Introduction to Math for Artificial Introduction

Orthonormal Basis and Eigenvectors

Andres Mendez-Vazquez

March 23, 2020

Outline

Cinvestav

Outline

Cinvestav

The Dot Product

Definition

The dot product of two vectors $\boldsymbol{v} = [v_1, v_2, ..., v_n]^T$ and $\boldsymbol{w} = [w_1, w_2, ..., w_n]^T$

$$\boldsymbol{v} \cdot \boldsymbol{w} = \sum_{i=1}^{n} v_i w_i$$

Cinvestav

Example!!! Splitting the Space?

For example, assume the following vector $\boldsymbol{w}$ and constant w_0

$$\boldsymbol{w} = (-1, 2)^T \text{ and } w_0 = 0$$

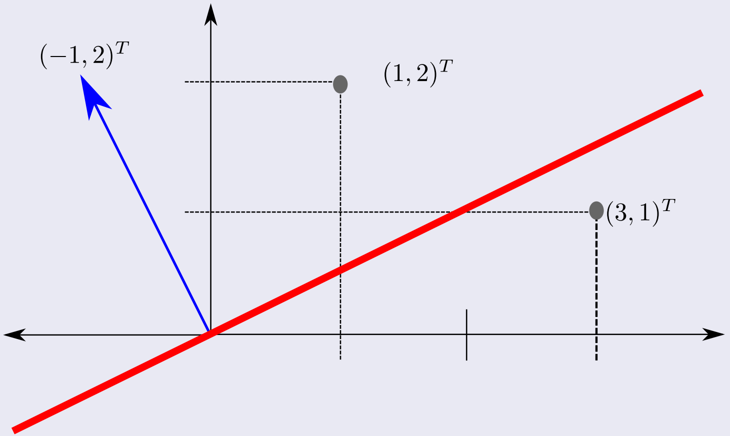
Hyperplane

Example!!! Splitting the Space?

For example, assume the following vector $\boldsymbol{w}$ and constant w_0

$$\boldsymbol{w} = (-1, 2)^T \text{ and } w_0 = 0$$

Hyperplane



Then, we have

The following results

$$g\left(\left(\begin{array}{c}1\\2\end{array}\right)\right) = (-1,2)\left(\begin{array}{c}1\\2\end{array}\right) = -1\times 1 + 2\times 2 = 3$$

$$g\left(\left(\begin{array}{c}3\\1\end{array}\right)\right) = (-1,2)\left(\begin{array}{c}3\\1\end{array}\right) = -1\times 3 + 2\times 1 = -1$$

YES!!! We have a positive side and a negative side!!!

This product is also know as the Inner Product

Where

An inner product $\langle \cdots, \cdots \rangle$ satisfies the following four properties ($\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}$ vectors and α a escalar):

1. $\langle \boldsymbol{u} + \boldsymbol{v}, \boldsymbol{w} \rangle = \langle \boldsymbol{u}, \boldsymbol{w} \rangle + \langle \boldsymbol{v}, \boldsymbol{w} \rangle$
2. $\langle \alpha \boldsymbol{v}, \boldsymbol{w} \rangle = \alpha \langle \boldsymbol{v}, \boldsymbol{w} \rangle$
3. $\langle \boldsymbol{v}, \boldsymbol{w} \rangle = \langle \boldsymbol{w}, \boldsymbol{v} \rangle$
4. $\langle \boldsymbol{v}, \boldsymbol{v} \rangle \geq 0$ and equal to zero if $\boldsymbol{v} = \boldsymbol{0}$.

Cinvestav

Outline

Cinvestav

The Norm as a dot product

We can define the longitude of a vector

$$\|\boldsymbol{v}\| = \sqrt{\boldsymbol{v} \cdot \boldsymbol{v}}$$

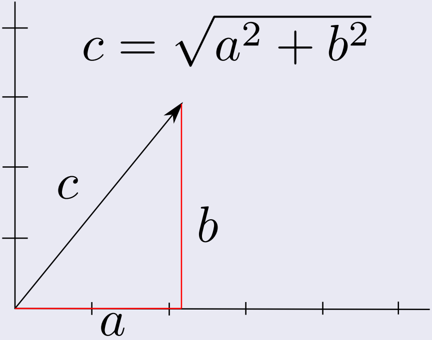
A nice way to think about the longitude of a vector

The Norm as a dot product

We can define the longitude of a vector

$$\|\boldsymbol{v}\| = \sqrt{\boldsymbol{v} \cdot \boldsymbol{v}}$$

A nice way to think about the longitude of a vector



Orthogonal Vectors

We have that

Two vectors are orthogonal when their dot product is zero:

$$\boldsymbol{v} \cdot \boldsymbol{w} = 0 \text{ or } \boldsymbol{v}^T \boldsymbol{w} = 0$$

Remark

We want orthogonal bases and orthogonal sub-spaces.

Orthogonal Vectors

We have that

Two vectors are orthogonal when their dot product is zero:

$$\boldsymbol{v} \cdot \boldsymbol{w} = 0 \text{ or } \boldsymbol{v}^T \boldsymbol{w} = 0$$

Remark

We want orthogonal bases and orthogonal sub-spaces.

Some stuff about Row and Null Space

Something Notable

Every row of A is perpendicular to every solution of $A\boldsymbol{x} = 0$

In a similar way

Every column of A is perpendicular to every solution of $A^T\boldsymbol{x} = 0$

Meaning

What are the implications for the Column and Row Space?

Some stuff about Row and Null Space

Something Notable

Every row of A is perpendicular to every solution of $A\boldsymbol{x} = 0$

In a similar way

Every column of A is perpendicular to every solution of $A^T\boldsymbol{x} = 0$

Meaning

What are the implications for the Column and Row Space?

Some stuff about Row and Null Space

Something Notable

Every row of A is perpendicular to every solution of $A\boldsymbol{x} = 0$

In a similar way

Every column of A is perpendicular to every solution of $A^T\boldsymbol{x} = 0$

Meaning

What are the implications for the Column and Row Space?

Implications

We have that under $Ax = b$

$$e = b - Ax$$

Remember

The error at the Least Squared Error.

Implications

We have that under $A\boldsymbol{x} = b$

$$e = b - A\boldsymbol{x}$$

Remember

The error at the Least Squared Error.

Orthogonal Spaces

Definition

Two sub-spaces V and W of a vector space are orthogonal if every vector $\boldsymbol{v} \in V$ is perpendicular to every vector $\boldsymbol{w} \in W$.

In mathematical notation

$$\boldsymbol{v}^T \boldsymbol{w} = \boldsymbol{0} \ \forall \boldsymbol{v} \in V \text{ and } \forall \boldsymbol{w} \in W$$

Orthogonal Spaces

Definition

Two sub-spaces V and W of a vector space are orthogonal if every vector $\boldsymbol{v} \in V$ is perpendicular to every vector $\boldsymbol{w} \in W$.

In mathematical notation

$$\boldsymbol{v}^T \boldsymbol{w} = \mathbf{0} \ \forall \boldsymbol{v} \in V \text{ and } \forall \boldsymbol{w} \in W$$

Cinvestav

Examples

At your Room

The floor of your room (extended to infinity) is a subspace V. The line where two walls meet is a subspace W (one-dimensional).

A more convoluted example

Two walls look perpendicular but they are not orthogonal sub-spaces!

Why?

Any Idea?

Cinvestav

Examples

At your Room

The floor of your room (extended to infinity) is a subspace V. The line where two walls meet is a subspace W (one-dimensional).

A more convoluted example

Two walls look perpendicular but they are not orthogonal sub-spaces!

Why?

Any Idea?

Examples

At your Room

The floor of your room (extended to infinity) is a subspace V. The line where two walls meet is a subspace W (one-dimensional).

A more convoluted example

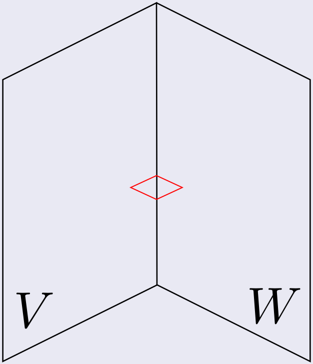
Two walls look perpendicular but they are not orthogonal sub-spaces!

Why?

Any Idea?

For Example

Something Notable



Yes!!

The Line Shared by the Two Planes in $\mathbb{R}^3$

Therefore!!!

We have then

Theorem

The Null Space $N(A)$ and the Row Space $C\left(A^T\right)$, as the column space of A^T, are orthogonal sub-spaces in $\mathbb{R}^n$

Proof

First, we have

$$Ax = \begin{bmatrix} A_1 \\ A_2 \\ \vdots \\ A_m \end{bmatrix} x = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Therefore

Rows in A are perpendicular to $x \Rightarrow$ Then x is also perpendicular to every combination of the rows.

We have then

Theorem

The Null Space $N(A)$ and the Row Space $C\left(A^T\right)$, as the column space of A^T, are orthogonal sub-spaces in $\mathbb{R}^n$

Proof

First, we have

$$A\boldsymbol{x} = \begin{bmatrix} A_1 \\ A_2 \\ \vdots \\ A_m \end{bmatrix} \boldsymbol{x} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Therefore

Rows in A are perpendicular to $x \Rightarrow$ Then x is also perpendicular to every combination of the rows.

We have then

Theorem

The Null Space $N(A)$ and the Row Space $C\left(A^T\right)$, as the column space of A^T, are orthogonal sub-spaces in $\mathbb{R}^n$

Proof

First, we have

$$A\boldsymbol{x} = \begin{bmatrix} A_1 \\ A_2 \\ \vdots \\ A_m \end{bmatrix} \boldsymbol{x} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Therefore

Rows in A are perpendicular to $\boldsymbol{x} \Rightarrow$ Then $\boldsymbol{x}$ is also perpendicular to every combination of the rows.

Then

Therefore

The whole row space is orthogonal to the $N(A)$

$A^T y$ are all the possible combinations of the row space!!!

Then

Therefore

The whole row space is orthogonal to the $N(A)$

Better proof $\boldsymbol{x} \in N(A)$- Hint What is $A^T\boldsymbol{y}$?

$$\boldsymbol{x}\left(A^T\boldsymbol{y}\right) = (A\boldsymbol{x})^T\boldsymbol{y} = \boldsymbol{0}^T\boldsymbol{y} = 0$$

$A^T y$ are all the possible combinations of the row space!!!

Cinvestav

Then

Therefore

The whole row space is orthogonal to the $N(A)$

Better proof $\boldsymbol{x} \in N(A)$- Hint What is $A^T\boldsymbol{y}$?

$$\boldsymbol{x}\left(A^T\boldsymbol{y}\right) = (A\boldsymbol{x})^T\boldsymbol{y} = \boldsymbol{0}^T\boldsymbol{y} = 0$$

$A^T\boldsymbol{y}$ are all the possible combinations of the row space!!!

Cinvestav

A little Bit of Notation

We use the following notation

$$N(A) \perp C\left(A^T\right)$$

Definition

The orthogonal complement of a subspace V contains every vector that is perpendicular to V.

This orthogonal subspace is denoted by

$$V^{\perp}$$

A little Bit of Notation

We use the following notation

$$N(A) \perp C\left(A^T\right)$$

Definition

The orthogonal complement of a subspace V contains every vector that is perpendicular to V.

