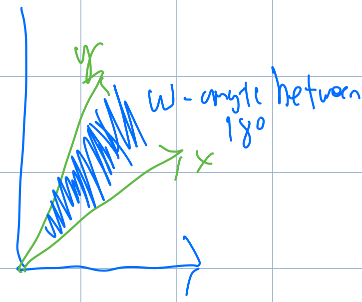
The angle between two vectors tells us how similar their orientations are.



$x = [1,1]^T \in \mathbb{R}^2$ $y = [1,2]^T \in \mathbb{R}^2$ use dot product:

$$\cos w = \frac{\langle x, y\rangle}{\sqrt{\langle x, x\rangle \langle y, y\rangle}} = \frac{x^T y}{\sqrt{x^T x y^T y}} = \frac{3}{\sqrt{10}}$$

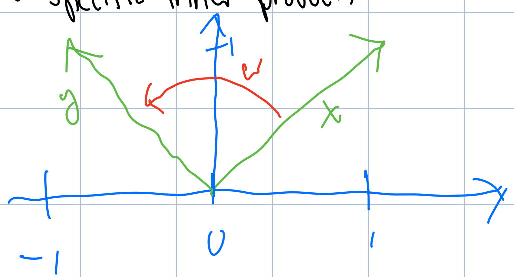
angle between: $\arccos\left(\frac{3}{\sqrt{10}}\right) \approx .32 \text{ rad} = 18 \text{ deg}$

inner product allows us to characterize vectors that are orthogonal.

def (Orthogonality): Two vectors are orthogonal iff their dot product $= 0$, and we write $x \perp y$. If additionally $\|x\| = 1 = \|y\|$ (the vectors are unit vectors), then x and y are orthonormal

$\Rightarrow$ the 0-vector is orthogonal to every vector in the vector space.

* Orthogonality is the generalization of the concept of perpendicularity to bilinear forms that do not have to be the dot product. Geometrically, we can think of orthogonal vectors as having a right angle with respect to a specific inner product.



$x = [1,1]^T$ $y = [-1,1] \in \mathbb{R}^2$

$\langle x, y\rangle = x^T \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} y$

w is given by: $\cos w = \frac{\langle x, y\rangle}{\|x\|\|y\|} = -\frac{1}{3}$

$w \approx 1.91 \text{ rad} \approx 109.5°$, x and y are not orthogonal.

Vectors that are orthogonal with respect to one inner product do not have to be orthogonal with respect to a different inner product.

Def (Orthogonal Matrix): A square matrix $A \in \mathbb{R}^{n \times n}$ is an orthogonal matrix iff its columns are orthonormal so that

$AA^T = I = A^T A$, implying $A^{-1} = A^T$, $\Rightarrow$ the inverse is obtained through transposition

The length of a vector x is not changed when transforming it using an orthogonal matrix A:

$$\|Ax\|^2 = (Ax)^T (Ax) = x^T A^T A x = x^T I x = x^T x = \|x\|^2$$

The angle between two vectors x, y, as measured by their inner product, is also unchanged.

$$\cos(\omega) = \frac{(Ax)^T(Ay)}{\|Ax\|\|Ay\|} = \frac{x^T A^T A y}{\sqrt{x^T A^T A x \, y^T A^T A y}} = \frac{x^T y}{\|x\|\|y\|}$$ $\Rightarrow$ gives the angle between

Orthogonal matrices A with $A^T = A^{-1}$ preserve both angles and distances.

Def (Orthonormal basis): Consider an n-dimensional vector space V and a basis $\{b_1, \dots, b_n\}$ of V.

$$\text{If } \overset{(a)}{\langle b_i, b_j \rangle = 0} \text{ for } i \neq j \quad \overset{(b)}{\langle b_i, b_i \rangle = 1}$$

for all $i, j = 1, \dots, n$ then the basis is called an orthonormal basis (ONB). If only (a) is satisfied, the basis is called an orthogonal basis.

The standard basis for a Euclidean vector space $\mathbb{R}^n$ is an orthonormal basis

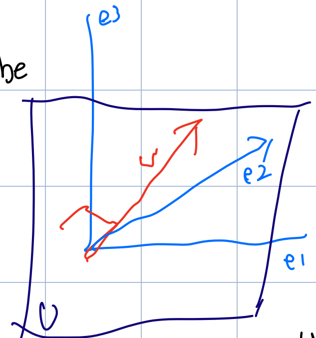
$b_1 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix}$, $b_2 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -1 \end{bmatrix}$ form an orthonormal basis since $b_1^T b_2 = 0$ and $\|b_1\| = 1 = \|b_2\|$

Consider a D-dimensional vector space V and an M-dimensional sub-space $U \subseteq V$. Then its orthogonal complement $U^\perp$ is a $(D-M)$-dimensional subspace of V and contains all vectors in V that are orthogonal to every vector in U.

