Chapter 4 Section 4.1

Take matrix A, $m \times n$, with rank r.

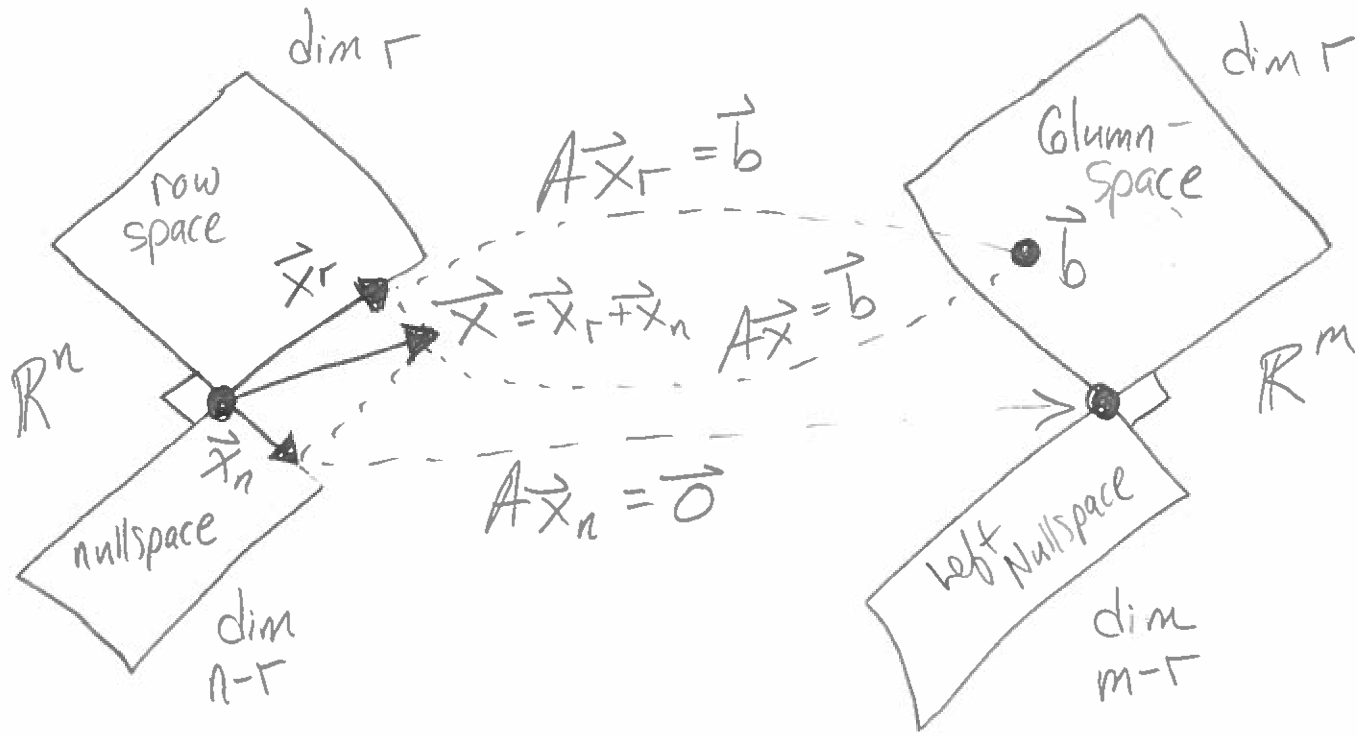
Rowspace$(A) \perp N(A)$ in $\mathbb{R}^n$

Columnspace$(A) \perp$ Left $N(A)$ in $\mathbb{R}^m$

- Every vector in $\mathbb{R}^n$ can be split into its rowspace part $\vec{x}_r$ and its nullspace part $\vec{x}_n$.
- $A\vec{x}$ takes us to $\vec{b}$, but so does $A\vec{x}_r$
- $A\vec{x}_n$ takes us to $\vec{0}$.
- Every vector $\vec{b}$ in the columnspace came from one and only one vector in the rowspace.
- Every A has an r by r invertible matrix hidden inside.
- Rowspace to Columnspace is invertible (if we ....)

The True Big Picture

When matrix A multiplies $\vec{x} = \vec{x}_r + \vec{x}_n$



Theorem: If $\vec{x}$ is in both W and $W^\perp$, then $\vec{x} = \vec{0}$.

Proof: Let $\vec{x} \in W$ and let $\vec{x} \in W^\perp$.

We know for any vector $\vec{w} \in W$ and $\vec{v} \in W^\perp$, that $\vec{w} \cdot \vec{v} = 0$.

Take $\vec{x}$ to be both vectors.

Then $\vec{x} \cdot \vec{x} = x_1^2 + x_2^2 + \cdots x_n^2 = 0$ iff all $x_i = 0$.

Thus, $\vec{x} = \vec{0}$.

<u>Example:</u> Find a basis for the orthogonal complement of the rowspace of A.

$$A = \begin{bmatrix} 1 & 0 & 2 \\ 1 & 1 & 4 \end{bmatrix} \sim \begin{array}{c} \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 2 \end{bmatrix} \\ \begin{array}{ccc} x_1 & x_2 & x_3 \\ P & P & F \end{array} \end{array} = \text{rref}(A).$$

$N(A)$ has basis $\vec{s}_1 = (-2, -2, 1)$

<u>Next</u>, take the vector $\vec{x} = (3, 3, 3)$ and split it into $\vec{x} = \vec{x}_r + \vec{x}_n$

We need $$\vec{x} = \begin{bmatrix} 3 \\ 3 \\ 3 \end{bmatrix} = \underbrace{c_1 \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} + c_2 \begin{bmatrix} 1 \\ 1 \\ 4 \end{bmatrix}}_{\vec{x}_r} + \underbrace{c_3 \begin{bmatrix} -2 \\ -2 \\ 1 \end{bmatrix}}_{\vec{x}_n}$$

(3 equations, 3 unknowns, 1 solution)

$c_1 = 0, \quad c_2 = 1, \quad c_3 = -1$

$$\vec{x} = \begin{bmatrix} 3 \\ 3 \\ 3 \end{bmatrix} = \underbrace{\begin{bmatrix} 1 \\ 1 \\ 4 \end{bmatrix}}_{\vec{x}_r} + \underbrace{\begin{bmatrix} 2 \\ 2 \\ -1 \end{bmatrix}}_{\vec{x}_n}$$

Example: The equation $x - 3y - 4z = 0$ describes a plane in $\mathbb{R}^3$.