This orthogonal subspace is denoted by

$$V^{\perp}$$

Cinvestav

A little Bit of Notation

We use the following notation

$$N(A) \perp C\left(A^T\right)$$

Definition

The orthogonal complement of a subspace V contains every vector that is perpendicular to V.

This orthogonal subspace is denoted by

$$V^{\perp}$$

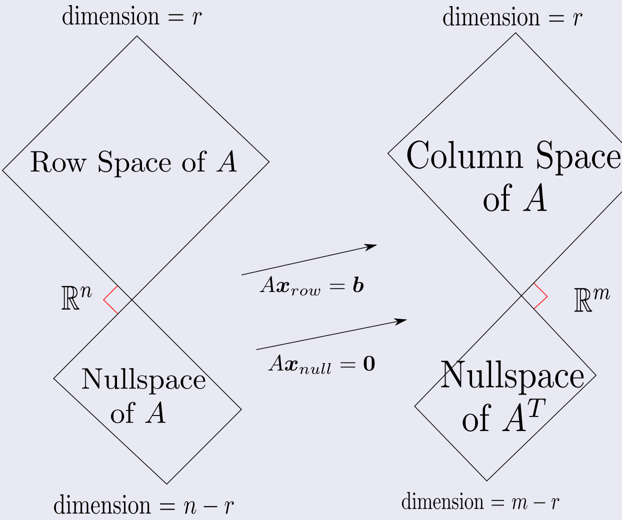
Thus, we have

Something Notable

By this definition, the nullspace is the orthogonal complement of the row space.

Look at this

The Orthogonality



Orthogonal Complements

Definition

- The orthogonal complement of a subspace V contains every vector that is perpendicular to V.
- This orthogonal subspace is denoted by $V^{\perp}$, pronounced "V prep".

Cinvestav

Orthogonal Complements

Definition

- The orthogonal complement of a subspace V contains every vector that is perpendicular to V.
- This orthogonal subspace is denoted by $V^{\perp}$, pronounced "V prep".

Something Notable

By this definition, the nullspace is the orthogonal complement of the row space.

Orthogonal Complements

Definition

- The orthogonal complement of a subspace V contains every vector that is perpendicular to V.
- This orthogonal subspace is denoted by $V^{\perp}$, pronounced "V prep".

Something Notable

By this definition, the nullspace is the orthogonal complement of the row space.

After All

Every x that is perpendicular to the rows satisfies $Ax = 0$.

Cinvestav

Orthogonal Complements

Definition

- The orthogonal complement of a subspace V contains every vector that is perpendicular to V.
- This orthogonal subspace is denoted by $V^{\perp}$, pronounced "V prep".

Something Notable

By this definition, the nullspace is the orthogonal complement of the row space.

After All

Every x that is perpendicular to the rows satisfies $Ax = \mathbf{0}$.

Cinvestav

Quite Interesting

We have the following

If v is orthogonal to the nullspace, it must be in the row space.

Therefore, we can build a new matrix

$$A' = \begin{bmatrix} A \\ v \end{bmatrix}$$

Problem

The row space starts to grow and can break the law $dim\left(R(A)\right) + dim\left(Ker\left(A\right)\right) = n.$

Cinvestav

Quite Interesting

We have the following

If $\boldsymbol{v}$ is orthogonal to the nullspace, it must be in the row space.

Therefore, we can build a new matrix

$$A' = \begin{bmatrix} A \\ \boldsymbol{v} \end{bmatrix}$$

Problem

The row space starts to grow and can break the law $dim\left(R(A)\right) + dim\left(Ker\left(A\right)\right) = n$.

Quite Interesting

We have the following

If $\boldsymbol{v}$ is orthogonal to the nullspace, it must be in the row space.

Therefore, we can build a new matrix

$$A' = \begin{bmatrix} A \\ \boldsymbol{v} \end{bmatrix}$$

Problem

The row space starts to grow and can break the law
$dim\left(R(A)\right) + dim\left(Ker\left(A\right)\right) = n.$

Additionally

The left nullspace and column space are orthogonal in $\mathbb{R}^m$

Basically, they are orthogonal complements.

As always

Their dimensions $dim\left(Ker\left(A^T\right)\right)$ and $dim\left(R\left(A^T\right)\right)$ add to the full dimension m.

Additionally

The left nullspace and column space are orthogonal in $\mathbb{R}^m$

Basically, they are orthogonal complements.

As always

Their dimensions $dim\left(Ker\left(A^T\right)\right)$ and $dim\left(R\left(A^T\right)\right)$ add to the full dimension m.

We have

Theorem

- The column space and row space both have dimension r.
  - The nullspaces have dimensions $n - r$ and $m - r$.

Theorem

The nullspace of A is the orthogonal complement of the row space $C\left(A^T\right)$ - $\mathbb{R}^n$.

Theorem

The null space of A^T is the orthogonal complement of the column space $C(A)$ - $\mathbb{R}^m$.

We have

Theorem

- The column space and row space both have dimension r.
  - The nullspaces have dimensions $n - r$ and $m - r$.

Theorem

The nullspace of A is the orthogonal complement of the row space $C\left(A^T\right)$ - $\mathbb{R}^n$.

Theorem

The null space of A^T is the orthogonal complement of the column space $C\left(A\right)$ - $\mathbb{R}^m$.

We have

Theorem

- The column space and row space both have dimension r.
  - The nullspaces have dimensions $n-r$ and $m-r$.

Theorem

The nullspace of A is the orthogonal complement of the row space $C\left(A^T\right)$ - $\mathbb{R}^n$.

Theorem

The null space of A^T is the orthogonal complement of the column space $C\left(A\right)$ - $\mathbb{R}^m$.

Cinvestav

Splitting the Vectors

The point of "complements"

$\boldsymbol{x}$ can be split into a row space component $\boldsymbol{x}_r$ and a nullspace component $\boldsymbol{x}_n$:

$$\boldsymbol{x} = \boldsymbol{x}_r + \boldsymbol{x}_n$$

Therefore

$$A\boldsymbol{x} = A\left[\boldsymbol{x}_r + \boldsymbol{x}_n\right] = A\boldsymbol{x}_r + A\boldsymbol{x}_n = A\boldsymbol{x}_r$$

Basically

Every vector goes to the column space.

Splitting the Vectors

The point of "complements"

$\boldsymbol{x}$ can be split into a row space component $\boldsymbol{x}_r$ and a nullspace component $\boldsymbol{x}_n$:

$$\boldsymbol{x} = \boldsymbol{x}_r + \boldsymbol{x}_n$$

Therefore

$$A\boldsymbol{x} = A\left[\boldsymbol{x}_r + \boldsymbol{x}_n\right] = A\boldsymbol{x}_r + A\boldsymbol{x}_n = A\boldsymbol{x}_r$$

Basically

Every vector goes to the column space.

Splitting the Vectors

The point of "complements"

$\boldsymbol{x}$ can be split into a row space component $\boldsymbol{x}_r$ and a nullspace component $\boldsymbol{x}_n$:

$$\boldsymbol{x} = \boldsymbol{x}_r + \boldsymbol{x}_n$$

Therefore

$$A\boldsymbol{x} = A\left[\boldsymbol{x}_r + \boldsymbol{x}_n\right] = A\boldsymbol{x}_r + A\boldsymbol{x}_n = A\boldsymbol{x}_r$$

Basically

Every vector goes to the column space.

Not only that

Every vector b in the column space

It comes from one and only one vector in the row space.

- It is also in the row space
- Given that the nullspace and the row space are orthogonal
- They only share the vector 0

Not only that

Every vector $\boldsymbol{b}$ in the column space

It comes from one and only one vector in the row space.

Proof

- If $A\boldsymbol{x}_r = A\boldsymbol{x}_r^{'} \longrightarrow$ the difference is in the nullspace $\boldsymbol{x}_r - \boldsymbol{x}_r^{'}$.
- It is also in the row space...
- Given that the nullspace and the row space are orthogonal.
- They only share the vector 0.

Not only that

Every vector $\boldsymbol{b}$ in the column space

It comes from one and only one vector in the row space.

Proof

- If $A\boldsymbol{x}_r = A\boldsymbol{x}_r^{'} \longrightarrow$ the difference is in the nullspace $\boldsymbol{x}_r - \boldsymbol{x}_r^{'}$.
- It is also in the row space...
- Given that the nullspace and the row space are orthogonal.
- They only share the vector 0.

Not only that

Every vector $\boldsymbol{b}$ in the column space

It comes from one and only one vector in the row space.

Proof

- If $A\boldsymbol{x}_r = A\boldsymbol{x}_r^{'} \longrightarrow$ the difference is in the nullspace $\boldsymbol{x}_r - \boldsymbol{x}_r^{'}$.
- It is also in the row space...
- Given that the nullspace and the row space are orthogonal.
- They only share the vector 0.

Cinvestav

Not only that

Every vector $\boldsymbol{b}$ in the column space

It comes from one and only one vector in the row space.

Proof

- If $A\boldsymbol{x}_r = A\boldsymbol{x}_r^{'} \longrightarrow$ the difference is in the nullspace $\boldsymbol{x}_r - \boldsymbol{x}_r^{'}$.
- It is also in the row space...
- Given that the nullspace and the row space are orthogonal.
- They only share the vector $\mathbf{0}$.

And From Here

Something Notable

There is a $r \times r$ invertible matrix there hiding inside A.

If we throwaway the two nullspaces

From the row space to the column space, A is invertible

And From Here

Something Notable

There is a $r \times r$ invertible matrix there hiding inside A.

If we throwaway the two nullspaces

From the row space to the column space, A is invertible

Example

We have the matrix after echelon reduced

$$A = \begin{pmatrix} 3 & 0 & 0 & 0 & 0 \\ 0 & 5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

You have the following invertible matrix

$$B = \begin{pmatrix} 3 & 0 \\ 0 & 5 \end{pmatrix}$$

Example

We have the matrix after echelon reduced

$$A = \begin{pmatrix} 3 & 0 & 0 & 0 & 0 \\ 0 & 5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

You have the following invertible matrix

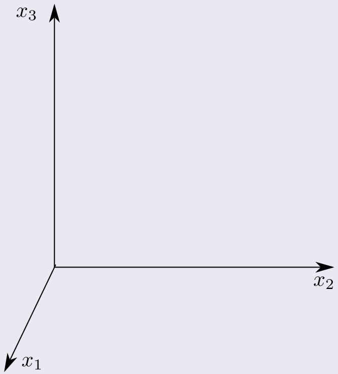
$$B = \begin{pmatrix} 3 & 0 \\ 0 & 5 \end{pmatrix}$$

Outline

Cinvestav

Assume that you are in $\mathbb{R}^3$

Something like



Simple but complex

A simple question

- What are the projections of $b = (2, 3, 4)$ onto the z axis and the xy plane?
- Can we use matrices to talk about these projections?

First

We must have a projection matrix P with the following property:

$$P^2 = P$$

Why?

Ideas?

Simple but complex

A simple question

- What are the projections of $b=(2,3,4)$ onto the z axis and the xy plane?
- Can we use matrices to talk about these projections?

First

We must have a projection matrix P with the following property:

$$P^2 = P$$

Why?

Ideas?

