28.2 As we saw, the reduction to Hessenberg form (via Givens rotations or Householder reflections) by unitary similarity transformations becomes, for a real, symmetric matrix, a reduction to ~~triangular form~~ tridiagonal form. This happens because symmetry is preserved by the unitary similarity transformation:

$$QAQ^* = H \to$$

$$\Rightarrow (Q^*)^* A^T Q^* = H^* -$$

$$\Rightarrow QAQ^* = H^* \Rightarrow H^* = H.$$

But, naturally, a symmetric, Hessenberg matrix is tridiagonal.

(b) Consider now the QR iteration for a symmetric, tridiagonal matrix.

We have $H \equiv H_0 = Q_1 R_1$

$$H_1 = R_1 Q_1$$

Clearly $Q_1^* H_0 = R_1 = H_1 Q_1^*$

$$\Rightarrow Q_1^* H_0 Q_1 = H_1$$

We already know that the QR iteration preserves the Hessenberg structure (shown in class – the book focuses entirely on symmetric matrices!). As above, since H_1 is symmetric and Hessenberg, it must be tridiagonal.

— • —

(a) Finally following the discussion in class, which ~~discussed~~ gave the construction of the QR algorithm for general matrices (after their reduction to Hessenberg form) in terms of Givens rotations, we recall that Q is also Hessenberg (and, in fact, ~~Q~~ has full upper triangle even if ~~A~~ H is tridiagonal), while R is upper triangular.

— • —

(c) Consider the first two steps of the QR factorization of a tridiagonal matrix:

$$\begin{pmatrix} \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} & \\ & I \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & 0 & 0 & 0 \\ a_{21} & a_{22} & a_{23} & 0 & 0 \\ 0 & a_{32} & a_{33} & a_{34} & 0 \\ 0 & 0 & * & * & * \\ & & \ddots & \ddots & \ddots \end{pmatrix} - \text{where}$$

$$= \begin{pmatrix} r_{11} & r_{12} & r_{13} & 0 & 0 \\ 0 & A_{22} & A_{23} & 0 & 0 \\ 0 & a_{32} & a_{33} & a_{34} & 0 \\ 0 & 0 & * & * & * \\ & & \ddots & \ddots & \ddots \end{pmatrix}$$

where: $r_{11} = \sqrt{a_{11}^2 + a_{21}^2}$, $\cos\theta = \frac{a_{11}}{r_{11}}$, $\sin\theta = \frac{a_{21}}{r_{11}}$

etc.

Notice that one more superdiagonal is created by the procedure!

It is easily seen that this happens at each subsequent step as well, so that <u>R is upper triangular with only three nonzero diagonals</u>

Each step requires $\begin{cases} 10 \text{ multiplies} \\ + 5 \text{ adds} \end{cases}$ for R (+ 2 divs, 2 squares, 1 add, 1 sqrt) to a total of $21n$ ops!

Set XII - Solutions

33.1 $\mathcal{K}_n = \langle b, Ab, A^2b, \dots, A^{n-1}b\rangle = \langle q_1, \dots, q_n\rangle \subseteq \mathbb{C}^m$

while $K_n = [b \,|\, Ab \,|\, \dots \,|\, A^{n-1}b] = Q_n R_n$.

Note that there is an obvious misprint in the problem statement: we must have either $x \in \mathcal{K}_n$ or $x \in R(K_n)$. Either way, $\exists\, c_i$:

$$x \in \sum_{i=1}^{n} c_i (A^{i-1}b) = \left(\sum_{i=1}^{n} c_i A^{i-1}\right) b =: p(A)\, b.$$

33.2 The Arnoldi iteration at step n gives

$$Aq_n = h_{1n} q_1 + \dots + h_{nn} q_n + h_{n+1,n} q_{n+1} \qquad \text{(eq. 33.4)}$$

If $h_{n+1,n} = 0$, then the iteration "closes":

$$Aq_n = \sum_{i=1}^{n} h_{in} q_i \Rightarrow Aq_n \in \langle q_1, \dots, q_n\rangle$$

Equivalently, since $\mathcal{K}_n = \langle b, Ab, \dots, A^{n-1}b\rangle = \langle q_1, \dots, q_n\rangle$

then $A^n b = A\left(\sum_{i=1}^{n} c_i q_i\right) = \sum_{i=1}^{n-1} c_i (Aq_i) + c_n A q_n \in \mathcal{K}_n$,

and $\mathcal{K}_{n+1} = \langle b, \dots, A^{n-1}b, A^n b\rangle = \langle b, \dots, A^{n-1}b\rangle = \mathcal{K}_n$.

