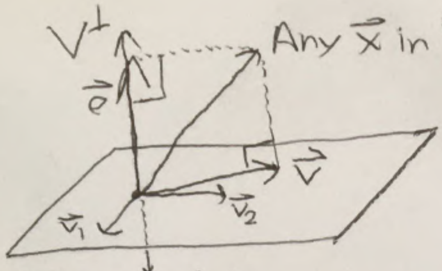
Last time: Orthogonal projection onto a subspace V of $\mathbb{R}^n$:



Any $\vec{x}$ in $\mathbb{R}^n$ $= \vec{v} + \vec{e}$

"Error" vector in $V^\perp$

"best approximation" vector in V

Subspace V,
Basis $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_m\}$

Projection matrix: $\vec{v} = P\vec{x}$

$P = P^T$ and $P^2 = P$

How to find P: Write $A = \begin{bmatrix} | & | & & | \\ \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_m \\ | & | & & | \end{bmatrix}$ ← $n \times m$ matrix, not usually invertible.

Then $\boxed{P = A(A^TA)^{-1}A^T}$

Example: V = plane in $\mathbb{R}^3$ with equation $x + 2y + 3z = 0$

① Find basis: $V = \text{all} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$ such that $x = -2y - 3z$

$$= \text{all} \begin{bmatrix} -2y-3z \\ y \\ z \end{bmatrix} = \text{all } y\begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} + z\begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix}$$

linearly independent spanning set, so basis for V

② $A = \begin{bmatrix} -2 & -3 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}$ Notice A is not invertible, but A^TA is!

③ $A^TA = \begin{bmatrix} -2 & 1 & 0 \\ -3 & 0 & 1 \end{bmatrix}\begin{bmatrix} -2 & -3 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 5 & 6 \\ 6 & 10 \end{bmatrix}$, $(A^TA)^{-1} = \dfrac{1}{\underbrace{50-36}_{\frac{1}{14}}}\begin{bmatrix} 10 & -6 \\ -6 & 5 \end{bmatrix}$

$$P = A(A^TA)^{-1}A^T = \begin{bmatrix} -2 & -3 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \frac{1}{14}\begin{bmatrix} 10 & -6 \\ -6 & 5 \end{bmatrix}\begin{bmatrix} -2 & 1 & 0 \\ -3 & 0 & 1 \end{bmatrix}$$

$$= \frac{1}{14}\begin{bmatrix} -2 & -3 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} -2 & 10 & -6 \\ -3 & -6 & 5 \end{bmatrix} = \frac{1}{14}\begin{bmatrix} 13 & -2 & -3 \\ -2 & 10 & -6 \\ -3 & -6 & 5 \end{bmatrix}$$

④ Let's project $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ onto V: $\vec{v} = P\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \frac{1}{14}\begin{bmatrix} 13 \\ -2 \\ -3 \end{bmatrix}$ Should look familiar?

Here, V is the orthogonal complement of the line from example ~~[illegible]~~ last time

So our new P is actually $I - (\text{old } P)$ for the line spanned by $\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$

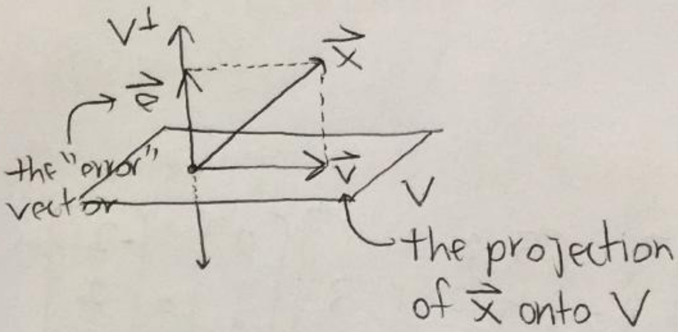
Question: What is the rank of our new P?

rank(P) = dim of C(P) ← all vectors like $P\vec{x}$ = all vectors in V

= dim V = 2

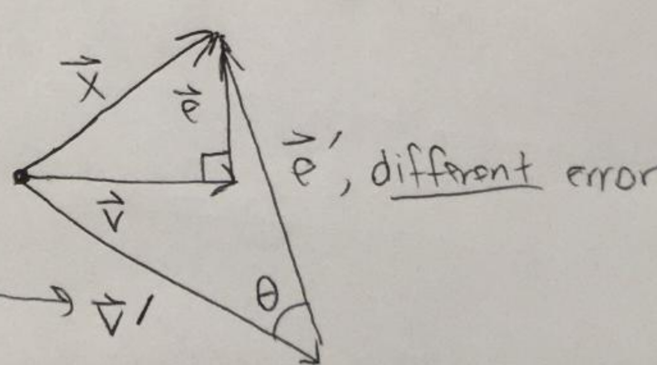
In general: rank of a projection matrix = dimension of the subspace you are projecting onto.