Cinvestav

Simple but complex

A simple question

- What are the projections of $b = (2, 3, 4)$ onto the z axis and the xy plane?
- Can we use matrices to talk about these projections?

First

We must have a projection matrix P with the following property:

$$P^2 = P$$

Why?

Ideas?

Then, the Projection $P\boldsymbol{b}$

First

When $\boldsymbol{b}$ is projected onto a line, its projection $\boldsymbol{p}$ is the part of $\boldsymbol{b}$ along that line.

Second

When $\boldsymbol{b}$ is projected onto a plane, its projection $\boldsymbol{p}$ is the part of the plane.

Then, the Projection $P\boldsymbol{b}$

First

When $\boldsymbol{b}$ is projected onto a line, its projection $\boldsymbol{p}$ is the part of $\boldsymbol{b}$ along that line.

Second

When $\boldsymbol{b}$ is projected onto a plane, its projection $\boldsymbol{p}$ is the part of the plane.

In our case

The Projection Matrices for the coordinate systems

$$P_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, P_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, P_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Example

We have the following vector $\boldsymbol{b} = (2, 3, 4)^T$

Onto the $\boldsymbol{z}$ axis:

$$P_1 \boldsymbol{b} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 2 \\ 3 \\ 4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 4 \end{pmatrix}$$

What about the plane xy

Any idea?

Example

We have the following vector $\boldsymbol{b} = (2,3,4)^T$

Onto the $\boldsymbol{z}$ axis:

$$P_1\boldsymbol{b} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 2 \\ 3 \\ 4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 4 \end{pmatrix}$$

What about the plane xy

Any idea?

Cinvestav

We have something more complex

Something Notable

$$P_4 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Then

$$P_4 b = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 2 \\ 3 \\ 4 \end{pmatrix} = \begin{pmatrix} 2 \\ 3 \\ 0 \end{pmatrix}$$

Cinvestav

We have something more complex

Something Notable

$$P_4 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Then

$$P_4 \boldsymbol{b} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 2 \\ 3 \\ 4 \end{pmatrix} = \begin{pmatrix} 2 \\ 3 \\ 0 \end{pmatrix}$$

Assume the following

We have that

$\boldsymbol{a}_1, \boldsymbol{a}_2, ..., \boldsymbol{a}_n$ in $\mathbb{R}^m$.

Assume they are linearly independent

They span a subspace, we want projections into the subspace

We want to project b into such subspace

How do we do it?

Assume the following

We have that

$\boldsymbol{a}_1, \boldsymbol{a}_2, ..., \boldsymbol{a}_n$ in $\mathbb{R}^m$.

Assume they are linearly independent

They span a subspace, we want projections into the subspace

We want to project b into such subspace

How do we do it?

Assume the following

We have that

$\boldsymbol{a}_1, \boldsymbol{a}_2, ..., \boldsymbol{a}_n$ in $\mathbb{R}^m$.

Assume they are linearly independent

They span a subspace, we want projections into the subspace

We want to project $\boldsymbol{b}$ into such subspace

How do we do it?

Cinvestav

This is the important part

Problem

Find the combination $\boldsymbol{p} = x_1\boldsymbol{a}_1 + x_2\boldsymbol{a}_2 + \cdots + x_n\boldsymbol{a}_n$ closest to vector $\boldsymbol{b}$.

Something Notable

With $n = 1$ (only one vector a_1) this projection onto a line.

This line is the column space of A

Basically the columns are spanned by a single column.

Cinvestav

This is the important part

Problem

Find the combination $\boldsymbol{p} = x_1\boldsymbol{a}_1 + x_2\boldsymbol{a}_2 + \cdots + x_n\boldsymbol{a}_n$ closest to vector $\boldsymbol{b}$.

Something Notable

With $n = 1$ (only one vector a_1) this projection onto a line.

This line is the column space of A

Basically the columns are spanned by a single column.

This is the important part

Problem

Find the combination $\boldsymbol{p} = x_1\boldsymbol{a}_1 + x_2\boldsymbol{a}_2 + \cdots + x_n\boldsymbol{a}_n$ closest to vector $\boldsymbol{b}$.

Something Notable

With $n = 1$ (only one vector a_1) this projection onto a line.

This line is the column space of A

Basically the columns are spanned by a single column.

In General

The matrix has n columns $\boldsymbol{a}_1, \boldsymbol{a}_2, ..., \boldsymbol{a}_n$

The combinations in $\mathbb{R}^m$ are vectors $A\boldsymbol{x}$ in the column space

We are looking for the particular combination

The nearest to the original b

$$\boldsymbol{p} = A\widehat{\boldsymbol{x}}$$

In General

The matrix has n columns $\boldsymbol{a}_1, \boldsymbol{a}_2, ..., \boldsymbol{a}_n$

The combinations in $\mathbb{R}^m$ are vectors $A\boldsymbol{x}$ in the column space

We are looking for the particular combination

The nearest to the original $\boldsymbol{b}$

$$\boldsymbol{p} = A\widehat{\boldsymbol{x}}$$

Cinvestav

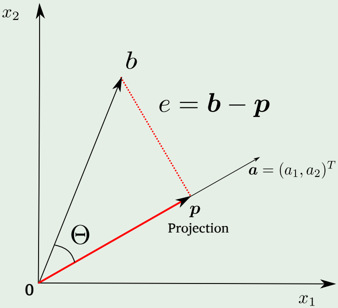
First

We look at the simplest case

The projection into a line...

With a little of Geometry

We have the following



Therefore

Using the fact that the projection is equal to

$$\boldsymbol{p} = x\boldsymbol{a}$$

Then, the error is equal to

$$\boldsymbol{e} = \boldsymbol{b} - x\boldsymbol{a}$$

We have that $\boldsymbol{a} \cdot \boldsymbol{e} = 0$

$$\boldsymbol{a} \cdot \boldsymbol{e} = \boldsymbol{a} \cdot (\boldsymbol{b} - x\boldsymbol{a}) = \boldsymbol{a} \cdot \boldsymbol{b} - x\boldsymbol{a} \cdot \boldsymbol{a} = 0$$

Therefore

Using the fact that the projection is equal to

$$\boldsymbol{p} = x\boldsymbol{a}$$

Then, the error is equal to

$$\boldsymbol{e} = \boldsymbol{b} - x\boldsymbol{a}$$

We have that $\boldsymbol{a} \cdot \boldsymbol{e} = 0$

$$\boldsymbol{a} \cdot \boldsymbol{e} = \boldsymbol{a} \cdot (\boldsymbol{b} - x\boldsymbol{a}) = \boldsymbol{a} \cdot \boldsymbol{b} - x\boldsymbol{a} \cdot \boldsymbol{a} = 0$$

Therefore

Using the fact that the projection is equal to

$$\boldsymbol{p} = x\boldsymbol{a}$$

Then, the error is equal to

$$\boldsymbol{e} = \boldsymbol{b} - x\boldsymbol{a}$$

We have that $\boldsymbol{a} \cdot \boldsymbol{e} = 0$

$$\boldsymbol{a} \cdot \boldsymbol{e} = \boldsymbol{a} \cdot (\boldsymbol{b} - x\boldsymbol{a}) = \boldsymbol{a} \cdot \boldsymbol{b} - x\boldsymbol{a} \cdot \boldsymbol{a} = \boldsymbol{0}$$

Therefore

We have that

$$x = \frac{\boldsymbol{a} \cdot \boldsymbol{b}}{\boldsymbol{a} \cdot \boldsymbol{a}} = \frac{\boldsymbol{a}^T \boldsymbol{b}}{\boldsymbol{a}^T \boldsymbol{a}}$$

Or something quite simple

$$p = \frac{\boldsymbol{a}^T \boldsymbol{b}}{\boldsymbol{a}^T \boldsymbol{a}} a$$

Therefore

We have that

$$x = \frac{\boldsymbol{a} \cdot \boldsymbol{b}}{\boldsymbol{a} \cdot \boldsymbol{a}} = \frac{\boldsymbol{a}^T \boldsymbol{b}}{\boldsymbol{a}^T \boldsymbol{a}}$$

Or something quite simple

$$\boldsymbol{p} = \frac{\boldsymbol{a}^T \boldsymbol{b}}{\boldsymbol{a}^T \boldsymbol{a}} \boldsymbol{a}$$

By the Law of Cosines

Something Notable

$$\|\boldsymbol{a}-\boldsymbol{b}\|^2 = \|\boldsymbol{a}\|^2 + \|\boldsymbol{b}\|^2 - 2\,\|\boldsymbol{a}\|\,\|\boldsymbol{b}\|\cos\Theta$$

We have

The following product

$$\boldsymbol{a}\cdot\boldsymbol{a}-2\boldsymbol{a}\cdot\boldsymbol{b}+\boldsymbol{b}\cdot\boldsymbol{b}=\left\|\boldsymbol{a}\right\|^{2}+\left\|\boldsymbol{b}\right\|^{2}-2\left\|\boldsymbol{a}\right\|\left\|\boldsymbol{b}\right\|\cos\Theta$$

Then

$$\boldsymbol{a}\cdot\boldsymbol{b}=\left\|\boldsymbol{a}\right\|\left\|\boldsymbol{b}\right\|\cos\Theta$$

We have

The following product

$$\boldsymbol{a} \cdot \boldsymbol{a} - 2\boldsymbol{a} \cdot \boldsymbol{b} + \boldsymbol{b} \cdot \boldsymbol{b} = \|\boldsymbol{a}\|^2 + \|\boldsymbol{b}\|^2 - 2\,\|\boldsymbol{a}\|\,\|\boldsymbol{b}\| \cos\Theta$$

Then

$$\boldsymbol{a} \cdot \boldsymbol{b} = \|\boldsymbol{a}\|\,\|\boldsymbol{b}\| \cos\Theta$$

With Length

Using the Norm

$$\|\boldsymbol{p}\| = \left|\frac{\boldsymbol{a}^T\boldsymbol{b}}{\boldsymbol{a}^T\boldsymbol{a}}\right| \|\boldsymbol{a}\| = \left|\frac{\|\boldsymbol{a}\| \, \|\boldsymbol{b}\| \cos\Theta}{\|\boldsymbol{a}\|^2}\right| \|\boldsymbol{a}\| = \|\boldsymbol{b}\| \, |\cos\Theta|$$

Example

Project

$$\boldsymbol{b} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \text{ onto } \boldsymbol{a} = \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}$$

Find

$$\boldsymbol{p} = x\boldsymbol{a}$$

Example

Project

$$\boldsymbol{b} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \text{ onto } \boldsymbol{a} = \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}$$

Find

$$\boldsymbol{p} = x\boldsymbol{a}$$

What about the Projection Matrix in general

We have

$$\boldsymbol{p} = \boldsymbol{a}x = \frac{\boldsymbol{a}\boldsymbol{a}^T\boldsymbol{b}}{\boldsymbol{a}^T\boldsymbol{a}} = P\boldsymbol{b}$$

Then

$$P = \frac{\boldsymbol{a}\boldsymbol{a}^T}{\boldsymbol{a}^T\boldsymbol{a}}$$

What about the Projection Matrix in general

We have

$$\boldsymbol{p} = \boldsymbol{a}x = \frac{\boldsymbol{a}\boldsymbol{a}^T\boldsymbol{b}}{\boldsymbol{a}^T\boldsymbol{a}} = P\boldsymbol{b}$$

Then

$$P = \frac{\boldsymbol{a}\boldsymbol{a}^T}{\boldsymbol{a}^T\boldsymbol{a}}$$

Example

Find the projection matrix for

$$\boldsymbol{b} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \text{ onto } \boldsymbol{a} = \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}$$

Cinvestav

What about the general case?

