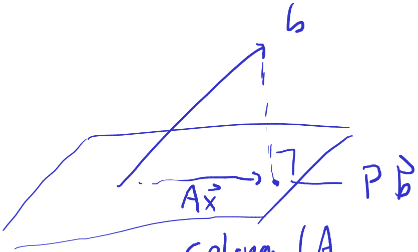
$$A\vec{x} \approx \vec{b}$$



colspan (A

$$A^TA\vec{x} = A^Tb$$

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

$$A\vec{x} = P\vec{b}$$

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

$\hookrightarrow P^2 = P$ ✓

$P^T = P$ ✓

Pseudoinverse

What is the pseudoinverse of A?

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

$$P = AA^+$$

What can we say about the condition number in the case of a tall-and-skinny, full-rank matrix? $\longrightarrow$ not full rank: $\kappa(A) = \infty$ (cond.)

$$\kappa(A) = \|A\| \|A^+\|$$

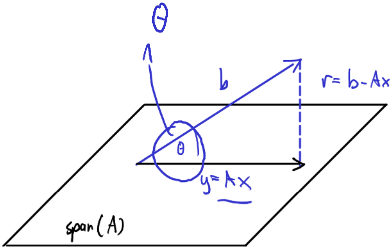
square: $\|A\| \|A^{-1}\| = \kappa(A)$

$\hookrightarrow$ same in square case

What does all this have to do with solving least squares problems?

$$\vec{x} = A^+ \vec{b}$$

Sensitivity and Conditioning of Least Squares



Relate $\|\boldsymbol{Ax}\|$ and $\|\boldsymbol{b}\|$ with $\$\theta$ via trig functions.

$$\cos\theta = \frac{\|Ax\|_2}{\|b\|_2}$$

Sensitivity and Conditioning of Least Squares (II)

Derive a conditioning bound for the least squares problem.

$$\vec{x} = A^+ \vec{b} \qquad \Delta \vec{x} = A^+ \Delta \vec{b} \qquad \|\cdot\| = \|\cdot\|_2$$

$$\frac{\|\Delta x\|}{\|x\|} \le \|A^+\| \frac{\|\Delta b\|}{\|x\|}$$

$$= \frac{\kappa(A)}{\|A\| \cancel{\|A^+\|}} \cancel{\|A^+\|} \frac{\|b\|}{\|b\|} \cdot \frac{\|\Delta b\|}{\|x\|}$$

$$= \kappa(A) \frac{\|b\|}{\|A\| \|x\|} \cdot \frac{\|\Delta b\|}{\|b\|} \le \kappa(A) \underbrace{\frac{\|b\|}{\|Ax\|}}_{\frac{1}{\cos\theta}} \cdot \frac{\|\Delta b\|}{\|b\|}$$

What values of θ are bad?

$$\theta \approx \pi/2$$

$$\frac{\|\Delta x\|}{\|x\|} \leq \kappa(A) \cdot \frac{1}{\cos\theta} \cdot \frac{\|\Delta b\|}{\|b\|}$$

Sensitivity and Conditioning of Least Squares (III)

Any comments regarding dependencies?

Here: conditioning of LS▯ depends on b!

What about changes in the matrix?

$$\frac{\|\Delta x\|}{\|x\|} \le \left[\operatorname{cond}(A) + \operatorname{cond}(A)^2 \cdot \tan(\theta)\right] \frac{\|\Delta A\|}{\|A\|}$$

Recap: Orthogonal Matrices

What's an *orthogonal (=orthonormal) matrix*?

One that satisfies $Q^T Q = I$ and $QQ^T = I$.

How do orthogonal matrices interact with the 2-norm?

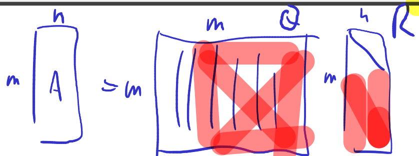
$$\|Q\boldsymbol{v}\|_2^2 = (Q\boldsymbol{v})^T(Q\boldsymbol{v}) = \boldsymbol{v}^T Q^T Q\boldsymbol{v} = \boldsymbol{v}^T\boldsymbol{v} = \|\boldsymbol{v}\|_2^2 .$$

Transforming Least Squares to Upper Triangular

full

Suppose we have $A = QR$, with Q square and orthogonal, and R upper triangular. This is called a QR factorization.
How do we transform the least squares problem $A\boldsymbol{x} \cong \boldsymbol{b}$ to one with an upper triangular matrix?

$$\|A\tilde{x} - \vec{b}\|_2$$
$$= \|Q^T(QR\tilde{x} - \vec{b})\|_2$$
$$= \|Rx - Q^T b\|_2 = \|Rx - Q^T b\|_2$$



"Full"

"economy"

Simpler Problems: Triangular

What do we win from transforming a least-squares system to upper triangular form?

$$\begin{pmatrix} R_{top} \\ 0 \end{pmatrix} \vec{x} \cong \begin{pmatrix} (Q^T \vec{b})_{top} \\ (Q^T b)_{bottom} \end{pmatrix}$$

How would we minimize the residual norm?

$$\|r\|_2^2 = \overbrace{\|(Q^T b)_{top} - R_{top}\vec{x}\|_2^2}^{T1} + \overbrace{\|(Q^T b)_{bottom}\|_2^2}^{T2}$$

$$\vec{x} = R_{top}^{-1}(Q^T b)_{top}$$

$$\Rightarrow T1 = 0 \qquad \Rightarrow \|r\|_2^2 = T2$$

Computing QR

- Gram-Schmidt
- Householder Reflectors
- Givens Rotations

Demo: Gram-Schmidt–The Movie [cleared]
Demo: Gram-Schmidt and Modified Gram-Schmidt [cleared]
Demo: Keeping track of coefficients in Gram-Schmidt [cleared]
Seen: Even modified Gram-Schmidt still unsatisfactory in finite precision arithmetic because of roundoff.

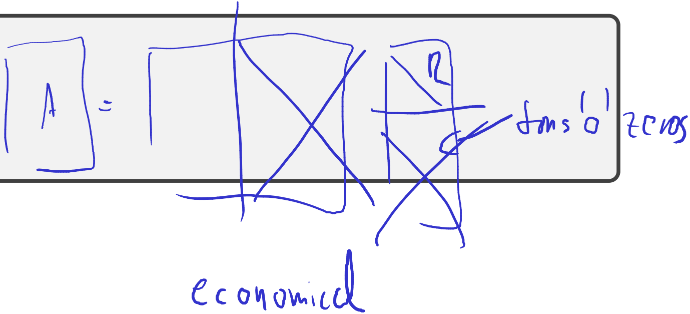
NOTE: Textbook makes further modification to 'modified' Gram-Schmidt:

- Orthogonalize *subsequent* rather than *preceding* vectors.
- Numerically: no difference, but sometimes algorithmically helpful.

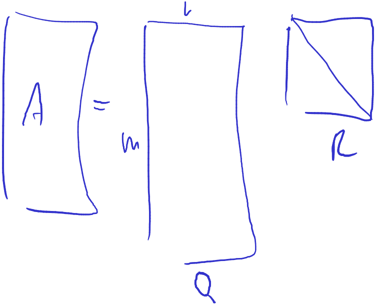
Economical/Reduced QR

Full

Is QR with square Q for $A \in \mathbb{R}^{m \times n}$ with $m > n$ efficient?



In-Class Activity: QR



$Q^T b$

In-class activity: QR

Householder Transformations

Find an *orthogonal* matrix Q to zero out the lower part of a vector $\boldsymbol{a}$.