- The plane is the nullspace of $A = \begin{bmatrix} 1 & -3 & -4 \end{bmatrix}$, since $A \begin{bmatrix} x \\ y \\ z \end{bmatrix} = 0$

- Basis for $N(A)$ is $\vec{s}_1 = \begin{bmatrix} 3 \\ 1 \\ 0 \end{bmatrix}$, $\vec{s}_2 = \begin{bmatrix} 4 \\ 0 \\ 1 \end{bmatrix}$

- Basis for $C(A^T)$ is $\vec{z} = \begin{bmatrix} 1 & -3 & -4 \end{bmatrix}$

- Take vector $\vec{v}$ in $\mathbb{R}^3$, $\vec{v} = (6, 4, 5)$ and split it into $\vec{v} = \vec{v_r} + \vec{v_n}$

$$\vec{v} = \begin{bmatrix} 6 \\ 4 \\ 5 \end{bmatrix} = \underbrace{c_1 \begin{bmatrix} 1 \\ -3 \\ -4 \end{bmatrix}}_{\vec{v_r}} + \underbrace{c_2 \begin{bmatrix} 3 \\ 1 \\ 0 \end{bmatrix} + c_3 \begin{bmatrix} 4 \\ 0 \\ 1 \end{bmatrix}}_{\vec{v_n}}$$

Solving gives $c_1 = -1$, $c_2 = 1$, $c_3 = 1$

$$\vec{v_r} = (-1, 3, 4)$$

$$\vec{v_n} = (3, 1, 0) + (4, 0, 1) = (7, 1, 1)$$

$$\vec{v} = \vec{v_r} + \vec{v_n} = (-1, 3, 4) + (7, 1, 1)$$

$$\vec{v} = (6, 4, 5)$$

Theorem: Every vector in $C(A)$ came from a unique vector in $C(A^T)$.

Proof: $A\vec{x}_r$ and $A\vec{x}_r'$ are both vectors in the columnspace of A

Suppose $A\vec{x}_r = A\vec{x}_r'$. (Show $\vec{x}_r = \vec{x}_r'$)

Then $A\vec{x}_r - A\vec{x}_r' = \vec{0}$

$$A(\vec{x}_r - \vec{x}_r') = \vec{0}$$

Now, vector $\vec{x}_r - \vec{x}_r'$ is in the nullspace of A.

But, vector $\vec{x}_r - \vec{x}_r'$ is also in the rowspace of A.

The only vector in both spaces is the zero vector,

so $\vec{x}_r - \vec{x}_r' = \vec{0}$ OR

$$\vec{x}_r = \vec{x}_r'.$$

Done.

Example: Suppose V is spanned by vectors $(1,2,2,3)$ and $(1,3,3,2)$. Find two vectors that span $V^\perp$.

Solution: If we had $A = \begin{bmatrix} 1 & 2 & 2 & 3 \\ 1 & 3 & 3 & 2 \end{bmatrix}$, then V is the rowspace of A. Find a basis for $N(A)$ and we're done.

$$A \sim \begin{bmatrix} 1 & 2 & 2 & 3 \\ 0 & 1 & 1 & -1 \end{bmatrix} \sim \begin{array}{c} \begin{bmatrix} 1 & 0 & 0 & 5 \\ 0 & 1 & 1 & -1 \end{bmatrix} \\ \begin{matrix} & & F & F \end{matrix} \end{array}$$

Take $\vec{s}_1 = (0, -1, 1, 0)$

$\vec{s}_2 = (-5, 1, 0, 1)$

Theorem: If $A^T A\vec{x} = \vec{0}$, then $A\vec{x} = \vec{0}$.

Proof: $A\vec{x}$ is in the nullspace of A^T (i.e. left nullspace). Also, $A\vec{x}$ is in the columnspace of A. The only vector in both $C(A)$ and $N(A^T)$ is the $\vec{0}$. Thus, $A\vec{x} = \vec{0}$.

Theorem: If $\vec{y} \perp \vec{u}$ and $\vec{y} \perp \vec{v}$, then $\vec{y} \perp (c\vec{u} + d\vec{v})$.

Proof: $\vec{y}^T\vec{u} = 0$ and $\vec{y}^T\vec{v} = 0$.

Then $\vec{y}^T(c\vec{u} + d\vec{v}) =$

$\vec{y}^T(c\vec{u}) + \vec{y}^T(d\vec{v}) =$

$c\vec{y}^T\vec{u} + d\vec{y}^T\vec{v} =$

$c(0) + d(0) = 0$.

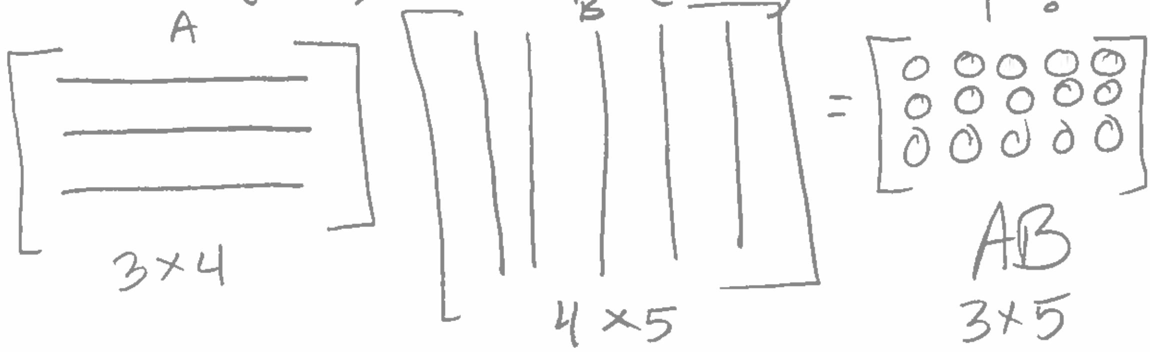
Example: A is 3 by 4

B is 4 by 5

$AB = 0$ matrix

Prove $\text{rank}(A) + \text{rank}(B) \leq 4$.

Proof:



* Columnspace(B) is contained in the Nullspace(A).

  (Null(A) could be bigger than the set of linear comb. of columns of B).