$U \cap U^\perp = \{0\}$ so any vector $x \in V$ can be

$$x = \sum_{m=1}^{M} \lambda_m b_m + \sum_{j=1}^{D-M} \psi_j b_j^\perp, \quad \lambda_m, \psi_j \in \mathbb{R}$$

$(b_1 \dots b_m)$ is a basis of U and $(b_1^\perp, \dots, b_{D-M}^\perp)$ is a basis of $U^\perp$

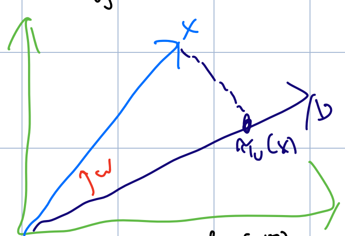


vector w is the normal vector of U.

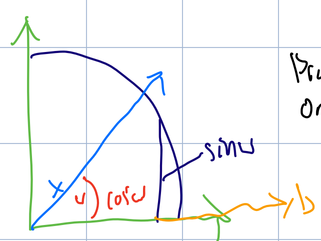
The orthogonal complement can be used to describe a 2D subspace in a 3D vector space. They can also be used to describe hyperplanes in n-dimensional vector and affine spaces.

Def (Projection) Let V be a vector space and $U \subseteq V$ a subspace of V. A linear mapping $\pi: V \to U$ is called a projection $\pi^2 = \pi \circ \pi = \pi$

Projection matrices P_π exhibit the property $P_\pi^2 = P_\pi$



The projection of $x \in \mathbb{R}^2$ onto a subspace U with basis vector b.



Projection of a two-dimensional vector x with $\|x\| = 1$ onto a one-dimensional subspace spanned by b.

- The projection $\pi_U(x)$ is closest to "x" where "closest" implies that the distance $\|x - \pi_U(x)\|$ is minimal. It follows that the segment $\pi_U(x) - x$ from $\pi_U(x)$ to x is orthogonal to U, thus the basis vector b of U. Orthogonality condition yields $\langle \pi_U(x) - x, b \rangle = 0$

- The projection $\pi_U(x)$ of x onto U must be an element of U, therefore, a multiple of the basis vector b that spans U. $\Rightarrow \pi_U(x) = \lambda b$ for some $\lambda \in \mathbb{R}$

In 3 steps, we determine the coordinate λ, the projection $\pi_U(x) \in U$, and the projection matrix P_π that maps any $x \in \mathbb{R}^n$ onto U:

1) Finding the coordinate λ: $\langle x - \pi_U(x), b\rangle = 0 \overset{\pi_U(x) = \lambda b}{\Leftrightarrow} \langle x - \lambda b, b\rangle = 0$

We exploit the bilinearity of the inner product:

$$\langle x, b\rangle - \lambda \langle b, b\rangle = 0 \Leftrightarrow \lambda = \frac{\langle x, b\rangle}{\langle b, b\rangle} = \frac{\langle b, x\rangle}{\|b\|^2}$$

$$\lambda = \frac{b^T x}{b^T b} = \frac{b^T x}{\|b\|^2}$$ if $\|b\| = 1$, then the coordinate λ of the projection is given by $b^T x$.

2) Finding the projection point $\pi_U(x) \in U$. Since $\pi_U(x) = \lambda b$:

$$\pi_U(x) = \lambda b = \frac{\langle x, b\rangle}{\|b\|^2} b = \frac{b^T x}{\|b\|^2} b,$$

We can compute the length of $\pi_U(x)$:

$$\|\pi_U(x)\| = \|\lambda b\| = |\lambda| \|b\|$$

Our projection is of length $|\lambda|$ times the length of $\|b\|$. $\Rightarrow \lambda$ is the coordinate of $\pi_U(x)$ with respect to basis vector b that spans 1-D subspace U.

Using dot product as the inner product

$$\|\pi_U(x)\| = \frac{|b^T x|}{\|b\|^2} \|b\| = |\cos\omega| \, \|x\| \, \|b\| \frac{\|b\|}{\|b\|^2} = |\cos\omega| \, \|x\|$$

ω is the angle between x and b.

