

Lecture 9: QR algorithm



School of Mathematical Sciences, Xiamen University

1. Subspace iteration (SI)

- Subspace iteration is also called *simultaneous iteration* (and sometimes *orthogonal iteration* or *block power iteration*).

Algorithm 1: Subspace iteration

Pick $\mathbf{Q}_n^{(0)} \in \mathbb{C}^{m \times n}$ with orthonormal columns
for $k = 1, 2, 3, \dots$,
 $\mathbf{Q}_n^{(k)} \mathbf{R}_n^{(k)} = \mathbf{A} \mathbf{Q}_n^{(k-1)}$ (QR factorization)
end

- Here is an informal analysis of this algorithm. Assume $\mathbf{A} = \mathbf{S} \mathbf{\Lambda} \mathbf{S}^{-1} \in \mathbb{C}^{m \times m}$ with

$$\mathbf{\Lambda} = \text{diag}\{\lambda_1, \lambda_2, \dots, \lambda_m\}$$

and

$$|\lambda_1| \geq \dots \geq |\lambda_n| > |\lambda_{n+1}| \geq \dots \geq |\lambda_m|.$$

If $n = 1$, then subspace iteration reduces to power iteration; see Lecture 8 for details.

Now we consider the case $n > 1$. Let

$$\mathbf{X}_n := [\mathbf{I}_n \quad \mathbf{0}] \mathbf{S}^{-1} \mathbf{Q}_n^{(0)},$$

and

$$\mathbf{X}_c := [\mathbf{0} \quad \mathbf{I}_{m-n}] \mathbf{S}^{-1} \mathbf{Q}_n^{(0)}.$$

Assume that \mathbf{X}_n has full rank (a generalization of the assumption $\alpha_1 \neq 0$ in power iteration). We have (the proof is left as an exercise)

$$\mathbf{Q}_n^{(k)} = \mathbf{A}^k \mathbf{Q}_n^{(0)} (\mathbf{R}_n^{(1)})^{-1} (\mathbf{R}_n^{(2)})^{-1} \cdots (\mathbf{R}_n^{(k)})^{-1}.$$

By $\mathbf{A} = \mathbf{S} \mathbf{\Lambda} \mathbf{S}^{-1}$, we have

$$\mathbf{Q}_n^{(k)} = \mathbf{S} \mathbf{\Lambda}^k \mathbf{S}^{-1} \mathbf{Q}_n^{(0)} (\mathbf{R}_n^{(1)})^{-1} (\mathbf{R}_n^{(2)})^{-1} \cdots (\mathbf{R}_n^{(k)})^{-1}.$$

Write $\mathbf{S} = [\mathbf{S}_n \quad \mathbf{S}_c]$ and $\mathbf{\Lambda} = \text{diag}\{\mathbf{\Lambda}_n, \mathbf{\Lambda}_c\}$. By $\mathbf{S}^{-1}\mathbf{Q}_n^{(0)} = \begin{bmatrix} \mathbf{X}_n \\ \mathbf{X}_c \end{bmatrix}$, we have

$$\begin{aligned} \mathbf{Q}_n^{(k)} &= [\mathbf{S}_n \quad \mathbf{S}_c] \begin{bmatrix} \mathbf{\Lambda}_n^k & \\ & \mathbf{\Lambda}_c^k \end{bmatrix} \begin{bmatrix} \mathbf{X}_n \\ \mathbf{X}_c \end{bmatrix} (\mathbf{R}_n^{(1)})^{-1} (\mathbf{R}_n^{(2)})^{-1} \cdots (\mathbf{R}_n^{(k)})^{-1} \\ &= [\mathbf{S}_n \quad \mathbf{S}_c] \begin{bmatrix} \mathbf{\Lambda}_n^k \mathbf{X}_n (\mathbf{R}_n^{(1)})^{-1} (\mathbf{R}_n^{(2)})^{-1} \cdots (\mathbf{R}_n^{(k)})^{-1} \\ \mathbf{\Lambda}_c^k \mathbf{X}_c (\mathbf{R}_n^{(1)})^{-1} (\mathbf{R}_n^{(2)})^{-1} \cdots (\mathbf{R}_n^{(k)})^{-1} \end{bmatrix}. \end{aligned}$$

Exercise: Prove that

$$\mathbf{S}_c \mathbf{\Lambda}_c^k \mathbf{X}_c (\mathbf{R}_n^{(1)})^{-1} (\mathbf{R}_n^{(2)})^{-1} \cdots (\mathbf{R}_n^{(k)})^{-1} \rightarrow \mathbf{0}$$

like $\left| \frac{\lambda_{n+1}}{\lambda_n} \right|^k$. Equivalently, we have the convergence result:

$$\mathbf{Q}_n^{(k)} - \mathbf{S}_n \mathbf{\Lambda}_n^k \mathbf{X}_n (\mathbf{R}_n^{(1)})^{-1} (\mathbf{R}_n^{(2)})^{-1} \cdots (\mathbf{R}_n^{(k)})^{-1} \rightarrow \mathbf{0},$$

which means that $\text{span}\{\mathbf{Q}_n^{(k)}\}$ converges to $\text{span}\{\mathbf{S}_n\}$.

