an arbitrary linear transformation.
- Let

  $$F\colon V \to \mathbb{R}^n, \quad v \mapsto [v]_{\mathcal{B}}$$

  and

  $$G\colon W \to \mathbb{R}^m, \quad w \mapsto [w]_{\mathcal{C}}$$

  be the coordinate mappings relative to $\mathcal{B}$ and $\mathcal{C}$, respectively.

These two maps are isomorphisms.

Part 1: Linear Transformations

- Recall: $F^{-1}\colon \mathbb{R}^n \to V$ is a linear transformation.
- We obtain a linear transformation

$$G \circ T \circ F^{-1}\colon \mathbb{R}^n \to \mathbb{R}^m.$$

- Therefore, we know that there exists a unique (!) $m \times n$ matrix A such that the linear transformation $G \circ T \circ F^{-1}\colon \mathbb{R}^n \to \mathbb{R}^m$ is given by $x \mapsto Ax$.
- We call A the **matrix** (or the *matrix representation*) of T relative to the bases $\mathcal{B}$ and $\mathcal{C}$ of V and W, respectively
- Notation: $\mathrm{M}_{\mathcal{C}\leftarrow\mathcal{B}}(T) := A$.

Part 1: Linear Transformations

Fact

The matrix of $T\colon V \to W$ relative to the bases $\mathcal{B} = (b_1, \ldots, b_n)$ and $\mathcal{C} = (c_1, \ldots, c_m)$ of V and W, respectively is the $m \times n$ matrix given by

$$\mathrm{M}_{\mathcal{C} \leftarrow \mathcal{B}}(T) = \left[\; [T(b_1)]_{\mathcal{C}} \cdots [T(b_n)]_{\mathcal{C}} \right].$$

Part 1: Linear Transformations

Example

Let $D\colon \mathbb{P}_3 \to \mathbb{P}_2, D(p(t)) = p'(t)$ be the linear transformation given by differentiation.
Choose bases $\mathcal{B} = (1, t, t^2, t^3)$ and $\mathcal{C} = (1, t, t^2)$ of $\mathbb{P}_3$ and $\mathbb{P}_2$, respectively.
What is the matrix of D relative to $\mathcal{B}$ and $\mathcal{C}$?

Part 1: Linear Transformations

Remark 1

For a linear transformation $T\colon V \to V$ and a given basis $\mathcal{B}$ of V, by the matrix of T relative to $\mathcal{B}$, we mean the matrix of T relative to $\mathcal{B}$ (in the domain) and $\mathcal{B}$ (in the codomain).

Part 1: Linear Transformations

Remark 2

- Having chosen (!) bases $\mathcal{B}$ and $\mathcal{C}$ as before, the operation

$$T \rightsquigarrow \mathrm{M}_{\mathcal{C}\leftarrow\mathcal{B}}(T)$$

  reduces essentially everything about linear transformations $V \to W$ to problems involving matrices.
- In particular, we can use matrix operations (e.g. row reduction) to study linear transformations!

Example

Write $A = \mathrm{M}_{\mathcal{C}\leftarrow\mathcal{B}}(T)$. Let $v \in V$. Then:

$$v \in \mathrm{Ker}\, T \Leftrightarrow [v]_{\mathcal{B}} \in \mathrm{Nul}\, A.$$

Part 1: Linear Transformations

Example

Write $A = \mathrm{M}_{\mathcal{C} \leftarrow \mathcal{B}}(T)$. Let $w \in W$.
Then:

$$w \in \mathrm{Ran}\, T \Leftrightarrow [w]_{\mathcal{C}} \in \mathrm{Col}\, A.$$

MA313
Week 9: (DRAFT) Inner Products and Orthogonality

Start of ...

PART 2: Inner Products

Inner products of vectors in $\mathbb{R}^n$

Part 2: Inner Products

Outlook

- We will now have a closer look at $\mathbb{R}^n$ from a geometric point of view.
- This will involve an additional structure on top of the vector space operations: **inner products**
- This leads us to some ideas in **data science**, particularly, linear ***least-squares problems***.

Part 2: Inner Products

An **inner product** is a function that maps a pair of vectors in $\mathbb{R}^n$ to a real number.

Definition (INNER PRODUCT)

The **inner product** (or **dot product**) of vectors u and v in $\mathbb{R}^n$ is the real number given by

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

Example

$$\begin{bmatrix} 2 \\ -5 \\ -1 \end{bmatrix} \cdot \begin{bmatrix} 3 \\ 2 \\ -3 \end{bmatrix} = (2)(3) + (-5)(2) + (-1)(-3)$$

$$= 6 - 10 + 3 = -1$$

Equivalent formulations

(i) The definition says that

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

(ii) More succinctly, this is $u \cdot v = \sum_{i=1}^{n} u_i v_i.$

(iii) From a practical point of view, $u \cdot v = u^T v$

This last view is crucial in many settings.