3) Finding the projection matrix P_π. $\exists P_\pi$ st $\pi_U(x) = P_\pi(x)$ with the dot product as inner product and

$$\pi_U(x) = \lambda b = b\lambda = b\frac{b^T x}{\|b\|^2} = \frac{bb^T}{\|b\|^2} x, \quad P_\pi = \frac{bb^T}{\|b\|^2}$$

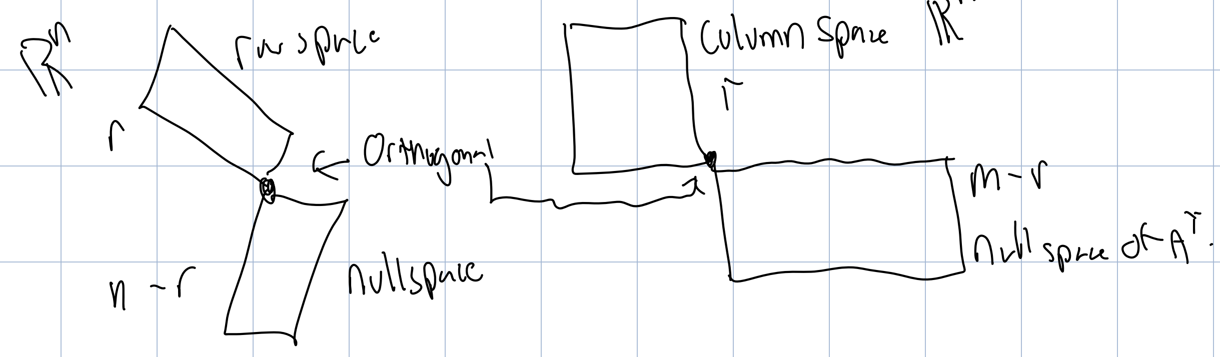
bb^T and P_π are symmetric of rank 1 and $\|b\|^2 = \langle b, b\rangle$ is scalar

The projection matrix P_π projects any vector $x \in \mathbb{R}^n$ onto the line through the origin with direction b.

$b = [1\ 2\ 2]$ $\quad P_\pi = \frac{bb^T}{b^Tb} = \frac{1}{9}\begin{bmatrix}1\\2\\2\end{bmatrix} = \frac{1}{9}\begin{bmatrix}1&2&2\\2&4&4\\2&4&4\end{bmatrix}\begin{bmatrix}1\\1\\1\end{bmatrix} = \frac{1}{9}\begin{bmatrix}5\\10\\10\end{bmatrix} \in \text{span}\begin{bmatrix}1\\2\\2\end{bmatrix}$

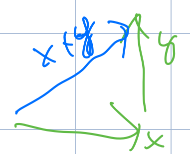
$\pi_u(x)$ is an eigenvector of P_π, and the corresponding eigenvalue is 1.

Null space $\perp$ row space



Orthogonal Vectors (perpendicular) - they form a right triangle

$\underbrace{x^Ty \text{ (row} \times \text{column)} = 0}_{\text{perpendicular test}}$



x^Tx

$x\begin{bmatrix}1\\2\\3\end{bmatrix},\ y\begin{bmatrix}2\\-1\\0\end{bmatrix},\ x+y = \begin{bmatrix}3\\1\\3\end{bmatrix}$

$x^2 = 14 \quad y^2 = 5 \quad \|x+y\|^2 = 19$

$\|x\|^2 + \|y\|^2 = \|x+y\|^2$ - only true when $x^Ty = 0$ (right triangle)

Pythagoras

$x^T\cancel{x} + y^Ty = (x+y)^T(x+y)$

↳ only true with right angle

dot product of orthogonal vectors = 0.

Subspace S is orthogonal to subspace T means: every vector in S is perpendicular to every vector in T.

row space is orthogonal to the null space $\quad$ why?

(x) dot product = 0 all zero.

$Ax = 0 \quad \begin{bmatrix}\text{row 1 of } A\\ \text{row 2 of } A\\ \vdots \\ \text{row } m \text{ of } A\end{bmatrix}\begin{bmatrix}x_1\\ \vdots\\ x_n\end{bmatrix} = \begin{bmatrix}0\\ \vdots\\ 0\end{bmatrix}$

$c_1(\text{row}_1)^Tx = 0 \quad c_2(\text{row } 2)^Tx = 0$

$(c_1\text{row}_1 + c_2\text{row}_2 + \dots)^Tx = 0$

Null space and row space are Orthogonal Complements in $\mathbb{R}^n$

$A = \begin{bmatrix}1&2&5\\2&4&10\end{bmatrix} \quad m=3 \quad r=1 \quad \begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix} = \begin{bmatrix}0\\0\end{bmatrix}$ $\quad$ dim $N(A) = 2$ $\quad$ Null Space contains all vectors $\perp$ row space

"solve" $Ax = b$ when there is no solution. $m > n$ more unknowns than equations.

