Math 5600 7/10/14

- State of the QR algorithm

$$H = U_0^T A U_0$$ upper Hessenberg

U_0 complex

- suppose H is unreduced

For $k = 1, 2, \ldots$

$$H - \mu I = UR$$ U orthogonal

$$H = RU + \mu I$$ R upper triangular

- If H is reduced work on the unreduced blocks of H.

- If μ is an eigenvalue it causes decoupling (in the $n,n-1$ position of H)

- But eigenvalues may be (conjugate) complex.

- In that case, how about a double shift?

- Suppose a_1 and a_2 are the possibly complex eigenvalues of

$$\begin{bmatrix} h_{mm} & h_{mn} \\ h_{nm} & h_{nn} \end{bmatrix} \qquad m = n-1$$

$$
\begin{aligned}
&(1) \quad & H - a_1 I &= U_1 R_1 \qquad \text{compute } U_1, R_1 \\
&(2) \quad & H_1 &= R_1 U_1 + a_1 I \\
&(3) \quad & H_1 - a_2 I &= U_2 R_2 \qquad \text{compute } U_2, R_2 \\
&(4) \quad & H_2 &= R_2 U_2 + a_2 I
\end{aligned}
$$

- This looks like it requires complex arithmetic.
- not quite
- It turns out that

$$H_2 = U_2^T U_1^T H U_1 U_2$$

where $(U_1 U_2)(R_2 R_1) = (H - a_1 I)(H - a_2 I)$

$$= H^2 - \underbrace{(a_1 + a_2)}_{\text{real}} I + \underbrace{a_1 a_2}_{\text{real}} I$$

QR factorization

$$= M$$

- M is a real matrix!
- So we could do the double shift by computing M and then the QR factorization of M

see Golub/van Loan p-387

verification:

$$H_1 - a_2 I = U_2 R_2 \quad \text{from } (3)$$

$$U_1 (H_1 - a_2 I) R_1 = U_1 U_2 R_2 R_1$$

$$U_1 H_1 R_1 - a_2 U_1 R_1 = U_1 U_2 R_2 R_1$$

$$U_1 \overset{(2)}{(R_1 U_1 + a_1 I)} R_1 - \overset{(1)}{a_2 (H - a_1 I)} = U_1 U_2 R_2 R_1$$

facto

$$U_1 R_1 (U_1 R_1 + a_1 I) - a_2 (H - a_1 I) = U_1 U_2 R_2 R_1$$

$$\overset{(1)}{(H - a_1 I)} (U_1 R_1 + a_1 I) - a_2 H + a_2 a_1 I = U_1 U_2 R_2 R_1$$

$$(H - a_1 I) \overset{(1)}{H} - a_2 H + a_1 a_2 I = U_1 U_2 R_2 R_1$$

$$M = (H - a_1 I)(H - a_2 I) = U_1 U_2 R_2 R_1$$

- However, computing the QR factorization of M would again require n^3 ops

- now (only now) things get complicated

- Francis QR steps

- see Golub/van Loan, Matrix Computations, 4th ed. ISBN 978-1-4214-0794-4

AEP, Chap. 6), and
atrix to Hessenberg
th. 40, 47–56.
enge because it is a
Out of Eigenvalue
trices to Condensed
teduction of a Non-
89–1211.
eduction to Hessen-
Upper Hessenberg,
' Parallel Comput.
out Conversion for
3, 819–821.
o a Similar Banded
Comput. 21, 223–
omposition for the
isive Schur decom-
to be very useful,
Linear Systems of
he Problem AX +
EE Trans. Autom.
trollability," IEEE
e Invariant Linear
in Algorithms and
tical Programming
ie Complex Schur
ue problem is dis-
trix Eigenvalues,"

4

7.5 The Practical QR Algorithm

We return to the Hessenberg QR iteration, which we write as follows:

$$\begin{array}{ll} H = U_0^T A U_0 & \text{(Hessenberg reduction)} \\ \textbf{for } k = 1, 2, \ldots & \\ \quad H = UR & \text{(QR factorization)} \\ \quad H = RU & \\ \textbf{end} & \end{array} \tag{7.5.1}$$

Our aim in this section is to describe how the H's converge to upper quasi-triangular form and to show how the convergence rate can be accelerated by incorporating *shifts*.