We have that

Find the combination $\boldsymbol{p} = x_1\boldsymbol{a}_1 + x_2\boldsymbol{a}_2 + \cdots + x_n\boldsymbol{a}_n$ closest to vector $\boldsymbol{b}$.

Now you need a vector

Find the vector $\boldsymbol{x}$,find the projection $\boldsymbol{p} = A\boldsymbol{x}$,find the matrix P.

Again, the error is perpendicular to the space

$$\boldsymbol{e} = \boldsymbol{b} - A\boldsymbol{x}$$

What about the general case?

We have that

Find the combination $\boldsymbol{p} = x_1\boldsymbol{a}_1 + x_2\boldsymbol{a}_2 + \cdots + x_n\boldsymbol{a}_n$ closest to vector $\boldsymbol{b}$.

Now you need a vector

Find the vector $\boldsymbol{x}$,find the projection $\boldsymbol{p} = A\boldsymbol{x}$,find the matrix P.

Again, the error is perpendicular to the space

$$\boldsymbol{e} = \boldsymbol{b} - A\boldsymbol{x}$$

What about the general case?

We have that

Find the combination $\boldsymbol{p} = x_1\boldsymbol{a}_1 + x_2\boldsymbol{a}_2 + \cdots + x_n\boldsymbol{a}_n$ closest to vector $\boldsymbol{b}$.

Now you need a vector

Find the vector $\boldsymbol{x}$,find the projection $\boldsymbol{p} = A\boldsymbol{x}$,find the matrix P.

Again, the error is perpendicular to the space

$$\boldsymbol{e} = \boldsymbol{b} - A\boldsymbol{x}$$

Cinvestav

Therefore

The error $\boldsymbol{e} = \boldsymbol{b} - A\boldsymbol{x}$

$$
\begin{aligned}
\boldsymbol{a}_1^T \left(\boldsymbol{b} - A\boldsymbol{x}\right) &= 0 \\
&\vdots \\
\boldsymbol{a}_n^T \left(\boldsymbol{b} - A\boldsymbol{x}\right) &= 0
\end{aligned}
$$

Or

$$
\begin{bmatrix} \boldsymbol{a}_1^T \\ \vdots \\ \boldsymbol{a}_n^T \end{bmatrix} \left[\boldsymbol{b} - A\boldsymbol{x}\right] = 0
$$

Cinvestav

Therefore

The error $\boldsymbol{e} = \boldsymbol{b} - A\boldsymbol{x}$

$$\begin{aligned} \boldsymbol{a}_1^T \left(\boldsymbol{b} - A\boldsymbol{x}\right) &= 0 \\ &\vdots \\ \boldsymbol{a}_n^T \left(\boldsymbol{b} - A\boldsymbol{x}\right) &= 0 \end{aligned}$$

Or

$$\begin{bmatrix} \boldsymbol{a}_1^T \\ \vdots \\ \boldsymbol{a}_n^T \end{bmatrix} \left[\boldsymbol{b} - A\boldsymbol{x}\right] = 0$$

Therefore

The Matrix with those rows is A^T

$$A^T\left(\boldsymbol{b} - A\boldsymbol{x}\right) = 0$$

Therefore

$$A^T\boldsymbol{b} - A^T A\boldsymbol{x} = 0$$

Or the most know form

$$\boldsymbol{x} = \left(A^T A\right)^{-1} A^T\boldsymbol{b}$$

Therefore

The Matrix with those rows is A^T

$$A^T\left(\boldsymbol{b} - A\boldsymbol{x}\right) = 0$$

Therefore

$$A^T\boldsymbol{b} - A^T A\boldsymbol{x} = 0$$

Or the most know form

$$\boldsymbol{x} = \left(A^T A\right)^{-1} A^T \boldsymbol{b}$$

Therefore

The Matrix with those rows is A^T

$$A^T\left(\boldsymbol{b} - A\boldsymbol{x}\right) = 0$$

Therefore

$$A^T\boldsymbol{b} - A^T A\boldsymbol{x} = 0$$

Or the most know form

$$\boldsymbol{x} = \left(A^T A\right)^{-1} A^T\boldsymbol{b}$$

Cinvestav

Therefore

The Projection is

$$\boldsymbol{p} = A\boldsymbol{x} = A\left(A^{T}A\right)^{-1}A^{T}\boldsymbol{b}$$

Therefore

$$P = A\left(A^{T}A\right)^{-1}A^{T}$$

Therefore

The Projection is

$$\boldsymbol{p} = A\boldsymbol{x} = A\left(A^T A\right)^{-1} A^T \boldsymbol{b}$$

Therefore

$$P = A\left(A^T A\right)^{-1} A^T$$

Cinvestav

The key step was $A^T[\boldsymbol{b} - A\boldsymbol{x}] = 0$

Linear algebra gives this "normal equation"

1. Our subspace is the column space of A.
2. The error vector $\boldsymbol{b} - A\boldsymbol{x}$ is perpendicular to that column space.
3. Therefore $\boldsymbol{b} - A\boldsymbol{x}$ is in the nullspace of A^T

When A has independent columns, $A^T A$ is invertible

Theorem

$A^T A$ is invertible if and only if Ahas linearly independent columns.

Proof

Consider the following

$$A^T A\boldsymbol{x} = 0$$

Here, $A\boldsymbol{x}$ is in the null space of A^T

- Remember the column space and null space of A^T are orthogonal complements.

And $A\boldsymbol{x}$ an element in the column space of A

$$A\boldsymbol{x} = 0$$

Proof

Consider the following

$$A^T A\boldsymbol{x} = 0$$

Here, Ax is in the null space of A^T

- Remember the column space and null space of A^T are orthogonal complements.

And Ax an element in the column space of A

$$Ax = 0$$

Proof

Consider the following

$$A^T A\boldsymbol{x} = 0$$

Here, Ax is in the null space of A^T

- Remember the column space and null space of A^T are orthogonal complements.

And Ax an element in the column space of A

$$A\boldsymbol{x} = 0$$

Cinvestav

Proof

If A has linearly independent columns

$$Ax = 0 \Longrightarrow x = 0$$

Then, the null space

$$Null\left(A^T A\right) = \{0\}$$

i.e $A^T A$ is full rank

- Then, $A^T A$ is invertible...

Cinvestav

Proof

If A has linearly independent columns

$$A\boldsymbol{x} = 0 \Longrightarrow \boldsymbol{x} = 0$$

Then, the null space

$$Null\left(A^T A\right) = \{0\}$$

i.e $A^T A$ is full rank

- Then, $A^T A$ is invertible...

Proof

If A has linearly independent columns

$$A\boldsymbol{x} = 0 \Longrightarrow \boldsymbol{x} = 0$$

Then, the null space

$$Null\left(A^T A\right) = \{0\}$$

i.e $A^T A$ is full rank

- Then, $A^T A$ is invertible...

Cinvestav

Finally

Theorem

- When A has independent columns, $A^T A$ is square, symmetric and invertible.

Example

Use Gauss-Jordan for finding if $A^T A$ is invertible

$$A = \begin{pmatrix} 1 & 2 \\ 1 & 2 \\ 0 & 0 \end{pmatrix}$$

Example

Given

$$A = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix} \text{ and } \boldsymbol{b} = \begin{pmatrix} 6 \\ 0 \\ 0 \end{pmatrix}$$

Find

x and p and P

Cinvestav

Example

Given

$$A = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix} \text{ and } \boldsymbol{b} = \begin{pmatrix} 6 \\ 0 \\ 0 \end{pmatrix}$$

Find

$\boldsymbol{x}$ and $\boldsymbol{p}$ and P

Outline

Cinvestav

Now, we always like to make our life easier

Something Notable

- Orthogonality makes easier to find $\boldsymbol{x}$, $\boldsymbol{p}$ and P.

For this, we will find the orthogonal vectors

- At the column space of A

Now, we always like to make our life easier

Something Notable

- Orthogonality makes easier to find $\boldsymbol{x}$, $\boldsymbol{p}$ and P.

For this, we will find the orthogonal vectors

- At the column space of A

Orthonormal Vectors

Definition

The vectors $\boldsymbol{q}_1, \boldsymbol{q}_2, ..., \boldsymbol{q}_n$ are orthonormal if

$$\boldsymbol{q}_i^T \boldsymbol{q}_j = \begin{cases} 0 & \text{when } i \neq j \\ 1 & \text{when } i = j \end{cases}$$

Then

- A matrix with orthonormal columns is assigned the special letter Q

Properties

- A matrix Q with orthonormal columns satisfies $Q^T Q = I$

Orthonormal Vectors

Definition

The vectors $\boldsymbol{q}_1, \boldsymbol{q}_2, ..., \boldsymbol{q}_n$ are orthonormal if

$$\boldsymbol{q}_i^T \boldsymbol{q}_j = \begin{cases} 0 & \text{when } i \neq j \\ 1 & \text{when } i = j \end{cases}$$

Then

- A matrix with orthonormal columns is assigned the special letter Q

Properties

- A matrix Q with orthonormal columns satisfies $Q^T Q = I$

Orthonormal Vectors

Definition

The vectors $\boldsymbol{q}_1, \boldsymbol{q}_2, ..., \boldsymbol{q}_n$ are orthonormal if

$$\boldsymbol{q}_i^T \boldsymbol{q}_j = \begin{cases} 0 & \text{when } i \neq j \\ 1 & \text{when } i = j \end{cases}$$

Then

- A matrix with orthonormal columns is assigned the special letter Q

Properties

- A matrix Q with orthonormal columns satisfies $Q^T Q = I$

Additionally

Given that

$$Q^T Q = I$$

Therefore

When Q is square, $Q^T Q = I$ means that $Q^T = Q^{-1}$: transpose = inverse.

Additionally

Given that

$$Q^T Q = I$$

Therefore

When Q is square, $Q^T Q = I$ means that $Q^T = Q^{-1}$: transpose = inverse.

Examples

Rotation

$$\begin{pmatrix} \cos\Theta & -\sin\Theta \\ \sin\Theta & \cos\Theta \end{pmatrix}$$

Permutation Matrix

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

Reflection

- Setting $Q = I - 2uu^T$ with u a unit vector.

Examples

Rotation

$$\begin{pmatrix} \cos\Theta & -\sin\Theta \\ \sin\Theta & \cos\Theta \end{pmatrix}$$

Permutation Matrix

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

Reflection

- Setting $Q = I - 2uu^T$ with u a unit vector.

Examples

Rotation

$$\begin{pmatrix} \cos\Theta & -\sin\Theta \\ \sin\Theta & \cos\Theta \end{pmatrix}$$

Permutation Matrix

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

Reflection

- Setting $Q = I - 2\boldsymbol{u}\boldsymbol{u}^T$ with $\boldsymbol{u}$ a unit vector.