Section 4.3 Least Squares Approximation



Now we want to emphasize: $\vec{v}$ is the vector in V that is closest to $\vec{x}$. It gives the smallest error, $\|\vec{e}\|$.

Why does $\vec{v}$ minimize the error? Consider any different approximation.



In the picture, it looks like $\|\vec{e}'\| > \|\vec{e}\|$ (worse approximation)
but let's prove it:

First: $\vec{x} = \vec{v}' + \vec{e}'$, so $\|\vec{e}'\| = \|\vec{x} - \vec{v}'\| = \|\vec{e} + (\vec{v} - \vec{v}')\|$

$\vec{x} = \vec{v} + \vec{e}$ also

Next: $\vec{v} - \vec{v}'$ is still a vector in V, and $\vec{e} \perp V$, so $\vec{e} \perp (\vec{v} - \vec{v}')$.

Pythagorean Theorem: $\|\vec{e}'\|^2 = \|\vec{e} + (\vec{v} - \vec{v}')\|^2$

$$= \|\vec{e}\|^2 + \|\vec{v} - \vec{v}'\|^2$$

This is a squared real number, has to be ≥ 0.

So $\|\vec{e}'\|$ is <u>smallest</u> when $\|\vec{v} - \vec{v}'\|$ is as small as possible. This happens when $\|\vec{v} - \vec{v}'\| = 0$, or $\vec{v} = \vec{v}'$ (and then $\vec{e} = \vec{e}'$).

So <u>smallest</u> error $\|\vec{e}\| \longleftrightarrow \vec{e} \perp \vec{v}$ (<u>orthogonal</u> projection)

Now: What is orthogonal projection good for?

<u>Linear</u> <u>regression</u> (and more general curve fitting)

<u>Example</u> Can we find a line that goes through the data points $(-1,3)$, $(0,0)$, $(1,1)$, and $(2,-3)$?

<u>Intuition</u>: Can always find a line through <u>two</u> points, but for <u>four</u> points? Probably not.

<u>Solution</u>: Turn this into a linear algebra problem: We want to find a line $y = C + Dt$ that contains the points, so:

$$\begin{aligned} C+(-1)D&=3\\ C+(0)D&=0\\ C+(1)D&=1\\ C+(2)D&=-3 \end{aligned} \longrightarrow \underbrace{\begin{bmatrix}1&-1\\1&0\\1&1\\1&2\end{bmatrix}}_{A}\begin{bmatrix}C\\D\end{bmatrix}=\begin{bmatrix}3\\0\\1\\-3\end{bmatrix}$$

What we are really asking is: Is $\vec{b}=\begin{bmatrix}3\\0\\1\\-3\end{bmatrix}$ in the column space $C(A)$?

Check:

$$\left[\begin{array}{cc|c}1&-1&3\\1&0&0\\1&1&1\\1&2&-3\end{array}\right] \xrightarrow[\text{Row 4}-\text{Row 1}]{\substack{\text{Row 2}-\text{Row 1}\\ \text{Row 3}-\text{Row 1}}} \left[\begin{array}{cc|c}1&-1&3\\0&1&-3\\0&2&-2\\0&3&-6\end{array}\right] \xrightarrow[\text{Row 4}-3\text{Row 2}]{\substack{\text{Row 1}+\text{Row 2}\\ \text{Row 3}-2\text{Row 2}}} \underbrace{\left[\begin{array}{cc|c}1&0&0\\0&1&-3\\0&0&4\\0&0&3\end{array}\right]}_{\text{Inconsistent equations, no solution}}$$

So $\vec{b}=\begin{bmatrix}3\\0\\1\\-3\end{bmatrix}$ is not in $C(A)$

$\rightarrow$ no line goes through the four points

What we can do instead: Find a vector $\vec{p}=C\begin{bmatrix}1\\1\\1\\1\end{bmatrix}+D\begin{bmatrix}-1\\0\\1\\2\end{bmatrix}$

In $C(A)$ that best approximates $\vec{b}$. The values of C and D for $\vec{p}$ will give us the line $y=C+Dt$ that best fits the four data points.

