Solving linear least squares problems..

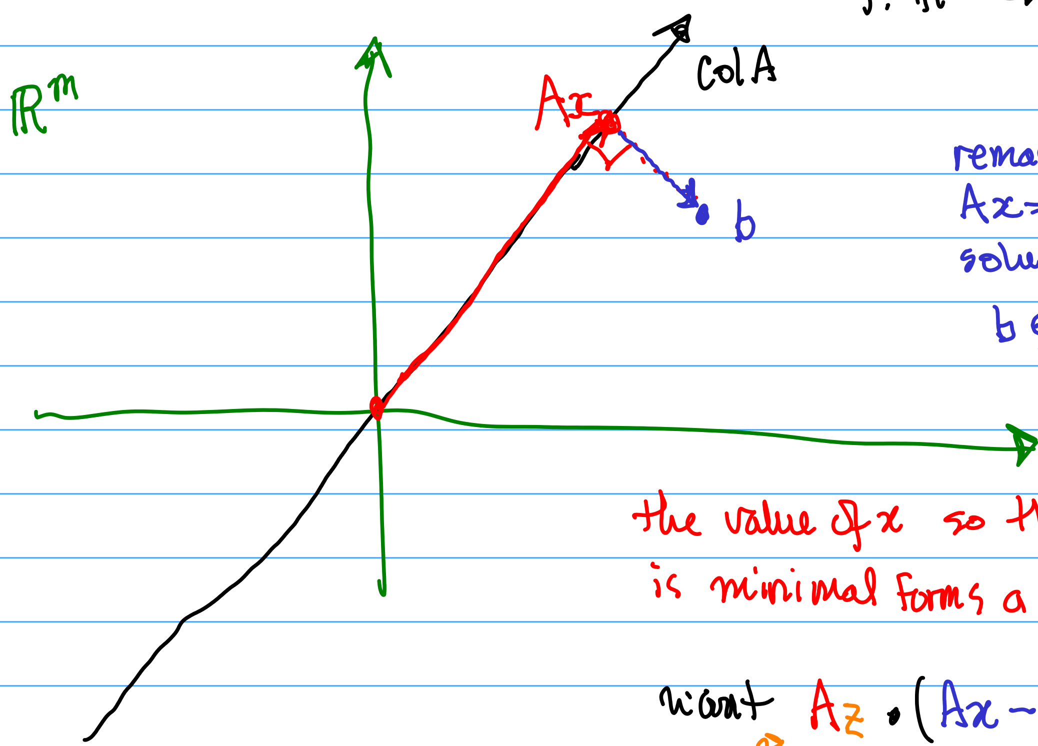
Idea: If $Ax = b$ doesn't have a solution, then instead minimize the residual error $\underbrace{\|Ax - b\|}_{\text{Euclidian norm}}$.

Review of linear algebra... $A \in \mathbb{R}^{m \times n}$

$$\text{Col}\,A = \{Ax : x \in \mathbb{R}^n\} = \text{range of } f \subseteq \mathbb{R}^m$$

where $f(x) = Ax$

$f : \mathbb{R}^n \to \mathbb{R}^m$



remark, since $Ax = b$ has no solution, then $b \notin \text{Col}\,A$.

the value of x so that $\|Ax - b\|$ is minimal forms a right angle.

want $Az \cdot (Ax - b) = 0$

Actually, Az here since the perpendicularity needs to hold for all vectors in Col A