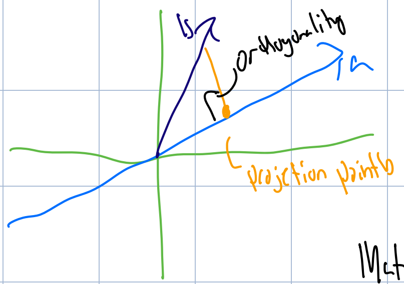
$A^TA \quad n\times m \ m\times n \quad n\times n$ $\quad$ Symmetric $(A^TA)^T = A^TA^{TT}$

$Ax = b \Rightarrow A^TA\hat{x} = A^Tb$

$\begin{bmatrix}1&1\\1&2\\1&5\end{bmatrix}\begin{bmatrix}1&1&1\\1&2&5\end{bmatrix} = \begin{bmatrix}3&8\\8&30\end{bmatrix}$

rank of A^TA = rank of A

A^TA is invertible exactly if A has independent columns.



Find the point on a that's closest to b.

$e = b - p$ $\quad p = xa \quad a^T(b - xa) = 0 \quad xa^Ta = a^Tb$

$x = \frac{a^Tb}{a^Ta} \qquad p = ax$

Matrix?

$p = a\frac{a^Tb}{a^Ta}$ $\quad$ proj p: Pb $\quad P = \frac{aa^T}{a^Ta}$ $\quad$ Col Space (P) = line through A $\quad$ rank$(P) = 1$

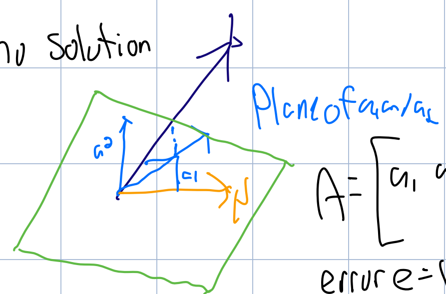
$$\boxed{P^T = P \qquad P^2 = P}$$

Two main properties of P: Symmetric $\quad$ it's square is itself

Why project? Because $Ax = b$ may have no solution

Solve $A\hat{x} = p$, projection of b onto col space

how do we project p onto the plane?



$A = \begin{bmatrix} a_1 & a_2 \end{bmatrix}$

error $e = b - p$ is $\perp$ to the plane

projection $p = \hat{x}_1 a_1 + \hat{x}_2 a_2 = A\hat{x}$

$p = A\hat{x}$ Find $\hat{x}$ $\quad b - A\hat{x}$ is perpendicular to the plane

$a_1^T(b - A\hat{x}) = 0 \quad a_2^T(b - A\hat{x}) = 0 \qquad \begin{bmatrix} a_1^T \\ a_2^T \end{bmatrix}(b - A\hat{x}) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

$A^T\overbrace{(b - A\hat{x})}^{e} = 0$ $\quad e$ in $N(A^T)$ $\quad e \perp$ Col Space (A) $\quad A^TA\hat{x} = A^Tb$

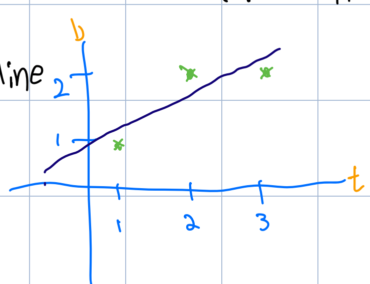
$\hat{x} = (A^TA)^{-1}A^Tb \qquad p = A\hat{x} = A(A^TA)^{-1}A^Tb \qquad \frac{aa^T}{a^Ta} \qquad$ Projection matrix $P = A(A^TA)^{-1}A^T$

$AA^{-1}(A^T)^{-1}A^T = I$ => wrong if A is not a square matrix. It would not have an inverse