Of course, $\vec{p}$ should be the orthogonal projection of $\vec{b}$ onto $C(A)$. Let's review how to find $\vec{p}$:

① $\vec{p}$ in $C(A)$ means $\vec{p}=C\begin{bmatrix}1\\1\\1\\1\end{bmatrix}+D\begin{bmatrix}-1\\0\\1\\2\end{bmatrix}=A\hat{x}$, $\hat{x}=\begin{bmatrix}C\\D\end{bmatrix}$

② $\vec{e}\perp C(A)$ means $\begin{bmatrix}1\\1\\1\\1\end{bmatrix}\cdot(\vec{b}-\vec{p})=0$

error, $\vec{b}-\vec{p}$

$\begin{bmatrix}-1\\0\\1\\2\end{bmatrix}\cdot(\vec{b}-\vec{p})=0$

So $\underbrace{\begin{bmatrix} 1 & 1 & 1 & 1 \\ -1 & 0 & 1 & 2 \end{bmatrix}}_{A^T}(\vec{b}-A\hat{x}) = \vec{0} \longrightarrow \boxed{(A^TA)\hat{x} = A^T\vec{b}}$

called the "normal equations" for $\hat{x} = \begin{bmatrix} C \\ D \end{bmatrix}$ (solution is the parameters C and D of the best fit line)

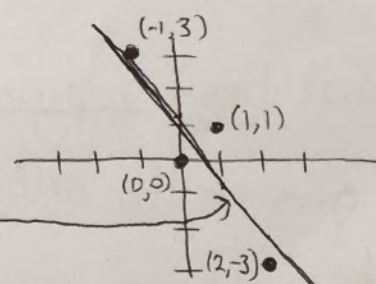
Specifically: $\underbrace{\begin{bmatrix} 1 & 1 & 1 & 1 \\ -1 & 0 & 1 & 2 \end{bmatrix}\begin{bmatrix} 1 & -1 \\ 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix}}\hat{x} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ -1 & 0 & 1 & 2 \end{bmatrix}\begin{bmatrix} 3 \\ 0 \\ +1 \\ -3 \end{bmatrix}$

$$\begin{bmatrix} 4 & 2 \\ 2 & 6 \end{bmatrix}\begin{bmatrix} C \\ D \end{bmatrix} = \begin{bmatrix} 1 \\ -8 \end{bmatrix}$$ ← "normal equations"

Solve: $\left[\begin{array}{cc|c} 4 & 2 & 1 \\ 2 & 6 & -8 \end{array}\right] \longrightarrow \left[\begin{array}{cc|c} 1 & 3 & -4 \\ 4 & 2 & 1 \end{array}\right] \xrightarrow[-4\text{Row1}]{\text{Row 2}} \left[\begin{array}{cc|c} 1 & 3 & -4 \\ 0 & -10 & 17 \end{array}\right] \longrightarrow$

$\left[\begin{array}{cc|c} 1 & 3 & -4 \\ 0 & 1 & -17/10 \end{array}\right] \xrightarrow[-3\text{Row 2}]{\text{Row 1}} \left[\begin{array}{cc|c} 1 & 0 & 11/10 \\ 0 & 1 & -17/10 \end{array}\right] \longrightarrow \begin{array}{l} C = 11/10 \\ D = -17/10 \end{array}$

So best fit line is $y = \frac{11}{10} - \frac{17}{10}t$



What is the "error" we have minimized here?

It is $\|\vec{e}\|$, which is $\|\vec{p}-\vec{b}\|$. Or, we are minimizing

$\|\vec{e}\|^2 = \|\vec{p}-\vec{b}\|^2 = \|A\hat{x}-\vec{b}\|^2$ "least squares"

$$\|A\hat{x}-\vec{b}\|^2 = \left\|\begin{bmatrix} 1 & -1 \\ 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} C \\ D \end{bmatrix} - \begin{bmatrix} 3 \\ 0 \\ 1 \\ -3 \end{bmatrix}\right\|^2 = \left\|\begin{bmatrix} C-D-3 \\ C \\ C+D-1 \\ C+2D+3 \end{bmatrix}\right\|^2$$

$= (C-D-3)^2 + C^2 + (C+D-1)^2 + (C+2D+3)^2$

squared difference between the line's y-values and the actual data points' y-values. We could also minimize this sum-of-squared-errors function using calculus:

Find C, D, such that $\frac{\partial}{\partial C}(\text{error function}) = 0$

$\frac{\partial}{\partial D}(\text{error function}) = 0$

Least-square lines in general: Find the best-fit line for the data points $(t_1, b_1), (t_2, b_2), \ldots, (t_n, b_n)$.

