QR Factorization (Heath §3.4.5)

... and least squares:

$$\underset{m\times n}{A} = \underset{m\times m}{Q} \underset{m\times n}{\begin{pmatrix} R \\ 0 \end{pmatrix}}, \quad R \in \mathbb{R}^{n\times n}$$

$$\|r\|_2^2 = \|b - Ax\|_2^2 = \left\| b - Q\begin{pmatrix} R \\ 0 \end{pmatrix} x \right\|_2^2 =$$

⟨multiply by Q^T⟩ $$= \left\| Q^T b - \begin{pmatrix} R \\ 0 \end{pmatrix} x \right\|_2^2$$

$$= \left\| \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} - \begin{pmatrix} R \\ 0 \end{pmatrix} x \right\|_2^2$$

$$= \|b_1 - Rx\|_2^2 + \|b_2\|_2^2$$

minimum occurs at $Rx = b_1$ (solve by back substitution)

residual is then s.t. $\|r\|_2^2 = \|b_2\|_2^2$

(7) (TAB Lect. 7) QR Factorization

One of most important algorithms in LA.

Consider spaces:

$$\mathrm{span}(a_1) \subseteq \mathrm{span}(a_1, a_2) \subseteq \dots \subseteq \mathrm{span}(a_1, a_2, \dots, a_n)$$

QR: make <u>orthonormal</u> <u>bases</u> for these <u>spaces</u>

$$\mathrm{span}(q_1, q_2, \dots, q_j) = \mathrm{span}(a_1, a_2, \dots, a_j)$$

→ <u>Condition</u>:

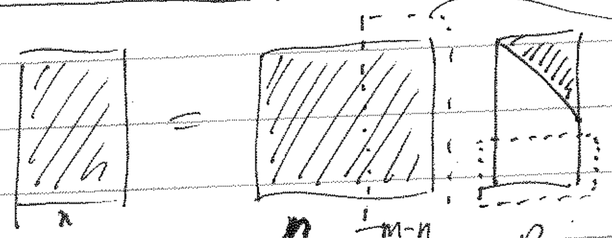
$$\begin{pmatrix} | & | & & | \\ a_1 & a_2 & \cdots & a_n \\ | & | & & | \end{pmatrix} = \begin{pmatrix} | & | & & | \\ q_1 & q_2 & \cdots & q_n \\ | & | & & | \end{pmatrix} \begin{pmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ & r_{22} & & \vdots \\ & & \ddots & \vdots \\ & & & r_{nn} \end{pmatrix}$$

equations written out:

$$\begin{cases} \vec{a}_1 = r_{11} \vec{q}_1 \\ \vec{a}_2 = r_{12} \vec{q}_1 + r_{22} \vec{q}_2 \\ \vdots \\ \vec{a}_n = r_{1n} \vec{q}_1 + \dots + r_{nn} \vec{q}_n \end{cases}$$

$$A = \hat{Q} \hat{R}.$$ "<u>Reduced</u> QR factorization of A"

<u>Full QR factorization</u>.



columns are ⊥ to Range(A)

$\mathrm{Range}(A)^\perp$

$\mathrm{Null}(A^T)$

A Q R

m×n m×m m×n

8 Gram-Schmidt Orthogonalization

$$v_j = a_j - (q_1^T a_j) q_1 - (q_2^T a_j) q_2 - \cdots - (q_{j-1}^T a_j) q_{j-1}$$

[remove components of a_j in direction of $q_1, \dots, q_{j-1}$]

$$q_j = \frac{v_j}{\|v_j\|}$$ [normalize]

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

$$q_2 = \frac{a_2 - r_{12} q_1}{r_{22}}$$

$\vdots$

$$q_n = \frac{a_n - r_{1n} q_1 - r_{2n} q_2 - \cdots - r_{n-1\,n} q_{n-1}}{r_{nn}}$$

$$\Rightarrow \begin{cases} r_{ij} = q_i^T a_j & (i \neq j) \\ |r_{jj}| = \| a_j - \sum_{i=1}^{j-1} r_{ij} q_i \|_2 \end{cases}$$

Note: r_{jj} could be +ive or -ive.

