Linear Systems of Eq.'s

Part 4: Orthogonalization & QR

Q ↑ unitary

Recap & Wrap Up GE

$$PA = LU$$

- Gaussian elimination (GE) triangularizes A via triangular transformations.
- Partial Pivoting suppresses growth in these triangular "elimination" matrices.
- GE is unstable in theory, but behaves like a backward stable alg. in practice.

There is a special class of matrices for which GE doesn't require pivoting, is provably backward stable, and 2x as fast.

Cholesky Factorization (brief sketch)

2 key properties required:

1) $A = A^*$ (Hermitian/Symmetric matrix)
   (Complex / Real)

2) $x^*Ax > 0$ when $x \neq 0$ (positive def.)

$\Rightarrow$ A is (Hermitian) symmetric positive definite (SPD).

Remark: Symmetric matrices have orthogonal eigenvectors and real eigenvalues, $A = V \Lambda V^*$. SPD matrices also have positive eigenvalues, and therefore full rank.

GE for SPD matrix (illustrative sketch)

$$A = \begin{bmatrix} 1 & w^* \\ w & K \end{bmatrix} \quad \underset{\text{GE}}{\overset{\text{1 step}}{\Rightarrow}} \quad A = \begin{bmatrix} 1 & 0 \\ w & I \end{bmatrix} \begin{bmatrix} 1 & w^* \\ 0 & K - ww^* \end{bmatrix}$$

Rather than move directly to 2^{nd} column, try to maintain symmetry: Apply same row operation again on columns of A.

$$A = \begin{bmatrix} 1 & 0 \\ w & I \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & \underbrace{K - ww^*}_{\substack{\text{repeat process} \\ \text{for this block}}} \end{bmatrix} \begin{bmatrix} 1 & w^* \\ 0 & I \end{bmatrix}$$

Cholesky continues by zeroing out columns and rows symmetrically until

$$\boxed{A = \underset{\text{lower triangular}}{R}\,\overset{\text{upper triangular}}{R^*} \qquad r_{ii} > 0}$$

Cholesky factorization

⇒ 2x fast to compute (work with half of matrix by symmetry)

⇒ No pivoting (largest entry always appears along diagonal)

⇒ Backward stable

Read more about Cholesky in LNT Lec. 23.

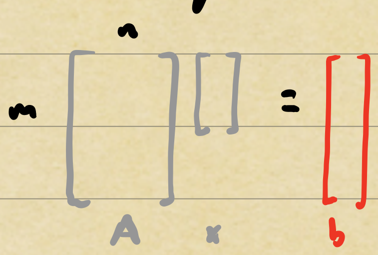
New topic

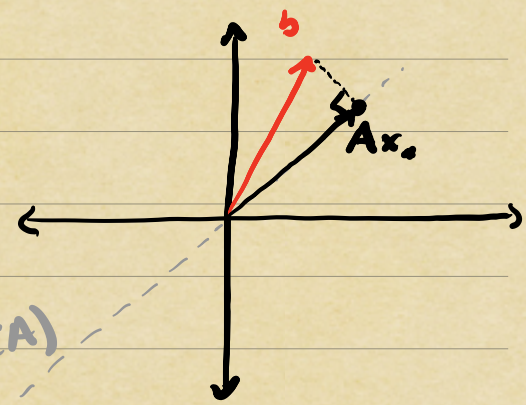
Least Squares & QR

$$A = \underset{\substack{\text{orthonormal}\\\text{columns}}}{Q}\,\underset{\text{upper triangular}}{R}$$

The columns of Q form an orthonormal basis for the column space of A.

Least squares in a picture:





$m > n \to$ overdetermined. Recall, if A has full column rank ($= n$, no zero singular vals) then there is no exact solution.

$$x_* = \underset{x \in \mathbb{R}^n}{\operatorname{argmin}} \|Ax - b\|_2$$

is the least-squares solution.

To find x_*, we have two steps:

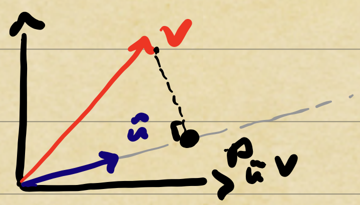
1) $d = Pb$
   P ← orthogonal projection onto range(A)

