

# Orthogonal Projection

Alvin Kim

July 19, 2024



## 1 Introduction

- Definitions
- Example

## 2 Formulas

- Projection onto a Vector
- Projection onto a Subspace
- The Projection Matrix
- The Error Matrix

## 3 Properties

- Projection Minimizes Error
- $P = P^T$
- $P = P^2$
- Sanity Check:  $E = E^2$

## 4 Special Cases

- Sanity Check:  $A$  has One Column
- $A$  is an Orthonormal Matrix
- Projection from  $\mathbb{R}^3$  onto the  $xz$ -plane
- More Special Cases

## 5 Conclusion

- Key Ideas
- Next Steps



# What is a projection?

$\text{proj}_S \vec{b}$ , the projection of vector  $\vec{b}$  onto subspace  $S$ , is the vector inside of  $S$  closest to  $\vec{b}$ .





# What is a projection?

$\text{proj}_S \vec{b}$ , the projection of vector  $\vec{b}$  onto subspace  $S$ , is the vector inside of  $S$  closest to  $\vec{b}$ .

## Note

In this presentation, projection refers to specifically orthogonal projection,  $\vec{p}$  refers to  $\text{proj}_S \vec{b}$ , and  $\vec{e}$  refers to the error vector  $\vec{e} = \vec{b} - \vec{p}$ . The error vector is also called the rejection.



# What is a projection?

$\text{proj}_S \vec{b}$ , the projection of vector  $\vec{b}$  onto subspace  $S$ , is the vector inside of  $S$  closest to  $\vec{b}$ .

Closest means that the error is orthogonal to  $S$ .



# What is a projection?

$\text{proj}_S \vec{b}$ , the projection of vector  $\vec{b}$  onto subspace  $S$ , is the vector inside of  $S$  closest to  $\vec{b}$ .

Closest means that the error is orthogonal to  $S$ .

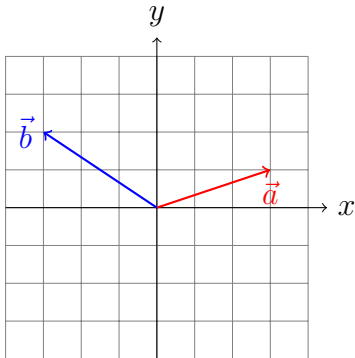
## Note

The dimension of subspace  $S$  can be anything from 0 to the dimension of the space we are working inside.



## Example

# Example

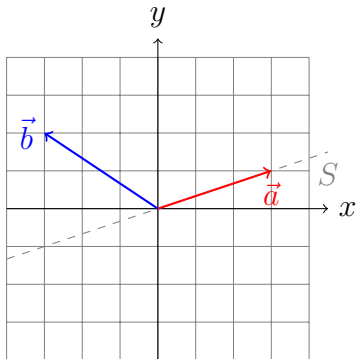


## Example

Draw  $\text{proj}_a b$ .



# Example



## Note

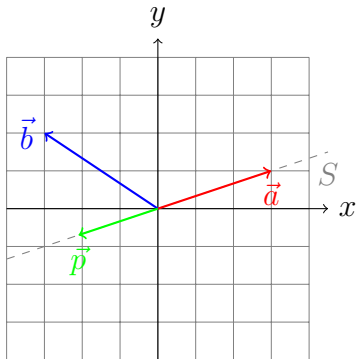
When projecting a vector onto another vector  $\vec{a}$ , the subspace  $S$  is the span of  $\vec{a}$ .





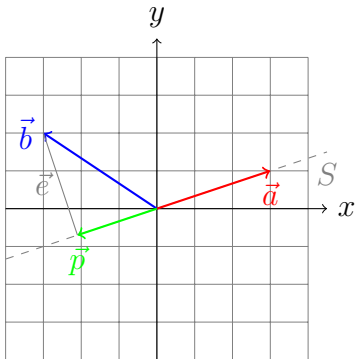
## Example

# Example





# Example



# Projection onto a Vector

Consider projecting  $\vec{b}$  onto  $\vec{a}$ .



