

Householder's method for QR factorization

Thursday, March 13, 2025 5:43 PM

Refs: ch 4.3 of Olver et al.
(APPM 3310 textbook), and also
see Golub and van Loan

Example Usage / Motivation

note we could use it to solve a square linear system $A \cdot \vec{x} = \vec{b}$.

How? $A = QR$, so

$$Q \cdot R \cdot \vec{x} = \vec{b} \Leftrightarrow R \vec{x} = Q^{-1} \vec{b}$$

$$= Q^T \vec{b} \text{ via orthogonality of } Q$$

Solving $R \vec{x} = Q^T \vec{b}$ via back-substitution since R is upper triangular.

When should you do this? It's a little slower than LU factorization,
but slightly better if A is "ill-conditioned"
to be defined in a numerics class.

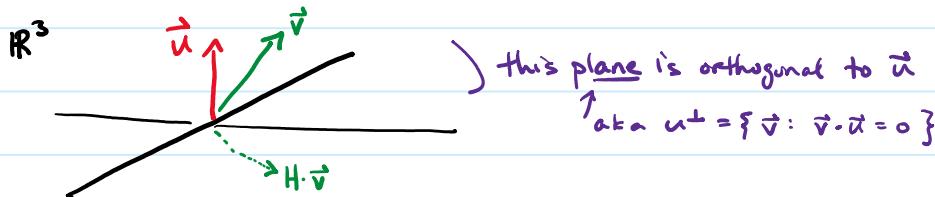
So, today: QR factorization of A if A is a bit ill-conditioned.

- Don't do Gram-Schmidt in this case
- Even modified Gram-Schmidt isn't great
- Book suggests "Householder's Method", which we'll discuss
 - ... but if A is ill-conditioned, this should be combined with a column pivoting strategy. i.e. re-order vectors

Householder Reflectors aka "elementary reflection matrix"

Given \vec{u} with $\|\vec{u}\| = 1$ (in Euclidean norm),

$H = I - 2 \cdot \vec{u} \vec{u}^T$ is a Householder reflection



$H \cdot \vec{v}$ reflects \vec{v} about the line/plane/hyperplane u^\perp

Note $H = H^T$ and

$$H \cdot H^T = (I - 2\vec{u}\vec{u}^T)(I - 2\vec{u}\vec{u}^T) = I - 4\vec{u}\vec{u}^T + 4\vec{u}\underbrace{\vec{u}^T\vec{u}}_{=1 \text{ since } \|\vec{u}\|=1}\vec{u}^T = I$$

... so H is orthogonal (and symmetric)

i.e. $H^{-1} = H$!

We can choose \vec{u} (hence H) to map a particular vector \vec{v} to a target output \vec{w}

as long as $\|\vec{v}\| = \|\vec{w}\|$:

Lemma 4.28: Let $\vec{v}, \vec{w} \in \mathbb{R}^n$ w/ $\|\vec{v}\| = \|\vec{w}\|$, then if $\vec{u} := \frac{\vec{v} - \vec{w}}{\|\vec{v} - \vec{w}\|}$ and $H = I - 2\vec{u}\vec{u}^T$,

then $H \cdot \vec{v} = \vec{w}$ (and since $H^2 = I$, $H\vec{w} = \vec{v}$)

proof: just calculate,

$$\begin{aligned} H \cdot \vec{v} &= (I - 2\vec{u}\vec{u}^T)\vec{v} = \vec{v} - \frac{2}{\|\vec{v} - \vec{w}\|^2} \cdot (\vec{v} - \vec{w}) \cdot (\vec{v} - \vec{w})^T \cdot \vec{v} \\ &= \vec{v} - \frac{2(\vec{v} - \vec{w})}{\|\vec{v}\|^2 - 2\langle \vec{v}, \vec{w} \rangle + \|\vec{w}\|^2} \cdot (\|\vec{v}\|^2 - \langle \vec{w}, \vec{v} \rangle) \\ &= \vec{v} - \frac{2(\|\vec{v}\|^2 - \langle \vec{w}, \vec{v} \rangle)}{2\|\vec{v}\|^2 - 2\langle \vec{v}, \vec{w} \rangle} \cdot (\vec{v} - \vec{w}) \quad \text{since } \|\vec{w}\|^2 = \|\vec{v}\|^2 \\ &= \vec{v} - (\vec{v} - \vec{w}) = \vec{w}. \quad \square \end{aligned}$$

Householder QR method

Goal is $A = QR$ or... $Q^T A = R$.

