Householder to get to Hessenberg, and Hermitian matrices

# Ch 9.4, 9.5: Householder's method

Thursday, March 13, 2025 5:43 PM

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

**Goal: compute Schur decomposition**

$A \in \mathbb{C}^{n\times n}$, $\exists$ unitary $U \in \mathbb{C}^{n\times n}$ ($U^{-1} = U^*$) and triangular $T \in \mathbb{C}^{n\times n}$ s.t.

$$A = UTU^*$$

Implication: $A \sim T$ so $\text{eigs}(A) = \text{eigs}(T)$, and $T$ is triangular so its eigs are the diagonal entries, thus trivial to find! (and if $A = A^*$ is Hermitian/self-adjoint or real symmetric, $T$ is diagonal and so columns of $U$ are eigenvectors)

But there's a catch! It's impossible to find the Schur decomp. for all $A$ using a direct (non-iterative) method

Why? There's no quintic formula: Abel's Thm says there's no equation (like the quadratic formula), involving $+, \times, \div, \sqrt{\ }$, that finds all roots of polynomials of degree 5 or higher [quartics: yes, Ferrari in 1540. Cubics: yes, Italians, 1500's, dramatic]

and, given $p(x) = c_0 + c_1 x + \ldots + c_{n-1}x^{n-1} + x^n$

define the companion matrix

$$C = \begin{bmatrix} 0 & & & & -c_0 \\ 1 & 0 & & & -c_1 \\ & 1 & \ddots & & \vdots \\ & & \ddots & 0 & -c_{n-2} \\ 0 & & & 1 & -c_{n-1} \end{bmatrix}$$

then the characteristic polynomial of $C$ is $p$.

So finding eigs of $C$ would find roots of $p$.

**Workaround:**

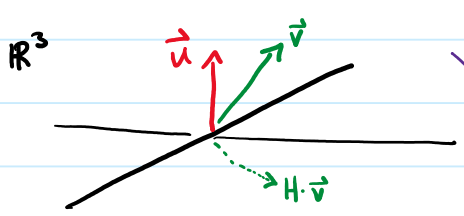
Use a direct method to make $A$ "nice", then do iterative method on this nicer version.

**Recall**

**Householder Reflectors** aka "elementary reflection matrix"

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

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




this plane is orthogonal to $\vec{u}$

aka $u^\perp = \{\vec{v} : \vec{v}\cdot\vec{u} = 0\}$

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

[Note: Burden and Faires discuss Householder in a very confusing way! Don't follow them]

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 $\|\vec{v}\| = \|\vec{w}\|$ : $\to$ in our

**Lemma 4.28**: Let $\vec{v}, \vec{w} \in \mathbb{R}^n$ w/ $\|\vec{v}\| = \|\vec{w}\|$, then if $\vec{u} := \dfrac{\vec{v} - \vec{w}}{\|\vec{v} - \vec{w}\|}$ 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,

$$H \cdot \vec{v} = (I - 2\vec{u}\vec{u}^T)\vec{v} = \vec{v} - \frac{2}{\|\vec{v}-\vec{w}\|^2} \cdot (\vec{v} - \vec{w}) \cdot (\vec{v} - \vec{w})^T \vec{v}$$

$$= \vec{v} - \frac{2 \, (\vec{v} - \vec{w})}{\|\vec{v}\|^2 - 2\langle \vec{v}, \vec{w} \rangle + \|\vec{w}\|^2} \cdot (\|\vec{v}\|^2 - \langle \vec{w}, \vec{v} \rangle)$$

$$= \vec{v} - \frac{2(\|\vec{v}\|^2 - \langle \vec{w}, \vec{v} \rangle)}{2\|\vec{v}\|^2 - 2\langle \vec{v}, \vec{w} \rangle} \cdot (\vec{v} - \vec{w}) \qquad \text{since } \|\vec{w}\|^2 = \|\vec{v}\|^2$$

$$= \vec{v} - (\vec{v} - \vec{w}) = \vec{w}. \qquad \square$$

## We already saw an application of Householder for QR

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**: $A = \begin{bmatrix} \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet \end{bmatrix}$ $\vec{v}_1$

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

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

zero-out

Then ...

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

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

still 0 new 0

etc.

Etc

Tuesday, December 2, 2025 9:15 AM

Could we use Householder matrices to find the Schur decomp.?

Afterall, we get a triangular matrix via QR

No. QR does $A = \underset{\text{orthogonal}: \checkmark}{Q}\ \underset{\text{triangular}: \checkmark}{R}$ $\Big\}$ A and R aren't similar, $\text{eigs}(A) \neq \text{eigs}(R)$

But we want $A \sim T$, i.e. $A = QTQ^T$. Not the same

Can we fix it?