(Also, since there is an “inner product” there should also be an “outer product”. More of that in 2 weeks).

Properties of Inner Products

For all $u, v, w \in \mathbb{R}^n$ and $c \in \mathbb{R}$:

- $u \cdot v = v \cdot u$. So $u^T v = v^T u$
- $(u + v) \cdot w = u \cdot w + v \cdot w$.
- $(cu) \cdot v = u \cdot (cv) = c(u \cdot v)$.
- $u \cdot u \geqslant 0$. And $u \cdot u = 0$ if and only if $u = 0$.

check $u = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$ $v = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$

$\Rightarrow cu = \begin{bmatrix} cu_1 \\ cu_2 \end{bmatrix}$ & $(cu) \cdot v = (cu)v_1 + (cu_2) \cdot v_2$

$= c(u_1 \cdot v_1) + c(u_2 v_2)$.

Properties of Inner Products

For all $u, v, w \in \mathbb{R}^n$ and $c \in \mathbb{R}$:

- $u \cdot v = v \cdot u$.
- $(u + v) \cdot w = u \cdot w + v \cdot w$.
- $(cu) \cdot v = u \cdot (cv) = c(u \cdot v)$.
- $u \cdot u \geqslant 0$. And $u \cdot u = 0$ if and only if $u = 0$.

$u = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}$ So $u \cdot u = (u_1)(u_1) + (u_2 u_2) + (u_3 u_3)$

$= u_1^2 + u_2^2 + u_3^2$.

But any $u_i^2 \geqslant 0$, since $u_i \in \mathbb{R}$.

So $u \cdot u \geqslant 0$. Also $u_1^2 + u_2^2 + u_3^2 = 0$

$\Rightarrow u_1 = 0$, $u_2 = 0$ & $u_3 = 0$.

Definition (LENGTH OF A VECTOR)

The **length** (or **Euclidean norm**) of a vector $v \in \mathbb{R}^n$ is

$$\|v\| = \sqrt{v \cdot v} = \sqrt{v_1^2 + \cdots + v_n^2} \geqslant 0.$$

Note: Scaling a vector scales its length:

$$\|cv\| = |c|\|v\| \quad \text{for all } c \in \mathbb{R} \quad \text{and } v \in \mathbb{R}^n.$$

Also, if $\|v\| = 0$ then $\sqrt{v \cdot v} = 0$

so v is zero vector.

 $\sqrt{x^2} = |x|$

Definition (LENGTH OF A VECTOR)

The **length** (or **Euclidean norm**) of a vector $v \in \mathbb{R}^n$ is

$$\|v\| = \sqrt{v \cdot v} = \sqrt{v_1^2 + \cdots + v_n^2} \geqslant 0.$$

Note: Scaling a vector scales its length:

$$\|cv\| = |c|\|v\| \quad \text{for all } c \in \mathbb{R} \quad \text{and } v \in \mathbb{R}^n.$$

$$\|cv\| = \sqrt{(cv_1)^2 + (cv_2)^2 + \cdots + (cv_n)^2}$$

$$= \sqrt{c^2v_1^2 + c^2v_2^2 + \cdots + c^2v_n^2}$$

$$= \sqrt{c^2(v_1 + v_2 + \cdots + v_n)}$$

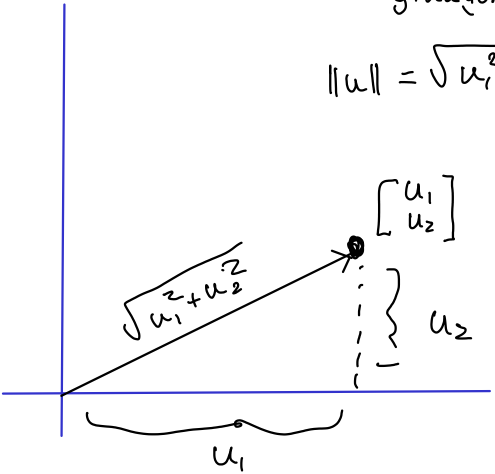
$$= |c|\,\|v\|.$$

Lengths in $\mathbb{R}^2$

$u = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$

Pythagoras Theorem.