Assume the t's are all different (so think of the b's as measurements that you make at different times.

We want a line $y = C + Dt$. At first we try to solve:

$$\begin{aligned} C + t_1 D &= b_1 \\ C + t_2 D &= b_2 \\ &\vdots \\ C + t_n D &= b_n \end{aligned} \longrightarrow \underbrace{\begin{bmatrix} 1 & t_1 \\ 1 & t_2 \\ \vdots & \vdots \\ 1 & t_n \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} C \\ D \end{bmatrix}}_{\hat{x}} = \underbrace{\begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}}_{\vec{b}}$$

Usually this is impossible, no solution, so instead project $\vec{b}$ onto $C(A)$ using the normal equations:

$$A\hat{x} = \vec{b} \xrightarrow[\text{by } A^T]{\text{Multiply}} (A^T A)\hat{x} = A^T \vec{b}$$

No solution

This is just a 2×2 matrix, so system isn't difficult to solve.

$$\begin{bmatrix} 1 & 1 & \cdots & 1 \\ t_1 & t_2 & \cdots & t_n \end{bmatrix} \begin{bmatrix} 1 & t_1 \\ 1 & t_2 \\ \vdots & \vdots \\ 1 & t_n \end{bmatrix} = \begin{bmatrix} n & t_1 + t_2 + \cdots + t_n \\ t_1 + t_2 + \cdots + t_n & t_1^2 + t_2^2 + \cdots + t_n^2 \end{bmatrix}$$

Also: $A^T\vec{b} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ t_1 & t_2 & \cdots & t_n \end{bmatrix}\begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} = \begin{bmatrix} b_1+b_2+\cdots+b_n \\ t_1b_1+t_2b_2+\cdots+t_nb_n \end{bmatrix}$

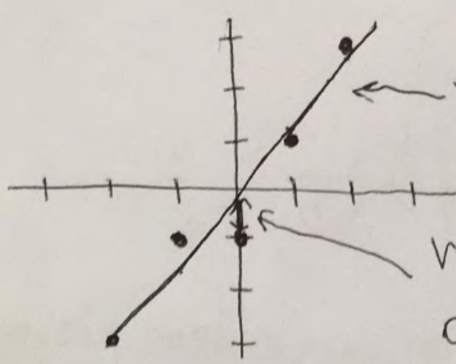
<u>Example</u>: Fit the data points $(-2,-3), (-1,-1), (0,-1), (1,1), (2,3)$.

Here, $A^TA = \begin{bmatrix} 5 & -2-1+0+1+2 \\ -2-1+0+1+2 & 4+1+0+1+4 \end{bmatrix} = \begin{bmatrix} 5 & 0 \\ 0 & 10 \end{bmatrix}$

Nice! This matrix is diagonal because the columns of A are $\perp$.

$A^T\vec{b} = \begin{bmatrix} -3-1-1+1+3 \\ 6+1+0+1+6 \end{bmatrix} = \begin{bmatrix} -1 \\ 14 \end{bmatrix}$

We need to solve the normal equations: $\begin{bmatrix} 5 & 0 \\ 0 & 10 \end{bmatrix}\begin{bmatrix} C \\ D \end{bmatrix} = \begin{bmatrix} -1 \\ 14 \end{bmatrix} \longrightarrow \begin{matrix} C = -1/5 \\ D = 7/5 \end{matrix}$



$\leftarrow y = -\frac{1}{5} + \frac{7}{5}t$

We are minimizing the sum of the squares of these errors.