7.5.1 Deflation

Without loss of generality we may assume that each Hessenberg matrix H in (7.5.1) is unreduced. If not, then at some stage we have

$$H = \begin{array}{c} \left[\begin{array}{cc} H_{11} & H_{12} \\ 0 & H_{22} \end{array}\right] \begin{array}{l} p \\ n-p \end{array} \\ \begin{array}{cc} \;\; p \quad & \quad n-p \end{array} \end{array}$$

where $1 \le p < n$ and the problem *decouples* into two smaller problems involving H_{11} and H_{22}. The term *deflation* is also used in this context, usually when $p = n - 1$ or $n - 2$.

In practice, decoupling occurs whenever a subdiagonal entry in H is suitably small. For example, if

$$|h_{p+1,p}| \;\le\; c\mathbf{u}(|h_{pp}| + |h_{p+1,p+1}|) \tag{7.5.2}$$

for a small constant c, then $h_{p+1,p}$ can justifiably be set to zero because rounding errors of order $\mathbf{u}\| H \|$ are typically present throughout the matrix anyway.

7.5.2 The Shifted QR Iteration

Let $\mu \in \mathbb{R}$ and consider the iteration:

$$\begin{array}{ll} H = U_0^T A U_0 & \text{(Hessenberg reduction)} \\ \textbf{for } k = 1, 2, \ldots & \\ \quad \text{Determine a scalar } \mu. & \\ \quad H - \mu I = UR & \text{(QR factorization)} \\ \quad H = RU + \mu I & \\ \textbf{end} & \end{array} \tag{7.5.3}$$

The scalar μ is referred to as a *shift*. Each matrix H generated in (7.5.3) is similar to A, since

$$RU + \mu I \;=\; U^T(UR + \mu I)U \;=\; U^T HU.$$

5

If we order the eigenvalues λ_i of A so that

$$|\lambda_1 - \mu| \geq \cdots \geq |\lambda_n - \mu|,$$

and μ is fixed from iteration to iteration, then the theory of §7.3 says that the pth subdiagonal entry in H converges to zero with rate

$$\left|\frac{\lambda_{p+1} - \mu}{\lambda_p - \mu}\right|^k.$$

Of course, if $\lambda_p = \lambda_{p+1}$, then there is no convergence at all. But if, for example, μ is much closer to λ_n than to the other eigenvalues, then the zeroing of the $(n, n-1)$ entry is rapid. In the extreme case we have the following:

Theorem 7.5.1. *Let μ be an eigenvalue of an n-by-n unreduced Hessenberg matrix H. If*

$$\tilde{H} = RU + \mu I,$$

where $H - \mu I = UR$ is the QR factorization of $H - \mu I$, then $\tilde{h}_{n,n-1} = 0$ and $\tilde{h}_{nn} = \mu$.

Proof. Since H is an unreduced Hessenberg matrix the first $n-1$ columns of $H - \mu I$ are independent, regardless of μ. Thus, if $UR = (H - \mu I)$ is the QR factorization then $r_{ii} \neq 0$ for $i = 1{:}n-1$. But if $H - \mu I$ is singular, then $r_{11} \cdots r_{nn} = 0$. Thus, $r_{nn} = 0$ and $\tilde{H}(n,:) = [\,0,\ \ldots,\ 0,\ \mu\,]$. $\square$

The theorem says that if we shift by an exact eigenvalue, then in exact arithmetic deflation occurs in one step.