- The first $1 \leq j \leq n$ columns of $\mathbf{Q}_n^{(k)}$ are *identical* to the iterates of the subspace iteration started with the first j columns of $\mathbf{Q}_n^{(0)}$.

In other words, subspace iteration is effectively running the algorithm for $j = 1, 2, \dots, n$ **all at the same time (simultaneous)**.

So if *all* the first $n + 1$ eigenvalues have distinct absolute values, i.e.,

$$|\lambda_1| > |\lambda_2| > \dots > |\lambda_n| > |\lambda_{n+1}|,$$

and if *all* the leading principal submatrices of

$$\mathbf{X}_n = [\mathbf{I}_n \quad \mathbf{0}] \mathbf{S}^{-1} \mathbf{Q}_n^{(0)}$$

have full rank, the same convergence analysis as before implies that $\text{span}\{\mathbf{Q}_j^{(k)}\}$ with $\mathbf{Q}_j^{(k)} := \mathbf{Q}_n^{(k)} \begin{bmatrix} \mathbf{I}_j \\ \mathbf{0} \end{bmatrix}$ converges to $\text{span}\{\mathbf{S}_j\}$

with $\mathbf{S}_j := \mathbf{S} \begin{bmatrix} \mathbf{I}_j \\ \mathbf{0} \end{bmatrix}$ for each $j = 1, 2, \dots, n$.

Theorem 1

Consider running subspace iteration on matrix $\mathbf{A} \in \mathbb{C}^{m \times m}$ with $n = m$ and $\mathbf{Q}_n^{(0)} = \mathbf{I}$. If $\mathbf{A} = \mathbf{S}\mathbf{\Lambda}\mathbf{S}^{-1}$ with

$$\mathbf{\Lambda} = \text{diag}\{\lambda_1, \lambda_2, \dots, \lambda_m\}, \quad |\lambda_1| > |\lambda_2| > \dots > |\lambda_m|,$$

and if all the leading principal submatrices of \mathbf{S}^{-1} have full rank, then $\mathbf{A}^{(k)} := (\mathbf{Q}_n^{(k)})^* \mathbf{A} \mathbf{Q}_n^{(k)}$ converges to a Schur form of \mathbf{A} . The eigenvalues will appear in decreasing order of absolute value.

Proof: See Demmel's book: Theorem 4.8, Page 158, [Applied numerical linear algebra](#).

- The entry $\mathbf{A}_{jj}^{(k)}$ converges to λ_j like $\max \left(\left| \frac{\lambda_{j+1}}{\lambda_j} \right|^k, \left| \frac{\lambda_j}{\lambda_{j-1}} \right|^k \right)$.
- The block $\mathbf{A}^{(k)}(j+1:m, 1:j)$ converges to zero like $\left| \frac{\lambda_{j+1}}{\lambda_j} \right|^k$.

2. “Pure” QR algorithm

Algorithm 2: “Pure” QR algorithm

$$\mathbf{A}^{(0)} = \mathbf{A}$$

for $k = 1, 2, 3, \dots$,

$$\mathbf{Q}^{(k)}\mathbf{R}^{(k)} = \mathbf{A}^{(k-1)} \quad (\text{QR factorization})$$

$$\mathbf{A}^{(k)} = \mathbf{R}^{(k)}\mathbf{Q}^{(k)}$$

end

Proposition 2

We have $\mathbf{A}^{(k)} = (\underline{\mathbf{Q}}^{(k)})^* \mathbf{A} \underline{\mathbf{Q}}^{(k)}$, where $\underline{\mathbf{Q}}^{(k)} := \mathbf{Q}^{(1)}\mathbf{Q}^{(2)} \dots \mathbf{Q}^{(k)}$.

