ROB 101 - Fall 2021

The Vector Space $\mathbb{R}^n$: Part 2

(Dot Product and Orthonormal Vectors)

October 13, 2021

UNIVERSITY OF MICHIGAN

- A second encounter with some of the essential concepts in Linear Algebra.
- A more abstract view of $\mathbb{R}^n$ as a vector space.

Outcomes

- A second encounter with some of the essential concepts in Linear Algebra.
- Gram Schmidt process for generating a basis consisting of orthogonal vectors.
- Orthogonal matrices: they have the magical property that their inverse is the matrix transpose.

Definition

Let $u \in \mathbb{R}^n$ and $v \in \mathbb{R}^n$ be column vectors,

$$u := \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}, \quad v := \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}.$$

The dot product of u and v is defined as

$$u \bullet v := \sum_{k=1}^{n} u_k v_k.$$

Remark

We note that

$$u^{\mathsf{T}} \cdot v = \begin{bmatrix} u_1 & u_2 & \cdots & u_n \end{bmatrix} \cdot \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = \sum_{k=1}^{n} u_k v_k =: u \bullet v.$$

Remark

We note that

$$u^{\mathsf{T}} \cdot v = \begin{bmatrix} u_1 & u_2 & \cdots & u_n \end{bmatrix} \cdot \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = \sum_{k=1}^{n} u_k v_k =: u \bullet v.$$

The dot product is also called the inner product.

Compute the dot product for

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

Compute the dot product for

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

$$u \bullet v = (1)(2) + (0)(4) + (3)(1) = 5$$

$$u^{\mathsf{T}} v = (1)(2) + (0)(4) + (3)(1) = 5.$$

You can use either notation.

Compute the inner product for

$$u = \begin{bmatrix} 1 \\ 0 \\ -1 \\ 0 \end{bmatrix}, \quad v = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix}.$$

Compute the inner product for

$$u = \begin{bmatrix} 1 \\ 0 \\ -1 \\ 0 \end{bmatrix}, \quad v = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix}.$$

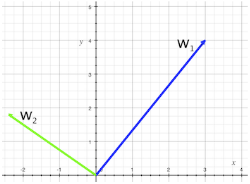
$$u \bullet v = (1)(0) + (0)(1) + (-1)(0) + (0)(1) = 0$$

$$u^{\mathsf{T}} v = (1)(0) + (0)(1) + (-1)(0) + (0)(1) = 0.$$

The inner product will provide us a generalization of a right angle ($90 \deg$ angle) between two vectors in $\mathbb{R}^n$.

$$w_1 \perp w_2 \iff w_1 \bullet w_2 = 0 \iff w_1^\mathsf{T} w_2 = 0$$

(Read it as: w_1 is orthogonal to w_2 if, and only if, their inner product is zero. Orthogonal means "at right angle")



Source: `https://study.com/academy/lesson/the-gram-schmidt-process-for-orthonormalizing-vectors.html`

Reading the values from the graph, we have

$$w_1 = \begin{bmatrix} 3 \\ 4 \end{bmatrix}, w_2 = \begin{bmatrix} -\frac{7}{3} \\ \frac{7}{4} \end{bmatrix}$$

$$\implies w_1 \bullet w_2 = w_1^{\mathsf{T}} w_2 = -3\frac{7}{3} + 4\frac{7}{4} = 0$$

Theorem (Pythagorean Theorem)

Suppose that $w_1 \perp w_2$*. Then,*

$$\|w_1 + w_2\|^2 = \|w_1\|^2 + \|w_2\|^2.$$

Proof.

Because $w_1 \perp w_2$, we know that $w_1 \bullet w_2 = 0$, which means that $w_1^\mathsf{T} \cdot w_2 = w_2^\mathsf{T} \cdot w_1 = 0$.

Proof.

Because $w_1 \perp w_2$, we know that $w_1 \bullet w_2 = 0$, which means that $w_1^\mathsf{T} \cdot w_2 = w_2^\mathsf{T} \cdot w_1 = 0$.

Recall that the norm-squared of a vector v is $\|v\|^2 = v^\mathsf{T} \cdot v$.

Proof.

Because $w_1 \perp w_2$, we know that $w_1 \bullet w_2 = 0$, which means that $w_1^\mathsf{T} \cdot w_2 = w_2^\mathsf{T} \cdot w_1 = 0$.