7.5.3 The Single-Shift Strategy

Now let us consider varying μ from iteration to iteration incorporating new information about $\lambda(A)$ as the subdiagonal entries converge to zero. A good heuristic is to regard h_{nn} as the best approximate eigenvalue along the diagonal. If we shift by this quantity during each iteration, we obtain the *single-shift QR iteration*:

$$\begin{array}{ll}
\textbf{for } k = 1, 2, \ldots & \\
\quad \mu = H(n,n) & \\
\quad H - \mu I = UR & \text{(QR factorization)} \\
\quad H = RU + \mu I & \\
\textbf{end} &
\end{array} \tag{7.5.4}$$

If the $(n, n-1)$ entry converges to zero, it is likely to do so at a quadratic rate. To see this, we borrow an example from Stewart (IMC, p. 366). Suppose H is an unreduced upper Hessenberg matrix of the form

$$H = \begin{bmatrix} \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times \\ 0 & 0 & \times & \times & \times \\ 0 & 0 & 0 & \epsilon & h_{nn} \end{bmatrix}$$

and that we perform one step of the single-shift QR algorithm, i.e.,

$$UR = H - h_{nn}$$
$$\tilde{H} = RU + h_{nn}I.$$

After $n-2$ steps in the orthogonal reduction of $H - h_{nn}I$ to upper triangular form we obtain a matrix with the following structure:

$$H = \begin{bmatrix} \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times \\ 0 & 0 & \times & \times & \times \\ 0 & 0 & 0 & a & b \\ 0 & 0 & 0 & \epsilon & 0 \end{bmatrix}.$$

It is not hard to show that

$$\tilde{h}_{n,n-1} = -\frac{\epsilon^2 b}{a^2+\epsilon^2}.$$

If we assume that $\epsilon \ll a$, then it is clear that the new $(n, n-1)$ entry has order ϵ^2, precisely what we would expect of a quadratically converging algorithm.

7.5.4 The Double-Shift Strategy

Unfortunately, difficulties with (7.5.4) can be expected if at some stage the eigenvalues a_1 and a_2 of

$$G = \begin{bmatrix} h_{mm} & h_{mn} \\ h_{nm} & h_{nn} \end{bmatrix}, \qquad m = n-1, \tag{7.5.5}$$

are complex for then h_{nn} would tend to be a poor approximate eigenvalue.

A way around this difficulty is to perform two single-shift QR steps in succession using a_1 and a_2 as shifts:

$$\begin{aligned} H - a_1 I &= U_1 R_1 \\ H_1 &= R_1 U_1 + a_1 I \\ H_1 - a_2 I &= U_2 R_2 \\ H_2 &= R_2 U_2 + a_2 I \end{aligned} \tag{7.5.6}$$

These equations can be manipulated to show that

$$(U_1 U_2)(R_2 R_1) = M \tag{7.5.7}$$

where M is defined by

$$M = (H - a_1 I)(H - a_2 I). \tag{7.5.8}$$

Note that M is a real matrix even if G's eigenvalues are complex since

$$M = H^2 - sH + tI$$

where

$$s = a_1 + a_2 = h_{mm} + h_{nn} = \operatorname{tr}(G) \in \mathbb{R}$$

and

$$t = a_1 a_2 = h_{mm}h_{nn} - h_{mn}h_{nm} = \det(G) \in \mathbb{R}.$$

Thus, (7.5.7) is the QR factorization of a real matrix and we may choose U_1 and U_2 so that $Z = U_1U_2$ is real orthogonal. It then follows that

$$H_2 = U_2^H H_1 U_2 = U_2^H(U_1^H H U_1)U_2 = (U_1U_2)^H H(U_1U_2) = Z^T HZ$$

is real.

Unfortunately, roundoff error almost always prevents an exact return to the real field. A real H_2 could be guaranteed if we

- explicitly form the real matrix $M = H^2 - sH + tI$,
- compute the real QR factorization $M = ZR$, and
- set $H_2 = Z^T HZ$.

But since the first of these steps requires $O(n^3)$ flops, this is not a practical course of action.