We did: $H_1 \cdot A = \begin{bmatrix} x & \cdots \\ 0 & x & \cdots \\ \vdots \\ 0 & x & \cdots \end{bmatrix}$

what if we look at

$$H_1 A H_1^T$$

Let's try it:

$$A = \begin{bmatrix} 1 & 3 & 6 \\ 2 & 4 & 7 \\ 2 & 5 & 8 \end{bmatrix} \qquad \vec{v} = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}, \quad \|\vec{v}\|_2 = \sqrt{1+4+4} = 3$$

target: $\vec{w} = 3 \cdot \vec{e}_1 = \begin{bmatrix} 3 \\ 0 \\ 0 \end{bmatrix}$

So $u = \dfrac{\vec{v}-\vec{w}}{\|\vec{v}-\vec{w}\|_2} = \begin{bmatrix} -2 \\ 2 \\ 2 \end{bmatrix} \cdot \dfrac{1}{\sqrt{12}}$ and $H_1 = I - 2\vec{u}\vec{u}^T$

$$H_1 \cdot A = A - \frac{2}{12}\begin{bmatrix} -2 \\ 2 \\ 2 \end{bmatrix}\begin{bmatrix} -2 & 2 & 2 \end{bmatrix}\begin{bmatrix} 1 & 3 & 6 \\ 2 & 4 & 7 \\ 2 & 5 & 8 \end{bmatrix} = \begin{bmatrix} 3 & 7 & 12 \\ 0 & 0 & 1 \\ 0 & 1 & 2 \end{bmatrix} \checkmark$$

I did on a computer to avoid mistakes

$$H_1 A H_1^T = \begin{bmatrix} 13\frac{2}{3} & -3\frac{2}{3} & 1\frac{1}{3} \\ \boxed{\begin{matrix} \frac{2}{3} \\ 2 \end{matrix}} & -\frac{2}{3} & \frac{1}{3} \\ & -1 & 0 \end{bmatrix}$$

We lost our zero pattern :(

We knew it was impossible via Abel's thm.

Tuesday, December 2, 2025 9:31 AM

The next best thing after triangular is Hessenberg

$$\begin{bmatrix} x & x & x & x \\ 0 & x & x & x \\ 0 & 0 & x & x \\ 0 & 0 & 0 & x \end{bmatrix} \qquad \begin{bmatrix} x & x & x & x \\ x & x & x & x \\ 0 & x & x & x \\ 0 & 0 & x & x \end{bmatrix}$$

"x" means anything

upper triangular — upper Hessenberg

1st off-diagonal below main diagonal is allowed to be nonzero

**Idea** Set $H_1 = \left[\begin{array}{c|c} 1 & 0 \\ \hline 0 & I_{n-1} - 2\vec{u}\vec{u}^T \end{array}\right]$ or, equivalently, $H_1 = I_n - 2\cdot\vec{u}\vec{u}^T$ and require $\vec{u} = \begin{bmatrix} 0 \\ \vdots \end{bmatrix}$ ← 0 in 1st entry

$$A = \begin{bmatrix} 1 & 3 & 6 \\ 2 & 4 & 7 \\ 2 & 5 & 8 \end{bmatrix}$$

$\vec{v} = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$, $\|\vec{v}\| = \sqrt{8}$, target is $\vec{w} = \begin{bmatrix} \sqrt{8} \\ 0 \end{bmatrix}$

so $\vec{u} = \dfrac{\vec{v}-\vec{w}}{\|\vec{v}-\vec{w}\|} = \begin{bmatrix} 2-\sqrt{8} \\ 2 \end{bmatrix} \cdot \dfrac{1}{\sqrt{(2-\sqrt{8})^2+4}} = \overbrace{\dfrac{1}{\sqrt{16-2\sqrt{8}}}}^{c} \cdot \begin{bmatrix} 2-\sqrt{8} \\ 2 \end{bmatrix}$

$4-2\sqrt{8}+8+4 = 16-2\sqrt{8}$

$c \cong \dfrac{1}{2.16} \cong .462$

Define $H_1 = I_3 - 2\cdot c^2 \begin{bmatrix} 0 \\ \vec{u} \end{bmatrix} \cdot [0 \ \ \vec{u}^T]$

$$= I_3 - 2c^2 \begin{bmatrix} 0 \\ 2-\sqrt{8} \\ 2 \end{bmatrix} [0 \quad 2-\sqrt{8} \quad 2]$$

so $H_1 \cdot A = \begin{bmatrix} 1 & 3 & 6 \\ 2.82 & 6.36 & 10.61 \\ 0 & -.71 & -.71 \end{bmatrix}$

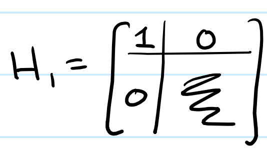
↖ as desired. so far, so good

what about $H_1 A H_1^T$? ($H_1^T = H_1$)

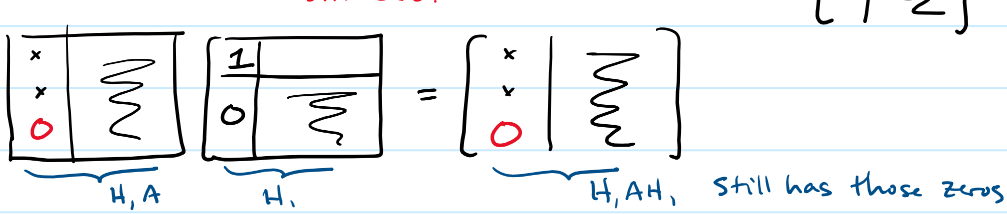
$$H_1 A H_1 = \begin{bmatrix} 1 & 6.36 & -2.12 \\ 2.82 & 12 & -3 \\ 0 & -1 & 0 \end{bmatrix}$$

↖ still zero!