Cinvestav

Finally

If Q has orthonormal columns

- The lengths are unchanged

How?

$$\|Qx\| = \sqrt{x^T Q^T Q x} = \sqrt{x^T x} = \|x\|$$

Finally

If Q has orthonormal columns

- The lengths are unchanged

How?

$$\|Q\boldsymbol{x}\| = \sqrt{\boldsymbol{x}^T Q^T Q \boldsymbol{x}} = \sqrt{\boldsymbol{x}^T \boldsymbol{x}} = \|\boldsymbol{x}\|$$

Remark

Something Notable

When the columns of A were a basis for the subspace.

All Formulas involve

$$A^T A$$

What happens when the basis vectors are orthonormal

$A^T A$ simplifies to $Q^T Q = I$

Cinvestav

Remark

Something Notable

When the columns of A were a basis for the subspace.

All Formulas involve

$$A^T A$$

What happens when the basis vectors are orthonormal

$A^T A$ simplifies to $Q^T Q = I$

Cinvestav

Remark

Something Notable

When the columns of A were a basis for the subspace.

All Formulas involve

$$A^T A$$

What happens when the basis vectors are orthonormal

$A^T A$ simplifies to $Q^T Q = I$

Therefore, we have

The following

$Ix = Q^T b$ and $p = Qx$ and $P = QIQ^T$

Not only that

The solution of $Qx = b$ is simply $x = Q^T b$

Therefore, we have

The following

$I\boldsymbol{x} = Q^T\boldsymbol{b}$ and $\boldsymbol{p} = Q\boldsymbol{x}$ and $P = QIQ^T$

Not only that

The solution of $Q\boldsymbol{x} = \boldsymbol{b}$ is simply $\boldsymbol{x} = Q^T\boldsymbol{b}$

Example

Given the following matrix

Verify that is a orthogonal matrix

$$\frac{1}{3}\begin{pmatrix} -1 & 2 & 2 \\ 2 & -1 & 2 \\ 2 & 2 & -1 \end{pmatrix}$$

We have that

Given that using orthonormal bases is good

How do we generate such basis given an initial basis?

Graham Schmidt Process

We begin with three linear independent vectors a, b and c

We have that

Given that using orthonormal bases is good

How do we generate such basis given an initial basis?

Graham Schmidt Process

We begin with three linear independent vectors $\boldsymbol{a}, \boldsymbol{b}$ and $\boldsymbol{c}$

Then

We can do the following

- Select $\boldsymbol{a}$ and rename it $\boldsymbol{A}$

Start with b and subtract its projection along a

$$B = b - \frac{A^T b}{A^T A} A$$

Properties

This vector B is what we have called the error vector e, perpendicular to a.

Then

We can do the following

- Select $\boldsymbol{a}$ and rename it $\boldsymbol{A}$

Start with $\boldsymbol{b}$ and subtract its projection along $\boldsymbol{a}$

$$\boldsymbol{B} = \boldsymbol{b} - \frac{\boldsymbol{A}^T \boldsymbol{b}}{\boldsymbol{A}^T \boldsymbol{A}} \boldsymbol{A}$$

Properties

This vector $\boldsymbol{B}$ is what we have called the error vector $\boldsymbol{e}$, perpendicular to $\boldsymbol{a}$.

Then

We can do the following

- Select $\boldsymbol{a}$ and rename it $\boldsymbol{A}$

Start with $\boldsymbol{b}$ and subtract its projection along $\boldsymbol{a}$

$$\boldsymbol{B} = \boldsymbol{b} - \frac{\boldsymbol{A}^T \boldsymbol{b}}{\boldsymbol{A}^T \boldsymbol{A}} \boldsymbol{A}$$

Properties

This vector $\boldsymbol{B}$ is what we have called the error vector $\boldsymbol{e}$, perpendicular to $\boldsymbol{a}$.

We can keep with such process

Now we do the same for the new $\boldsymbol{c}$

$$\boldsymbol{C} = \boldsymbol{c} - \frac{\boldsymbol{A}^T \boldsymbol{c}}{\boldsymbol{A}^T \boldsymbol{A}} \boldsymbol{A} - \frac{\boldsymbol{B}^T \boldsymbol{c}}{\boldsymbol{B}^T \boldsymbol{B}} \boldsymbol{B}$$

Normalize them

To obtain the final result!!!

$$q_1 = \frac{\boldsymbol{A}}{\|\boldsymbol{A}\|}, q_1 = \frac{\boldsymbol{B}}{\|\boldsymbol{B}\|}, q_3 = \frac{\boldsymbol{C}}{\|\boldsymbol{C}\|}$$

Cinvestav

We can keep with such process

Now we do the same for the new $\boldsymbol{c}$

$$\boldsymbol{C} = \boldsymbol{c} - \frac{\boldsymbol{A}^T \boldsymbol{c}}{\boldsymbol{A}^T \boldsymbol{A}} \boldsymbol{A} - \frac{\boldsymbol{B}^T \boldsymbol{c}}{\boldsymbol{B}^T \boldsymbol{B}} \boldsymbol{B}$$

Normalize them

To obtain the final result!!!

$$q_1 = \frac{\boldsymbol{A}}{\|\boldsymbol{A}\|}, q_1 = \frac{\boldsymbol{B}}{\|\boldsymbol{B}\|}, q_3 = \frac{\boldsymbol{C}}{\|\boldsymbol{C}\|}$$

We can keep with such process

Now we do the same for the new $\boldsymbol{c}$

$$\boldsymbol{C} = \boldsymbol{c} - \frac{\boldsymbol{A}^T \boldsymbol{c}}{\boldsymbol{A}^T \boldsymbol{A}} \boldsymbol{A} - \frac{\boldsymbol{B}^T \boldsymbol{c}}{\boldsymbol{B}^T \boldsymbol{B}} \boldsymbol{B}$$

Normalize them

To obtain the final result!!!

$$q_1 = \frac{\boldsymbol{A}}{\|\boldsymbol{A}\|}, q_1 = \frac{\boldsymbol{B}}{\|\boldsymbol{B}\|}, q_3 = \frac{\boldsymbol{C}}{\|\boldsymbol{C}\|}$$

Cinvestav

Example

Suppose the independent non-orthogonal vectors $\boldsymbol{a}, \boldsymbol{b}$ and $\boldsymbol{c}$

$$\boldsymbol{a} = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \boldsymbol{c} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \boldsymbol{d} = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$$

Then

- Do the procedure...

Cinvestav

Example

Suppose the independent non-orthogonal vectors $\boldsymbol{a}, \boldsymbol{b}$ and $\boldsymbol{c}$

$$\boldsymbol{a} = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \boldsymbol{c} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \boldsymbol{d} = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$$

Then

- Do the procedure...

We have the following process

We begin with a matrix A

$$A = [\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}]$$

We ended with the following matrix

$$Q = [q_1, q_2, q_3]$$

How are these matrices related?

- There is a third matrix!!!

$$A = QR$$

We have the following process

We begin with a matrix A

$$A = [\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}]$$

We ended with the following matrix

$$Q = [q_1, q_2, q_3]$$

How are these matrices related?

- There is a third matrix!!!

$$A = QR$$

We have the following process

We begin with a matrix A

$$A = [\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}]$$

We ended with the following matrix

$$Q = [q_1, q_2, q_3]$$

How are these matrices related?

- There is a third matrix!!!

$$A = QR$$

Cinvestav

Notice the following

Something Notable

- The vectors $\boldsymbol{a}$ and $\boldsymbol{A}$ and q_1 are all along a single line.

Then

The vectors $\boldsymbol{a}, \boldsymbol{b}$ and $\boldsymbol{A}, \boldsymbol{B}$ and q_1, q_2 are all in the same plane.

Further

The vectors $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}$ and $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{B}$ and q_1, q_2, q_2 are all in the same subspace.

Notice the following

Something Notable

- The vectors $\boldsymbol{a}$ and $\boldsymbol{A}$ and q_1 are all along a single line.

Then

The vectors $\boldsymbol{a}, \boldsymbol{b}$ and $\boldsymbol{A}, \boldsymbol{B}$ and q_1, q_2 are all in the same plane.

Further

The vectors $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}$ and $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{B}$ and q_1, q_2, q_2 are all in the same subspace.

Cinvestav

Notice the following

Something Notable

- The vectors $\boldsymbol{a}$ and $\boldsymbol{A}$ and q_1 are all along a single line.

Then

The vectors $\boldsymbol{a}, \boldsymbol{b}$ and $\boldsymbol{A}, \boldsymbol{B}$ and q_1, q_2 are all in the same plane.

Further

The vectors $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}$ and $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{B}$ and q_1, q_2, q_2 are all in the same subspace.

Cinvestav

Therefore

It is possible to see that

$$\boldsymbol{a}_1, \boldsymbol{a}_2, ..., \boldsymbol{a}_k$$

They are combination of $q_1, q_2, ..., q_k$

$$[\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3] = [q_1, q_2, q_3] \begin{bmatrix} q_1^T \boldsymbol{a} & q_1^T \boldsymbol{b} & q_1^T \boldsymbol{c} \\ 0 & q_2^T \boldsymbol{b} & q_2^T \boldsymbol{c} \\ 0 & 0 & q_3^T \boldsymbol{c} \end{bmatrix}$$

Therefore

It is possible to see that

$$\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_k$$

They are combination of $q_1, q_2, \ldots, q_k$

$$[\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3] = [q_1, q_2, q_3] \begin{bmatrix} q_1^T \boldsymbol{a} & q_1^T \boldsymbol{b} & q_1^T \boldsymbol{c} \\ 0 & q_2^T \boldsymbol{b} & q_2^T \boldsymbol{c} \\ 0 & 0 & q_3^T \boldsymbol{c} \end{bmatrix}$$

Gram-Schmidt

From linear independent vectors $\boldsymbol{a}_1, \boldsymbol{a}_2, ..., \boldsymbol{a}_n$

Gram-Schmidt constructs orthonormal vectors $q_1, q_2, ..., q_n$ that when used as column vectors in a matrix Q

These matrices satisfy

$$A = QR$$

Properties

Then $R = Q^T A$ is a upper triangular matrix because later $q's$ are orthogonal to earlier $a's$.

Gram-Schmidt

From linear independent vectors $\boldsymbol{a}_1, \boldsymbol{a}_2, ..., \boldsymbol{a}_n$

Gram-Schmidt constructs orthonormal vectors $q_1, q_2, ..., q_n$ that when used as column vectors in a matrix Q

These matrices satisfy

$$A = QR$$

Properties

Then $R = Q^T A$ is a upper triangular matrix because later $q's$ are orthogonal to earlier $a's$.

Cinvestav

Gram-Schmidt

From linear independent vectors $\boldsymbol{a}_1, \boldsymbol{a}_2, ..., \boldsymbol{a}_n$

Gram-Schmidt constructs orthonormal vectors $q_1, q_2, ..., q_n$ that when used as column vectors in a matrix Q

These matrices satisfy

$$A = QR$$

Properties

Then $R = Q^T A$ is a upper triangular matrix because later $q's$ are orthogonal to earlier $a's$.

Therefore

Any $m \times n$ matrix A with linear independent columns can be factored into QR

- The $m \times n$ matrix Q has orthonormal columns.
- The square matrix R is upper triangular with positive diagonal.