7.5.5 The Double-Implicit-Shift Strategy

Fortunately, it turns out that we can implement the double-shift step with $O(n^2)$ flops by appealing to the implicit Q theorem of §7.4.5. In particular we can effect the transition from H to H_2 in $O(n^2)$ flops if we

- compute Me_1, the first column of M;
- determine a Householder matrix P_0 such that $P_0(Me_1)$ is a multiple of e_1;
- compute Householder matrices $P_1, \ldots, P_{n-2}$ such that if

$$Z_1 = P_0P_1 \cdots P_{n-2},$$

  then $Z_1^T HZ_1$ is upper Hessenberg and the first columns of Z and Z_1 are the same.

Under these circumstances, the implicit Q theorem permits us to conclude that, if $Z^T HZ$ and $Z_1^T HZ_1$ are both unreduced upper Hessenberg matrices, then they are essentially equal. Note that if these Hessenberg matrices are not unreduced, then we can effect a decoupling and proceed with smaller unreduced subproblems.

Let us work out the details. Observe first that P_0 can be determined in $O(1)$ flops since $Me_1 = [x,\ y,\ z,\ 0, \ldots, 0]^T$ where

$$\begin{aligned} x &= h_{11}^2 + h_{12}h_{21} - sh_{11} + t, \\ y &= h_{21}(h_{11} + h_{22} - s), \\ z &= h_{21}h_{32}. \end{aligned}$$

Since a similarity transformation with P_0 only changes rows and columns 1, 2, and 3, we see that

$$P_0 H P_0 = \begin{bmatrix} \times & \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times & \times \\ 0 & 0 & 0 & \times & \times & \times \\ 0 & 0 & 0 & 0 & \times & \times \end{bmatrix}.$$

Now the mission of the Householder matrices $P_1, \ldots, P_{n-2}$ is to restore this matrix to upper Hessenberg form. The calculation proceeds as follows:

$$\begin{bmatrix} \times & \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times & \times \\ 0 & 0 & 0 & \times & \times & \times \\ 0 & 0 & 0 & 0 & \times & \times \end{bmatrix} \xrightarrow{P_1} \begin{bmatrix} \times & \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times & \times \\ 0 & 0 & 0 & 0 & \times & \times \end{bmatrix} \xrightarrow{P_2}$$

$$\begin{bmatrix} \times & \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times & \times \\ 0 & 0 & \times & \times & \times & \times \\ 0 & 0 & \times & \times & \times & \times \\ 0 & 0 & \times & \times & \times & \times \end{bmatrix} \xrightarrow{P_3} \begin{bmatrix} \times & \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times & \times \\ 0 & 0 & \times & \times & \times & \times \\ 0 & 0 & 0 & \times & \times & \times \\ 0 & 0 & 0 & \times & \times & \times \end{bmatrix} \xrightarrow{P_4} \begin{bmatrix} \times & \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times & \times \\ 0 & 0 & \times & \times & \times & \times \\ 0 & 0 & 0 & \times & \times & \times \\ 0 & 0 & 0 & 0 & \times & \times \end{bmatrix}.$$

Each P_k is the identity with a 3-by-3 or 2-by-2 Householder somewhere along its diagonal, e.g.,

$$P_1 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & \times & \times & \times & 0 & 0 \\ 0 & \times & \times & \times & 0 & 0 \\ 0 & \times & \times & \times & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \quad P_2 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & \times & \times & \times & 0 \\ 0 & 0 & \times & \times & \times & 0 \\ 0 & 0 & \times & \times & \times & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix},$$

$$P_3 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & \times & \times & \times \\ 0 & 0 & 0 & \times & \times & \times \\ 0 & 0 & 0 & \times & \times & \times \end{bmatrix}, \quad P_4 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & \times & 0 & 0 \\ 0 & 0 & 0 & \times & \times & \times \\ 0 & 0 & 0 & 0 & \times & \times \end{bmatrix}.$$

The applicability of Theorem 7.4.3 (the implicit Q theorem) follows from the observation that $P_k e_1 = e_1$ for $k = 1{:}n-2$ and that P_0 and Z have the same first column. Hence, $Z_1 e_1 = Z e_1$, and we can assert that Z_1 essentially equals Z provided that the upper Hessenberg matrices $Z^T H Z$ and $Z_1^T H Z_1$ are each unreduced.