# Projection onto a Vector

Consider projecting  $\vec{b}$  onto  $\vec{a}$ .

What we know:

$$\vec{e} = \vec{b} - \vec{p}$$

$$\vec{p} = c\vec{a}$$

$$\vec{e} \cdot \vec{a} = 0$$

# Projection onto a Vector

From that:

Consider projecting  $\vec{b}$  onto  $\vec{a}$ .  
What we know:

$$\vec{e} = \vec{b} - \vec{p}$$

$$\vec{p} = c\vec{a}$$

$$\vec{e} \cdot \vec{a} = 0$$

$$\vec{e} = \vec{b} - c\vec{a}$$

$$(\vec{b} - c\vec{a}) \cdot \vec{a} = 0$$

$$\vec{b} \cdot \vec{a} - c\vec{a} \cdot \vec{a} = 0$$

$$\vec{b} \cdot \vec{a} = c\vec{a} \cdot \vec{a}$$

$$c = \frac{\vec{a} \cdot \vec{b}}{\vec{a} \cdot \vec{a}}$$

$$\vec{p} = \frac{\vec{a} \cdot \vec{b}}{\vec{a} \cdot \vec{a}} \vec{a}$$

# Projection onto a Vector

From that:

Consider projecting  $\vec{b}$  onto  $\vec{a}$ .

What we know:

$$\vec{e} = \vec{b} - \vec{p}$$

$$\vec{p} = c\vec{a}$$

$$\vec{e} \cdot \vec{a} = 0$$

$$\vec{e} = \vec{b} - c\vec{a}$$

$$(\vec{b} - c\vec{a}) \cdot \vec{a} = 0$$

$$\vec{b} \cdot \vec{a} - c\vec{a} \cdot \vec{a} = 0$$

$$\vec{b} \cdot \vec{a} = c\vec{a} \cdot \vec{a}$$

$$c = \frac{\vec{a} \cdot \vec{b}}{\vec{a} \cdot \vec{a}}$$

$$\vec{p} = \frac{\vec{a} \cdot \vec{b}}{\vec{a} \cdot \vec{a}} \vec{a}$$

# Projection onto a Vector

From that:

Consider projecting  $\vec{b}$  onto  $\vec{a}$ .  
What we know:

$$\vec{e} = \vec{b} - \vec{p}$$

$$\vec{p} = c\vec{a}$$

$$\vec{e} \cdot \vec{a} = 0$$

$$\vec{e} = \vec{b} - c\vec{a}$$

$$(\vec{b} - c\vec{a}) \cdot \vec{a} = 0$$

$$\vec{b} \cdot \vec{a} - c\vec{a} \cdot \vec{a} = 0$$

$$\vec{b} \cdot \vec{a} = c\vec{a} \cdot \vec{a}$$

$$c = \frac{\vec{a} \cdot \vec{b}}{\vec{a} \cdot \vec{a}}$$

$$\vec{p} = \frac{\vec{a} \cdot \vec{b}}{\vec{a} \cdot \vec{a}} \vec{a}$$

# Projection onto a Vector

From that:

Consider projecting  $\vec{b}$  onto  $\vec{a}$ .  
What we know:

$$\vec{e} = \vec{b} - \vec{p}$$

$$\vec{p} = c\vec{a}$$

$$\vec{e} \cdot \vec{a} = 0$$

$$\vec{e} = \vec{b} - c\vec{a}$$

$$(\vec{b} - c\vec{a}) \cdot \vec{a} = 0$$

$$\vec{b} \cdot \vec{a} - c\vec{a} \cdot \vec{a} = 0$$

$$\vec{b} \cdot \vec{a} = c\vec{a} \cdot \vec{a}$$

$$c = \frac{\vec{a} \cdot \vec{b}}{\vec{a} \cdot \vec{a}}$$

$$\vec{p} = \frac{\vec{a} \cdot \vec{b}}{\vec{a} \cdot \vec{a}} \vec{a}$$



# Projection onto a Vector

From that:

Consider projecting  $\vec{b}$  onto  $\vec{a}$ .  
What we know:

$$\vec{e} = \vec{b} - \vec{p}$$

$$\vec{p} = c\vec{a}$$

$$\vec{e} \cdot \vec{a} = 0$$

$$\vec{e} = \vec{b} - c\vec{a}$$

$$(\vec{b} - c\vec{a}) \cdot \vec{a} = 0$$

$$\vec{b} \cdot \vec{a} - c\vec{a} \cdot \vec{a} = 0$$

$$\vec{b} \cdot \vec{a} = c\vec{a} \cdot \vec{a}$$

$$c = \frac{\vec{a} \cdot \vec{b}}{\vec{a} \cdot \vec{a}}$$

$$\vec{p} = \frac{\vec{a} \cdot \vec{b}}{\vec{a} \cdot \vec{a}} \vec{a}$$

# Projection onto a Vector

From that:

Consider projecting  $\vec{b}$  onto  $\vec{a}$ .  
What we know:

$$\vec{e} = \vec{b} - \vec{p}$$

$$\vec{p} = c\vec{a}$$

$$\vec{e} \cdot \vec{a} = 0$$

$$\vec{e} = \vec{b} - c\vec{a}$$

$$(\vec{b} - c\vec{a}) \cdot \vec{a} = 0$$

$$\vec{b} \cdot \vec{a} - c\vec{a} \cdot \vec{a} = 0$$

$$\vec{b} \cdot \vec{a} = c\vec{a} \cdot \vec{a}$$

$$c = \frac{\vec{a} \cdot \vec{b}}{\vec{a} \cdot \vec{a}}$$

$$\vec{p} = \frac{\vec{a} \cdot \vec{b}}{\vec{a} \cdot \vec{a}} \vec{a}$$



# Project $\vec{b}$ onto $S$



# Project $\vec{b}$ onto $S$

Suppose that  $\vec{a}_1, \vec{a}_2, \dots, \vec{a}_n$  are a basis of  $S$ , let

$$A = \begin{bmatrix} | & | & & | \\ \vec{a}_1 & \vec{a}_2 & \dots & \vec{a}_n \\ | & | & & | \end{bmatrix}$$

and let  $\vec{p} = A\vec{c}$ . We also know:

$$\vec{e} = \vec{b} - \vec{p}$$

$$C(A)^\perp = N(A^T)$$

# Project $\vec{b}$ onto $S$

Suppose that  $\vec{a}_1, \vec{a}_2, \dots, \vec{a}_n$  are a basis of  $S$ , let

$$A = \begin{bmatrix} | & | & & | \\ \vec{a}_1 & \vec{a}_2 & \dots & \vec{a}_n \\ | & | & & | \end{bmatrix}$$

and let  $\vec{p} = A\vec{c}$ . We also know:

$$\begin{aligned}\vec{e} &= \vec{b} - \vec{p} \\ A^T \vec{e} &= \vec{0}\end{aligned}$$

# Project $\vec{b}$ onto $S$

Suppose that  $\vec{a}_1, \vec{a}_2, \dots, \vec{a}_n$  are a basis of  $S$ , let

$$A = \begin{bmatrix} | & | & \dots & | \\ \vec{a}_1 & \vec{a}_2 & \dots & \vec{a}_n \\ | & | & & | \end{bmatrix}$$

and let  $\vec{p} = A\vec{c}$ . We also know:

$$\begin{aligned}\vec{e} &= \vec{b} - \vec{p} \\ A^T \vec{e} &= \vec{0}\end{aligned}$$

Then,

$$\vec{e} = \vec{b} - A\vec{c}$$

$$A^T(\vec{b} - A\vec{c}) = \vec{0}$$

$$A^T \vec{b} - A^T A \vec{c} = \vec{0}$$

$$A^T \vec{b} = A^T A \vec{c}$$

$$(A^T A)^{-1} A^T \vec{b} = \vec{c}$$

$$\boxed{\vec{c} = (A^T A)^{-1} A^T \vec{b}}$$

$$\boxed{\vec{p} = A(A^T A)^{-1} A^T \vec{b}}$$

# The Projection Matrix

Recall that for some matrix  $A$ , the projection  
 $\vec{p} = A(A^T A)^{-1} A^T \vec{b}$ .