"Classical Gram-Schmidt" (unstable)

Existence: All matrices have QR factorization.

$QRx = b$

$Rx = Q^T b$

2x ops as Gaussian Elimination

~~Yes 32 anxiety / ? 37 food / tackling / ? 39 education / spending~~

~~? 39~~

(a) Gram-Schmidt Orthogonalization
(T&B, Lec. 8)

" triangular orthogonalization "

Modified Gram-Schmidt
Use a sequence of orthogonal projections rather than 1

Classical (A.7.1)

for $j = 1 \dots n$

$v_j = a_j$

for $i = 1 \dots j-1$

$\boxed{r_{ij} = q_i^T a_j}$

$v_j = v_j - r_{ij} q_i$

$r_{jj} = \|v_j\|_2$

normalize $q_j = v_j / r_{jj}$

can replace $v \leftarrow q$

requires separate storage for A, Q, & R

Modified (A.8.1)

for $i = 1$ to n

$v_i = a_i$

for $i = 1 \dots n$

$r_{ii} = \|v_i\|$

$q_i = v_i / r_{ii}$

for $j = i+1 \dots n$

$\boxed{r_{ij} = q_i^T v_j}$

$v_j = v_j - r_{ij} q_i$

requires separate storage for Q, R

(but explicit representation of Q)

$$\begin{pmatrix} | & | & & | \\ a_1 & a_2 & \cdots & a_n \\ | & | & & | \end{pmatrix} = \begin{pmatrix} | & | & & | \\ q_1 & q_2 & \cdots & q_n \\ | & | & & | \end{pmatrix} \begin{pmatrix} r_{11} & r_{12} & r_{13} & \cdots & r_{1n} \\ & r_{22} & r_{23} & & \vdots \\ & & r_{33} & & \vdots \\ & & & \ddots & \\ & & & & r_{nn} \end{pmatrix}$$

$v_j \leftarrow (I - q_1 q_1^T - q_2 q_2^T \cdots - q_{j-1} q_{j-1}^T) v_j$

$v_j \leftarrow (I - q_n q_n^T) \cdots (I - q_2 q_2^T) (I - q_1 q_1^T) v_j$

one flop = $+, -, \times, \div, \sqrt{}$

$\sim 2mn^2$ flops $\quad m \times n$ factorization

m mult
m-1 add m mult
m subt

$\sim 4m$

$2n^2m$

j^{th} step

$$\begin{pmatrix} | & | & | & | \\ a_1 & a_2 \cdots a_j & \cdots & a_n \\ | & | & | & | \end{pmatrix} \qquad \begin{pmatrix} | & | & & | \\ q_1 & q_2 & \cdots & q_{j-1} \\ | & | & & | \end{pmatrix}$$

$$\begin{pmatrix} r_{11} & r_{12} & \cdots & r_{1,j-1} \\ & r_{22} & & \vdots \\ & & & r_{j-1,j-1} \end{pmatrix}$$

$$\downarrow$$

① v_j

② orthogonalize to previous
[loop over $i = 1, \ldots, j-1$]

$$r_{ij} = \vec{q}_i^{\,T} \vec{a}_j$$

(j^{th} <u>column</u> of r to define ~~q_j~~ $\vec{a}_j$ (rows $\to$ linear comb. of q_i's))

$$v_j = v_j - r_{ij} \vec{q}_i$$

③ normalize to get q_j

$$r_{jj} = \pm \| \vec{v}_j \|$$

$$q_j = v_j / r_{jj}$$

Classical Gram-Schmidt

(0) $\begin{pmatrix} | & & | \\ a_1 & \cdots & a_n \\ | & & | \end{pmatrix} \longrightarrow \begin{pmatrix} | & & | \\ v_1 & \cdots & v_n \\ | & & | \end{pmatrix}$