(a) When $h_{n+1,n} = 0$, equation (33.13) simplifies to

$$AQ_n = Q_n H_n \quad (H_n \text{ as defined in eq. } 33.9)$$

If the full Hessenberg form contains a zero subdiagonal entry, $h_{n+1,n} = 0$, then the eigenvalue problem for the matrix A is reduced to eigenvalue problems for two smaller matrices, of sizes $(n \times n)$ and $(m-n) \times (m-n)$:

$$H = \begin{pmatrix} \underbrace{\begin{bmatrix} * & \cdots & * \\ * & \ddots & \vdots \\ 0 & * & * \end{bmatrix}}_{B_1} & \begin{bmatrix} * & \cdots & * \\ \vdots & & \vdots \\ * & \cdots & * \end{bmatrix} \\ 0 & \underbrace{\begin{bmatrix} * & \cdots & * \\ * & \ddots & \vdots \\ & * & * \end{bmatrix}}_{B_2} \end{pmatrix}$$

Clearly, B_1 can be diagonalized separately (or triangularized for Schur form) by using similarity transforms with matrices

of the form $S_1 = {}_n\begin{pmatrix} \begin{bmatrix} * & \cdots & * \\ \vdots & & \vdots \\ * & \cdots & * \end{bmatrix} & 0 \\ 0 & \begin{matrix} 1 & & \\ & \ddots & \\ & & 1 \end{matrix} \end{pmatrix}$

and analogous result holds for B_2.

(b) Let $y \in \mathcal{K}_n \Rightarrow y = \sum_{i=1}^{n} c_i q_i$ or $y = \sum_{i=1}^{n} b_i (A^{i-1} b)$

Can use either form:

(i) $Ay = \left(\sum_{i=1}^{n-1} c_i A q_i \right) + c_n A q_n$

But $Aq_i = Q_n h_{*,i}$, $i = 1, n-1$ and $A q_n = Q_n h_{*,n}$

so $Ay = Q_n H_n \hat{c} \in \mathcal{K}_n$

or (ii) $Ay = \sum_{i=1}^{n-1} c_i A^{i+1} b + c_n A^n b \in \mathcal{K}_n$ since $A^n b \in \mathcal{K}_n$

(the $A^i b \in \mathcal{K}_n$, $i = 1, .., n-1$ by definition of $\mathcal{K}_n$).

(c) Obvious: since $A^n b \in \mathcal{K}_n \Rightarrow$

$$A^n b = \sum_{i=1}^{n}{}' d_i A^{i-1} b = p_{n-1}^{(0)}(A)\, b$$

Then $A^{n+1} b = \sum_{i=1}^{n-1}{}' d_i A^i b + d_n A^n b = p_{n-1}^{(0)}(A) b + q_{n-1}(A) b =: p_{n-1}^{(1)}(A) b$

and $A^{n+1} b \in \mathcal{K}_n \Rightarrow \mathcal{K}_{n+2} = \langle b, \dots, A^n b, A^{n+1} b \rangle = \langle b, A^{n-1} b \rangle = \mathcal{K}_n$ and the same is true for any $\mathcal{K}_\ell$, $\ell > n$.

(d) This was already shown in (a), as part of the discussion of the structure of the Hessenberg reduction.