* $\dim(C(B)) \leq \dim(N(A))$

  $\text{rank}(B) \leq 4 - \text{rank}(A)$

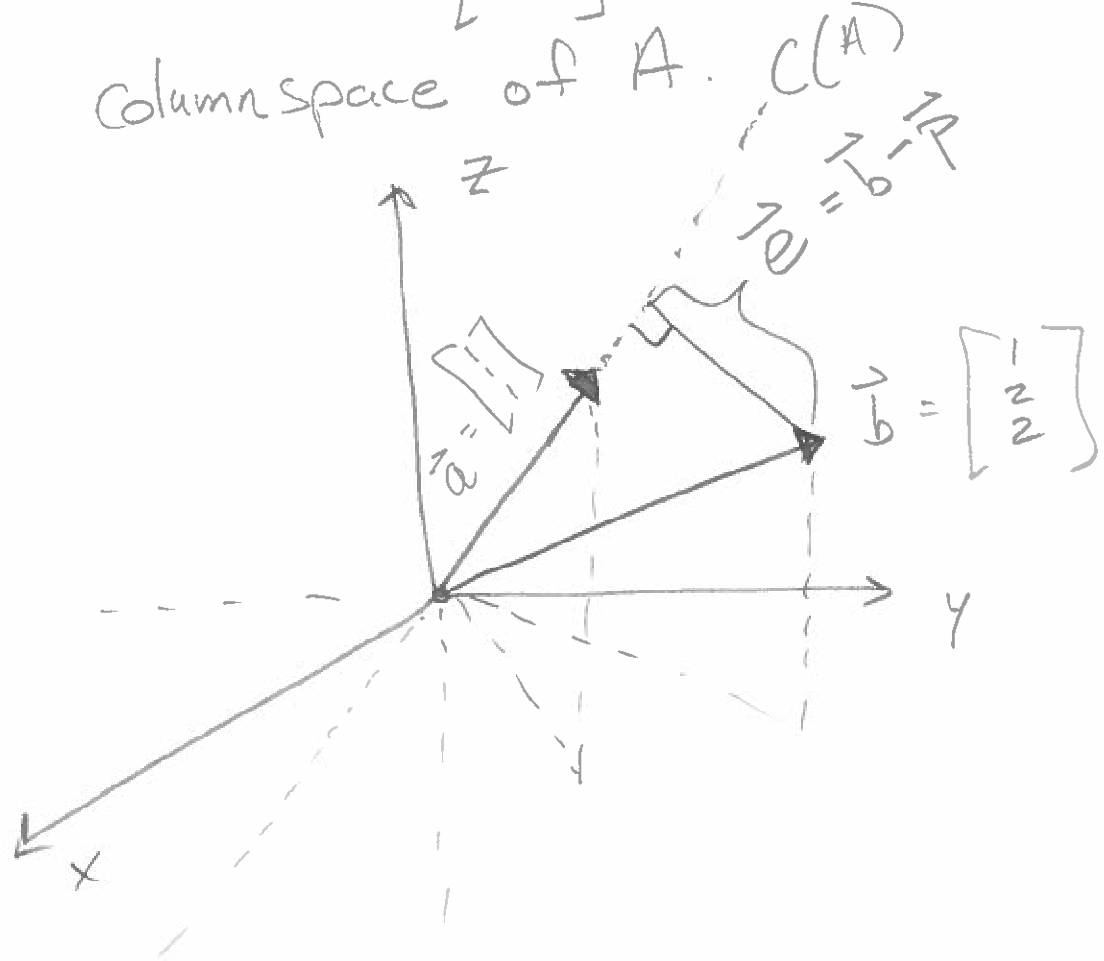
  $\text{rank}(A) + \text{rank}(B) \leq 4$.

Section 4.2 Projections

* We can project vectors onto any of the subspaces of $\mathbb{R}^m$. We need a <u>basis</u> to project onto.
* Put the basis vectors into matrix A (columns must be independent).

<u>Example</u>: Projecting onto the line $A = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$.

Project $\vec{b} = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}$ onto the columnspace of A.



Projection vector : $\vec{p} = \hat{x}\vec{a}$

Error vector : $\vec{e} = \vec{b} - \vec{p}$

$$\vec{e} = \vec{b} - \hat{x}\vec{a}$$

Rule : error vector $\vec{e}$ is perp to $\vec{a}$

$$\vec{a} \perp \vec{e} \quad \text{or} \quad \vec{a} \cdot \vec{e} = 0$$

$$\vec{a}^T \vec{e} = 0$$

$$\vec{a}^T (\vec{b} - \hat{x}\vec{a}) = 0$$

$$\vec{a}^T \vec{b} - \hat{x}\vec{a}^T \vec{a} = 0$$

This is the multiplier times vector $\vec{a}$ $\longrightarrow$ $\hat{x} = \dfrac{\vec{a}^T \vec{b}}{\vec{a}^T \vec{a}}$

So $\vec{p} = \hat{x}\vec{a} = \left(\dfrac{\vec{a}^T \vec{b}}{\vec{a}^T \vec{a}}\right)\vec{a}$

To get the Matrix P that works for any $\vec{b}$, do some creative housekeeping $\rightarrow$

$P\vec{b} = \vec{p}$ so...

$$\vec{p} = \vec{a}\left(\frac{\vec{a}^T \vec{b}}{\vec{a}^T \vec{a}}\right) = \left(\frac{\vec{a}\,\vec{a}^T}{\vec{a}^T \vec{a}}\right)\vec{b} = P\vec{b}$$

With numbers: $\vec{a} = (1,1,1)$
$\vec{b} = (1,2,2)$

$$\hat{x} = \frac{\vec{a}^T\vec{b}}{\vec{a}^T\vec{a}} = \frac{5}{3}$$

$$\vec{p} = \hat{x}\vec{a} = \frac{5}{3}\begin{bmatrix}1\\1\\1\end{bmatrix} = \begin{bmatrix}5/3\\5/3\\5/3\end{bmatrix}$$

$$\vec{e} = \vec{b} - \vec{p} = \begin{bmatrix}1\\2\\2\end{bmatrix} - \begin{bmatrix}5/3\\5/3\\5/3\end{bmatrix} = \begin{bmatrix}-2/3\\1/3\\1/3\end{bmatrix}$$

$$P = \frac{\vec{a}\vec{a}^T}{\vec{a}^T\vec{a}} = \frac{\begin{bmatrix}1\\1\\1\end{bmatrix}\begin{bmatrix}1 & 1 & 1\end{bmatrix}}{\begin{bmatrix}1 & 1 & 1\end{bmatrix}\begin{bmatrix}1\\1\\1\end{bmatrix}} = \frac{1}{3}\begin{bmatrix}1 & 1 & 1\\1 & 1 & 1\\1 & 1 & 1\end{bmatrix}$$

Check that $\vec{e} \perp \vec{a}$:

$$(-2/3, 1/3, 1/3)\cdot(1,1,1) = 0 \checkmark$$

check that $P\vec{b} = \vec{p}$

$$\frac{1}{3}\begin{bmatrix}1 & 1 & 1\\1 & 1 & 1\\1 & 1 & 1\end{bmatrix}\begin{bmatrix}1\\2\\2\end{bmatrix} = \frac{1}{3}\begin{bmatrix}5\\5\\5\end{bmatrix} = \begin{bmatrix}5/3\\5/3\\5/3\end{bmatrix}$$

If you project the projection, you don't change anything on 2nd time

i.e. $P(P\vec{b}) = P(\vec{p}) = \vec{p}$

Matrix P is <u>idempotent</u> : $P^2 = P$

$$P^2 = \left(\frac{\vec{a}\vec{a}^T}{\vec{a}^T\vec{a}}\right)\left(\frac{\vec{a}\vec{a}^T}{\vec{a}^T\vec{a}}\right) = \frac{\vec{a}\vec{a}^T}{\vec{a}^T\vec{a}} \checkmark$$

With numbers: $P^2 = \frac{1}{3}\begin{bmatrix}1&1&1\\1&1&1\\1&1&1\end{bmatrix}\frac{1}{3}\begin{bmatrix}1&1&1\\1&1&1\\1&1&1\end{bmatrix}$

$$= \frac{1}{9}\begin{bmatrix}3&3&3\\3&3&3\\3&3&3\end{bmatrix} = P \checkmark$$

* We can project onto the orthogonal complement by taking

$$P^*_{\text{ORTH. COMPLEMENT}} = I - P$$

* With numbers: Project $\vec{b} = \begin{bmatrix}1\\2\\2\end{bmatrix}$ onto the orthogonal complement of the line $\vec{a} = \begin{bmatrix}1\\1\\1\end{bmatrix}$.

$$I-P = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - \frac{1}{3}\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 2/3 & -1/3 & -1/3 \\ -1/3 & 2/3 & -1/3 \\ -1/3 & -1/3 & 2/3 \end{bmatrix}$$

$$(I-P)\vec{b} = \begin{bmatrix} -2/3 \\ 1/3 \\ 1/3 \end{bmatrix}.$$

* BTW, our two projections should be perpendicular!

$$(5/3, 5/3, 5/3)^T(-2/3, 1/3, 1/3) = 0 \checkmark$$

Projection onto Subspaces

* Want to solve $A\vec{x} = \vec{b}$, but RHS $\vec{b} \notin$ columnspace of A!
* <u>Best alternative</u>: Find the projection $\vec{p}$, it will be the vector closest to $\vec{b}$. Solve that system instead!!