i^{th} step

(1) normalize

$$r_{ii} = \|\vec{v}_i\|$$

$$q_i = \vec{v}_i / r_{ii}$$

(2) orthogonalize all ~~a~~ remaining data w.r.t. q_i
[loop over $j = i+1 \ldots n$]

$$r_{ij} = \vec{q}_i^{\,T} \vec{v}_j$$

$$\vec{v}_j = \vec{v}_j - r_{ij}\vec{q}_i$$

Modified Gram-Schmidt

Orthogonalization Methods (Heath §3.5)

goal: introduce 0's into the matrix with orthogonal transform rather than Gauss transforms

commonly used: Householder Reflections (transforms)
H

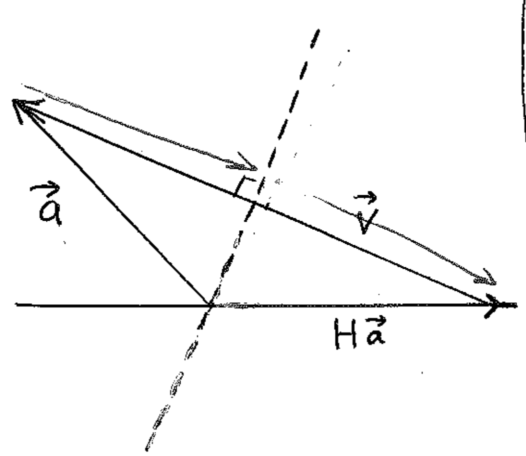
$$H\vec{a} = \begin{pmatrix} \alpha \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \qquad \alpha = \pm \|\vec{a}\|_2$$

(*) How do you choose between $\pm$?

We want

$$H\vec{a} = \vec{a} - 2\frac{\vec{v}\vec{v}^T\vec{a}}{\vec{v}^T\vec{v}}$$

$$= \underbrace{\left(I - 2\frac{\vec{v}\vec{v}^T}{\vec{v}^T\vec{v}}\right)}_{=H}\vec{a}$$

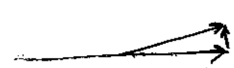


$$\vec{v} = \vec{a} - H\vec{a}$$

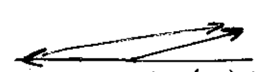
$$\vec{v} = \vec{a} - \alpha\vec{e}_1$$

(*) To avoid cancellation, choose

$$\alpha = -\operatorname{sign}(a_1)\|a\|_2$$



$\alpha = \operatorname{sign}(a_1)\|a\|_2$
potential for cancellation error

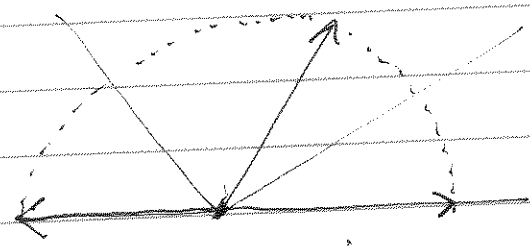


$\alpha = -\operatorname{sign}(a_1)\|a\|_2$
less cancellation error

– use $a \leftarrow \dfrac{a}{\max|a_i|}$ to avoid overflow / underflow

Practical:

- $\alpha = -\operatorname{sign}(a_1)\|a\|_2$ avoid cancellation
- use $a \leftarrow \dfrac{a}{\max|a_i|}$ avoid overflow & underflow



Example $a = \begin{pmatrix}2\\1\\2\end{pmatrix}$ ① $\alpha e_1 = -\sqrt{4+1+4}\; e_1 = -3e_1$

② $v = a - \alpha e_1 \Rightarrow v = a + 3e_1 = \begin{pmatrix}5\\1\\2\end{pmatrix}$

$H = I - 2\dfrac{vv^T}{v^Tv}$ $\quad vv^T = (a+3e_1)(a+3e_1)^T = aa^T + 3e_1a^T + 3ae_1^T + 9e_1e_1^T$