Error: $\|\vec{e}\| = \|A\hat{x} - \vec{b}\| = \left\| \begin{bmatrix} 1 & -2 \\ 1 & -1 \\ 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} -1/5 \\ 7/5 \end{bmatrix} - \begin{bmatrix} -3 \\ -1 \\ -1 \\ 1 \\ 3 \end{bmatrix} \right\| = \left\| \begin{bmatrix} -3-(-3) \\ -8/5+1 \\ -1/5+1 \\ 6/5-1 \\ 13/5-3 \end{bmatrix} \right\|$

$= \left\| \begin{bmatrix} 0 \\ -3/5 \\ 4/5 \\ 1/5 \\ -2/5 \end{bmatrix} \right\| = \sqrt{0 + \frac{9}{25} + \frac{16}{25} + \frac{1}{25} + \frac{4}{25}} = \sqrt{\frac{30}{25}} = \sqrt{\frac{6}{5}}$ (Not too big)

Why do we require t's to be different?
This guarantees that columns of A aren't multiples of each other (so they are independent).
Then A^TA is <u>invertible</u>, and there is <u>only one</u> solution for $\hat{x}$.

We can also find best fits using other types of functions:
Try to fit our original data set to a parabola. Do $(-1,3)$, $(0,0)$, $(1,1)$, $(2,-3)$ fit on $y = C + Dx + Ex^2$?

Try to solve:

$$\begin{aligned} C - D + E &= 3 \\ C + 0D + 0E &= 0 \\ C + D + E &= 1 \\ C + 2D + 4E &= -3 \end{aligned} \rightsquigarrow \underset{A}{\begin{bmatrix} 1 & -1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \end{bmatrix}} \underset{\hat{x}}{\begin{bmatrix} C \\ D \\ E \end{bmatrix}} = \underset{\vec{b}}{\begin{bmatrix} 3 \\ 0 \\ 1 \\ -3 \end{bmatrix}}$$

Probably no solution still, so try to project onto $C(A)$ instead:
Same normal equations $A^TA\hat{x} = A^T\vec{b}$

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ -1 & 0 & 1 & 2 \\ 1 & 0 & 1 & 4 \end{bmatrix} \begin{bmatrix} 1 & -1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \end{bmatrix} = \begin{bmatrix} 4 & 2 & 6 \\ 2 & 6 & 8 \\ 6 & 8 & 18 \end{bmatrix} \qquad \begin{bmatrix} 1 & 1 & 1 & 1 \\ -1 & 0 & 1 & 2 \\ 1 & 0 & 1 & 4 \end{bmatrix} \begin{bmatrix} 3 \\ 0 \\ 1 \\ -3 \end{bmatrix} = \begin{bmatrix} 1 \\ -8 \\ -8 \end{bmatrix}$$

Solve:

$$\left[\begin{array}{ccc|c} 4 & 2 & 6 & 1 \\ 2 & 6 & 8 & -8 \\ 6 & 8 & 18 & -8 \end{array}\right] \xrightarrow[\text{Row 3} - \frac{3}{2}\text{Row 1}]{\text{Row 2} - \frac{1}{2}\text{Row 1}} \left[\begin{array}{ccc|c} 4 & 2 & 6 & 1 \\ 0 & 5 & 5 & -17/2 \\ 0 & 5 & 9 & -19/2 \end{array}\right] \longrightarrow \left[\begin{array}{ccc|c} 4 & 2 & 6 & 1 \\ 0 & 5 & 5 & -17/2 \\ 0 & 0 & 4 & -1 \end{array}\right]$$

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & 27/20 \\ 0 & 1 & 0 & -29/20 \\ 0 & 0 & 1 & -1/4 \end{array}\right] \longleftarrow \left[\begin{array}{ccc|c} 1 & 1/2 & 0 & 5/8 \\ 0 & 1 & 0 & -29/20 \\ 0 & 0 & 1 & -1/4 \end{array}\right] \longleftarrow \left[\begin{array}{ccc|c} 1 & 1/2 & 3/2 & 1/4 \\ 0 & 1 & 1 & -17/10 \\ 0 & 0 & 1 & -1/4 \end{array}\right]$$

Best fit parabola: $y = \frac{27}{20} - \frac{29}{20}t - \frac{1}{4}t^2$

Compare with line: $y = \frac{11}{10} - \frac{17}{10}t$

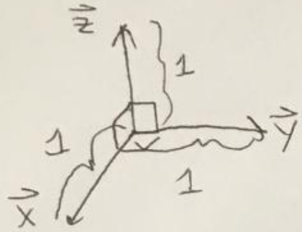
Parabola should have less error since $C(A)$ bigger $\rightarrow$ easier to get a good approximation.