No: $A\vec{x} = \vec{b}$ YES: $A\hat{x} = \vec{p}$

Requirements: Columns of A must be linearly independent

Solution: Solve for the projection:

$$\vec{p} = \hat{x}_1 \vec{a}_1 + \hat{x}_2 \vec{a}_2 + \cdots + \hat{x}_n \vec{a}_n = A\hat{x}$$

with $\hat{x} = (A^T A)^{-1} A^T b$, $P = A(A^T A)^{-1} A^T$

The method finds ① $\hat{x}$ (The best comb. of columns of A)

② $\vec{p} = A\hat{x}$ (The projection)

3. P (The projection matrix)

Key Idea: The error vector is perpendicular to the columns of matrix A.

$\vec{e} = \vec{b} - \vec{p} = \vec{b} - A\hat{x}$ is perpendicular to all columns $\vec{a}_i$ in A.

$$\vec{a}_1^T(\vec{b}-A\hat{x}) = 0$$
$$\vec{a}_2^T(\vec{b}-A\hat{x}) = 0$$
$$\vdots \quad \vdots \quad \vdots \quad \vdots$$
$$\vec{a}_n^T(\vec{b}-A\hat{x}) = 0$$

The left nullspace is perpendicular to the columnspace!

OR

$$\underset{n\times m}{\begin{bmatrix} \vec{a}_1^T \\ \vdots \\ \vec{a}_n^T \end{bmatrix}} \underset{m\times 1}{\begin{bmatrix} \vec{b}-A\hat{x} \end{bmatrix}} = \underset{n\times 1}{\begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}}$$

$$A^T(\vec{b}-A\hat{x}) = A^T\vec{b} - A^TA\hat{x} = \vec{0}$$

OR $\quad A^TA\hat{x} = A^T\vec{b}$

$$\hat{x} = (A^TA)^{-1}A^T\vec{b}$$

and (A^TA) is invertible exactly when columns of A are in~~dependent~~.

Example: Project $\vec{b} = (2,3,4)$ onto columnspace of $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$

Note: $A\vec{x} = \vec{b}$ not solvable, $\vec{b} \notin C(A)$.

$$\hat{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = (A^T A)^{-1} A^T \vec{b}$$

$$A^T A = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}$$

$$(A^T A)^{-1} = \frac{1}{1} \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix}$$

$$(A^T A)^{-1} A^T = \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

$$\hat{x} = (A^T A)^{-1} A^T \vec{b} = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix} = \begin{bmatrix} -1 \\ 3 \end{bmatrix} = \begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \end{bmatrix}$$

The projection vector $\vec{p}$ is the closest vector to $\vec{b}$

$$\vec{p} = A\hat{x} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} -1 \\ 3 \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \\ 0 \end{bmatrix}$$

Projection matrix P:

$$P = A(A^TA)^{-1}A^T = \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

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

This P projects any $\vec{b} \in \mathbb{R}^3$ onto $C(A)$.

The error vector $\vec{e} = \vec{b} - \vec{p} = \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix} - \begin{bmatrix} 2 \\ 3 \\ 0 \end{bmatrix}$

$$= \begin{bmatrix} 0 \\ 0 \\ 4 \end{bmatrix}$$

is perp. to $C(A)$.

Example: (128.)

Project the vector $\vec{b} = (1,0,1,0,2)$ onto the plane in $\mathbb{R}^5$ spanned by $\vec{v}_1 = (0,0,7,7,7)$ and $\vec{v}_2 = (2,4,-2,-2,-2)$.

Report $\hat{x}$, $\vec{p}$, P, and $\vec{e}$.

$$A = \underset{5\times 2}{\begin{bmatrix} 0 & 2 \\ 0 & 4 \\ 7 & -2 \\ 7 & -2 \\ 7 & -2 \end{bmatrix}}$$ has independent columns ✓

$$A^TA = \underset{2\times 2}{\begin{bmatrix} 147 & -42 \\ -42 & 32 \end{bmatrix}}, \quad (A^TA)^{-1} = \begin{bmatrix} 8/735 & 1/70 \\ 1/70 & 1/20 \end{bmatrix}$$

$$P = A(A^TA)^{-1}A^T = \underset{5\times 5}{\begin{bmatrix} 1/5 & 2/5 & 0 & 0 & 0 \\ 2/5 & 4/5 & 0 & 0 & 0 \\ 0 & 0 & 1/3 & 1/3 & 1/3 \\ 0 & 0 & 1/3 & 1/3 & 1/3 \\ 0 & 0 & 1/3 & 1/3 & 1/3 \end{bmatrix}}$$

$$\vec{p} = P\vec{b} = P\begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 2 \end{bmatrix} = \begin{bmatrix} 1/5 \\ 2/5 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

$$\hat{x} = \begin{bmatrix} 6/35 \\ 1/10 \end{bmatrix}$$

$$\vec{e} = \vec{b} - \vec{p} = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 2 \end{bmatrix} - \begin{bmatrix} 1/5 \\ 2/5 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 4/5 \\ -2/5 \\ 0 \\ -1 \\ 1 \end{bmatrix}$$

is perp. to the columns of A.

LAST Example: Project $\vec{b} = \begin{bmatrix} 4 \\ 4 \\ 6 \end{bmatrix}$ onto $C(A)$, $A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 0 & 1 \end{bmatrix}$

Note: $\vec{b} \in C(A)$

Use TI to get $P = A(A^TA)^{-1}A^T$

$$P = \begin{bmatrix} 1/2 & 1/2 & 0 \\ 1/2 & 1/2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