Recall (orthogonal matrix) \cdot (orthogonal matrix) = (orthogonal matrix)

so we'll build Q^T as the product of Householder reflectors

Step 1: $A = \begin{bmatrix} \bullet & \bullet & \bullet & \bullet \\ \hline \vec{v}_1 & & & \end{bmatrix}$

Choose \vec{u}_1 (hence H_1) so that

$$H_1 \cdot A = \begin{bmatrix} \bullet & \bullet & \bullet & \bullet \\ \hline \vec{0} & \vec{0} & \vec{0} & \vec{0} \end{bmatrix}$$

'Zero-out'

Then...

Choose \vec{u}_2 (hence H_2) so that

$$H_2 \cdot (H_1 \cdot A) = \begin{bmatrix} \bullet & \bullet & \bullet & \bullet \\ \hline \vec{0} & \vec{0} & \vec{0} & \vec{0} \end{bmatrix}$$

still 0 new 0

etc.

So for step 1, want to turn $\vec{v}_1 = \begin{bmatrix} \vdots \\ \vdots \\ \vdots \end{bmatrix}$ into $\vec{\omega}_1 = \begin{bmatrix} \vdots \\ 0 \\ 0 \\ 0 \end{bmatrix}$ anything.

So... define $\vec{\omega}_1 = \|\vec{v}_1\| \cdot \vec{e}_1$ has the right zeros, and $\|\vec{\omega}_1\| = \|\vec{v}_1\|$

hence use our lemma: (choose $\vec{u}_1 = \frac{\vec{v}_1 - \vec{\omega}_1}{\|\vec{v}_1 - \vec{\omega}_1\|}$ so that $H\vec{v}_1 = \vec{\omega}_1$)

$$\text{So, set } \vec{u}_1 = \frac{\vec{v}_1 - \|\vec{v}_1\| \cdot \vec{e}_1}{\|\vec{v}_1 - \|\vec{v}_1\| \cdot \vec{e}_1\|}$$

note: if numerator (hence also denominator) is zero, it means $\vec{v}_1 = c \cdot \vec{e}_1$ for some c , so in that case choose $\vec{u}_1 = \vec{0}$, i.e. $H = I$
i.e. $\vec{\omega}_1 := \vec{v}_1 - \|\vec{v}_1\| \cdot \vec{e}_1$, $\vec{u}_1 = \vec{\omega}_1 / \|\vec{\omega}_1\|$

thus

$$H_1 \cdot A = \begin{bmatrix} r_{11} & \bullet & \bullet & \bullet \\ 0 & \bullet & \bullet & \bullet \\ 0 & \bullet & \bullet & \bullet \\ 0 & \bullet & \bullet & \bullet \end{bmatrix} \quad \text{not the same as in original matrix}$$

Step 2 Want to turn \vec{v}_2 into $\begin{bmatrix} \vdots \\ 0 \\ 0 \\ 0 \end{bmatrix}$ and not ruin 1st column.
two ways to think of it:

Way #1: define $H_2 = \begin{bmatrix} 1 & 0 \\ 0 & I - 2\vec{u}_1\vec{u}_1^\top \end{bmatrix}$