But usually, you would only fit with a parabola if you expect y to depend quadratically on t. (Such as: y = distance traveled in gravitational free fall near the Earth's surface.)

Section 4.4 Orthonormal Bases and Gram-Schmidt

Some of the best bases have orthogonal unit vectors.

Example Standard xyz basis of $\mathbb{R}^3$ is easy to work with.



$$\vec{x} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \vec{y} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \vec{z} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

In general, we say that vectors $\vec{q}_1, \vec{q}_2, \dots, \vec{q}_n$ (in $\mathbb{R}^m$) are orthonormal if they are unit vectors and are all orthogonal to each other:

$$\vec{q}_i^{\,T} \vec{q}_j = \begin{cases} 0 \text{ when } i \neq j \leftarrow \text{orthogonal when they are different} \\ 1 \text{ when } i = j \leftarrow \text{each squared length is } 1 \end{cases}$$

What happens if we put orthonormal vectors into a matrix?

$$Q = \begin{bmatrix} \vec{q}_1 & \vec{q}_2 & \cdots & \vec{q}_n \end{bmatrix}$$ It's $m \times n$ (doesn't have to be square).

notation for a matrix with orthonormal columns.

Look at: $$Q^T Q = \begin{bmatrix} - & \vec{q}_1 & - \\ - & \vec{q}_2 & - \\ & \vdots & \\ - & \vec{q}_n & - \end{bmatrix} \begin{bmatrix} | & | & & | \\ \vec{q}_1 & \vec{q}_2 & \cdots & \vec{q}_n \\ | & | & & | \end{bmatrix} = \begin{array}{r} \\ i\text{th} \rightarrow \\ \\ \end{array} \begin{bmatrix} & | & \\ - & \vec{q}_i^{\,T} \vec{q}_j & - \\ & | & \end{bmatrix}$$
$\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \uparrow$ jth

The (i,j)-entry is 1 if $i=j$ and 0 otherwise.

$\longrightarrow Q^T Q = n \times n$ identity matrix I. Very nice!

If Q is square ($m=n$), we call Q and orthogonal matrix

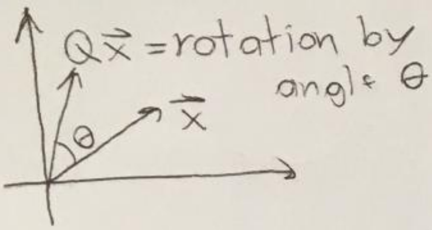
For an orthogonal matrix Q, $Q^{-1} = Q^T$ since $Q^T Q = I$.

So orthogonal matrices are invertible.

Examples 2×2 orthogonal matrices.

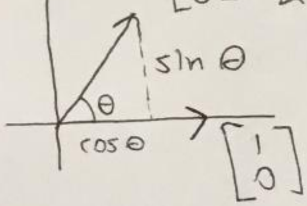
There are two different types:

(1) Rotation matrices

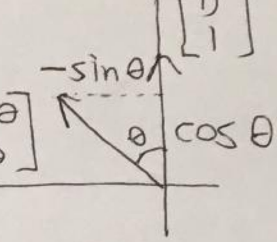


We can find the entries of $Q=\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ in terms of θ by finding what Q does to $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$.

$Q\begin{bmatrix} 1 \\ 0 \end{bmatrix}=\begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix}$



$Q\begin{bmatrix} 0 \\ 1 \end{bmatrix}=\begin{bmatrix} -\sin\theta \\ \cos\theta \end{bmatrix}$



So: $Q\begin{bmatrix} 1 \\ 0 \end{bmatrix}=\begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix}$ $\qquad Q\begin{bmatrix} 0 \\ 1 \end{bmatrix}=\begin{bmatrix} -\sin\theta \\ \cos\theta \end{bmatrix}$

$\quad \|$ $\qquad\qquad\qquad\qquad \|$

$\begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix}=\begin{bmatrix} a \\ c \end{bmatrix}$ $\qquad \begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} 0 \\ 1 \end{bmatrix}=\begin{bmatrix} b \\ d \end{bmatrix}$