$$(Az)^T (Ax - b) = 0$$

$$z^T A^T (Ax - b) = 0$$

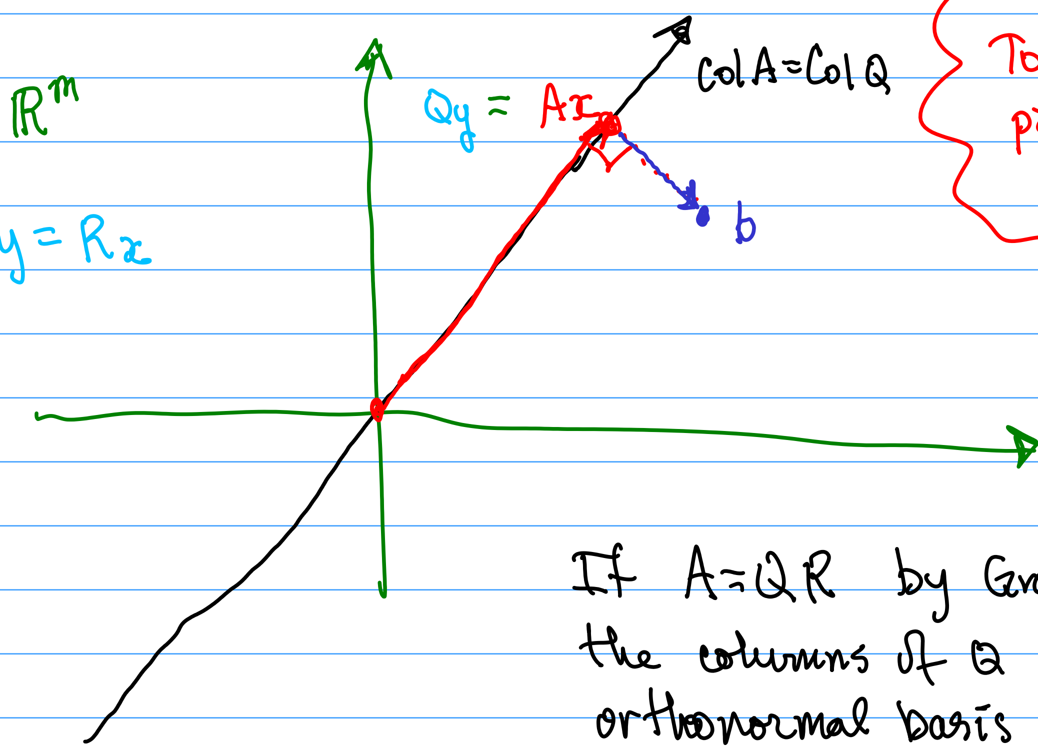
Therefore

$$z^T\left(A^TAx - A^Tb\right) = 0$$

$$z \cdot \left(\underbrace{A^TAx - A^Tb}\right) = 0 \quad \text{for all } z \in \mathbb{R}^n$$

If $A^TAx = A^Tb$ then the above holds...
these are called the normal equations...

Solve them? $\boxed{x = (A^TA)^{-1}A^Tb}$ is the solution...



where $y = Rx$

To draw this picture need col A

If $A = QR$ by Gram-Schmidt, the columns of Q form an orthonormal basis for col A.

Col A = Col Q

Again, perpendicularity must hold for all vectors in Col Q so z here

$$Qz \cdot (Qy - b) = z^TQ^T(Qy - b)$$

$$= z^T(\cancel{Q^TQ}y - Q^Tb) = z \cdot (y - Q^Tb)$$

$y = Rx$

$= z \cdot (Rx - Q^T b) = 0$ for all $z \in \mathbb{R}^n$

if $Rx = Q^T b$ then above holds..

Idea is to solve $Rx = Q^T b$ to find the minimizer of $\|Ax - b\|$.

Note: one could define

$$J(x) = \|Ax - b\|^2 = (Ax - b) \cdot (Ax - b)$$

and since $J(x)$ is differentiable one could find the minimum by solving $J'(x) = 0$.

To solve $Rx = Q^T b$ we need to find the factorization $A = QR$ which means Gram-Schmidt in Math 330 and something else next time...

Gram-schmidt $\quad A = \left[a_1 \,\middle|\, a_2 \,\middle|\, \cdots \,\middle|\, a_n \right]$

$t_1 = a_1 \qquad\qquad q_1 = \dfrac{t_1}{\|t_1\|}$

$t_2 = a_2 - (q_1 \cdot a_2) q_1 \qquad\qquad q_2 = \dfrac{t_2}{\|t_2\|}$

$t_3 = a_3 - (q_1 \cdot a_3) q_1 - (q_2 \cdot a_3) q_2 \qquad\qquad q_3 = \dfrac{t_3}{\|t_3\|}$

$\vdots$

$$t_n = a_n - (q_1 \cdot a_n) q_1 - \cdots - (q_{n-1} \cdot a_n) q_{n-1} \qquad q_n = \frac{t_n}{\|t_n\|}.$$

Problem: no (or little) flexibility in this algorithm. But I need flexibility to control propagation of rounding error.

Note solving $Rx = Q^T b$ is a good way to solve $Ax = b$ even when $Ax = b$ has a solution, because minimizing the residual error $\|Ax - b\|$ also minimizes rounding error.

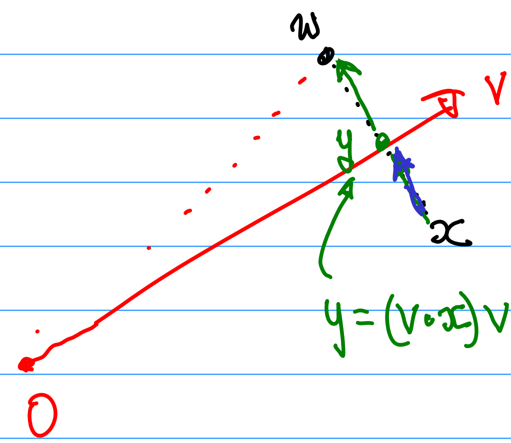
Why is this not ideal? $x = (A^T A)^{-1} A^T b$. Because the matrix $A^T A$ may not be well conditioned...
Note if $A \in \mathbb{R}^{m \times n}$ and $Ax = b$ has no solution, the usual reason is because $m >> n$. This is overdetermined system of linear equations. But then $\underbrace{A^T}_{n \times m} \underbrace{A}_{m \times n} \in \mathbb{R}^{n \times n}$ which is very small in size compared to A. That's another reason the normal equations are not well behaved with rounding error...

Idea Find $A = QR$ using a different algorithm...

① Method of reflections
② Method of rotations...

Note reflections and rotations are orthogonal transformations, that is they are like the Q matrix.

Let v be a unit vector



reflect x about v gives w

$$w = y + w - y$$

$$= y + y - x$$

$$= 2y - x$$

$$= 2(v \cdot x)v - x$$

$$= 2(v^T x)v - x$$

since $v^T x$ is a scalar

$$= 2v\, v^T x - x$$

$$= (2vv^T - I)x$$

Define $H_v x = -w = (I - 2vv^T)x$

$$H_v = I - 2vv^T$$

Next time we use H_v to find QR decomposition of A in a way that behaves well with respect to rounding.