We must not forget why this is useful for least squares

- $A^T A = R^T Q^T Q R = R^T R$

Least Squared Simplify to

$R^T R x = R^T Q^T b$ or $Rx = Q^T b$ or $x = R^{-1} Q^T b$

Therefore

Any $m \times n$ matrix A with linear independent columns can be factored into QR

- The $m \times n$ matrix Q has orthonormal columns.
- The square matrix R is upper triangular with positive diagonal.

We must not forget why this is useful for least squares

- $A^T A = R^T Q^T Q R = R^T R$

Least Squared Simplify to

$R^T R x = R^T Q^T b$ or $Rx = Q^T b$ or $x = R^{-1} Q^T b$

Therefore

Any $m \times n$ matrix A with linear independent columns can be factored into QR

- The $m \times n$ matrix Q has orthonormal columns.
- The square matrix R is upper triangular with positive diagonal.

We must not forget why this is useful for least squares

- $A^T A = R^T Q^T Q R = R^T R$

Least Squared Simplify to

$R^T R\boldsymbol{x} = R^T Q^T \boldsymbol{b}$ or $R\boldsymbol{x} = Q^T \boldsymbol{b}$ or $\boldsymbol{x} = R^{-1} Q^T \boldsymbol{b}$

Cinvestav

Algorithm

Basic Gram-Schmidt

1. for $j = 1$ to n
2. $\quad \boldsymbol{v} = A(:, j)$
3. $\quad$ for $i = 1$ to $j - 1$
4. $\quad\quad R(i, j) = Q(:, i)^T \boldsymbol{v}$
5. $\quad\quad \boldsymbol{v} = \boldsymbol{v} - R(i, j)\, Q(:, i)$
6. $\quad R(j, j) = \|\boldsymbol{v}\|$
7. $\quad Q(:, j) = \frac{\boldsymbol{v}}{R(j,j)}$

Cinvestav

Outline

A as a change factor

Most vectors change direction when multiplied against a random A

$$A\boldsymbol{v} \longrightarrow \boldsymbol{v}'$$

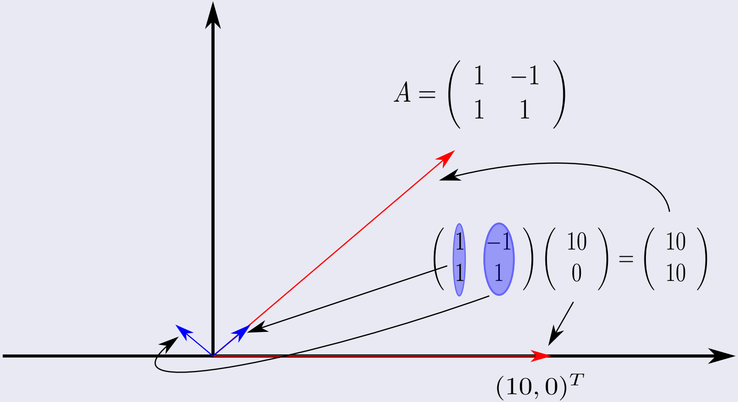
Example

A as a change factor

Most vectors change direction when multiplied against a random A

$$A\boldsymbol{v} \longrightarrow \boldsymbol{v}'$$

Example



However

There is a set of special vectors called eigenvectors

$$A\boldsymbol{v} = \lambda \boldsymbol{v}$$

- Here, the eigenvalue is λ and the eigenvector is $\boldsymbol{v}$.

Definition

- If T is a linear transformation from a vector space V over a field F, $T : V \longrightarrow V$, then $v \neq 0$ is an eigenvector of T if $T(v)$ is a scalar multiple of v.

Something quite interesting

- Such linear transformations can be expressed by matrices A, $T(v) = Av$

However

There is a set of special vectors called eigenvectors

$$A\boldsymbol{v} = \lambda \boldsymbol{v}$$

- Here, the eigenvalue is λ and the eigenvector is $\boldsymbol{v}$.

Definition

- If T is a linear transformation from a vector space V over a field F, $T : V \longrightarrow V$, then $\boldsymbol{v} \neq 0$ is an eigenvector of T if $T(\boldsymbol{v})$ is a scalar multiple of $\boldsymbol{v}$.

Something quite interesting

- Such linear transformations can be expressed by matrices A, $T(\boldsymbol{v}) = A\boldsymbol{v}$

However

There is a set of special vectors called eigenvectors

$$Av = \lambda v$$

- Here, the eigenvalue is λ and the eigenvector is v.

Definition

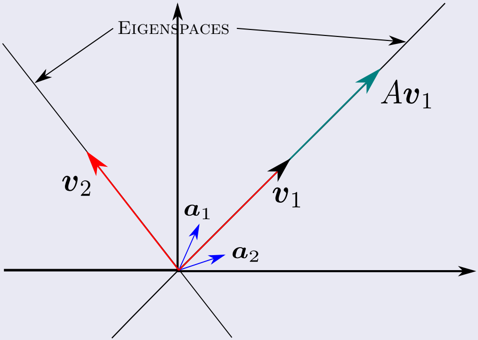
- If T is a linear transformation from a vector space V over a field F, $T : V \longrightarrow V$, then $v \neq 0$ is an eigenvector of T if $T(v)$ is a scalar multiple of v.

Something quite interesting

- Such linear transformations can be expressed by matrices A, $T(v) = Av$

A little bit of Geometry

Points in a direction in which it is stretched by the transformation A



Implications

You can see the eigenvalues as the vector of change by the mapping

$$T(\boldsymbol{v}) = A\boldsymbol{v}$$

Therefore, for an Invertible Square Matrix A

- If your rank is $n \Rightarrow$ if you have $\{v_1, v_2, v_3, ..., v_n\}$ eigenvalues

Implications

You can see the eigenvalues as the vector of change by the mapping

$$T(\boldsymbol{v}) = A\boldsymbol{v}$$

Therefore, for an Invertible Square Matrix A

- If your rank is $n \Rightarrow$ if you have $\{v_1, v_2, v_3, ..., v_n\}$ eigenvalues

Outline

Cinvestav

A simple case

Given a vector $\boldsymbol{v} \in V$

- We then apply the linear transformation sequentially:

$$\boldsymbol{v}, A\boldsymbol{v}, A^2\boldsymbol{v}...$$

For example

$$A = \begin{pmatrix} 0.7 & 0.3 \\ 0.3 & 0.7 \end{pmatrix}$$

A simple case

Given a vector $\boldsymbol{v} \in V$

- We then apply the linear transformation sequentially:

$$\boldsymbol{v}, A\boldsymbol{v}, A^2\boldsymbol{v}...$$

For example

$$A = \begin{pmatrix} 0.7 & 0.3 \\ 0.3 & 0.7 \end{pmatrix}$$

Cinvestav

We have the following sequence

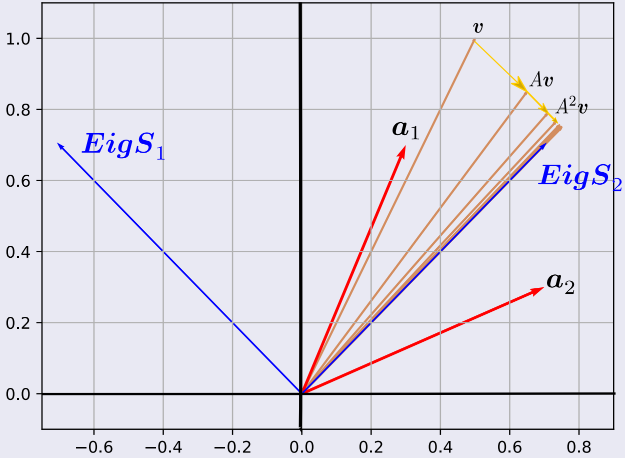
As you can see

$$\boldsymbol{v} = \begin{pmatrix} 0.5 \\ 1 \end{pmatrix}, A\boldsymbol{v} = \begin{pmatrix} 0.65 \\ 0.85 \end{pmatrix}, ..., A^k\boldsymbol{v} = \begin{pmatrix} 0.75 \\ 0.75 \end{pmatrix}, ...$$

Cinvestav

Geometrically

We have



Notably

We have that

- The eigenvalue λ tells whether the special vector $\boldsymbol{v}$ is stretched or shrunk or reversed or left unchanged-when it is multiplied by A.

Something Notable

1. Eigenvalues can repeat!!!
2. Eigenvalues can be positive or negative
3. Eigenvalues could be 0

Properties

The eigenvectors make up the nullspace of $(A - \lambda I)$.

Notably

We have that

- The eigenvalue λ tells whether the special vector $\boldsymbol{v}$ is stretched or shrunk or reversed or left unchanged-when it is multiplied by A.

Something Notable

1. Eigenvalues can repeat!!!
2. Eigenvalues can be positive or negative
3. Eigenvalues could be 0

Properties

The eigenvectors make up the nullspace of $(A - \lambda I)$.

Notably

We have that

- The eigenvalue λ tells whether the special vector $\boldsymbol{v}$ is stretched or shrunk or reversed or left unchanged-when it is multiplied by A.

Something Notable

1. Eigenvalues can repeat!!!
2. Eigenvalues can be positive or negative
3. Eigenvalues could be 0

Properties

The eigenvectors make up the nullspace of $(A - \lambda I)$.

Cinvestav

Outline

Cinvestav

An Intuition

Imagine that A is a symmetric real matrix

- Then, we have that $A\boldsymbol{v}$ is a mapping

What happens to the unitary circle?

$$\left\{\boldsymbol{v}|\boldsymbol{v}^T\boldsymbol{v}=1\right\}$$

An Intuition

Imagine that A is a symmetric real matrix

- Then, we have that $A\boldsymbol{v}$ is a mapping

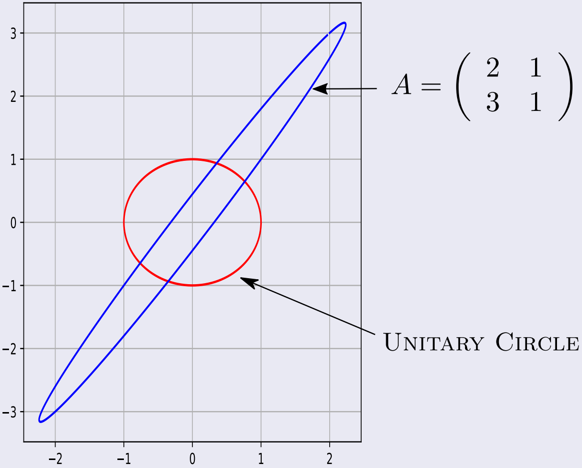
What happens to the unitary circle?

$$\left\{\boldsymbol{v}|\boldsymbol{v}^T\boldsymbol{v}=1\right\}$$

Cinvestav

We have something like

A modification of the distances



If we get the Q matrix

We go back to the unitary circle

- A is a modification of distances

Therefore

- Our best bet is to build A with specific properties at hand...

If we get the Q matrix

We go back to the unitary circle

- A is a modification of distances

Therefore

- Our best bet is to build A with specific properties at hand...

Outline

Therefore

Relation with invertibility

- What if $(A - \lambda I)\,\boldsymbol{v} = 0$?

What if $v \neq 0$?

- Then, columns $A - \lambda I$ are not linear independents.