$P^T = P \qquad P^2 = P \qquad (\cancel{A(A^TA)^{-1}}A^T)^2, \; A^T P^2 = P$ ✓

Least squares fitting by a line

$(1,1), (2,2), (3,2)$



$b = C + Dt$

$C + D = 1$
$C + 2D = 2$
$C + 3D = 2$

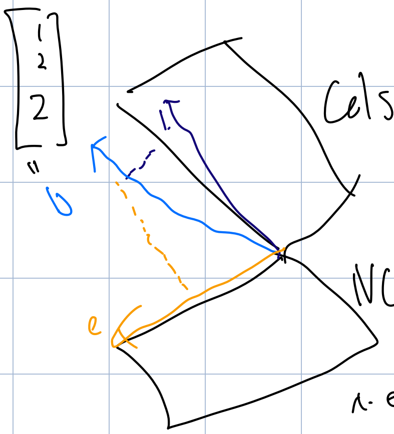
$$\underset{A}{\begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix}} \underset{x}{\begin{bmatrix} C \\ D \end{bmatrix}} = \underset{b}{\begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}}$$

(no solution. look for best solution)

$A^TA\hat{x} = A^Tb$ => best solution

Proj matrix: $P = A(A^TA)^{-1}A^T$ $\quad$ If b in column space $Pb = b$ $\quad Ax = b$

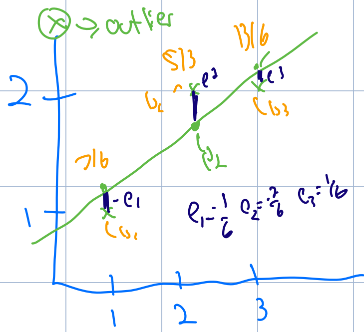
If $b \perp$ column space $Pb = 0$ (nullspace of a transpose)



$p+e = (I-P)b$

I proj onto $\perp$ space

e - error vector



$y = C + Dt$ $\quad Ax = b$

$C + D = 1$ $\quad \begin{bmatrix}1&1\\1&2\\1&3\end{bmatrix}\begin{bmatrix}C\\D\end{bmatrix}\begin{bmatrix}1\\2\\3\end{bmatrix}$

$C + 2D = 2$ $\quad$ no solution, how do

$C + 3D = 2$ $\quad$ we get "best" solution?

least squares solution. Minimize $\|Ax-b\|^2 = \|e\|^2$

overall error: $e_1^2 + e_2^2 + e_3^2 \Rightarrow$ linear regression

$p_1 - p_3$ - points on line of best fit

Find $\hat{x} = \begin{bmatrix}\hat{c}\\\hat{d}\end{bmatrix}$, p $\quad$ most important equation in stats

$\boxed{A^TA\hat{x} = A^Tb}$

$p = A\hat{x}$

$\underbrace{\begin{bmatrix}1&1&1\\1&2&3\end{bmatrix}\begin{bmatrix}1&1\\1&2\\1&3\end{bmatrix} = \left[\begin{array}{cc|c}3&6&5\\6&14&11\end{array}\right]}_{\text{normal equations}}$

$3C + 6D = 5$
$\underline{6C + 14D = 11}$
$2D = 1$
$D = \frac{1}{2}$ $\quad C = \frac{2}{3}$

$(C+D-1)^2 + (C+2D-2)^2 + (C+3D-2)^2$ (terms labelled e_1, e_2, e_3)

Best line: $y = \frac{2}{3} + \frac{1}{2}t$ $\qquad$ error vector: $[-\frac{1}{6}\ \ \frac{2}{6}\ \ -\frac{1}{6}]$ $\qquad$ $p = [\frac{7}{6}\ \ \frac{10}{6}\ \ \frac{13}{6}]$

$p + e = b$: $\begin{bmatrix}7/6 - 1/6\\10/6 + 2/6\\13/6 - 1/6\end{bmatrix}\ \ \begin{bmatrix}1\\2\\2\end{bmatrix}$ $\qquad e \perp p$

If A has independent columns, then A^TA is invertible. To prove: x must be $\begin{bmatrix}0\\0\end{bmatrix}$

Suppose $A^TAx = 0$ $\quad$ Brilliant Idea $- x^TA^TAx = 0 = (Ax)^T(Ax) \Rightarrow Ax = 0$

$\Rightarrow A$ has independent columns $\Rightarrow x = 0$

Columns definitely independent if they are $\underbrace{\text{perpendicular unit vectors}}_{\text{orthonormal vectors}}$, $\begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix}$ $\begin{bmatrix}1\\0\\0\end{bmatrix}$ $\begin{bmatrix}0\\1\\0\end{bmatrix}$ $\begin{bmatrix}0\\0\\1\end{bmatrix}$

Orthonormal vectors: $q_i^Tq_j = \begin{cases}0 & \text{if } i \neq j\\1 & \text{if } i = j\end{cases}$ $\qquad$ Orthogonal basis: $q_1 \dots q_n$