③ $\boxed{H = I - \dfrac{2(a+3e_1)(a+3e_1)^T}{(a+3e_1)^T(a+3e_1)}}$ $\quad \|v\| = \sqrt{25+1+4} = \sqrt{30}$

$$= I - \frac{2}{30}\begin{pmatrix}5\\1\\2\end{pmatrix}(5\ 1\ 2)$$

$$\cancel{H = I - \frac{1}{15}\begin{pmatrix}25 & 5 & 10\\ 5 & 1 & 2\\ 10 & 2 & 4\end{pmatrix} = \begin{pmatrix}-\frac{2}{3} & -\frac{1}{3} & -\frac{2}{3}\\ -\frac{1}{3} & \frac{14}{15} & -\frac{2}{15}\\ -\frac{2}{3} & -\frac{2}{15} & \frac{11}{15}\end{pmatrix}}$$

④ confirm $Ha = \left(I - \dfrac{2vv^T}{v^Tv}\right)a = a - \dfrac{2v(v^Ta)}{v^Tv}$

$$= \begin{pmatrix}2\\1\\2\end{pmatrix} - \frac{2}{30}\begin{pmatrix}5\\1\\2\end{pmatrix}\cancel{(10+1+4)} = \begin{pmatrix}-3\\0\\0\end{pmatrix} \checkmark$$

$H_k = \left(\begin{array}{c|c} I_k & \\ \hline & H' \end{array}\right)$ $\quad k = 0, \dots$

define v in this subspace

$\boxed{v = \begin{pmatrix}0\\a_2\end{pmatrix} - \alpha e_k}$

Annihilate all subdiagonal entries of an $m \times n$ matrix A

$$\underbrace{H_n \cdots H_2 H_1}_{Q^T} A = \begin{pmatrix} R \\ 0 \end{pmatrix} \Rightarrow A = QR$$

"Householder QR"

~~Givens~~

Note: <u>Computing Hx</u>:

$$Hx = \left(I - \frac{2vv^T}{v^Tv} \right) x = x - \left(\frac{2v^Tx}{v^Tv} \right) v$$

cheaper than matrix-vector mult.
only need to know <u>v</u>.

implementation:
- R stored in upper tri. portion of A
- <u>vectors</u>, v, stored in lower tri. portion of A
(+1 vector)

<u>See Example 3.5 in book</u>: Householder QR on 5×3
3.8 " : Gram-Schmidt

Rank Deficiency

- so far assumed A full column rank, if not...
- QR still exists, but
  - R singular
  - $A^T A$ singular
  - L.S. solution not unique
- can usually avoid this in experiment design
- otherwise, commonly...
- pick LS solution $\vec{x}$ with min. norm
  - QR w/ column pivoting, or
  - SVD

Column Pivoting

$\mathrm{rank}(A) = k < n$

$$Q^T A P = \begin{pmatrix} R & S \\ 0 & 0 \end{pmatrix}, \quad R \text{ } k \times k \text{, nonsingular}$$

solve

$$Ry = c, \quad c = \text{first } k \text{ comp. of } Q^T b$$

$$x = P \begin{pmatrix} y \\ 0 \end{pmatrix}$$

$$Q^T A P = \begin{pmatrix} R & S \\ 0 & 0 \end{pmatrix}$$

$$\min \| b - Ax \|_2$$

$$= \min \| \cancel{Q} Q^T b - Q^T A P \underbrace{P^T x}_{y} \|_2 = \circledast$$

$$Q^T A P y = \begin{pmatrix} R & S \\ 0 & 0 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = R y_1 + S y_2$$

$$\circledast = \left\| \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} - \begin{pmatrix} R & S \\ 0 & 0 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \right\|_2$$

$$y_2 = 0$$

$$R y_1 = b_1 \rightsquigarrow y_1$$

$$x = \cancel{R} P \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$$