Then

- $A - \lambda I$ is not invertible...

Therefore

Relation with invertibility

- What if $(A - \lambda I)\,\boldsymbol{v} = 0$?

What if $\boldsymbol{v} \neq 0$?

- Then, columns $A - \lambda I$ are not linear independents.

Then

- $A - \lambda I$ is not invertible...

Therefore

Relation with invertibility

- What if $(A - \lambda I)\, \boldsymbol{v} = 0$?

What if $\boldsymbol{v} \neq 0$?

- Then, columns $A - \lambda I$ are not linear independents.

Then

- $A - \lambda I$ is not invertible...

Cinvestav

Also for Determinants

If $A - \lambda I$ is not invertible

- $det\left(A - \lambda I\right) = 0 \leftarrow$ How?

Theorem

- A square matrix is invertible if and only if its determinant is non-zero.

Proof $\Longrightarrow$

- We know for Jordan-Gauss that an invertible matrix can be reduced to the identity by elementary matrix operations

$$A = E_1 E_2 \cdots E_k$$

Also for Determinants

If $A - \lambda I$ is not invertible

- $det\left(A - \lambda I\right) = 0 \leftarrow$ How?

Theorem

- A square matrix is invertible if and only if its determinant is non-zero.

Proof $\Longrightarrow$

- We know for Jordan-Gauss that an invertible matrix can be reduced to the identity by elementary matrix operations

$$A = E_1 E_2 \cdots E_k$$

Also for Determinants

If $A - \lambda I$ is not invertible

- $det\left(A - \lambda I\right) = 0 \leftarrow$ How?

Theorem

- A square matrix is invertible if and only if its determinant is non-zero.

Proof $\Longrightarrow$

- We know for Jordan-Gauss that an invertible matrix can be reduced to the identity by elementary matrix operations

$$A = E_1 E_2 \cdots E_k$$

Cinvestav

Furthermore

We have then

- $det\,(A) = det\,(E_1) \cdots det\,(E_k)$

An interesting thing is that, for example

- Let A be a $K \times K$ matrix. Let E be an elementary matrix obtained by multiplying a row of the $K \times K$ identity matrix I by a constant $c \neq 0$. Then $det\,(E) = c$.

Furthermore

We have then

- $det\left(A\right) = det\left(E_1\right)\cdots det\left(E_k\right)$

An interesting thing is that, for example

- Let A be a $K \times K$ matrix. Let E be an elementary matrix obtained by multiplying a row of the $K \times K$ identity matrix I by a constant $c \neq 0$. Then $det\left(E\right) = c$.

The same for the other elementary matrices

Then, $det\left(A\right) = det\left(E_1\right) \cdots det\left(E_k\right) \neq 0$

- Now, the return is quite simple

Then, $A - \lambda I$ is not invertible

- $det\left(A - \lambda I\right) = 0$

The same for the other elementary matrices

Then, $det\,(A) = det\,(E_1) \cdots det\,(E_k) \neq 0$

- Now, the return is quite simple

Then, $A - \lambda I$ is not invertible

- $det\,(A - \lambda I) = 0$

Now, for eigenvalues

Theorem

The number λ is an eigenvalue $\iff$ $(A - \lambda I)$ is not invertible i.e. singular.

$\Longrightarrow$

- The number λ is an eigenvalue $\Rightarrow$ then $\exists v$ such that $(A - \lambda I)\, v = 0$

The columns of $A - \lambda I$

- They are linear dependent so $(A - \lambda I)$ is not invertible
- What about $\Longleftarrow$?

Now, for eigenvalues

Theorem

The number λ is an eigenvalue $\Longleftrightarrow$ $(A - \lambda I)$ is not invertible i.e. singular.

$\Longrightarrow$

- The number λ is an eigenvalue $\Rightarrow$ then $\exists \boldsymbol{v}$ such that $(A - \lambda I)\,\boldsymbol{v} = 0$

The columns of $A - \lambda I$

- They are linear dependent so $(A - \lambda I)$ is not invertible
- What about $\Longleftarrow$?

Now, for eigenvalues

Theorem

The number λ is an eigenvalue $\iff (A - \lambda I)$ is not invertible i.e. singular.

$\Longrightarrow$

- The number λ is an eigenvalue $\Rightarrow$ then $\exists \boldsymbol{v}$ such that $(A - \lambda I)\,\boldsymbol{v} = 0$

The columns of $A - \lambda I$

- They are linear dependent so $(A - \lambda I)$ is not invertible
- What about $\Longleftarrow$?

Cinvestav

Outline

Cinvestav

Now, How do we find eigenvalues and eigenvectors?

Ok, we know that for each eigenvalue there is an eigenvector

- We have seen that they represent the stretching of the vectors

How do we get such eigenvalues

- Basically, use the fact that if $\lambda \Rightarrow det\left[A - \lambda I\right] = 0$

In this way

- **We obtain a polynomial know as characteristic polynomial.**

Now, How do we find eigenvalues and eigenvectors?

Ok, we know that for each eigenvalue there is an eigenvector

- We have seen that they represent the stretching of the vectors

How do we get such eigenvalues

- Basically, use the fact that if $\lambda \Rightarrow det\left[A - \lambda I\right] = 0$

In this way

- We obtain a polynomial know as characteristic polynomial.

Now, How do we find eigenvalues and eigenvectors?

Ok, we know that for each eigenvalue there is an eigenvector

- We have seen that they represent the stretching of the vectors

How do we get such eigenvalues

- Basically, use the fact that if $\lambda \Rightarrow det\left[A - \lambda I\right] = 0$

In this way

- **We obtain a polynomial know as characteristic polynomial.**

Characteristic Polynomial

Then get the root of the polynomial i.e.

- Values of λ that make

$$p(\lambda) = a_o + a_1\lambda + a_2\lambda + \cdots + a_n\lambda^n = 0$$

Then, once you have the eigenvalues

- For each eigenvalue λ solve

$$(A - \lambda I)\, v = 0 \text{ or } Av = \lambda v$$

It is quite simple

- But a lot of theorems to get here!!!

Characteristic Polynomial

Then get the root of the polynomial i.e.

- Values of λ that make

$$p(\lambda) = a_o + a_1\lambda + a_2\lambda + \cdots + a_n\lambda^n = 0$$

Then, once you have the eigenvalues

- For each eigenvalue λ solve

$$(A - \lambda I)\boldsymbol{v} = 0 \text{ or } A\boldsymbol{v} = \lambda\boldsymbol{v}$$

It is quite simple

- But a lot of theorems to get here!!!

Cinvestav

Characteristic Polynomial

Then get the root of the polynomial i.e.

- Values of λ that make

$$p(\lambda) = a_o + a_1\lambda + a_2\lambda + \cdots + a_n\lambda^n = 0$$

Then, once you have the eigenvalues

- For each eigenvalue λ solve

$$(A - \lambda I)\boldsymbol{v} = 0 \text{ or } A\boldsymbol{v} = \lambda\boldsymbol{v}$$

It is quite simple

- But a lot of theorems to get here!!!

Cinvestav

Example

Given

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$$

Find

Its eigenvalues and eigenvectors.

Cinvestav

Example

Given

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$$

Find

Its eigenvalues and eigenvectors.

Summary

To solve the eigenvalue problem for an $n \times n$ matrix, follow these steps

1. Compute the determinant of $A - \lambda I$.
2. Find the roots of the polynomial $det\left(A - \lambda I\right) = 0$.
3. For each eigenvalue solve $\left(A - \lambda I\right)\boldsymbol{v} = 0$ to find the eigenvector $\boldsymbol{v}$.

Some Remarks

Something Notable

If you add a row of A to another row, or exchange rows, the eigenvalues usually change.

Nevertheless

1. The product of the n eigenvalues equals the determinant.
2. The sum of the n eigenvalues equals the sum of the n diagonal entries.

Some Remarks

Something Notable

If you add a row of A to another row, or exchange rows, the eigenvalues usually change.

Nevertheless

1. The product of the n eigenvalues equals the determinant.
2. The sum of the n eigenvalues equals the sum of the n diagonal entries.

Cinvestav

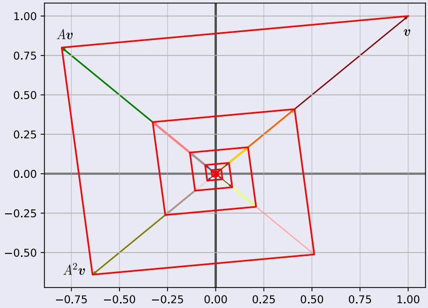
Outline

Cinvestav

They impact many facets of our life!!!

Example, given the composition of the linear function

$$A = \begin{pmatrix} 0 & -0.8 \\ 0.8 & 0 \end{pmatrix} \quad Eig = \begin{pmatrix} 0.707 & 0.707 \\ -0.707i & -0.707i \end{pmatrix}$$



Then, for recurrent systems

Something like

$$\boldsymbol{v}_{n+1} = A\boldsymbol{v}_n + \boldsymbol{b}$$

Making $b = 0$

$$\boldsymbol{v}_{n+1} = A\boldsymbol{v}_n$$

The eigenvalues are telling us if the recurrent system converges or not

- For example if we modify the matrix A.

Then, for recurrent systems

Something like

$$\boldsymbol{v}_{n+1} = A\boldsymbol{v}_n + \boldsymbol{b}$$

Making $\boldsymbol{b} = 0$

$$\boldsymbol{v}_{n+1} = A\boldsymbol{v}_n$$

The eigenvalues are telling us if the recurrent system converges or not

- For example if we modify the matrix A.

Cinvestav

Then, for recurrent systems

Something like

$$\boldsymbol{v}_{n+1} = A\boldsymbol{v}_n + \boldsymbol{b}$$

Making $\boldsymbol{b} = 0$

$$\boldsymbol{v}_{n+1} = A\boldsymbol{v}_n$$

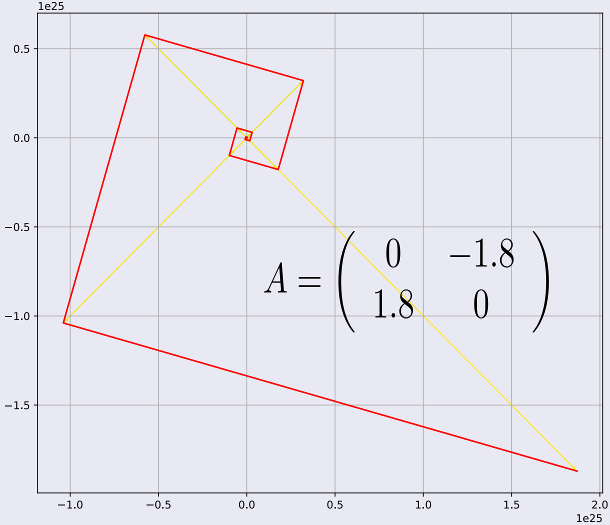
The eigenvalues are telling us if the recurrent system converges or not

- For example if we modify the matrix A.

