

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.

$$A = U T U^*$$

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 method \downarrow non-iterative

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 $p(x) = c_0 + c_1 x + \dots + c_{n-1} x^{n-1} + x^n$

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",

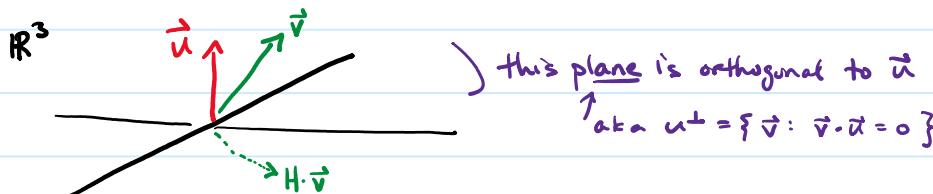
then do iterative method on this nicer version. Direct method:

Recall

Householder Reflectors aka "elementary reflection matrix"

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

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



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

Terminates in a finite, predetermined number of iterations. In exact arithmetic, this is the exact answer. Ex: Gaussian Elim., Quadratic Formula

Iterative Method:

Needs to iterate. More accuracy requires more iterations.

Ex: Newton's Method

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 $\|\vec{v}\| = \|\vec{w}\|$: → in over

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,

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

We already saw an application of Householder for QR

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

Recall (orthogonal matrix) · (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.

Etc

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

Afterall, we get a triangular matrix via QR

No. QR does $A = Q R \leftarrow$ triangular ✓ } A and R aren't similar,
 $\text{eigs}(A) \neq \text{eigs}(R)$

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

Can we fix it?

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

what if we look at

$$H_1 A H_1^T$$

Let's try it:

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

$$\text{so } u = \frac{\vec{v} - \vec{w}}{\|\vec{v} - \vec{w}\|_2} = \begin{bmatrix} -2 \\ 2 \\ 2 \end{bmatrix} \cdot \frac{1}{\sqrt{12}} \quad \text{and } H = I - 2\vec{u}\vec{u}^T$$

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

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

¹ I did on a computer to avoid mistakes
 We lost our zero pattern ;)

We knew it was impossible via Abel's thm.

The next best thing after triangular is Hessenberg

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

"x" means anything
upper triangular

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

upper Hessenberg

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

Idea Set $H_1 = \begin{bmatrix} 1 & 0 & \dots & 0 \\ \vdots & \ddots & & 0 \\ 0 & I_{n-1} - 2\vec{u}\vec{u}^T \end{bmatrix}$ or, equivalently, $H_1 = I_n - 2\vec{u}\vec{u}^T$
and require $\vec{u} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$ in 1st entry

$$A = \begin{bmatrix} 1 & 3 & 6 \\ 2 & 4 & 7 \\ 2 & 5 & 8 \end{bmatrix} \quad \vec{v} = \begin{bmatrix} 2 \\ 2 \end{bmatrix}, \|\vec{v}\| = \sqrt{8}, \text{ target is } \vec{w} = \begin{bmatrix} \sqrt{8} \\ 0 \end{bmatrix}$$

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

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

$$c \approx \frac{1}{2.16} \approx .462$$

$$\begin{aligned} \text{Define } H_1 &= I_3 - 2 \cdot c^2 \begin{bmatrix} 0 \\ \vec{u} \end{bmatrix} \cdot [0 \ \vec{u}^T] \\ &= I_3 - 2c^2 \begin{bmatrix} 0 \\ 2 - \sqrt{8} \\ 2 \end{bmatrix} [0 \ 2 - \sqrt{8} \ z] \end{aligned}$$

$$\text{so } H_1 \cdot A = \begin{bmatrix} 1 & 3 & 6 \\ 2.82 & 6.36 & 10.61 \\ 0 & -0.71 & -0.71 \end{bmatrix}$$

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

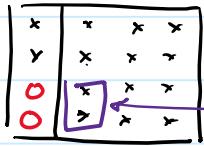
$$H_1 = \begin{bmatrix} 1 & 0 \\ 0 & \Xi \end{bmatrix} \quad \text{Structure was key}$$

$$\begin{bmatrix} x & \Xi \\ x & \Xi \\ 0 & \Xi \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & \Xi \end{bmatrix} = \begin{bmatrix} x & \Xi \\ x & \Xi \\ 0 & \Xi \end{bmatrix}$$

$H_1 A H_1$, still has those zeros

Now just continue:

$$A_1 = H_1 A H_1 \quad (\text{so } A_1 \sim A, \text{ they have same eigs})$$



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

$$H_2 = \begin{bmatrix} I_2 & & & \\ & \vdots & & 0 \\ & \cdots & \cdots & \cdots \\ 0 & & I_{n-2} - 2\vec{u}\vec{u}' & \end{bmatrix}$$

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

Still have zeros due to block structure of H_2

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

\uparrow orthogonal/unitary
product of orthog./unitary matrices is still orthog./unitary: if $H_i^{-1} = H_i^\top$
 $H_2^{-1} = H_2^\top$

then

$$(H_1 H_2)^{-1} = H_2^{-1} \cdot H_1^{-1} = H_2^\top \cdot H_1^\top = (H_1 H_2)^\top \checkmark$$

Summarize:

Householder QR

$$R = H_{n-1} \cdot H_{n-2} \cdot \dots \cdot H_1 \cdot A$$

\uparrow
upper triangular

$$A = Q R$$

Householder Hessenberg

$$S = H_{n-2} \cdot H_{n-3} \cdot \dots \cdot H_1 \cdot A \cdot H_1 \cdot \dots \cdot H_{n-3} \cdot H_{n-2}$$

\uparrow
upper Hessenberg

$$A = Q S Q^\top$$

so $A \sim S$, have same eigenvalues

Observation: if $A = A^\top$ then $Q S Q^\top = (Q S Q^\top)^\top = Q S^\top Q^\top \Rightarrow S = S^\top \Rightarrow S$ is tridiagonal

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

\uparrow
our book focuses on $A = A^\top$ case, so it focuses on S tridiag.

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

Suppose we do the QR algorithm

Assume A is Hessenberg or tridiagonal.

QR iteration: $A_1 = A$ (or "S" in our old notation)
for $k = 1, 2, \dots$

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

$$Q_k R_k = A_k \quad \text{QR factorization}$$

$$A_{k+1} = R_k Q_k \quad \text{multiply}$$

You can show (we want) that if A_k is upper Hessenberg, then

all A_k ($k=1, 2, 3, \dots$) 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 $O(n^2)$

(using Givens rotations, see book for details)

So,

QR iteration on an arbitrary square matrix costs #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 #iterations $\cdot O(n^2)$.
Much faster!