§ 9.3 Reduction to Hessenberg or tridiagonal form

<u>Motivation</u>: We would like to somehow compute Schur facto of A by successive multiplications by orthogonal matrices: $(AQ = QT)$

$$T_j = Q_j^* Q_{j-1}^* \cdots Q_2^* Q_1^* A Q_1 Q_2 \cdots Q_j$$

so that $T_j \to T = (\nabla)$ as $j \to \infty$.

Usually this is split into 2 phases:

$$\underset{A \neq A^*}{\begin{bmatrix} \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \end{bmatrix}} \xrightarrow[\text{form}]{\substack{①\\ \text{Reduction to}\\ \text{Hessenberg}}} \underset{H = (\nabla)}{\begin{bmatrix} \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ & \times & \times & \times & \times \\ & & \times & \times & \times \\ & & & \times & \times \end{bmatrix}} \xrightarrow{②} \underset{T = (\nabla)}{\begin{bmatrix} \times & \times & \times & \times & \times \\ & \times & \times & \times & \times \\ & & \times & \times & \times \\ & & & \times & \times \\ & & & & \times \end{bmatrix}}$$

$O(m^3)$ $\qquad$ $O(m^3)$ (usually)

② could run forever, however convergence is usually obtained in m steps and each step in ② takes $O(m^2)$ flops.

If $A = A^*$ (A Hermitian)

$$\begin{bmatrix} \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \end{bmatrix} \xrightarrow[\substack{\text{tridiag.}\\ \text{form}\\ O(m^3)}]{\substack{①\\ \text{Reduction to}}} \begin{bmatrix} \times & \times & & & \\ \times & \times & \times & & \\ & \times & \times & \times & \\ & & \times & \times & \times \\ & & & \times & \times \end{bmatrix} \xrightarrow[O(m^2)]{②} \begin{bmatrix} \times & & & & \\ & \times & & & \\ & & \times & & \\ & & & \times & \\ & & & & \times \end{bmatrix}$$

② each step takes m flops (if no eigenvectors are needed)
$\sim m$ steps total $= O(m^2$ flops$)$

In Burden & Faires: ① = § 9.3 Householder's Method
② = § 9.4 The QR algorithm

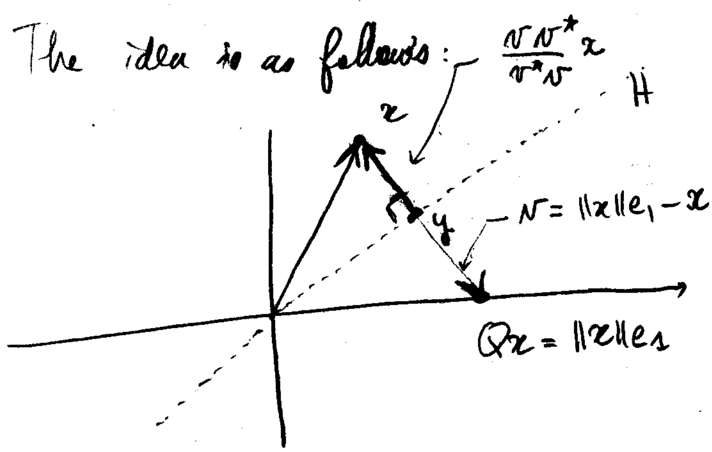
Phase ① can be made using orthogonal transformations called <u>Householder reflectors</u> (which are extremely useful in linear algebra ...)

Householder reflectors are orthogonal transformations that zero out part of a vector.

Given $x = \begin{bmatrix} x \\ x \\ x \\ \vdots \\ x \end{bmatrix}$ we can build a Q with $Q^*Q = I$ s.t.

$$Qx = \begin{bmatrix} \|x\| \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \|x\| e_1.$$

The idea is as follows:



The $\perp$ proj of x onto v is:

$$\frac{vv^*}{v^*v}x$$

Thus:

$$y = x - \frac{vv^*}{v^*v}x$$

and:

$$\boxed{Qx = x - 2\frac{vv^*}{v^*v}x}$$

Another geometric interpretation:

y is projection $\perp$ of x onto H

Qx is reflection of x accross H.

It's amazing that a simple factor of 2 converts a projector (not full rank) into an orthogonal matrix!!

$$\boxed{Q = I - 2\frac{vv^*}{v^*v}}$$

Q is orthogonal:

$$Q^*Q = \left(I - \frac{2vv^*}{v^*v}\right)\left(I - \frac{2vv^*}{v^*v}\right)$$

$$= I - 4\frac{vv^*}{v^*v} + 4\left(\frac{vv^*}{v^*v}\right)\left(\frac{vv^*}{v^*v}\right) = I.$$