Recall that the norm-squared of a vector v is $\|v\|^2 = v^\mathsf{T} \cdot v$.

$$\begin{aligned}
\|w_1 + w_2\|^2 := & (w_1 + w_2)^\mathsf{T} \cdot (w_1 + w_2) \\
= & w_1^\mathsf{T} \cdot (w_1 + w_2) + w_2^\mathsf{T} \cdot (w_1 + w_2) \\
= & w_1^\mathsf{T} \cdot w_1 + w_1^\mathsf{T} \cdot w_2 + w_2^\mathsf{T} \cdot w_1 + w_2^\mathsf{T} \cdot w_2 \\
= & \underbrace{w_1^\mathsf{T} \cdot w_1}_{||w_1||^2} + \underbrace{w_1^\mathsf{T} \cdot w_2}_{0} + \underbrace{w_2^\mathsf{T} \cdot w_1}_{0} + \underbrace{w_2^\mathsf{T} \cdot w_2}_{||w_2||^2} \\
= & \|w_1\|^2 + \|w_2\|^2.
\end{aligned}$$

$\square$

Determine which pairs of vectors, if any, are orthogonal

$$u = \begin{bmatrix} 2 \\ 1 \\ -1 \end{bmatrix}, v = \begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix}, w = \begin{bmatrix} -5 \\ 0 \\ 1 \end{bmatrix}.$$

Example

$$u^\mathsf{T} \cdot v = \begin{bmatrix} 2 & 1 & -1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix} = (2)(1) + (1)(3) + (-1)(5) = 0$$

$$u^\mathsf{T} \cdot w = \begin{bmatrix} 2 & 1 & -1 \end{bmatrix} \cdot \begin{bmatrix} -5 \\ 0 \\ 1 \end{bmatrix} = (2)(-5) + (1)(0) + (-1)(1) = -11$$

$$v^\mathsf{T} \cdot w = \begin{bmatrix} 1 & 3 & 5 \end{bmatrix} \cdot \begin{bmatrix} -5 \\ 0 \\ 1 \end{bmatrix} = (1)(-5) + (3)(0) + (5)(1) = 0.$$

Example

$$u^{\mathsf{T}} \cdot v = \begin{bmatrix} 2 & 1 & -1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix} = 0$$

$$u^{\mathsf{T}} \cdot w = \begin{bmatrix} 2 & 1 & -1 \end{bmatrix} \cdot \begin{bmatrix} -5 \\ 0 \\ 1 \end{bmatrix} = -11$$

$$v^{\mathsf{T}} \cdot w = \begin{bmatrix} 1 & 3 & 5 \end{bmatrix} \cdot \begin{bmatrix} -5 \\ 0 \\ 1 \end{bmatrix} = 0.$$

Hence, $u \perp v$, $u \not\perp w$, and $v \perp w$. In words, u is orthogonal to v, u is not orthogonal to w, and v is orthogonal to w.

Definition

A set of vectors $\{v_1, v_2, \ldots, v_n\}$ is orthogonal if, for all $1 \le i, j \le n$, and $i \neq j$

$$v_i \bullet v_j = 0.$$

We can also write this as $v_i^\mathsf{T} v_j = 0$ or $v_i \perp v_j$.

Orthogonal and Orthonormal Vectors

Definition

A set of vectors $\{v_1, v_2, \ldots, v_n\}$ is orthonormal if,

- they are orthogonal, and
- for all $1 \leq i \leq n$, $\|v_i\| = 1$.

Scale the vector w so that its norm becomes one,

$$w = \begin{bmatrix} -5 \\ 0 \\ 1 \end{bmatrix}.$$

We need to form $\frac{w}{\|w\|}$ $(\|w\| \neq 0)$, which gives

$$\tilde{w} := \frac{1}{\|w\|} \cdot w = \frac{1}{\sqrt{26}} \begin{bmatrix} -5 \\ 0 \\ 1 \end{bmatrix}.$$

Given the orthogonal set $\{u, v\}$, make it an orthonormal set.

$$u = \begin{bmatrix} 2 \\ 1 \\ -1 \end{bmatrix}, v = \begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix}.$$

We need to normalize their lengths to one. We compute

$$\|u\| = \sqrt{(2)^2 + (1)^2 + (-1)^2} = \sqrt{6}$$
$$\|v\| = \sqrt{(1)^2 + (3)^2 + (5)^2} = \sqrt{35}$$