2) Solve $A x_* = d$

Orthogonal projections

To project a vector v orthogonally onto the line in the direction of unit vector $\hat{u}$,

$$P_{\hat{u}} v = (\hat{u}^* v)\, \hat{u}$$



If we have a vector u in the direction of $\hat{u}$, we can always just normalize:

$$P_{\hat{u}} v = \left(\frac{u^* v}{\|u\|}\right) \frac{u}{\|u\|}$$

To project orthogonally onto a <u>plane</u> P spanned by orthonormal basis $\{\hat{u}_1, \hat{u}_2\}$,

$$P_P v = (\hat{u}_1^* v)\hat{u}_1 + (\hat{u}_2^* v)\hat{u}_2$$

$$= (\hat{u}_1 \hat{u}_1^* + \hat{u}_2 \hat{u}_2^*) v$$

$$= \underbrace{\begin{bmatrix} | & | \\ \hat{u}_1 & \hat{u}_2 \\ | & | \end{bmatrix} \begin{bmatrix} - & \hat{u}_1^* & - \\ - & \hat{u}_2^* & - \end{bmatrix}}_{\text{Structure of OP onto span of ONB } \hat{u}_1, \hat{u}_2} \begin{bmatrix} | \\ v \\ | \end{bmatrix}$$

<u>Orthogonal projection onto range(A)</u>

If columns of $Q = \begin{bmatrix} | & & | \\ q_1 & \cdots & q_n \\ | & & | \end{bmatrix}$ form an ONB

for range(A), then we can solve Least-Squares problems

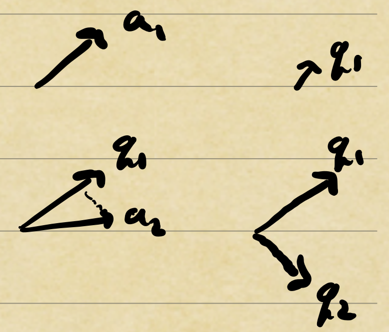
$$1) \quad d = \underbrace{\begin{bmatrix} | & & | \\ q_1 & \cdots & q_n \\ | & & | \end{bmatrix}}_{Q} \underbrace{\begin{bmatrix} - & q_1^* & - \\ & \vdots & \\ - & q_n^* & - \end{bmatrix}}_{Q^*} \begin{bmatrix} | \\ b \\ | \end{bmatrix}$$

2) Solve $Ax_* = d$

How to compute ONB for range(A)?

Gram-Schmidt

Idea: orthogonalize columns of A sequentially.



$$q_1 = \frac{a_1}{r_{11}}$$

$$q_2 = \frac{1}{r_{22}}\left[a_2 - \underbrace{(q_1^* a_2)\, q_1}_{P_{q_1} a_2 = r_{12}}\right]$$

$$q_3 = \frac{1}{r_{33}}\left[a_3 - \underbrace{(q_1^* a_3)}_{r_{13}} q_1 - \underbrace{(q_2^* a_3)}_{r_{23}} q_2\right]$$

$\vdots$

$$q_n = \frac{1}{r_{nn}}\left[a_n - \sum_{i=1}^{n-1} \underbrace{(q_i^* a_n)}_{r_{in}} q_i\right]$$

$$r_{jj} = \left\| a_j - \sum_{i=1}^{j-1} r_{ij} q_i \right\|_2$$

In matrix notation:

$$\begin{bmatrix} | & | & & | \\ a_1 & a_2 & \cdots & a_n \\ | & | & & | \end{bmatrix} \begin{bmatrix} r_{11}^{-1} & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{bmatrix} = \begin{bmatrix} | & | & & | \\ q_1 & a_2 & \cdots & a_n \\ | & | & & | \end{bmatrix}$$

$$\begin{bmatrix} | & | & & | \\ q_1 & a_2 & \cdots & a_n \\ | & | & & | \end{bmatrix} \begin{bmatrix} 1 & r_{12}r_{22}^{-1} & & & \\ 0 & r_{22}^{-1} & & & \\ \vdots & & 1 & & \\ & & & \ddots & \\ 0 & & & & 1 \end{bmatrix} = \begin{bmatrix} | & | & | & & | \\ q_1 & q_2 & a_3 & \cdots & a_n \\ | & | & | & & | \end{bmatrix}$$

$$\vdots$$

$$A R_1 R_2 \cdots R_n = \begin{bmatrix} | & & | \\ q_1 & \cdots & q_n \\ | & & | \end{bmatrix}$$

Sequence of Upper Triangular transformations from the <u>right</u> orthogonalizes the columns of A.

Inverting: $R = R_n^{-1} \cdots R_2^{-1} R_1^{-1}$

$$\boxed{\begin{array}{c} \begin{bmatrix} | & & | \\ a_1 & \cdots & a_n \\ | & & | \end{bmatrix} = \begin{bmatrix} | & & | \\ q_1 & \cdots & q_n \\ | & & | \end{bmatrix} \begin{bmatrix} \diagdown & | & | & | \\ & \ddots & | & | \\ & & & | \end{bmatrix} \\ A \quad = \quad Q \quad R \end{array}}$$

QR factorization

Orthogonalizing columns of A with (classical) Gram-Schmidt runs into some stability issues. We'll look at two alternative algorithms next time.