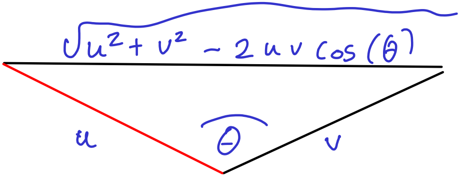
$\|u\| = \sqrt{u_1^2 + u_2^2}.$



 $\|u\|^2 = u \cdot u$

Law of cosines

$$\|u - v\|^2 = \|u\|^2 + \|v\|^2 - 2\|u\|\|v\| \cos \vartheta.$$



Note that

$$-\|u\| \cdot \|v\| \cos\theta = \tfrac{1}{2}\left(\|u - v\|^2 - \|u\|^2 - \|v\|^2\right).$$

$$= \tfrac{1}{2}\left((u-v)\cdot(u-v) - u\cdot u - v \cdot v\right)$$

$$= \tfrac{1}{2}\left(u\cdot u - v\cdot u - u\cdot v + v\cdot v - u\cdot u - v\cdot v\right)$$

$$= \tfrac{1}{2}\left(-u\cdot v - u\cdot v\right) = -(u\cdot v).$$

Law of cosines

$$\|u - v\|^2 = \|u\|^2 + \|v\|^2 - 2\|u\|\|v\| \cos \vartheta.$$

So

$$- \|u\| \cdot \|v\| \cos \theta = - u \cdot v$$

So $$\cos \theta = \frac{u \cdot v}{\|u\| \cdot \|v\|}.$$

Finished here Tues.

Definition (ANGLE BETWEEN VECTORS)

Let $u, v \in \mathbb{R}^n$ both be non-zero.

Then the **angle** $\measuredangle(u, v) \in [0, \pi]$ between u and v is defined by

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

Definition (UNIT VECTOR)

A **unit vector** in $\mathbb{R}^n$ is a vector v with $\|v\| = 1$.

For any non-zero $v \in \mathbb{R}^n$, the vector $\frac{1}{\|v\|}v$ is a unit vector "in the same direction" as v. This process is called *normalizing* the vector.

Example

$$\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{bmatrix}$$

MA313
Week 9: (DRAFT) Inner Products and Orthogonality

Start of ...

PART 3: Orthogonality

Part 4: Orthogonality

Definition (ORTHOGONAL VECTORS)

We say that $u, v \in \mathbb{R}^n$ are **orthogonal** if $u \cdot v = 0$.

Notation: $u \perp v \Leftrightarrow u \cdot v = 0$

Example

$$\begin{bmatrix} 1 \\ -2 \end{bmatrix} \cdot \begin{bmatrix} 6 \\ 3 \end{bmatrix}$$

Part 4: Orthogonality

Fact

If $u \neq 0 \neq v$, then $u \perp v$ if and only if $\measuredangle(u, v) = \frac{\pi}{2} = 90°$.

Example

Let $v = \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$ and $w = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$.

(i) Show that $v \perp w$.

(ii) Give an example of another vector that is linearly independent of v and w, for which is orthogonal to v.

Pythagorean Theorem in $\mathbb{R}^n$

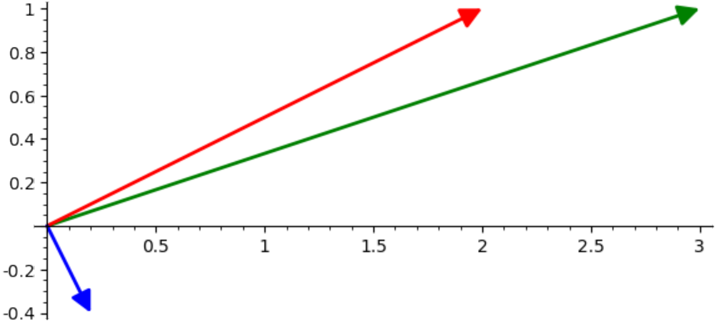
If $u \perp v$, then $\|u+v\|^2 = \|u\|^2 + \|v\|^2$.

It will often be *extremely* useful to by able to construct a vector that is orthogonal to some given one.

Fact

If u and v are vectors in $\mathbb{R}^n$, then $w = u - \dfrac{u \cdot v}{v \cdot v} v$ is orthogonal to v.

Example: If $u = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$ and $v = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ then $w = \begin{bmatrix} \frac{1}{5} \\ -\frac{2}{5} \end{bmatrix}$.



MA313
Week 9: (DRAFT) Inner Products and Orthogonality

Start of ...

PART 4: Cauchy-Schwarz Inequality

Part 5: Cauchy-Schwarz Inequality

The Cauchy-Schwarz inequality

Let $u, v \in \mathbb{R}^n$. Then

$$|u \cdot v| \leqslant \|u\|\|v\|$$

with equality if and only if u and v are linearly dependent.

If $u \neq 0 \neq v$, then

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

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

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

actually makes sense!