## Observation

For a specific  $A$ ,  $\vec{p}$  is an unchanging linear transformation of  $\vec{b}$ . So, we can represent the linear transformation with a matrix  $P$ .

# The Projection Matrix

Recall that for some matrix  $A$ , the projection  
 $\vec{p} = A(A^T A)^{-1} A^T \vec{b}$ .





# The Projection Matrix

Recall that for some matrix  $A$ , the projection

$$\vec{p} = A(A^T A)^{-1} A^T \vec{b}.$$

Let  $\boxed{P = A(A^T A)^{-1} A^T}$ . Then,  $\vec{p} = P\vec{b}$ .

## Note

$A$  must have linearly independent columns.

# The Error Matrix

The error vector is  $\vec{b}$  projected onto  $S^\perp$ .



# The Error Matrix

The error vector is  $\vec{b}$  projected onto  $S^\perp$ . Since finding the error vector is also a projection, there should be an error matrix  $E$  for some  $P$  such that  $\vec{e} = E\vec{b}$ .



# The Error Matrix

The error vector is  $\vec{b}$  projected onto  $S^\perp$ . Since finding the error vector is also a projection, there should be an error matrix  $E$  for some  $P$  such that  $\vec{e} = E\vec{b}$ .

$$\vec{e} = \vec{b} - \vec{p}$$

$$\vec{e} = I\vec{b} - P\vec{b}$$

$$\vec{e} = (I - P)\vec{b}$$

$$\boxed{E = I - P}$$



# Projection Minimizes Error

Prove that for a vector  $\vec{x} \in S$  where  $\vec{x} \neq \vec{p}$  and  $\vec{d} = \vec{b} - \vec{x}$ , that  $\|\vec{d}\| > \|\vec{e}\|$ .

Hint

$$\|\vec{d}\| > \|\vec{e}\| \iff d^2 > e^2$$

# Projection Minimizes Error

Prove that for a vector  $\vec{x} \in S$  where  $\vec{x} \neq \vec{p}$  and  $\vec{d} = \vec{b} - \vec{x}$ , that  $\|\vec{d}\| > \|\vec{e}\|$ .

$$d^2 = (\vec{b} - \vec{x})^2$$





# Projection Minimizes Error

Prove that for a vector  $\vec{x} \in S$  where  $\vec{x} \neq \vec{p}$  and  $\vec{d} = \vec{b} - \vec{x}$ , that  $\|\vec{d}\| > \|\vec{e}\|$ .

$$\vec{d}^2 = (\vec{b} + \vec{0} - \vec{x})^2$$



# Projection Minimizes Error

Prove that for a vector  $\vec{x} \in S$  where  $\vec{x} \neq \vec{p}$  and  $\vec{d} = \vec{b} - \vec{x}$ , that  $\|\vec{d}\| > \|\vec{e}\|$ .

$$\vec{d}^2 = (\vec{b} - \vec{p} + \vec{p} - \vec{x})^2$$



# Projection Minimizes Error

Prove that for a vector  $\vec{x} \in S$  where  $\vec{x} \neq \vec{p}$  and  $\vec{d} = \vec{b} - \vec{x}$ , that  $\|\vec{d}\| > \|\vec{e}\|$ .

$$\vec{d}^2 = (\vec{b} - \vec{p} + \vec{p} - \vec{x})^2$$



# Projection Minimizes Error

Prove that for a vector  $\vec{x} \in S$  where  $\vec{x} \neq \vec{p}$  and  $\vec{d} = \vec{b} - \vec{x}$ , that  $\|\vec{d}\| > \|\vec{e}\|$ .