(e) If A is nonsingular then $\cancel{\{Aq_i\}_{i=1}^n}$ are ~~independent:~~ $\cancel{\sum_i^n c_i A q_i = 0 \Rightarrow}$ the n vectors $\{Ab, A(Ab), \dots, A(A^{n-1}b)\}$ are independent: indeed

$$\sum_{i=1}^n c_i A^i b = A\left\{\sum_{i=1}^n c_i A^{i-1} b\right\} = 0 \xRightarrow{(A \text{ non-singular})} \sum_{i=1}^n c_i A^{i-1} b = 0$$

But then $c_i = 0$, $i = 1, \dots, n$ since the vectors $\{A^{i-1} b\}_{i=1}^{n-1}$ are independent. Since $\mathcal{K}_n$ invariant, the vectors $\{A^i b\}_{i=1}^n$ are all in $\mathcal{K}_n$, and thus they form a basis.

Let $b = \sum_{i=1}^n d_i A^i b \Rightarrow A x = \sum_{i=1}^n d_i A^i b$

$$\Rightarrow x = \sum_{i=1}^n d_i A^{i-1} b \in \mathcal{K}_n.$$

34.1 $A \in \mathbb{C}^{m\times m}$, $b \in \mathbb{C}^m$, $p \in P^n$ (monic polys., highest coeff. $=1$)

Computing $p(A)b$

$$p(z) = c_0 + z(c_1 + \cdots + z(c_{m-1} + z)\ldots)$$

(a)
```
P = A + c(m-1)I
for i = 2:m
    P = AP + c(m-i)I
end
```

$\longrightarrow p(A) = c_0 I + A(c_1 I + \cdots + A(c_{m-1}I + A)$

(Note: the summation only touches diagonal entries)

Total cost: $2(m-1)*m^3 + 2m(m-1) = \Theta(m^4)$

(b) Now:
```
pb = Ab + c(m-1)b
for i = 2:m
    pb = A*pb + c(m-i)b
end
```

Cost: $(m-1)*2m^2 + 2m(m-1) = \Theta(m^3)$ $(\sim 2m^3)$

(c) "Obvious method": (i) need $A^k b$, $k = 1, \ldots, m-1$

Total cost: $(m-1)*2m^2$ (successively compute terms)

(ii) $c_k A^k b$: $m*m = m^2$

(iii) Additions : $m*m = m^2$

Total: $\Theta(m^3)$ $(\sim 2m^3)$

i.e. (b)–(c) cost the same!

Set $\overline{\underline{XII}}$

35.4 To minimize:

$\|r_n\| = \|\tilde{H}_n y - \|b\| e_1\|$: need QR factorization of $\tilde{H}_n$. This is accomplished by $(n-1)$ Givens rotations, as follows:

$$\begin{pmatrix} \begin{pmatrix} c_1 & -s_1 \\ s_1 & c_1 \end{pmatrix} & \\ & I_{n-2} \end{pmatrix} \begin{pmatrix} h_{11} & h_{12} & \cdots & h_{1n} \\ h_{21} & h_{22} & \cdots & h_{2n} \\ & h_{32} & \ddots & \vdots \\ & & h_{n+1,n} & h_{nn} \end{pmatrix}$$

c_1, s_1 chosen to eliminate h_{21}:

$$s_1 = \frac{h_{21}}{\sqrt{h_{11}^2 + h_{21}^2}}, \quad c_1 = -\frac{h_{11}}{\sqrt{h_{11}^2 + h_{21}^2}}.$$

$h_{21}' = s_1 h_{11} + c_1 h_{21} = 0 \Rightarrow$

Proceeding similarly, we arrive at $H_n = Q_n R_n$ where $Q_n = G_{n-1}^T \cdots G_1^T$.

Each step is $O(n)$ (total n^2); then

$$Q_n^*(H_n y = \|b\| e_1) \Rightarrow R_n y = \underbrace{(Q_n^* e_1)\|b\|}_{O(n)}$$

just steal 1^{st} column

Solve by backsubstitution in $O(n^2)$.

Going to H_{n+1}, can reuse all the same G_i;

Additional work: $(G_{n-1} \cdots G_1) h_{(n+1)}$ ($n+1$ column) is $O(n)$

Eliminating $h_{n+1,n}$ requires one additional G, G_n ($O(n)$).