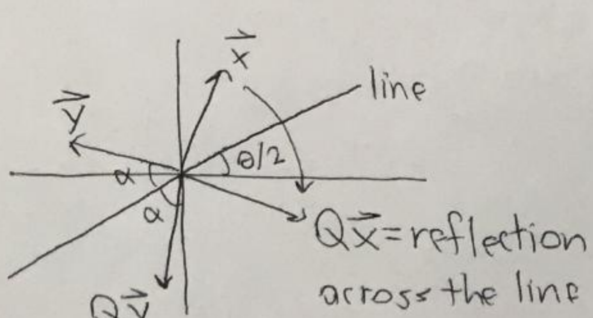
$\rightarrow$ $\boxed{Q=\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}}$ rotates $\vec{x}$ by θ <u>counterclockwise</u>

Is Q really orthogonal?

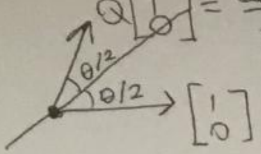
$\begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix}\cdot\begin{bmatrix} -\sin\theta \\ \cos\theta \end{bmatrix}=0$ ✓ $\qquad \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix}\cdot\begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix}=\cos^2\theta+\sin^2\theta=\underline{1}$ ✓

$\begin{bmatrix} -\sin\theta \\ \cos\theta \end{bmatrix}\cdot\begin{bmatrix} -\sin\theta \\ \cos\theta \end{bmatrix}=(-\sin\theta)^2+\cos^2\theta=\underline{1}$ ✓ $\quad$ Yes! Q is orthogonal

(2) Reflection matrices

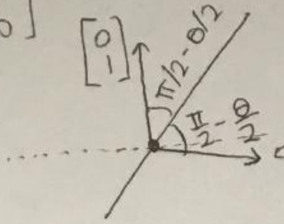


What does Q do to $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$?



$Q\begin{bmatrix}1\\0\end{bmatrix} = \cancel{\text{clockwise}}$ counterclockwise rotation by θ $= \begin{bmatrix}\cos\theta\\ \sin\theta\end{bmatrix}$ as before

$\begin{bmatrix}1\\0\end{bmatrix}$



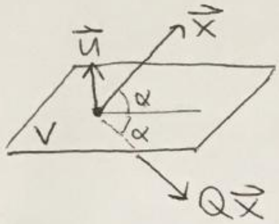
$Q\begin{bmatrix}0\\1\end{bmatrix} =$ clockwise rotation by $\pi-\theta$

= counterclockwise rotation by $\theta-\pi$

$= \begin{bmatrix}-\sin(\theta-\pi)\\ \cos(\theta-\pi)\end{bmatrix} = \begin{bmatrix}+\sin\theta\\ -\cos\theta\end{bmatrix}$

So $Q = \begin{bmatrix}\cos\theta & \sin\theta\\ \sin\theta & -\cos\theta\end{bmatrix}$ reflects $\vec{x}$ across the $\theta/2$ line.

Reflection in higher dimensions ($\mathbb{R}^n$):



V = "hyperplane" in $\mathbb{R}^n$ ($(n-1)$-dimensional subspace)

$\vec{u}$ = unit vector $\perp$ to V (basis for $V^\perp$)

Q = matrix that reflects $\vec{x}$ across the hyperplane.

Claim: $Q = I - 2\vec{u}\vec{u}^T$

Why does this work? Check that Q does what we expect to $\vec{u}$ and to vectors in V

① It sends $\vec{u}$ to $-\vec{u}$:

$$Q\vec{u} = (I-2\vec{u}\vec{u}^T)\vec{u} = I\vec{u} - 2\vec{u}\underbrace{(\vec{u}^T\vec{u})}_{1} = \vec{u}-2\vec{u} = -\vec{u} \checkmark$$

② It doesn't change vectors in V:

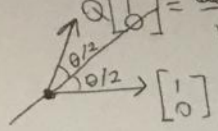
$$Q\vec{v} = (I-2\vec{u}\vec{u}^T)\vec{v} = I\vec{v} - 2\vec{u}\underbrace{(\vec{u}^T\vec{v})}_{0 \text{ because } \vec{u} \text{ is } \perp \text{ to } V} = \vec{v} \checkmark$$

Is this Q really an orthogonal matrix?