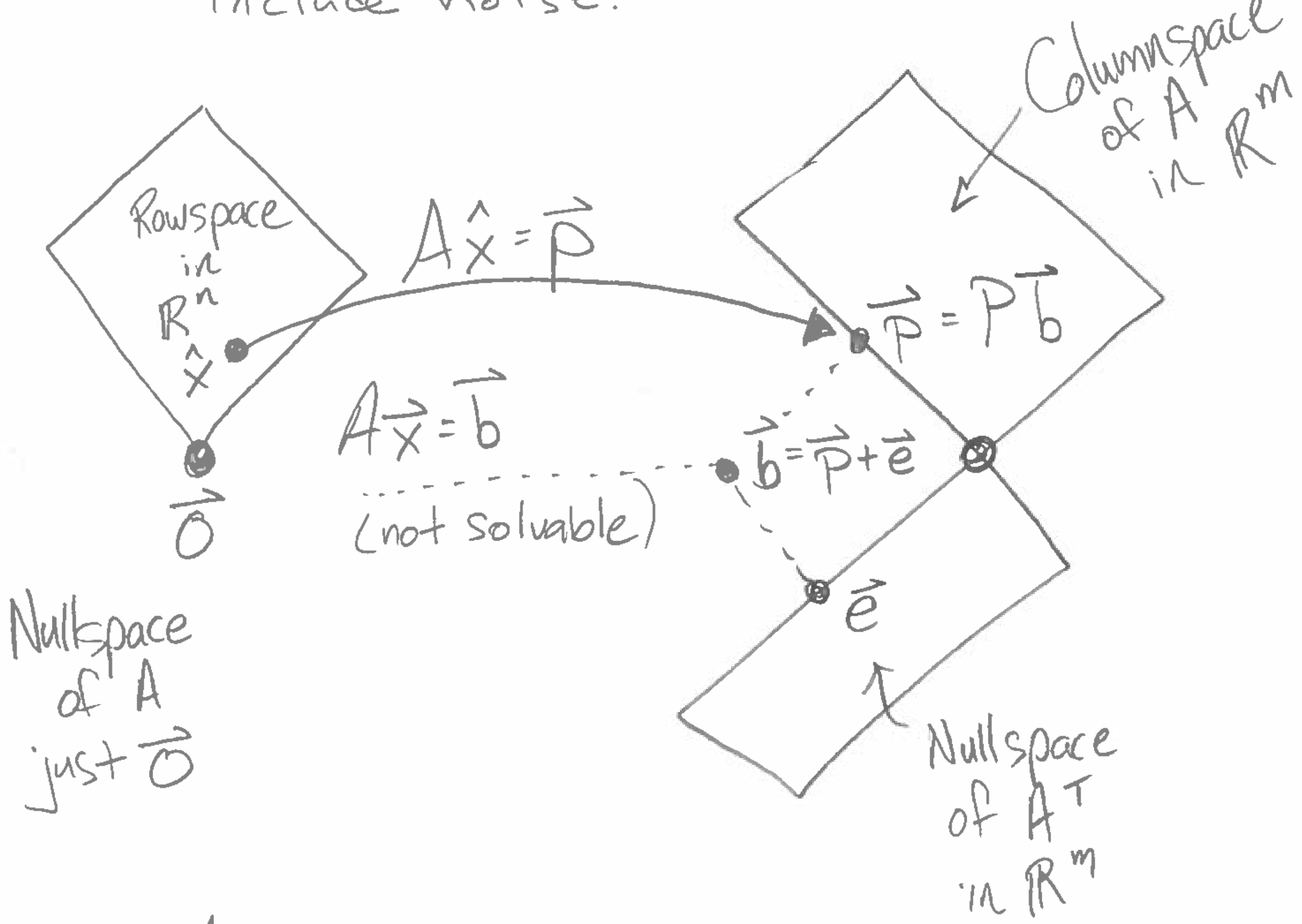
Projection is $\vec{p} = P\vec{b} = P\begin{bmatrix} 4 \\ 4 \\ 6 \end{bmatrix} = \begin{bmatrix} 4 \\ 4 \\ 6 \end{bmatrix}$

So $\vec{b}$ didn't move.

Other vectors not on the plane will move!!

Section 4.3 Least Squares

- $A\vec{x} = \vec{b}$ has no solution (too many equations)
- Cannot get error $\vec{e} = \vec{b} - \vec{p}$ down to zero (our measurements include noise.



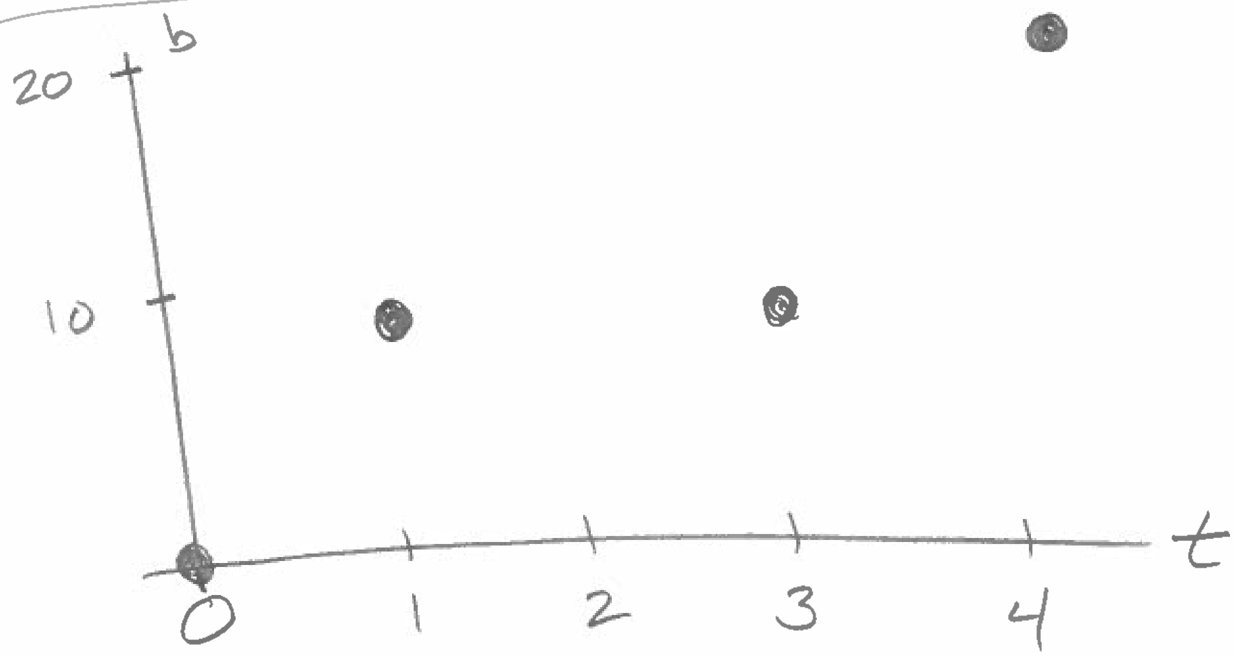
- A is m×n with indep. columns
- $\vec{b} \notin C(A)$.
- Before, $\vec{x}$ went to $\vec{b} = A\vec{x}$ and we split $\vec{x} = \vec{x}_r + \vec{x}_n$
- Now, split $\vec{b} = \vec{p} + \vec{e}$

Normal Equations:

$A^T A \hat{x} = A^T \vec{b}$ will give the "least squares" solution (best fit) (minimizes error $\vec{e} = \vec{b} - \vec{p}$).

Example: We have four data points (t, b)
$(0,0), (1,8), (3,8), (4,20)$
(Not on a straight line)

Each measurement is an estimate of the true equation $C + Dt = b$
with t = indep. variable $(y = mx + b)$
b = dep. variable.