The Triangle inequality

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

Definition (DISTANCE between vectors)

The **distance** between vectors $u, v \in \mathbb{R}^n$ is

$$\mathrm{dist}(u, v) := \|u - v\|.$$

Example

$$\mathrm{dist}\left(\begin{bmatrix}3\\2\\1\end{bmatrix}, \begin{bmatrix}1\\2\\3\end{bmatrix}\right)$$

This definition of distance make sense

Let $u, v, w \in \mathbb{R}^n$. Then:

- $\operatorname{dist}(u, v) = 0$ if and only if $u = v$.
- $\operatorname{dist}(u, v) = \operatorname{dist}(v, u)$.
- $\operatorname{dist}(u, w) \leqslant \operatorname{dist}(u, v) + \operatorname{dist}(v, w)$.

Exercises

The dimension of a vector space

These exercises are taken from Section 4.6 and 6.1 of the textbook.

Matrices and linear transformations

1. Let us define the linear transformation

$$T\colon M_{2\times 2} \to M_{2\times 2}, \quad A \mapsto A + A^\top$$

   1.1 Show that

$$\mathcal{B} = \left(\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right).$$

   is a basis of $M_{2\times 2}$.

   1.2 Find the matrix of T relative to $\mathcal{B}$.

Exercises

2. Find the matrix of the linear transformation

$$T\colon \mathbb{P}_2 \to \mathbb{P}_2, \quad p(t) \mapsto p(t) + p'(t) + p''(t)$$

relative to the basis $(1, t, t^2)$ of $\mathbb{P}_2$.

3. Find the matrix of the linear transformation

$$T\colon \mathbb{P}_3 \to \mathbb{R}, \quad p(t) \mapsto \int_0^1 p(x)\mathrm{d}x$$

relative to the bases $(1, t, t^2, t^3)$ and (1) of $\mathbb{P}_3$ and $\mathbb{R}$, respectively.

Inner product, length, and orthogonality

4. *6.1.1–6.1.8* Let

$$u = \begin{bmatrix} -1 \\ 2 \end{bmatrix}, \quad v = \begin{bmatrix} 4 \\ 6 \end{bmatrix}, \quad w = \begin{bmatrix} 3 \\ -1 \\ -5 \end{bmatrix}, \quad x = \begin{bmatrix} 6 \\ -2 \\ 3 \end{bmatrix}.$$

Compute

Exercises

4.1 $u \cdot u$, $v \cdot u$, and $\frac{u \cdot v}{u \cdot u}$.

4.2 $w \cdot w$, $x \cdot w$, and $\frac{x \cdot w}{w \cdot w}$.

4.3 $\frac{1}{w \cdot w} w$.

4.4 $\frac{1}{u \cdot u} u$.

4.5 $\frac{u \cdot v}{v \cdot v} v$.

4.6 $\frac{x \cdot w}{x \cdot x} x$.

4.7 $\|w\|$.

4.8 $\|x\|$.

5. *6.1.9–6.1.12* Find a unit vector in the direction of each of the following

$$\begin{bmatrix} -30 \\ 40 \end{bmatrix}, \quad \begin{bmatrix} -6 \\ 4 \\ -3 \end{bmatrix}, \quad \begin{bmatrix} 7/4 \\ 1/2 \\ 1 \end{bmatrix}, \quad \begin{bmatrix} 8/3 \\ 2 \end{bmatrix}.$$

6. *6.1.13* Find the distance between $x = \begin{bmatrix} 10 \\ -3 \end{bmatrix}$ and $y = \begin{bmatrix} -1 \\ -5 \end{bmatrix}$.

7. *6.1.14* Find the distance between $x = \begin{bmatrix} 0 \\ -5 \\ 2 \end{bmatrix}$ and $y = \begin{bmatrix} -4 \\ -1 \\ 8 \end{bmatrix}$.

Exercises

8. *6.1.15–18* In each of the following four cases, determine whether the given pair of vectors are orthogonal.

   8.1

   $$a = \begin{bmatrix} 8 \\ -5 \end{bmatrix}, \quad b = \begin{bmatrix} -2 \\ -3 \end{bmatrix},$$

   8.2

   $$u = \begin{bmatrix} 12 \\ 3 \\ -5 \end{bmatrix}, \quad v = \begin{bmatrix} 2 \\ -3 \\ 3 \end{bmatrix},$$

   8.3

   $$u = \begin{bmatrix} 3 \\ 2 \\ -5 \\ 0 \end{bmatrix}, \quad v = \begin{bmatrix} -4 \\ 1 \\ -2 \\ 6 \end{bmatrix},$$

Exercises

8.4

$$y = \begin{bmatrix} -3 \\ 7 \\ 4 \\ 0 \end{bmatrix}, \quad z = \begin{bmatrix} 1 \\ -8 \\ 15 \\ -7 \end{bmatrix}.$$