The implicit determination of H_2 from H outlined above was first described by Francis (1961) and we refer to it as a *Francis QR step*. The complete Francis step is summarized as follows:

Algorithm 7.5.1 (Francis QR step) Given the unreduced upper Hessenberg matrix $H \in \mathbb{R}^{n\times n}$ whose trailing 2-by-2 principal submatrix has eigenvalues a_1 and a_2, this algorithm overwrites H with $Z^T HZ$, where Z is a product of Householder matrices and $Z^T(H - a_1I)(H - a_2I)$ is upper triangular.

$m = n - 1$
{Compute first column of $(H - a_1I)(H - a_2I)$}
$s = H(m, m) + H(n, n)$
$t = H(m, m)\cdot H(n, n) - H(m, n)\cdot H(n, m)$
$x = H(1, 1)\cdot H(1, 1) + H(1, 2)\cdot H(2, 1) - s\cdot H(1, 1) + t$
$y = H(2, 1)\cdot (H(1, 1) + H(2, 2) - s)$
$z = H(2, 1)\cdot H(3, 2)$
for $k = 0{:}n - 3$
  $[v, \beta] = \text{house}([x\ y\ z]^T)$
  $q = \max\{1, k\}$.
  $H(k + 1{:}k + 3, q{:}n) = (I - \beta vv^T)\cdot H(k + 1{:}k + 3, q{:}n)$
  $r = \min\{k + 4, n\}$
  $H(1{:}r, k + 1{:}k + 3) = H(1{:}r, k + 1{:}k + 3)\cdot (I - \beta vv^T)$
  $x = H(k + 2, k + 1)$
  $y = H(k + 3, k + 1)$
  if $k < n - 3$
    $z = H(k + 4, k + 1)$
  end
end
$[v, \beta] = \text{house}([\ x\ y\ ]^T)$
$H(n - 1{:}n, n - 2{:}n) = (I - \beta vv^T)\cdot H(n - 1{:}n, n - 2{:}n)$
$H(1{:}n, n - 1{:}n) = H(1{:}n, n - 1{:}n)\cdot (I - \beta vv^T)$

This algorithm requires $10n^2$ flops. If Z is accumulated into a given orthogonal matrix, an additional $10n^2$ flops are necessary.

7.5.6 The Overall Process

Reduction of A to Hessenberg form using Algorithm 7.4.2 and then iteration with Algorithm 7.5.1 to produce the real Schur form is the standard means by which the dense unsymmetric eigenproblem is solved. During the iteration it is necessary to monitor the subdiagonal elements in H in order to spot any possible decoupling. How this is done is illustrated in the following algorithm:

Algorithm 7.5.2 (QR Algorithm) Given $A \in \mathbb{R}^{n \times n}$ and a tolerance tol greater than the unit roundoff, this algorithm computes the real Schur canonical form $Q^T AQ = T$. If Q and T are desired, then T is stored in H. If only the eigenvalues are desired, then diagonal blocks in T are stored in the corresponding positions in H.

Use Algorithm 7.4.2 to compute the Hessenberg reduction

$H = U_0^T A U_0$ where $U_0 = P_1 \cdots P_{n-2}$.

If Q is desired form $Q = P_1 \cdots P_{n-2}$. (See §5.1.6.)

until $q = n$

Set to zero all subdiagonal elements that satisfy:

$|h_{i,i-1}| \le \text{tol}\cdot(|h_{ii}| + |h_{i-1,i-1}|)$.