$A\vec{x} = \vec{b}$ is built from our four equations:

$$C + Dt_1 = b_1$$
$$C + Dt_2 = b_2$$
$$C + Dt_3 = b_3$$
$$C + Dt_4 = b_4$$

with unknowns C, D

$$\underset{4\times 2}{\begin{bmatrix} 1 & t_1 \\ 1 & t_2 \\ 1 & t_3 \\ 1 & t_4 \end{bmatrix}} \underset{2\times 1}{\begin{bmatrix} C \\ D \end{bmatrix}} = \underset{4\times 1}{\begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \end{bmatrix}}$$

We have $A = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 3 \\ 1 & 4 \end{bmatrix}$, $\vec{b} = \begin{bmatrix} 0 \\ 8 \\ 8 \\ 20 \end{bmatrix}$, $\vec{b} \notin C(A)$

$A^T A \hat{x} = A^T \vec{b}$ or

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

$$A^T A = \begin{bmatrix} 4 & 8 \\ 8 & 26 \end{bmatrix}, \quad A^T \vec{b} = \begin{bmatrix} 36 \\ 112 \end{bmatrix}$$

$$\hat{x} = \begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \end{bmatrix} = \begin{bmatrix} C \\ D \end{bmatrix} = \begin{bmatrix} 4 & 8 \\ 8 & 26 \end{bmatrix}^{-1} \begin{bmatrix} 36 \\ 112 \end{bmatrix} = \begin{bmatrix} 1 \\ 4 \end{bmatrix}$$

And the least squares equation is

$\hat{b} = C + Dt$ (Graph it)

$\hat{b} = 1 + 4t$

Find your four predicted $\hat{b}_i$:

$$A\hat{x} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 3 \\ 1 & 4 \end{bmatrix} \begin{bmatrix} 1 \\ 4 \end{bmatrix} = \begin{bmatrix} 1 \\ 5 \\ 13 \\ 17 \end{bmatrix} = \vec{p} = \begin{bmatrix} \hat{b}_1 \\ \hat{b}_2 \\ \hat{b}_3 \\ \hat{b}_4 \end{bmatrix}$$

Find your four errors (residuals) e_i:

$$\vec{e} = \vec{b} - \vec{p} = \begin{bmatrix} 0 \\ 8 \\ 8 \\ 20 \end{bmatrix} - \begin{bmatrix} 1 \\ 5 \\ 13 \\ 17 \end{bmatrix} = \begin{bmatrix} -1 \\ 3 \\ -5 \\ 3 \end{bmatrix} = \begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \end{bmatrix}$$

<u>Notice</u>: $\sum e_i = 0$

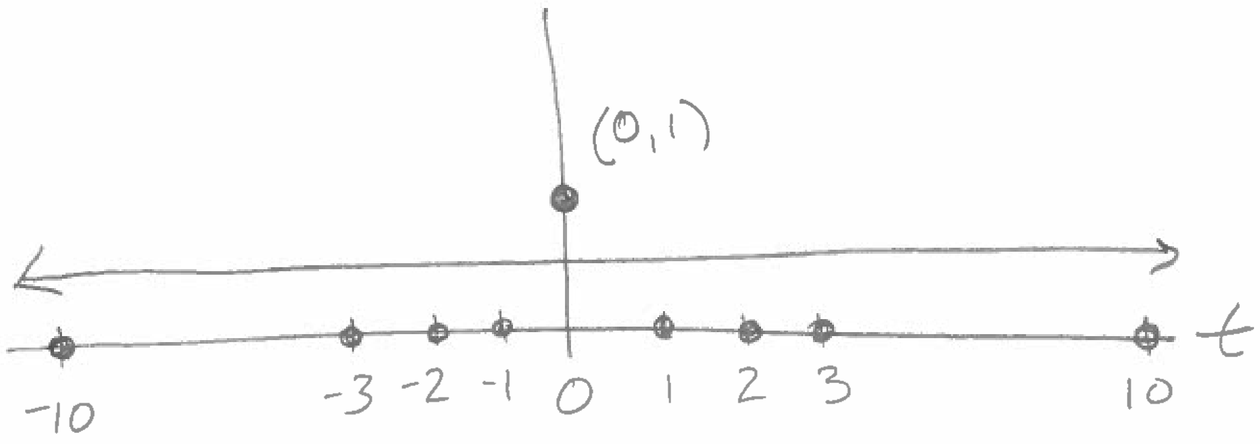
<u>Notice</u>: $\vec{e}^T \vec{a}_1 = 0$, $\vec{e}^T \vec{a}_2 = 0$

Shortest Distance from $\vec{b}$ to $C(A)$ is $\|\vec{e}\| \approx 6.63$

Example: Suppose we take 21 measurements at equally spaced times $t = -10, -9, \ldots, -1, 0, 1, \ldots, 9, 10$

All measurements are $b_i = 0$ except $b_{11} = 1$ at middle time $t = 0$.

(a.) Fit the least squares equation, Solve for $\hat{C}, \hat{D}$ for straight line $C + Dt = b$.



Slope should be zero.

Knowing t tells us nothing about the values of b!

$$A = \begin{bmatrix} 1 & -10 \\ 1 & -9 \\ 1 & -8 \\ \vdots & \vdots \\ 1 & 8 \\ 1 & 9 \\ 1 & 10 \end{bmatrix}, \quad \vec{b} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \qquad \hat{x} = \begin{bmatrix} \hat{C} \\ \hat{D} \end{bmatrix}$$

21×2 $\qquad$ 21×1 $\qquad$ 2×1

$\vec{b} \notin C(A)$

$(A^T A)\hat{x} = A^T\vec{b}$ is

$$\begin{bmatrix} 21 & 0 \\ 0 & 770 \end{bmatrix} \begin{bmatrix} \hat{C} \\ \hat{D} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