Cinvestav

For example

Here, iterations send the system to the infinity



In another Example

Imagine the following example

1. F represents the number of foxes in a population
2. R represents the number of rabits in a population

Then, if we have that

- The number of rabbits is related to the number of foxes in the following way
  - At each time you have three times the number of rabbits minus the number of foxes

In another Example

Imagine the following example

1. F represents the number of foxes in a population
2. R represents the number of rabits in a population

Then, if we have that

- The number of rabbits is related to the number of foxes in the following way
  - At each time you have three times the number of rabbits minus the number of foxes

Cinvestav

Therefore

We have the following relation

$$\frac{dR}{dt} = 3R - 1F$$
$$\frac{dF}{dt} = 1F$$

Or as a matrix operations

$$\begin{pmatrix} R' \\ F' \end{pmatrix} = \begin{pmatrix} 3 & -1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} R \\ F \end{pmatrix}$$

Therefore

We have the following relation

$$\frac{dR}{dt} = 3R - 1F$$
$$\frac{dF}{dt} = 1F$$

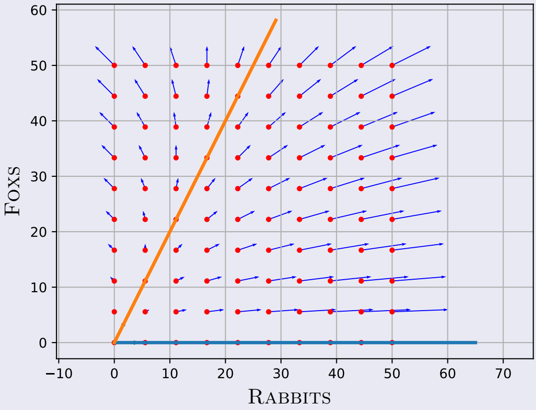
Or as a matrix operations

$$\begin{pmatrix} R' \\ F' \end{pmatrix} = \begin{pmatrix} 3 & -1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} R \\ F \end{pmatrix}$$

Cinvestav

Geometrically

As you can see through the eigenvalues we have a stable population



Therefore

We can try to cast our problems as system of equations

- Solve by methods found in linear algebra

Then, using properties of the eigenvectors

- We can look at sought properties that we would like to have

Cinvestav

Therefore

We can try to cast our problems as system of equations

- Solve by methods found in linear algebra

Then, using properties of the eigenvectors

- We can look at sought properties that we would like to have

Outline

Assume a matrix A

Definition

- An $n \times n$ matrix A is diagonalizable is called diagonalizable if there exists an invertible matrix P such that $P^{-1}AP$ is a diagonal matrix.

Some remarks

- Is every diagonalizable matrix invertible?

Assume a matrix A

Definition

- An $n \times n$ matrix A is diagonalizable is called diagonalizable if there exists an invertible matrix P such that $P^{-1}AP$ is a diagonal matrix.

Some remarks

- Is every diagonalizable matrix invertible?

Cinvestav

Nope

Given the structure

$$P^{-1}AP = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}$$

Then using the determinant

$$det\left[P^{-1}AP\right] = det\left[P\right]^{-1} det\left[A\right] det\left[P\right] = det\left[A\right] = \prod_{i=}^{n} \lambda_i$$

if one of the eigenvalues of A is zero

- The determinant of A is zero, and hence A is not invertible.

Nope

Given the structure

$$P^{-1}AP = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}$$

Then using the determinant

$$det\left[P^{-1}AP\right] = det\left[P\right]^{-1} det\left[A\right] det\left[P\right] = det\left[A\right] = \prod_{i=}^{n} \lambda_i$$

if one of the eigenvalues of A is zero

- The determinant of A is zero, and hence A is not invertible.

Nope

Given the structure

$$P^{-1}AP = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}$$

Then using the determinant

$$det\left[P^{-1}AP\right] = det\left[P\right]^{-1} det\left[A\right] det\left[P\right] = det\left[A\right] = \prod_{i=}^{n} \lambda_i$$

if one of the eigenvalues of A is zero

- The determinant of A is zero, and hence A is not invertible.

Actually

Theorem

- A diagonal matrix is invertible if and only if its eigenvalues are nonzero.

Is Every Invertible Matrix Diagonalizable?

- Consider the matrix:

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

The determinant of A is 1, hence A is invertible (Characteristic Polynomial)

$$p(\lambda) = det\left[A - \lambda I\right] = (1-t)^2$$

Actually

Theorem

- A diagonal matrix is invertible if and only if its eigenvalues are nonzero.

Is Every Invertible Matrix Diagonalizable?

- Consider the matrix:

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

The determinant of A is 1, hence A is invertible (Characteristic Polynomial)

$$p(\lambda) = det[A - \lambda I] = (1 - t)^2$$

Actually

Theorem

- A diagonal matrix is invertible if and only if its eigenvalues are nonzero.

Is Every Invertible Matrix Diagonalizable?

- Consider the matrix:

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

The determinant of A is 1, hence A is invertible (Characteristic Polynomial)

$$p(\lambda) = det[A - \lambda I] = (1 - t)^2$$

Cinvestav

Therefore, you have a repetition in the eigenvalue

Thus, the geometric multiplicity of the eigenvalue 1 is 1, $\begin{pmatrix} 1 & 0 \end{pmatrix}^T$

- Since the geometric multiplicity is strictly less than the algebraic multiplicity, the matrix A is defective and not diagonalizable.

Why?

- Let us to look at the eigenvectors for this answer

Therefore, you have a repetition in the eigenvalue

Thus, the geometric multiplicity of the eigenvalue 1 is 1, $\begin{pmatrix} 1 & 0 \end{pmatrix}^T$

- Since the geometric multiplicity is strictly less than the algebraic multiplicity, the matrix A is defective and not diagonalizable.

Why?

- Let us to look at the eigenvectors for this answer

Relation with Eigenvectors

Suppose that the $n \times n$ matrix A has n linearly independent eigenvectors

$$\boldsymbol{v}_1, \boldsymbol{v}_2, ..., \boldsymbol{v}_n$$

Put them into an eigenvector matrix P

$$P = \begin{bmatrix} \boldsymbol{v}_1 & \boldsymbol{v}_2 & \dots & \boldsymbol{v}_n \end{bmatrix}$$

Cinvestav

Relation with Eigenvectors

Suppose that the $n \times n$ matrix A has n linearly independent eigenvectors

$$\boldsymbol{v}_1, \boldsymbol{v}_2, ..., \boldsymbol{v}_n$$

Put them into an eigenvector matrix P

$$P = \begin{bmatrix} \boldsymbol{v}_1 & \boldsymbol{v}_2 & ... & \boldsymbol{v}_n \end{bmatrix}$$

We have

What if we apply it to the canonical basis elements?

$$P\left(\boldsymbol{e}_i\right) = \boldsymbol{v}_i$$

Then apply this to the matrix A

$$AP\left(\boldsymbol{e}_i\right) = \lambda_i \boldsymbol{v}_i$$

Finally

$$P^{-1}AP\left(\boldsymbol{e}_i\right) = \lambda_i \boldsymbol{e}_i$$

Cinvestav

We have

What if we apply it to the canonical basis elements?

$$P\left(\boldsymbol{e}_i\right) = \boldsymbol{v}_i$$

Then apply this to the matrix A

$$AP\left(\boldsymbol{e}_i\right) = \lambda_i \boldsymbol{v}_i$$

Finally

$$P^{-1}AP\left(\boldsymbol{e}_i\right) = \lambda_i \boldsymbol{e}_i$$

We have

What if we apply it to the canonical basis elements?

$$P\left(\boldsymbol{e}_i\right) = \boldsymbol{v}_i$$

Then apply this to the matrix A

$$AP\left(\boldsymbol{e}_i\right) = \lambda_i \boldsymbol{v}_i$$

Finally

$$P^{-1}AP\left(\boldsymbol{e}_i\right) = \lambda_i \boldsymbol{e}_i$$

Therefore

e_i is the set of eigenvectors of $P^{-1}AP$

$$I = \begin{bmatrix} \boldsymbol{e}_1 & \boldsymbol{e}_2 & \cdots & \boldsymbol{e}_n \end{bmatrix}$$

Then

$$P^{-1}AP = P^{-1}API = \begin{bmatrix} \lambda_1 \boldsymbol{e}_1 & \lambda_2 \boldsymbol{e}_2 & \cdots & \lambda_n \boldsymbol{e}_n \end{bmatrix}$$

Cinvestav

Therefore

$\boldsymbol{e}_i$ is the set of eigenvectors of $P^{-1}AP$

$$I = \begin{bmatrix} \boldsymbol{e}_1 & \boldsymbol{e}_2 & \cdots & \boldsymbol{e}_n \end{bmatrix}$$

Then

$$P^{-1}AP = P^{-1}API = \begin{bmatrix} \lambda_1\boldsymbol{e}_1 & \lambda_2\boldsymbol{e}_2 & \cdots & \lambda_n\boldsymbol{e}_n \end{bmatrix}$$

Therefore

We have that

$$P^{-1}AP = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & \vdots \\ \vdots & 0 & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{bmatrix} = D$$

Therefore

We can see the diagonalization as a decomposition A

$$P\left[P^{-1}AP\right] = IDP$$

In a similar way

$$A = PDP^{-1}$$

Therefore

Only if we have n linearly independent eigenvectors (Different Eigenvalues), we can diagonalize it.

Therefore

We can see the diagonalization as a decomposition A

$$P\left[P^{-1}AP\right] = IDP$$

In a similar way

$$A = PDP^{-1}$$

Therefore

Only if we have n linearly independent eigenvectors (Different Eigenvalues), we can diagonalize it.

Cinvestav

Therefore

We can see the diagonalization as a decomposition A

$$P\left[P^{-1}AP\right] = IDP$$

In a similar way

$$A = PDP^{-1}$$

Therefore

Only if we have n linearly independent eigenvectors (Different Eigenvalues), we can diagonalize it.

Outline

Cinvestav

Some Interesting Properties

What is A^2

- Assuming $n \times n$ matrix that can be diagonlized.

Quite simple

$$A^k = S\Lambda^K S^{-1}$$

What happens if for all $|\lambda_i| < 1$

$$A^k \rightarrow 0 \text{ when } k \longrightarrow \infty$$

Some Interesting Properties

What is A^2

- Assuming $n \times n$ matrix that can be diagonlized.

Quite simple

$$A^k = S\Lambda^K S^{-1}$$

What happens if for all $|\lambda_i| < 1$

$$A^k \rightarrow 0 \text{ when } k \longrightarrow \infty$$

Some Interesting Properties

What is A^2

- Assuming $n \times n$ matrix that can be diagonlized.

Quite simple

$$A^k = S\Lambda^K S^{-1}$$

What happens if for all $|\lambda_i| < 1$

$$A^k \to 0 \text{ when } k \longrightarrow \infty$$

Cinvestav

Some Basic Properties of the Symmetric Matrices

Symmetric Matrix

1. A symmetric matrix has only real eigenvalues.
2. The eigenvectors can be chosen orthonormal.

Spectral Theorem

Theorem

- Every symmetric matrix has the factorization $A = Q\Lambda Q^T$ with the real eigenvalues in Λ and orthonormal eigenvectors $P = Q$.

Proof

- A direct proof from the previous ideas.

Spectral Theorem

Theorem

- Every symmetric matrix has the factorization $A = Q\Lambda Q^T$ with the real eigenvalues in Λ and orthonormal eigenvectors $P = Q$.

Proof

- A direct proof from the previous ideas.