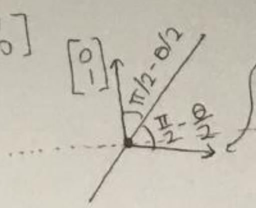
Yes! $Q^TQ = (I-2\vec{u}\vec{u}^T)^T(I-2\vec{u}\vec{u}^T) = (I-2\vec{u}\vec{u}^T)(I-2\vec{u}\vec{u}^T)$

$$= I - 2\vec{u}\vec{u}^T - 2\vec{u}\vec{u}^T + 4\vec{u}\underbrace{\vec{u}^T\vec{u}}_{1 \text{ because } \|\vec{u}\|=1}\vec{u}^T$$

$$= I \checkmark$$



$Q\begin{bmatrix}1\\0\end{bmatrix} =$ ~~clockwise~~ counterclockwise rotation by θ $= \begin{bmatrix}\cos\theta\\ \sin\theta\end{bmatrix}$ as before



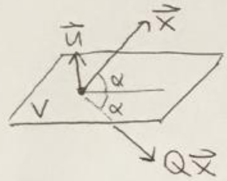
$Q\begin{bmatrix}0\\1\end{bmatrix} =$ clockwise rotation by $\pi - \theta$

$=$ counterclockwise rotation by $\theta - \pi$

$$= \begin{bmatrix}-\sin(\theta-\pi)\\ \cos(\theta-\pi)\end{bmatrix} = \begin{bmatrix}+\sin\theta\\ -\cos\theta\end{bmatrix}$$

So $Q = \begin{bmatrix}\cos\theta & \sin\theta\\ \sin\theta & -\cos\theta\end{bmatrix}$ reflects $\vec{x}$ across the $\theta/2$ line.

Reflection in higher dimensions ($\mathbb{R}^n$):



V = "hyperplane" in $\mathbb{R}^n$ ($(n-1)$-dimensional subspace)

$\vec{u}$ = unit vector $\perp$ to V (basis for $V^\perp$)

Q = matrix that reflects $\vec{x}$ across the hyperplane.

Claim: $Q = I - 2\vec{u}\vec{u}^T$

Why does this work? Check that Q does what we expect to $\vec{u}$ and to vectors in V

① It sends $\vec{u}$ to $-\vec{u}$:

$$Q\vec{u} = (I - 2\vec{u}\vec{u}^T)\vec{u}$$
$$= I\vec{u} - 2\vec{u}\underbrace{(\vec{u}^T\vec{u})}_{1}$$
$$= \vec{u} - 2\vec{u} = -\vec{u} \checkmark$$

② It doesn't change vectors in V:

$$Q\vec{v} = (I - 2\vec{u}\vec{u}^T)\vec{v}$$
$$= I\vec{v} - 2\vec{u}\underbrace{(\vec{u}^T\vec{v})}_{\text{0 because } \vec{u} \text{ is } \perp \text{ to } V} = \vec{v} \checkmark$$

Is this Q really an orthogonal matrix?

Yes! $Q^TQ = (I - 2\vec{u}\vec{u}^T)^T(I - 2\vec{u}\vec{u}^T) = (I - 2\vec{u}\vec{u}^T)(I - 2\vec{u}\vec{u}^T)$

$$= I - 2\vec{u}\vec{u}^T - 2\vec{u}\vec{u}^T + 4\vec{u}\underbrace{\vec{u}^T\vec{u}}_{1 \text{ because } \|\vec{u}\| = 1}\vec{u}^T$$
$$= I \checkmark$$

Example Find Q for reflection in the plane V with equation $x + 2y + 3z = 0$ in $\mathbb{R}^3$.

Perpendicular vector: $\begin{bmatrix}1\\2\\3\end{bmatrix} \rightsquigarrow \vec{u} = \frac{1}{\sqrt{1+4+9}}\begin{bmatrix}1\\2\\3\end{bmatrix} = \frac{1}{\sqrt{14}}\begin{bmatrix}1\\2\\3\end{bmatrix}$

So $Q = I - 2\vec{u}\vec{u}^T = I - \frac{2}{14}\begin{bmatrix}1\\2\\3\end{bmatrix}\begin{bmatrix}1 & 2 & 3\end{bmatrix}$

$$= \begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix} - \frac{1}{7}\begin{bmatrix}1&2&3\\2&4&6\\3&6&9\end{bmatrix} = \frac{1}{7}\begin{bmatrix}6&-2&-3\\-2&3&-6\\-3&-6&-2\end{bmatrix}$$