has $\hat{C} = \frac{1}{21}$

$\hat{D} = 0$

line of best fit is

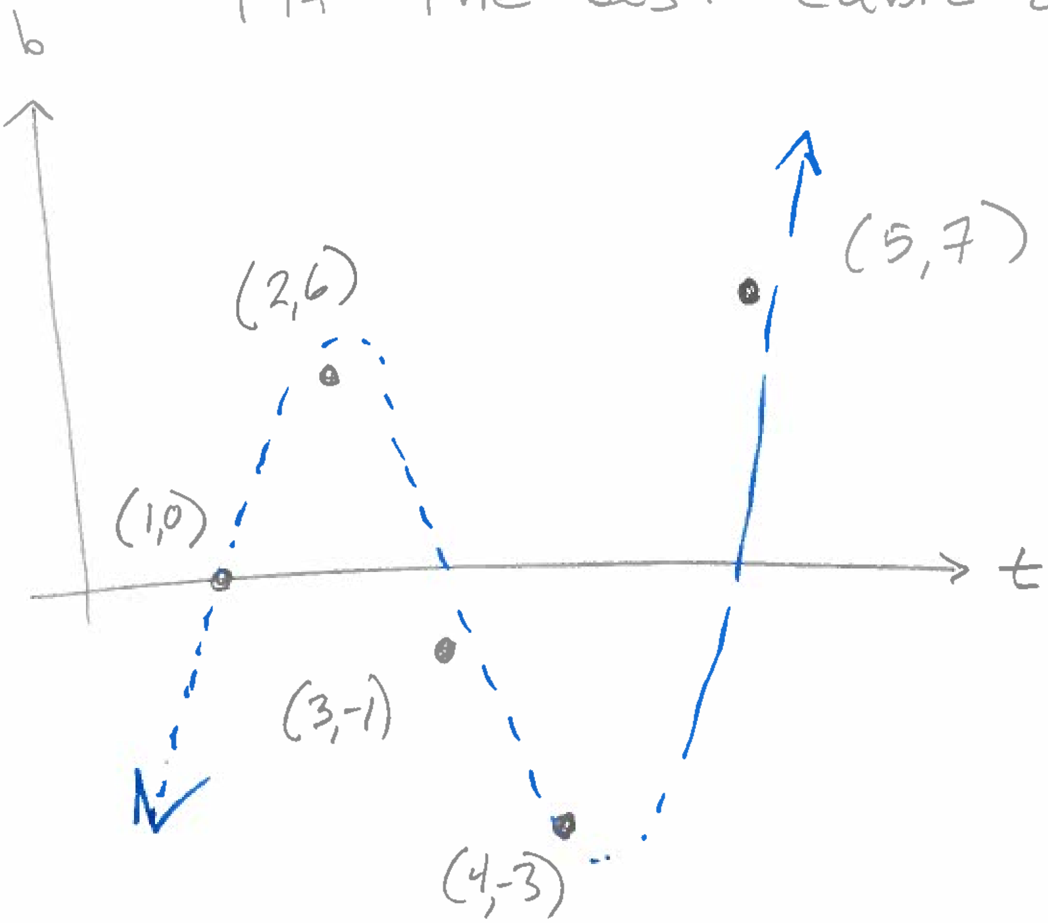
$$\hat{b} = \frac{1}{21} + 0t$$

$$\hat{b} = \frac{1}{21}$$

Example: A certain phenomenom is known to be cubic.

We have measurements at $(1,0)$, $(2,6)$, $(3,-1)$, $(4,-3)$, $(5,7)$

Fit the best cubic approx.



Equation: $C + Dt + Et^2 + Ft^3 = b$

$$A = \underbrace{\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 4 & 8 \\ 1 & 3 & 9 & 27 \\ 1 & 4 & 16 & 64 \\ 1 & 5 & 25 & 125 \end{bmatrix}}_{5\times 4}, \quad \vec{b} = \underbrace{\begin{bmatrix} 0 \\ 6 \\ -1 \\ -3 \\ 7 \end{bmatrix}}_{5\times 1}, \quad \hat{x} = \underbrace{\begin{bmatrix} \hat{C} \\ \hat{D} \\ \hat{E} \\ \hat{F} \end{bmatrix}}_{4\times 1}$$

TI

$$\hat{x} = (A^TA)^{-1}A^T\vec{b} = \begin{bmatrix} -141/5 \\ 926/21 \\ -499/28 \\ 25/12 \end{bmatrix}$$

The equation of best fit is:

$$-141/5 + \left(\frac{926}{21}\right)t - \left(\frac{499}{28}\right)t^2 + \left(\frac{25}{12}\right)t^3 = \hat{b}$$

The projection vector is

$$\vec{p} = A\hat{x} = \begin{bmatrix} 11/70 \\ 188/35 \\ -2/35 \\ -127/35 \\ 501/70 \end{bmatrix} \approx \begin{bmatrix} 0.157 \\ 5.371 \\ -0.057 \\ -3.629 \\ 7.157 \end{bmatrix}$$

Recall

$$\vec{b} = \begin{bmatrix} 0 \\ 6 \\ -1 \\ -3 \\ 7 \end{bmatrix}$$

$$\vec{e} = \vec{b} - \vec{p} = \begin{bmatrix} -11/70 \\ 22/35 \\ -33/35 \\ 22/35 \\ -11/70 \end{bmatrix},$$

$\vec{e} \perp \vec{a}_i$ for $i = 1, 2, 3, 4$

$\|\vec{e}\| \approx 1.315$ is a minimum.

$$\sum e_i = 0$$

Section 4.4 Orthonormal Bases, Gram-Schmidt.

The "best" basis for a vectorspace has mutually perpendicular basis vectors of length 1.

Our previous projections did not use $\vec{q}_1, \vec{q}_2, \ldots, \vec{q}_n$ (orthonormal) basis vectors.

$\vec{q}_1, \vec{q}_2, \ldots, \vec{q}_n$ are orthonormal if

$$\vec{q}_i^{\,T}\vec{q}_j = \begin{cases} 0 & \text{if } i \neq j \text{ (perp.)} \\ 1 & \text{if } i = j \text{ (unit)} \end{cases}$$

- Matrix Q is not necessarily square, but $Q^TQ = I$
- If Q is square, $Q^TQ = I$ gives $Q^T = Q^{-1}$

If we choose Q instead of A to do projections:

$$A^T A \hat{x} = A^T \vec{b} \quad \text{becomes}$$

$$Q^T Q \hat{x} = Q^T \vec{b}$$

$$\hat{x} = Q^T \vec{b}$$

$$\vec{p} = A\hat{x} = Q\hat{x} = QQ^T \vec{b}$$

and QQ^T is the projection matrix.

(There is nothing to invert, A^{-1} not required to be computed).

* Finding Q has a cost, so either do work finding Q, or do work solving $\hat{x} = (A^T A)^{-1} A^T \vec{b}$.

* Q is more mathematically elegant,

* Orthogonal matrices preserve length and angles

$\|Q\vec{x}\| = \|\vec{x}\|$ for all $\vec{x}$

$(Q\vec{x})^T(Q\vec{y}) = \vec{x}^T Q^T Q \vec{y} = \vec{x}^T\vec{y}$ for all $\vec{x}, \vec{y}$.

Example: Famous $Q = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$ in $\mathbb{R}^2$

$$I = Q^TQ = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$P = QQ^T = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}\begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Take $\vec{x} = \begin{bmatrix} 3 \\ 3 \end{bmatrix}$, $\|\vec{x}\| = \sqrt{9+9} = \sqrt{18} = 3\sqrt{2}$

$$Q\vec{x} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}\begin{bmatrix} 3 \\ 3 \end{bmatrix} = \begin{bmatrix} 3\cos\theta - 3\sin\theta \\ 3\sin\theta + 3\cos\theta \end{bmatrix}$$

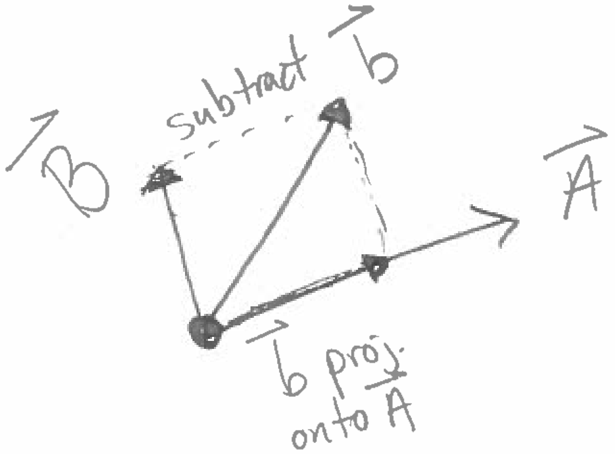
$$\|Q\vec{x}\| = \sqrt{9(\cos^2\theta - 2\sin\theta\cos\theta + \sin^2\theta) + 9(\cos^2\theta + 2\sin\theta\cos\theta + \sin^2\theta)} = 3\sqrt{2}$$

Gram-Schmidt : How to get Q.

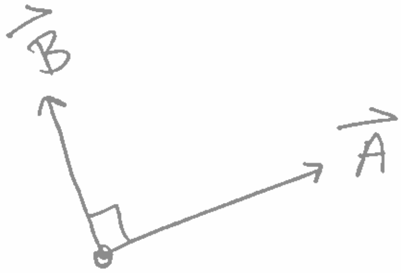
Example:

Start with three independent vectors $\vec{a}, \vec{b}, \vec{c}$ in $\mathbb{R}^3$.