Find the largest nonnegative q and the smallest non-negative p such that

$$H = \begin{array}{cc} \left[\begin{array}{ccc} H_{11} & H_{12} & H_{13} \\ 0 & H_{22} & H_{23} \\ 0 & 0 & H_{33} \end{array}\right] & \begin{array}{c} p \\ n-p-q \\ q \end{array} \\ \begin{array}{ccc} p & n-p-q & q \end{array} & \end{array}$$

where H_{33} is upper quasi-triangular and H_{22} is unreduced.

if $q < n$

Perform a Francis QR step on H_{22}: $H_{22} = Z^T H_{22} Z$.

if Q is required

$Q = Q \cdot \text{diag}(I_p, Z, I_q)$

$H_{12} = H_{12} Z$

$H_{23} = Z^T H_{23}$

end

end

end

Upper triangularize all 2-by-2 diagonal blocks in H that have real eigenvalues and accumulate the transformations (if necessary).

This algorithm requires $25n^3$ flops if Q and T are computed. If only the eigenvalues are desired, then $10n^3$ flops are necessary. These flops counts are very approximate and are based on the empirical observation that on average only two Francis iterations are required before the lower 1-by-1 or 2-by-2 decouples.

The roundoff properties of the QR algorithm are what one would expect of any orthogonal matrix technique. The computed real Schur form $\hat{T}$ is orthogonally similar to a matrix near to A, i.e.,

$$Q^T(A + E)Q = \hat{T}$$

where $Q^T Q = I$ and $\| E \|_2 \approx \mathbf{u} \| A \|_2$. The computed $\hat{Q}$ is almost orthogonal in the sense that $\hat{Q}^T \hat{Q} = I + F$ where $\| F \|_2 \approx \mathbf{u}$.

The order of the eigenvalues along $\hat{T}$ is somewhat arbitrary. But as we discuss in §7.6, any ordering can be achieved by using a simple procedure for swapping two adjacent diagonal entries.

11

7.5.7 Balancing

Finally, we mention that if the elements of A have widely varying magnitudes, then A should be *balanced* before applying the QR algorithm. This is an $O(n^2)$ calculation in which a diagonal matrix D is computed so that if

$$D^{-1}AD = [\, c_1 \,|\cdots|\, c_n \,] = \begin{bmatrix} r_1^T \\ \vdots \\ r_n^T \end{bmatrix}$$

then $\| r_i \|_\infty \approx \| c_i \|_\infty$ for $i = 1{:}n$. The diagonal matrix D is chosen to have the form

$$D = \mathrm{diag}(\beta^{i_1}, \ldots, \beta^{i_n})$$

where β is the floating point base. Note that $D^{-1}AD$ can be calculated without roundoff. When A is balanced, the computed eigenvalues are usually more accurate although there are exceptions. See Parlett and Reinsch (1969) and Watkins(2006).

Problems

P7.5.1 Show that if $\tilde{H} = Q^T HQ$ is obtained by performing a single-shift QR step with

$$H = \begin{bmatrix} w & x \\ y & z \end{bmatrix},$$

then $|\tilde{h}_{21}| \le |y^2 x|/[(w-z)^2 + y^2]$.

P7.5.2 Given $A \in \mathbb{R}^{2\times 2}$, show how to compute a diagonal $D \in \mathbb{R}^{2\times 2}$ so that $\| D^{-1}AD \|_F$ is minimized.

P7.5.3 Explain how the single-shift QR step $H - \mu I = UR$, $\tilde{H} = RU + \mu I$ can be carried out implicitly. That is, show how the transition from H to $\tilde{H}$ can be carried out without subtracting the shift μ from the diagonal of H.

P7.5.4 Suppose H is upper Hessenberg and that we compute the factorization $PH = LU$ via Gaussian elimination with partial pivoting. (See Algorithm 4.3.4.) Show that $H_1 = U(P^T L)$ is upper Hessenberg and similar to H. (This is the basis of the *modified LR algorithm*.)

P7.5.5 Show that if $H = H_0$ is given and we generate the matrices H_k via $H_k - \mu_k I = U_k R_k$, $H_{k+1} = R_k U_k + \mu_k I$, then $(U_1 \cdots U_j)(R_j \cdots R_1) = (H - \mu_1 I)\cdots(H - \mu_j I)$.