We need to normalize their lengths to one. We compute

$$\|u\| = \sqrt{(2)^2 + (1)^2 + (-1)^2} = \sqrt{6}$$
$$\|v\| = \sqrt{(1)^2 + (3)^2 + (5)^2} = \sqrt{35}$$

and thus

$$\left\{ \tilde{u} := \frac{1}{\sqrt{6}} \begin{bmatrix} 2 \\ 1 \\ -1 \end{bmatrix}, \tilde{v} := \frac{1}{\sqrt{35}} \begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix} \right\}$$

is an orthonormal set of vectors.

Orthonormal Vectors are Linearly Independent

Fact

For a set of vectors $\{v_1, v_2, \ldots, v_k\}$ *in* $\mathbb{R}^n$*, the following statements are true:*

1. $\{v_1, v_2, \ldots, v_k\}$ *orthonormal implies it is linearly independent.*

2. $\{v_1, v_2, \ldots, v_k\}$ *orthogonal and for all* i*,* $v_i \neq 0_{n\times 1}$*, together imply that the set is linearly independent.*

Fact

An $n \times m$ rectangular matrix Q is orthonormal:

- *if $n > m$ (tall matrix), its columns are orthonormal vectors, which is equivalent to $Q^{\mathsf{T}} \cdot Q = I_m$; and*

Fact

An $n \times m$ rectangular matrix Q is orthonormal:

- *if $n > m$ (tall matrix), its columns are orthonormal vectors, which is equivalent to $Q^{\mathsf{T}} \cdot Q = I_m$; and*
- *if $n < m$ (wide matrix), its rows are orthonormal vectors, which is equivalent to $Q \cdot Q^{\mathsf{T}} = I_n$.*

Fact

An $n \times m$ rectangular matrix Q is orthonormal:

- *if $n > m$ (tall matrix), its columns are orthonormal vectors, which is equivalent to $Q^{\mathsf{T}} \cdot Q = I_m$; and*
- *if $n < m$ (wide matrix), its rows are orthonormal vectors, which is equivalent to $Q \cdot Q^{\mathsf{T}} = I_n$.*
- *A square $n \times n$ matrix is orthogonal if $Q^{\mathsf{T}} \cdot Q = I_n$ and $Q \cdot Q^{\mathsf{T}} = I_n$, and hence, $Q^{-1} = Q^{\mathsf{T}}$.*

Remark

- *For a square matrix,* $n = m$*,*
  $(Q^{\mathsf{T}} \cdot Q = I_n) \iff (Q \cdot Q^{\mathsf{T}} = I_n) \iff (Q^{-1} = Q^{\mathsf{T}})$.

- *For a tall matrix,* $n > m$*,*
  $(Q^{\mathsf{T}} \cdot Q = I_m) \not\Longrightarrow (Q \cdot Q^{\mathsf{T}} = I_n)$.

- *For a wide matrix,* $m > n$*,*
  $(Q \cdot Q^{\mathsf{T}} = I_n) \not\Longrightarrow (Q^{\mathsf{T}} \cdot Q = I_m)$.

Constructing Orthonormal Vectors: the Gram-Schmidt Process

Suppose that that the set of vectors $\{u_1, u_2, \ldots, u_m\}$ is linearly independent and you generate a new set of vectors by

$$v_1 = u_1$$

Constructing Orthonormal Vectors: the Gram-Schmidt Process

Suppose that that the set of vectors $\{u_1, u_2, \ldots, u_m\}$ is linearly independent and you generate a new set of vectors by

$$v_1 = u_1$$

$$v_2 = u_2 - \left(\frac{u_2 \bullet v_1}{v_1 \bullet v_1}\right) v_1$$

Constructing Orthonormal Vectors: the Gram-Schmidt Process

Remark

$\left(\frac{u \bullet v}{v \bullet v}\right) v$ projects the vector u orthogonally onto the line spanned by vector v.

$$\left(\frac{u \bullet v}{v \bullet v}\right) v = \left(\frac{u \bullet v}{\|v\|^2}\right) v = \left(u \bullet \frac{v}{\|v\|}\right) \frac{v}{\|v\|}$$

Constructing Orthonormal Vectors: the Gram-Schmidt Process

Remark

$\left(\frac{u \bullet v}{v \bullet v}\right) v$ projects the vector u orthogonally onto the line spanned by vector v.

$$\left(\frac{u \bullet v}{v \bullet v}\right) v = \left(\frac{u \bullet v}{\|v\|^2}\right) v = \left(u \bullet \frac{v}{\|v\|}\right) \frac{v}{\|v\|}$$

- $e := \frac{v}{\|v\|}$ *is the unit vector along* v.