(work in a smaller space)

$$\text{where } \vec{u}_2 = \frac{\vec{v}_2 - \|\vec{v}_2\| \cdot \vec{e}_1}{\|\vec{v}_2 - \|\vec{v}_2\| \cdot \vec{e}_1\|}$$

$$\vec{e}_1 \in \mathbb{R}^{n-1}$$

block matrix, i.e. at step K

\vec{V}_2 are rows 2 through n of 2nd column of $\underbrace{H_1 \cdot A}_{\text{not of } A}$

$$\begin{array}{|c|c|} \hline & 0 \\ \hline 0 & \downarrow \\ \hline \end{array} \quad \begin{array}{l} (K-1) \times (K-1) \text{ identity} \\ I - 2\vec{u}_k\vec{u}_k^\top \end{array}$$

$$U_k \in \mathbb{R}^{n-k+1}$$

Way #2: (equivalent) define $\tilde{V}_2 = \begin{bmatrix} 0 \\ \vec{v}_2 \end{bmatrix}$ append zeros (at step K, this is $K-1$ zeros), $H_2 = I - 2\vec{u}_2\vec{u}_2^\top$

(work in original space but append zeros)

$$\text{where } \vec{u}_2 = \frac{\tilde{V}_2 - \|\tilde{V}_2\| \cdot \vec{e}_2}{\|\tilde{V}_2 - \|\tilde{V}_2\| \cdot \vec{e}_2\|} \quad \vec{e}_2 \in \mathbb{R}^n \quad \text{i.e. } \vec{\omega}_2 = \tilde{V}_2 - \|\tilde{V}_2\| \cdot \vec{e}_2, \vec{u}_2 = \vec{\omega}_2 / \|\vec{\omega}_2\|$$

either way, we know have $H_2 \cdot (H_1 \cdot A) = \begin{bmatrix} r_{11} & r_{12} & \bullet & \bullet \\ 0 & r_{22} & \bullet & \bullet \\ 0 & 0 & \bullet & \bullet \\ 0 & 0 & \bullet & \bullet \end{bmatrix}$

Steps 3...n-1 follow same pattern

→ we don't need step "n" since nothing left to do then
 at the end of step $n-1$

we have $H_{n-1} H_{n-2} \cdots H_2 H_1 A = Q^T \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ r_{21} & & & \vdots \\ \vdots & & & r_{nn} \\ 0 & & & \end{bmatrix} = R$

so $Q = H_1^T H_2^T \cdots H_{n-1}^T$
 $= H_1 \cdot H_2 \cdots H_{n-1}$ since $H_i = H_i^T$

product of orthogonal matrices is orthogonal ✓

* numerically, we don't always need the matrix Q , just to be able to apply $Q^T \vec{v}$ for some \vec{v} , so in that case we use the $\{\vec{u}_i\}$ vectors for an efficient implementation. (Not important for our class)

Ex. 4.29 Do Householder QR on $A = \begin{bmatrix} 1 & 1 & 2 \\ 1 & 0 & -2 \\ -1 & 2 & 3 \end{bmatrix}$

Step 1: $\tilde{v}_1 = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}$, $\|\tilde{v}_1\| = \sqrt{3}$, $\tilde{w}_1 = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} - \frac{1}{\sqrt{3}} \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}$, $\tilde{u}_1 = \frac{\tilde{w}_1}{\|\tilde{w}_1\|}$ (messy)

$$H_1 = \begin{bmatrix} 1 & & \\ & 1 & \\ & & 1 \end{bmatrix} - 2 \cdot \tilde{u}_1 \tilde{u}_1^T \quad \text{so } H_1 A = \begin{bmatrix} 1.7 & -0.57 & -1.73 \\ 0 & 2.15 & 3.1 \\ 0 & -0.15 & -2.1 \end{bmatrix}$$

Comment:

To multiply $H_1 \cdot A$, do not form H_1 .