Notes and References for §7.5

Historically important papers associated with the QR iteration include:

H. Rutishauser (1958). "Solution of Eigenvalue Problems with the LR Transformation," *Nat. Bur. Stand. App. Math. Ser. 49*, 47–81.
J.G.F. Francis (1961). "The QR Transformation: A Unitary Analogue to the LR Transformation, Parts I and II" *Comput. J. 4*, 265–72, 332–345.
V.N. Kublanovskaya (1961). "On Some Algorithms for the Solution of the Complete Eigenvalue Problem," *Vychisl. Mat. Mat. Fiz 1(4)*, 555–570.
R.S. Martin and J.H. Wilkinson (1968). "The Modified LR Algorithm for Complex Hessenberg Matrices," *Numer. Math. 12*, 369–376.
R.S. Martin, G. Peters, and J.H. Wilkinson (1970). "The QR Algorithm for Real Hessenberg Matrices," *Numer. Math. 14*, 219–231.

For a general insight, we recommend:

D.S. Watkins (1982). "Understanding the QR Algorithm," *SIAM Review 24*, 427–440.
D.S. Watkins (1993). "Some Perspectives on the Eigenvalue Problem," *SIAM Review 35*, 430–471.

D.S. Watkins (2008). "The QR Algorithm Revisited," *SIAM Review 50*, 133–145.
D.S. Watkins (2011). "Francis's Algorithm," *Amer. Math. Monthly 118*, 387–403.

Papers concerned with the convergence of the method, shifting, deflation, and related matters include:

P.A. Businger (1971). "Numerically Stable Deflation of Hessenberg and Symmetric Tridiagonal Matrices, *BIT 11*, 262–270.
D.S. Watkins and L. Elsner (1991). "Chasing Algorithms for the Eigenvalue Problem," *SIAM J. Matrix Anal. Applic. 12*, 374–384.
D.S. Watkins and L. Elsner (1991). "Convergence of Algorithms of Decomposition Type for the Eigenvalue Problem," *Lin. Alg. Applic. 143*, 19–47.
J. Erxiong (1992). "A Note on the Double-Shift QL Algorithm," *Lin. Alg. Applic. 171*, 121–132.
A.A. Dubrulle and G.H. Golub (1994). "A Multishift QR Iteration Without Computation of the Shifts," *Numer. Algorithms 7*, 173–181.
D.S. Watkins (1996). "Forward Stability and Transmission of Shifts in the QR Algorithm," *SIAM J. Matrix Anal. Applic. 16*, 469–487.
D.S. Watkins (1996). "The Transmission of Shifts and Shift Blurring in the QR algorithm," *Lin. Alg. Applic. 241-3*, 877–896.
D.S. Watkins (1998). "Bulge Exchanges in Algorithms of QR Type," *SIAM J. Matrix Anal. Applic. 19*, 1074–1096.
R. Vandebril (2011). "Chasing Bulges or Rotations? A Metamorphosis of the QR-Algorithm" *SIAM. J. Matrix Anal. Applic. 32*, 217–247.

Aspects of the balancing problem are discussed in:

E.E. Osborne (1960). "On Preconditioning of Matrices," *J. ACM 7*, 338–345.
B.N. Parlett and C. Reinsch (1969). "Balancing a Matrix for Calculation of Eigenvalues and Eigenvectors," *Numer. Math. 13*, 292–304.
D.S. Watkins (2006). "A Case Where Balancing is Harmful," *ETNA 23*, 1–4.

Versions of the algorithm that are suitable for companion matrices are discussed in:

D.A. Bini, F. Daddi, and L. Gemignani (2004). "On the Shifted QR iteration Applied to Companion Matrices," *ETNA 18*, 137–152.
M. Van Barel, R. Vandebril, P. Van Dooren, and K. Frederix (2010). "Implicit Double Shift QR-Algorithm for Companion Matrices," *Numer. Math. 116*, 177–212.