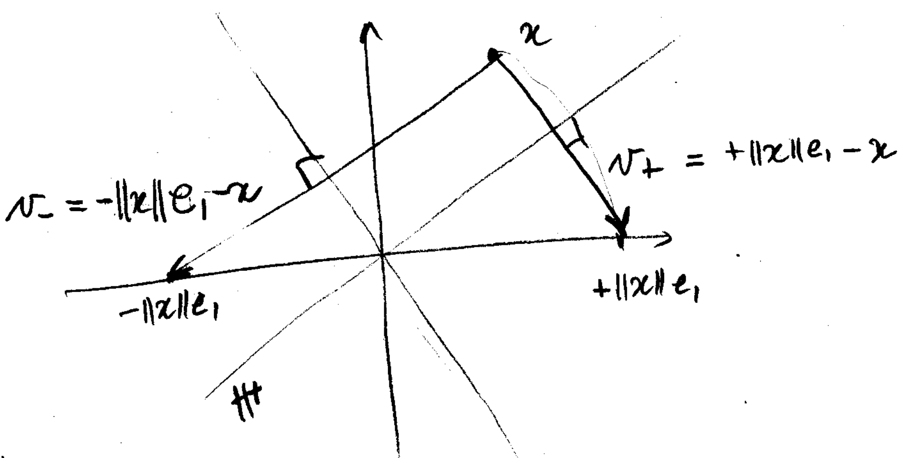
More generally we can introduce 0 anywhere in the vector.
A useful pattern is:

$$Q_k = \begin{array}{c} \overset{m}{\leftrightarrow} \\ \begin{bmatrix} I & 0 \\ 0 & Q \end{bmatrix} \end{array}, \text{ where } Q \text{ is a Householder refl.}$$

Q_k will leave untouched the first m entries:

$$x = \begin{bmatrix} x \\ x \\ x \\ x \\ x \\ x \\ x \end{bmatrix} \begin{matrix} \}m \\ \\ \}n-m \end{matrix} \longrightarrow Q_k x = \begin{bmatrix} x \\ x \\ x \\ x \\ 0 \\ \vdots \\ 0 \end{bmatrix} \} \text{unchanged}$$

There is a slight detail with HH reflectors: we could have chosen Q s.t. $Q_- x = -\|x\| e_1$, so which HH refl is better:
$Q_+ x = +\|x\| e_1$ or the former?



To avoid numerical cancellation in first component one usually takes $v = -\mathrm{sgn}(x_1)\|x\|e_1 - x$

But since $Q = I - 2\dfrac{vv^*}{v^*v}$ we can eliminate the $-$ sign:

$$\boxed{v = \mathrm{sgn}(x_1)\|x\|e_1 - x}.$$

<u>Note</u>
- each application of HH refl Q (or Q^*) is $O(n^2)$.
- we don't need to store Q, only v. Therefore successive applications of different HH refl $Q_1, Q_2 \dots Q_j$ can be stored with j vectors only.

Back to original problem:

$$\underset{A}{\begin{bmatrix} \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \end{bmatrix}} \longrightarrow \underset{H}{\begin{bmatrix} \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ & \times & \times & \times & \times \\ & & \times & \times & \times \\ & & & \times & \times \end{bmatrix}}$$

how do we introduce zeros?

A first idea:

$$Q_1^* A = \begin{bmatrix} \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times \end{bmatrix} \xrightarrow{\times Q_1} \overset{Q_1^* A Q_1}{\begin{bmatrix} \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \end{bmatrix}}$$

each col is lin comb of columns of $Q_1^* A$ so zeros disappear.

does not work!

A better idea: be less ambitious:

leave unchanged

$$\underset{A}{\begin{bmatrix} \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \end{bmatrix}} \xrightarrow{Q_1^* \cdot} \underset{Q_1^* A}{\begin{bmatrix} \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times \end{bmatrix}} \xrightarrow{\cdot Q_1} \underset{Q_1^* A Q_1}{\begin{bmatrix} \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times \end{bmatrix}}$$

$$\xrightarrow{Q_2^* \cdot} \underset{Q_2^* Q_1^* A Q_1}{\begin{bmatrix} \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ & \times & \times & \times & \times \\ & 0 & \times & \times & \times \\ & 0 & \times & \times & \times \end{bmatrix}} \xrightarrow{\cdot Q_2} \begin{bmatrix} \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ & \times & \times & \times & \times \\ & & \times & \times & \times \\ & & \times & \times & \times \end{bmatrix}$$

unchanged

etc...

at the end we get (after $n-2$ steps)

$$H = \begin{bmatrix} \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ & \times & \times & \times & \times \\ & & \times & \times & \times \\ & & & \times & \times \end{bmatrix} = Q_{n-2}^* Q_{n-1}^* \cdots Q_1^* A Q_1 Q_2 \cdots Q_{n-1}$$

Householder reduction to Hessenberg form

for $k = 1 \ldots m-2$

$$x = A(k+1:m,\, k)$$

$$v_k = \operatorname{sign}(x(1))\, \|x\|_2\, e_1 + x$$

$$v_k = v_k / \|v_k\|_2$$

$$A(k+1:m,\, k:m) = A(k+1:m,\, k:m) - 2 v_k \left(v_k^* A(k+1:m,\, k:m)\right)$$

$$A(1:m,\, k+1:m) = A(1:m,\, k+1:m) - 2\left(A(1:m,\, k+1:m)\, v_k\right) v_k^*$$

When $A = A^*$ Hermitian, we can do similar procedure.

$Q_1^* A Q_1$ will also be symmetric so any zeros we introduced in columns of A will also appear as zeros in corresp. row.

Applying HH refl on both sides has exactly same cost, so reduction to tridiagonal form is relatively cheap.