$$\vec{d}^2 = (\vec{e} + \vec{p} - \vec{x})^2$$





# Projection Minimizes Error

Prove that for a vector  $\vec{x} \in S$  where  $\vec{x} \neq \vec{p}$  and  $\vec{d} = \vec{b} - \vec{x}$ , that  $\|\vec{d}\| > \|\vec{e}\|$ .

$$\vec{d}^2 = \vec{e}^2 + 2\vec{e} \cdot (\vec{p} - \vec{x}) + (\vec{p} - \vec{x})^2$$



# Projection Minimizes Error

Prove that for a vector  $\vec{x} \in S$  where  $\vec{x} \neq \vec{p}$  and  $\vec{d} = \vec{b} - \vec{x}$ , that  $\|\vec{d}\| > \|\vec{e}\|$ .

$$\vec{d}^2 = \vec{e}^2 + 2\vec{e} \cdot (\vec{p} - \vec{x}) + (\vec{p} - \vec{x})^2$$

# Projection Minimizes Error

Prove that for a vector  $\vec{x} \in S$  where  $\vec{x} \neq \vec{p}$  and  $\vec{d} = \vec{b} - \vec{x}$ , that  $\|\vec{d}\| > \|\vec{e}\|$ .

$$d^2 = e^2 + (\vec{p} - \vec{x})^2$$



# Projection Minimizes Error

Prove that for a vector  $\vec{x} \in S$  where  $\vec{x} \neq \vec{p}$  and  $\vec{d} = \vec{b} - \vec{x}$ , that  $\|\vec{d}\| > \|\vec{e}\|$ .

$$\begin{aligned}\vec{d}^2 &= \vec{e}^2 + (\vec{p} - \vec{x})^2 \\ \vec{d}^2 - \vec{e}^2 &= (\vec{p} - \vec{x})^2\end{aligned}$$



# Projection Minimizes Error

Prove that for a vector  $\vec{x} \in S$  where  $\vec{x} \neq \vec{p}$  and  $\vec{d} = \vec{b} - \vec{x}$ , that  $\|\vec{d}\| > \|\vec{e}\|$ .

$$\begin{aligned}\vec{d}^2 &= \vec{e}^2 + (\vec{p} - \vec{x})^2 \\ \vec{d}^2 - \vec{e}^2 &= (\vec{p} - \vec{x})^2 > 0\end{aligned}$$



# Projection Minimizes Error

Prove that for a vector  $\vec{x} \in S$  where  $\vec{x} \neq \vec{p}$  and  $\vec{d} = \vec{b} - \vec{x}$ , that  $\|\vec{d}\| > \|\vec{e}\|$ .

$$\begin{aligned}\vec{d}^2 &= \vec{e}^2 + (\vec{p} - \vec{x})^2 \\ \vec{d}^2 - \vec{e}^2 &= (\vec{p} - \vec{x})^2 > 0 \\ \vec{d}^2 &> \vec{e}^2\end{aligned}$$







$$P = P^T$$

Prove that  $P = P^T$ .



$$P = P^T$$

Prove that  $P = P^T$ .

Hint

Start with  $P = A(A^T A)^{-1} A^T$ .



$$P = P^T$$

Prove that  $P = P^T$ .

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

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

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

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

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

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

$$P^T = P$$



$$P = P^2$$

Prove that  $P = P^2$ .



$$P = P^2$$

Prove that  $P = P^2$ .

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

$$P^2 = (A(A^T A)^{-1} A^T)^2$$

$$P^2 = (A(A^T A)^{-1} A^T)(A(A^T A)^{-1} A^T)$$

$$P^2 = A(A^T A)^{-1} (A^T A) (A^T A)^{-1} A^T$$

$$P^2 = A(A^T A)^{-1} I A^T$$

$$P^2 = A(A^T A)^{-1} A^T$$

$$P^2 = P$$

Sanity Check:  $E = E^2$

Prove that  $E = E^2$ .