Papers that are concerned with the high-performance implementation of the QR iteration include:

Z. Bai and J.W. Demmel (1989). "On a Block Implementation of Hessenberg Multishift QR Iteration," *Int. J. High Speed Comput. 1*, 97–112.
R.A. Van De Geijn (1993). "Deferred Shifting Schemes for Parallel QR Methods," *SIAM J. Matrix Anal. Applic. 14*, 180–194.
D.S. Watkins (1994). "Shifting Strategies for the Parallel QR Algorithm," *SIAM J. Sci. Comput. 15*, 953–958.
G. Henry and R. van de Geijn (1996). "Parallelizing the QR Algorithm for the Unsymmetric Algebraic Eigenvalue Problem: Myths and Reality," *SIAM J. Sci. Comput. 17*, 870–883.
Z. Bai, J. Demmel, J. Dongarra, A. Petitet, H. Robinson, and K. Stanley (1997). "The Spectral Decomposition of Nonsymmetric Matrices on Distributed Memory Parallel Computers," *SIAM J. Sci. Comput. 18*, 1446–1461.
G. Henry, D.S. Watkins, and J. Dongarra (2002). "A Parallel Implementation of the Nonsymmetric QR Algorithm for Distributed Memory Architectures," *SIAM J. Sci. Comput. 24*, 284–311.
K. Braman, R. Byers, and R. Mathias (2002). "The Multishift QR Algorithm. Part I: Maintaining Well-Focused Shifts and Level 3 Performance," *SIAM J. Matrix Anal. Applic. 23*, 929–947.
K. Braman, R. Byers, and R. Mathias (2002). "The Multishift QR Algorithm. Part II: Aggressive Early Deflation," *SIAM J. Matrix Anal. Applic. 23*, 948–973.
M.R. Fahey (2003). "Algorithm 826: A Parallel Eigenvalue Routine for Complex Hessenberg Matrices," *ACM Trans. Math. Softw. 29*, 326–336.
D. Kressner (2005). "On the Use of Larger Bulges in the QR Algorithm," *ETNA 20*, 50–63.
D. Kressner (2008). "The Effect of Aggressive Early Deflation on the Convergence of the QR Algorithm," *SIAM J. Matrix Anal. Applic. 30*, 805–821.

From: Frank Uhlig <uhligfd@auburn.edu>
Date: Wed, 25 Mar 2009 08:48:14 -0500
Subject: John Francis of QR found

John Francis and 50 years of QR

John Francis submitted his first QR paper almost 50 years ago in October 1959. By 1962 he had left the NA field. When his algorithm was judged one of the top ten algorithms of the 20th century in 2000 by Jack Dongarra and Francis Sullivan, nobody alive in the mathematics community had ever seen John Francis or knew where or if he lived. Gene Golub and Frank Uhlig independently tracked John Francis down, joined forces, and visited and interviewed him over the last couple of years.

When first contacted, John Francis had no idea about QR's impact. He is 74 years old now and well. Re QR, he remembers his math and computational work of 50 years ago clearly. John Francis will be the lead-off speaker at a mini symposium, held in his honor, at the 23rd Biennial Conference on Numerical Analysis, June 23rd - 26th 2009 in Glasgow to which everyone is cordially invited.

For a detailed list of invited speakers,
see http://www.maths.strath.ac.uk/naconf/minisymposia

Notes about QR algorithm

- it's sophisticated.
- However, it grew out of simple ideas in natural steps

  This is a typical procedure in mathematics (and elsewhere)

- Ingredients:
  - work on subproblems by embedding in identity (conceptually)
  - power iteration
  - orthogonal iteration
  - many simultaneous orthogonal iterations
  - shift of origin
  - real Schur form
  - Use of Householder reflections and Givens rotations
  - $O(n^2)$ effort per step
  - real arithmetic
  - start with upper Hessenberg
  - keep it upper Hessenberg
  - Implicit Q Theorem.