Gram-Schmidt: $A \to Q$ $\qquad$ Orthogonal matrix: Q: Square matrix

$Q = [q_1 \dots q_n]$ $\quad Q^TQ = \begin{bmatrix}q_1\\\vdots\\q_n\end{bmatrix}[q_1 \dots q_n] = \begin{bmatrix}1&0&\cdots&0\\0&1&\cdots&0\\0&0&\ddots&0\\0&0&\cdots&1\end{bmatrix}$ $\quad Q^TQ = I$ $\quad$ If Q is square, then

* a matrix Q must be square to be considered an orthogonal matrix. $\quad Q^TQ = I$ tells us $Q^T = Q^{-1}$

Examples:

perm $Q = \begin{bmatrix}0&0&1\\1&0&0\\0&1&0\end{bmatrix}\begin{bmatrix}0&1&0\\0&0&1\\1&0&0\end{bmatrix} = I$ (Q^T over the second matrix) $\qquad Q = \begin{bmatrix}\cos\theta & -\sin\theta\\\sin\theta & \cos\theta\end{bmatrix}$ $\qquad Q = \frac{1}{\sqrt{2}}\begin{bmatrix}1&1\\1&-1\end{bmatrix}$ $\quad Q = \frac{1}{2}\begin{bmatrix}1&1&1&1\\1&-1&1&-1\\1&1&-1&-1\\1&-1&-1&1\end{bmatrix}$

$Q = \frac{1}{3}\begin{bmatrix} 1 & -2 & 2 \\ 2 & -1 & -2 \\ 2 & 2 & 1 \end{bmatrix}$ orthonormal basis $\mathbb{R}^3$

Q has orthonormal columns. Project onto its column space.

$P = Q\underbrace{(Q^TQ)^{-1}}_{\text{Identity}}Q^T = QQ^T \; (= I \text{ if } Q \text{ is square})$

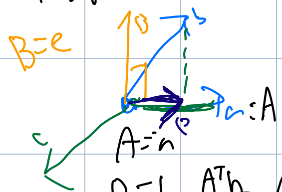
$(QQ^T)(QQ^T) = QQ^T$

$A^TA\hat{x} = A^Tb$ now A is Q. $Q^TQ\hat{x} = Q^Tb \Rightarrow I\hat{x} = Q^Tb = \hat{x} = Q^Tb$

$\boxed{\hat{x}_i = \hat{q}_i^T b}$

Gram-Schmidt: goal: to get from any two vectors a, b to orthogonal vectors A, B to orthonormal q_1, q_2

independent vectors a, b



$A, B \quad q_1 = \frac{A}{\|A\|} \quad q_2 = \frac{B}{\|B\|}$

$a, b, c, \; A, B, C, \; q_1 = \frac{A}{\|A\|} \quad q_2 = \frac{B}{\|B\|} \quad q_3 = \frac{C}{\|C\|}$

$B = b - \frac{A^Tb}{A^TA}A \quad A \perp B \quad A^TB = A^T\left(b - \frac{A^Tb}{A^TA}A\right) = 0$

$C = c - \frac{A^Tc}{A^TA}A - \frac{B^Tc}{B^TB}B \quad C \perp A, \; C \perp B$

$a = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \quad b = \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} \quad B = \underbrace{\begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}}_{b} - \frac{3}{3}\underbrace{\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}}_{A} = \underbrace{\begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}}_{B}$

check that $A \perp B$ (dot)
$[0 - 1 + 1] = 0$ ✓

$Q = [q_1 \; q_2] = \begin{bmatrix} 1/3 & 0 \\ 1/3 & -1/\sqrt{2} \\ 1/3 & -1/\sqrt{2} \end{bmatrix} \quad A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 1 & 2 \end{bmatrix} \Rightarrow Q = \begin{bmatrix} 1/3 & 0 \\ 1/3 & -1/\sqrt{2} \\ 1/3 & -1/\sqrt{2} \end{bmatrix} \quad \operatorname{col}(A) = \operatorname{col}(Q)$

$\underline{A = LU}$ extension of elim

$\underline{A = QR}$ - expression of Gram-Schmidt $\begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} q_1 \\ q_2 \end{bmatrix}\begin{bmatrix} a_1^Tq_1 & * \\ a_1^Tq_2 & * \end{bmatrix}$

R ~ upper triangular

$a_1^Tq_2 = 0$ bc $a_1 \perp q_2$