Instead: $(I - 2 \cdot \tilde{u}_1 \tilde{u}_1^T) \cdot A = A - 2 \cdot \tilde{u}_1 \cdot (\tilde{u}_1^T A)$

compute this first!

$$\boxed{\quad} - \boxed{\quad} = \boxed{\quad}$$

Step 2 $\tilde{v}_2 = \begin{bmatrix} 0 \\ 2.15 \\ -0.15 \end{bmatrix}$ $\|\tilde{v}_2\| = 2.16$, $\tilde{w}_2 = \begin{bmatrix} 0 \\ 2.15 \\ -0.15 \end{bmatrix} - 2.16 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$

$$\tilde{u}_2 = \tilde{w}_2 / \|\tilde{w}_2\|, H_2 = I - 2 \cdot \tilde{u}_2 \tilde{u}_2^T$$

so $H_2 H_1 A = \begin{bmatrix} 1.7 & -0.57 & -1.73 \\ 0 & 2.16 & 3.2 \\ 0 & 0 & 1.8 \end{bmatrix} = R$ Same as in $H_1 \cdot A$

$Q = H_1 H_2 = \begin{bmatrix} 0.57 & 0.61 & 0.53 \\ 0.57 & 0.15 & -0.80 \\ -0.57 & 0.77 & -0.26 \end{bmatrix}$ Done!

Some remarks on computational complexity

Square case $m = n$

$$n \begin{bmatrix} A \end{bmatrix} = \begin{bmatrix} Q \end{bmatrix} \begin{bmatrix} R \end{bmatrix}$$

Gram-Schmidt and Modified Gram-Schmidt

take $O(n^3)$ timeHouseholder takes $O(n^3)$ time to create R and all the reflection vectors $\vec{u}_1, \dots, \vec{u}_{n-1}$

$$Q = H_1 \cdot H_2 \cdot \dots \cdot H_{n-1}, \quad H_k = I - 2\vec{u}_k \vec{u}_k^\top$$

Creating Q from $\{\vec{u}_1, \dots, \vec{u}_{n-1}\}$ takes $O(n^3)$ flops tooTall case $m > n$

$$\textcircled{1} \quad m \begin{bmatrix} A \end{bmatrix} = m \begin{bmatrix} Q \end{bmatrix} \begin{bmatrix} R \end{bmatrix}$$

Gram-Schmidt doesn't compute this directly, you'd have to do postprocessing

Householder takes $O(mn^2)$ time to create R and all the reflection vectors $\vec{u}_1, \dots, \vec{u}_{n-1}$ Alternative: Keep Q implicitto multiply $Q \cdot \vec{b}$,

$$\text{use } Q \vec{b} = H_1 \cdot H_2 \cdot \dots \cdot H_{n-1} \cdot \vec{b}$$

$$\underbrace{\qquad}_{O(m)} \underbrace{\qquad}_{O(m)} \underbrace{\qquad}_{O(m)}$$

{ but creating Q from $\{\vec{u}_1, \dots, \vec{u}_{n-1}\}$ takes $O(m^2n)$ time. **Expensive!**Similarly for $Q^T \vec{b}$ { so $O(mn)$ flops and we didn't have to spend $O(m^2n)$ flops building Q explicitly

(2) "thin QR"

$$m \begin{bmatrix} A \end{bmatrix} = m \begin{bmatrix} \tilde{Q} \end{bmatrix} \begin{bmatrix} \tilde{R} \end{bmatrix}$$

Gram-Schmidt and Modified Gram-Schmidt

take $O(mn^2)$ timeHouseholder doesn't give \tilde{Q} but it gives $Q = [\tilde{Q} : \tilde{Q}^\top]$ from case (1)so if you want $\vec{y} = \tilde{Q}^T \vec{b}$

$$\text{just do } \vec{y} = Q^T \vec{b} = \begin{bmatrix} \tilde{Q}^T \vec{b} \\ \tilde{Q}^T \vec{b} \end{bmatrix} \text{ and } \vec{y} = \begin{bmatrix} \vec{y}_1 \\ \vec{y}_2 \end{bmatrix} \text{ so return } \vec{y}_1$$

so $O(mn)$ to compute $\tilde{Q}^T \vec{b}$ i.e. $\vec{y}_1 = y[1:n]$ plus $O(mn^2)$ from case (1)