Good! we didn't ruin our structure




structure was key



Now just continue:

$A_1 = H_1 A H_1$ (so $A_1 \sim A$, they have same eigs)

$$\underbrace{\begin{bmatrix} \times & \times & \times & \times \\ \times & \times & \times & \times \\ 0 & \boxed{\times} & \times & \times \\ 0 & \boxed{\times} & \times & \times \end{bmatrix}}_{A_1}$$

source $\vec{v}$, target is $\vec{w} = \|\vec{v}\| \cdot \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $\vec{u} = \frac{\vec{v} - \vec{w}}{\|\vec{v} - \vec{w}\|}$ ← or $\vec{w} - \vec{v}$, choose whichever is more stable, see book

$$H_2 = \left[\begin{array}{c|c} I_2 & 0 \\ \hline 0 & I_{n-2} - 2 \cdot \vec{u}\vec{u}' \end{array}\right]$$

so $H_2 A_1 = \begin{bmatrix} \times & \times & \times & \times \\ \times & \times & \times & \times \\ 0 & \times & \times & \times \\ 0 & 0 & \times & \times \end{bmatrix}$ and $H_2 A_1 H_2 = \begin{bmatrix} \times & \times & \times & \times \\ \times & \times & \times & \times \\ 0 & \times & \times & \times \\ 0 & 0 & \times & \times \end{bmatrix}$

Still have zeros due to block structure of $H_2$

... do this $(n-2)$ times to get $A_{n-2} = \underbrace{H_{n-2} \cdot H_{n-1} \cdot \ldots \cdot H_1}_{} \cdot A \cdot H_1 \cdot H_2 \cdot \ldots \cdot H_{n-2}$

↑ orthogonal/unitary

product of orthog./unitary matrices is still orthog./unitary: if $H_1^{-1} = H_1^T$, $H_2^{-1} = H_2^T$

then $(H_1 H_2)^{-1} = H_2^{-1} \cdot H_1^{-1} = H_2^T \cdot H_1^T = (H_1 H_2)^T$ ✓

Summarize:

**Householder QR**

$R = \underbrace{H_{n-1} \cdot H_{n-2} \cdot \ldots \cdot H_1}_{Q^T} \cdot A$

↑ upper triangular

$A = Q R$

**Householder Hessenberg**

$S = \underbrace{H_{n-2} \cdot H_{n-3} \cdot \ldots \cdot H_1}_{Q^T} A \cdot \underbrace{H_1 \cdot \ldots \cdot H_{n-3} \cdot H_{n-2}}_{Q}$

↑ upper Hessenberg

$A = Q S Q^T$

so $A \sim S$, have same eigenvalues

**Observation**: if $A = A^T$ then $QSQ^T = (QSQ^T)^T = Q S^T Q^T \Rightarrow S = S^T \Rightarrow$ <u>$S$ is tridiagonal</u>

↑ our book focuses on $A = A^T$ case, so it focuses on $S$ tridiag.

- For a symmetric (or Hermitian) matrix, we can't "diagonalize" it with a direct method, but we <u>can</u> "tridiagonalize" it.

Tuesday, December 2, 2025 4:19 PM

What good is a Hessenberg (or tridiag.) matrix?

Suppose we do the QR algorithm

Assume $A$ is Hessenberg or tridiagonal.

⚠ "QR iteration" isn't the same as the "QR factorization / decomposition"

$\boxed{\text{QR iteration}}$: $A_1 = A$ (or "S" in our old notation)

for $k = 1, 2, \ldots$

$Q_k R_k = A_k$ QR factorization

$A_{k+1} = R_k Q_k$ multiply

You can show (we won't) that if $A_1$ is upper Hessenberg, then all $A_k$ ($k=1,2,3,\ldots$) stay upper Hessenberg

Why is that helpful?

QR of a $n \times n$ matrix is $O(n^3)$

but QR of an upper Hessenberg matrix is $O(n)$

(using Givens rotations, see book for details)

So,

QR iteration on an arbitrary square matrix costs $\#\text{iterations} \cdot O(n^3)$

The Householder method does:

1. $O(n^3)$ to reduce $A$ to upper Hessenberg
2. then QR iteration cost is $\#\text{iterations} \cdot O(n^3)$.

Much faster!