### Reminder

$E$  is also a projection matrix and  $E = I - P$ .

$$E = I - P$$

$$E^2 = (I - P)^2$$

$$E^2 = I - 2IP + P^2$$

$$E^2 = I - 2P + P$$

$$E^2 = I - P$$



# Sanity Check: $A$ has One Column

Suppose  $A = [\vec{a}]$ . Then,  $\vec{p} = A(A^T A)^{-1} A^T b$  and  $\vec{p} = \frac{\vec{a} \cdot \vec{b}}{\vec{a} \cdot \vec{a}} \vec{a}$ .



$A$  is an Orthonormal Matrix

# $A$ is an Orthonormal Matrix

## Note

An orthonormal matrix has unit columns all orthogonal to each other. In other words, an orthonormal matrix  $Q$  will satisfy  $Q^T Q = I$ .





# $A$ is an Orthonormal Matrix

## Note

An orthonormal matrix has unit columns all orthogonal to each other. In other words, an orthonormal matrix  $Q$  will satisfy  $Q^T Q = I$ .

Then,  $P = A(A^T A)^{-1} A^T = A I A^T = A A^T$ .

# $A$ is an Orthonormal Matrix

## Note

An orthonormal matrix has unit columns all orthogonal to each other. In other words, an orthonormal matrix  $Q$  will satisfy  $Q^T Q = I$ .

Then,  $P = A(A^T A)^{-1} A^T = A I A^T = A A^T$ .

Additionally, the coefficients vector,  $\vec{c} = (A^T A)^{-1} A^T \vec{b}$ , will just be  $\vec{c} = A^T \vec{b}$ .

# Projection from $\mathbb{R}^3$ onto the $xz$ -plane



Projection from  $\mathbb{R}^3$  onto the  $xz$ -plane

# Projection from $\mathbb{R}^3$ onto the $xz$ -plane

Hint

What are a basis of the  $xz$ -plane?



# Projection from $\mathbb{R}^3$ onto the $xz$ -plane

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



Projection from  $\mathbb{R}^3$  onto the  $xz$ -plane

# Projection from $\mathbb{R}^3$ onto the $xz$ -plane

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

**Note** $A$  is orthonormal.

Projection from  $\mathbb{R}^3$  onto the  $xz$ -plane

# Projection from $\mathbb{R}^3$ onto the $xz$ -plane

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

$$P = AA^T = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$



Projection from  $\mathbb{R}^3$  onto the  $xz$ -plane

# Projection from $\mathbb{R}^3$ onto the $xz$ -plane

$$P = AA^T = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

## Example

$$P \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x \\ 0 \\ z \end{bmatrix}$$



# More Special Cases (Practice)

- Find the projection matrix from  $\mathbb{R}^3$  onto the y-axis.
- What would  $\vec{p}$  be if  $\vec{b} \perp S$ ? What if  $\vec{b} \in S$ ?



# Key Ideas

- The orthogonal projection of  $\vec{b}$  onto subspace  $S$ ,  $\vec{p} = \text{proj}_S \vec{b} \in S$ , makes the error  $\vec{e} = \vec{b} - \vec{p}$  orthogonal to  $S$ .
- The projection also minimizes the error.
- $P = A(A^T A)^{-1}A^T$
- $P = P^2$
- Other useful equations:  $\vec{p} = Pb$ ,  $E = I - P$ ,  
 $\vec{e} = (A^T A)^{-1}A^T b$ ,  $P = P^T$ ,  $c = \frac{\vec{a} \cdot \vec{b}}{\vec{a} \cdot \vec{a}}$ ,  $\vec{p} = \frac{\vec{a} \cdot \vec{b}}{\vec{a} \cdot \vec{a}} \vec{a}$ ,  $\vec{p} = c\vec{a}$ ,  
 $\vec{p} = A\vec{c}$

# Next Steps

- Applications
  - Graphics
  - Least Squares Regression
  - Orthonormal and Orthogonal Matrices
- Further Learning
  - Practice
  - *Introduction to Linear Algebra* 6th Edition Chapter 4.2 by Gilbert Strang