* Vector $\vec{a}$ goes this way
* No change to $\vec{a}$. Call it $\vec{A}$.
* Force $\vec{b}$ to be perpendicular to $\vec{A}$.



Now we have:



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

(subtract off the projection.)

* Finally, take $\vec{c}$ and subtract off two projections

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

* Now $\vec{A} \perp \vec{B} \perp \vec{C}$ and we have orthogonality.
* Last step → make all vectors unit by dividing by each length

---

<u>Example:</u> Find an orthonormal basis for $\mathbb{R}^3$ with $\vec{a} = (4,0,0)$,
$\vec{b} = (2,2,0)$,
$\vec{c} = (3,8,10)$

$$\vec{A} = \begin{bmatrix} 4 \\ 0 \\ 0 \end{bmatrix}. \quad \text{Now}$$

$$\vec{B} = \begin{bmatrix} 2 \\ 2 \\ 0 \end{bmatrix} - \frac{\vec{A}^T \vec{b}}{\vec{A}^T \vec{A}} \vec{A} = \begin{bmatrix} 2 \\ 2 \\ 0 \end{bmatrix} - \frac{8}{16} \begin{bmatrix} 4 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 2 \\ 0 \end{bmatrix}.$$

$$\vec{C} = \begin{bmatrix} 3 \\ 8 \\ 10 \end{bmatrix} - \frac{\vec{A}^T \vec{C}}{\vec{A}^T \vec{A}} \vec{A} - \frac{\vec{B}^T \vec{C}}{\vec{B}^T \vec{B}} \vec{B}$$

$$= \begin{bmatrix} 3 \\ 8 \\ 10 \end{bmatrix} - \frac{12}{16} \begin{bmatrix} 4 \\ 0 \\ 0 \end{bmatrix} - \frac{16}{4} \begin{bmatrix} 0 \\ 2 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 10 \end{bmatrix}$$

Now make unit:

$$\vec{A}^* = \frac{\vec{A}}{\|\vec{A}\|} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad \vec{C}^* = \frac{\vec{C}}{\|\vec{C}\|} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix},$$

$$\vec{B}^* = \frac{\vec{B}}{\|\vec{B}\|} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

Example: Find an orthonormal basis for the columnspace of

$$D = \begin{bmatrix} 1 & -2 \\ 1 & 0 \\ 1 & 1 \\ 1 & 3 \end{bmatrix}$$

Solution: Use Gram-Schmidt.

By the way, $\|(1,1,1,1)\| = 2$, $\|(-2,0,1,3)\| = \sqrt{14}$

$$\cos\theta = \frac{-2+1+3}{2 \times \sqrt{14}} \approx -\frac{1}{\sqrt{14}}$$

$$\theta \approx 105.5^\circ \text{ in } \mathbb{R}^4.$$

$$\vec{a} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \vec{A}, \quad \vec{b} = \begin{bmatrix} -2 \\ 0 \\ 1 \\ 3 \end{bmatrix}$$

144

$$\vec{B} = \begin{bmatrix} -2 \\ 0 \\ 1 \\ 3 \end{bmatrix} - \frac{2}{4}\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -5/2 \\ -1/2 \\ 1/2 \\ 5/2 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} -5 \\ -1 \\ 1 \\ 5 \end{bmatrix}$$

Make unit: $\vec{A}^* = \begin{bmatrix} 1/2 \\ 1/2 \\ 1/2 \\ 1/2 \end{bmatrix}$

$$\vec{B}^* = \frac{\vec{B}}{\|\vec{B}\|} = \frac{1}{\sqrt{52}}\begin{bmatrix} -5 \\ -1 \\ 1 \\ 5 \end{bmatrix}$$

Continued: Project $\vec{d} = (-4, -3, 3, 0)$ onto the columnspace of D.

Old Way: $\vec{p} = D\hat{x} = D(D^TD)^{-1}D^T\vec{d}$

TI $\quad \vec{p} = \begin{bmatrix} -3.5 \\ -1.5 \\ -0.5 \\ 1.5 \end{bmatrix} = \begin{bmatrix} -7/2 \\ -3/2 \\ -1/2 \\ 3/2 \end{bmatrix}$

New Way: $\vec{p} = QQ^T\vec{d} =$

$$Q = \begin{bmatrix} 1/2 & -5/\sqrt{52} \\ 1/2 & -1/\sqrt{52} \\ 1/2 & 1/\sqrt{52} \\ 1/2 & 5/\sqrt{52} \end{bmatrix} \qquad \frac{1}{26}\begin{bmatrix} 19 & 9 & 4 & -6 \\ 9 & 7 & 6 & 4 \\ 4 & 6 & 7 & 9 \\ -6 & 4 & 9 & 19 \end{bmatrix}\begin{bmatrix} -4 \\ -3 \\ 3 \\ 0 \end{bmatrix} =$$

$A = QR$ (Gram-Schmidt)

* Since the columns of Q are linear combinations of the columns of A, there is a third matrix R connecting the two.

$$A = QR$$
$$Q^T A = Q^T Q R$$
$$Q^T A = R$$

* Any $m \times n$ matrix with independent columns can be factored as $A = QR$

* In least squares we get:

$$A^T A \hat{x} = A^T \vec{b}$$
$$(QR)^T (QR) \hat{x} = (QR)^T \vec{b}$$
$$R^T Q^T Q R \hat{x} = R^T Q^T \vec{b}$$
$$R^T R \hat{x} = R^T Q^T \vec{b}$$
$$R \hat{x} = Q^T \vec{b}$$
$$\hat{x} = R^{-1} Q^T \vec{b}$$

To find upper triangular R, entry i,j of R is row i of Q^T times column j of A.

<u>Example:</u> We had $D = \begin{bmatrix} 1 & -2 \\ 1 & 0 \\ 1 & 1 \\ 1 & 3 \end{bmatrix}$

and $Q = \begin{bmatrix} 1/2 & -5/\sqrt{52} \\ 1/2 & -1/\sqrt{52} \\ 1/2 & 1/\sqrt{52} \\ 1/2 & 5/\sqrt{52} \end{bmatrix}$

Write $D = QR$

$\sqrt{13}$

$$R = \underset{2\times 4}{Q^T} \underset{4\times 2}{D} = \begin{bmatrix} 2 & 1 \\ 0 & 26/\sqrt{52} \end{bmatrix}$$

$\sqrt{13}$

$$D = \underset{4\times 2}{\begin{bmatrix} 1 & -2 \\ 1 & 0 \\ 1 & 1 \\ 1 & 3 \end{bmatrix}} = \underset{Q}{\underset{4\times 2}{\begin{bmatrix} 1/2 & -5/\sqrt{52} \\ 1/2 & -1/\sqrt{52} \\ 1/2 & 1/\sqrt{52} \\ 1/2 & 5/\sqrt{52} \end{bmatrix}}} \underset{R}{\underset{2\times 2}{\begin{bmatrix} 2 & 1 \\ 0 & 26/\sqrt{52} \end{bmatrix}}}$$