Constructing Orthonormal Vectors: the Gram-Schmidt Process

Remark

$\left(\frac{u \bullet v}{v \bullet v}\right) v$ projects the vector u orthogonally onto the line spanned by vector v.

$$\left(\frac{u \bullet v}{v \bullet v}\right) v = \left(\frac{u \bullet v}{\|v\|^2}\right) v = \left(u \bullet \frac{v}{\|v\|}\right) \frac{v}{\|v\|}$$

- $e := \frac{v}{\|v\|}$ *is the unit vector along v.*
- $\left(u \bullet \frac{v}{\|v\|}\right) \frac{v}{\|v\|} = (u \bullet e)\, e$.

Constructing Orthonormal Vectors: the Gram-Schmidt Process

Suppose that that the set of vectors $\{u_1, u_2, \ldots, u_m\}$ is linearly independent and you generate a new set of vectors by

$$v_1 = u_1, \qquad e_1 = \frac{v_1}{\|v_1\|}$$

$$v_2 = u_2 - \left(\frac{u_2 \bullet v_1}{v_1 \bullet v_1}\right) v_1, \qquad e_2 = \frac{v_2}{\|v_2\|}$$

Constructing Orthonormal Vectors: the Gram-Schmidt Process

Suppose that that the set of vectors $\{u_1, u_2, \ldots, u_m\}$ is linearly independent and you generate a new set of vectors by

$$v_1 = u_1, \qquad e_1 = \frac{v_1}{\|v_1\|}$$

$$v_2 = u_2 - \left(\frac{u_2 \bullet v_1}{v_1 \bullet v_1}\right) v_1, \qquad e_2 = \frac{v_2}{\|v_2\|}$$

$$v_3 = u_3 - \left(\frac{u_3 \bullet v_1}{v_1 \bullet v_1}\right) v_1 - \left(\frac{u_3 \bullet v_2}{v_2 \bullet v_2}\right) v_2, \qquad e_3 = \frac{v_3}{\|v_3\|}$$

$$\vdots$$

Constructing Orthonormal Vectors: the Gram-Schmidt Process

Suppose that that the set of vectors $\{u_1, u_2, \ldots, u_m\}$ is linearly independent and you generate a new set of vectors by

$$
\begin{aligned}
v_1 &= u_1, \qquad e_1 = \frac{v_1}{\|v_1\|} \\
v_2 &= u_2 - \left(\frac{u_2 \bullet v_1}{v_1 \bullet v_1}\right) v_1, \qquad e_2 = \frac{v_2}{\|v_2\|} \\
v_3 &= u_3 - \left(\frac{u_3 \bullet v_1}{v_1 \bullet v_1}\right) v_1 - \left(\frac{u_3 \bullet v_2}{v_2 \bullet v_2}\right) v_2, \qquad e_3 = \frac{v_3}{\|v_3\|} \\
&\vdots \\
v_k &= u_k - \sum_{i=1}^{k-1} \left(\frac{u_k \bullet v_i}{v_i \bullet v_i}\right) v_i, \qquad e_k = \frac{v_k}{\|v_k\|} \quad \text{(General Step)}
\end{aligned}
$$

Then the set of vectors $\{v_1, v_2, \ldots, v_m\}$ is

- orthogonal, meaning, $i \neq j \implies v_i \bullet v_j = 0$
- span preserving, meaning that, for all $1 \leq k \leq m$,

$$\operatorname{span}\{v_1, v_2, \ldots, v_k\} = \operatorname{span}\{u_1, u_2, \ldots, u_k\},$$

- and linearly independent.

Remark

The unit vectors $\{e_1, e_2, \ldots, e_m\}$ *form orthonormal set.*

See Example 9.17 in ROB 101 Book. Try to visualize your 3D vectors before and after applying the Gram-Schmidt process.

Next Time

- The Vector Space $\mathbb{R}^n$: Part 2 (QR Factorization)
- Read Chapter 9 of ROB 101 Book