Proof. By $\mathbf{Q}^{(k)}\mathbf{R}^{(k)} = \mathbf{A}^{(k-1)}$, we have $\mathbf{R}^{(k)} = (\mathbf{Q}^{(k)})^* \mathbf{A}^{(k-1)}$. Then,

$$\begin{aligned} \mathbf{A}^{(k)} &= \mathbf{R}^{(k)}\mathbf{Q}^{(k)} = (\mathbf{Q}^{(k)})^* \mathbf{A}^{(k-1)}\mathbf{Q}^{(k)} \\ &= (\mathbf{Q}^{(k)})^* (\mathbf{Q}^{(k-1)})^* \mathbf{A}^{(k-2)}\mathbf{Q}^{(k-1)}\mathbf{Q}^{(k)} \\ &= (\mathbf{Q}^{(k)})^* \dots (\mathbf{Q}^{(1)})^* \mathbf{A}^{(0)}\mathbf{Q}^{(1)} \dots \mathbf{Q}^{(k)} = (\underline{\mathbf{Q}}^{(k)})^* \mathbf{A} \underline{\mathbf{Q}}^{(k)}. \quad \square \end{aligned}$$

Proposition 3

We have (a QR factorization of \mathbf{A}^k)

$$\mathbf{A}^k = \underline{\mathbf{Q}}^{(k)} \underline{\mathbf{R}}^{(k)},$$

where $\underline{\mathbf{Q}}^{(k)} := \mathbf{Q}^{(1)} \mathbf{Q}^{(2)} \dots \mathbf{Q}^{(k)}$, and $\underline{\mathbf{R}}^{(k)} := \mathbf{R}^{(k)} \mathbf{R}^{(k-1)} \dots \mathbf{R}^{(1)}$.

Proof.

We use induction. For $k = 1$, $\mathbf{A} = \mathbf{A}^{(0)} = \mathbf{Q}^{(1)} \mathbf{R}^{(1)} = \underline{\mathbf{Q}}^{(1)} \underline{\mathbf{R}}^{(1)}$. Now we prove the case $k > 1$ with the assumption $\mathbf{A}^{k-1} = \underline{\mathbf{Q}}^{(k-1)} \underline{\mathbf{R}}^{(k-1)}$. By Proposition 2, we have $\mathbf{A}^{(k-1)} = (\underline{\mathbf{Q}}^{(k-1)})^* \mathbf{A} \underline{\mathbf{Q}}^{(k-1)}$, which implies $\mathbf{A} \underline{\mathbf{Q}}^{(k-1)} = \underline{\mathbf{Q}}^{(k-1)} \mathbf{A}^{(k-1)}$. Then we have

$$\begin{aligned} \mathbf{A}^k &= \mathbf{A} \mathbf{A}^{k-1} = \mathbf{A} \underline{\mathbf{Q}}^{(k-1)} \underline{\mathbf{R}}^{(k-1)} = \underline{\mathbf{Q}}^{(k-1)} \mathbf{A}^{(k-1)} \underline{\mathbf{R}}^{(k-1)} \\ &= \underline{\mathbf{Q}}^{(k-1)} \mathbf{Q}^{(k)} \mathbf{R}^{(k)} \underline{\mathbf{R}}^{(k-1)} = \underline{\mathbf{Q}}^{(k)} \underline{\mathbf{R}}^{(k)}. \end{aligned}$$

This completes the proof. □

- Connection with power iteration: By $\mathbf{A}^k = \underline{\mathbf{Q}}^{(k)} \underline{\mathbf{R}}^{(k)}$, the first column of $\underline{\mathbf{Q}}^{(k)}$ is the result of applying k steps of power iteration on \mathbf{A} to the starting vector \mathbf{e}_1 . We have $\mathbf{A}_{11}^{(k)} \rightarrow \lambda_1$.
- Connection with inverse iteration: By $\underline{\mathbf{Q}}^{(k)} = (\mathbf{A}^*)^{-k} (\underline{\mathbf{R}}^{(k)})^*$, the last column of $\underline{\mathbf{Q}}^{(k)}$ is the result of applying k steps of inverse iteration on \mathbf{A}^* to the starting vector \mathbf{e}_m . We have $\mathbf{A}_{mm}^{(k)} \rightarrow \lambda_m$.

Theorem 4

If \mathbf{A} is diagonalizable with $\mathbf{A} = \mathbf{S}\mathbf{\Lambda}\mathbf{S}^{-1}$,

$$\mathbf{\Lambda} = \text{diag}\{\lambda_1, \lambda_2, \dots, \lambda_m\}, \quad |\lambda_1| > |\lambda_2| > \dots > |\lambda_m|,$$

and if all the leading principal submatrices of \mathbf{S}^{-1} have full rank, then $\mathbf{A}^{(k)}$ computed by “pure” QR algorithm converges to a Schur form of \mathbf{A} . The eigenvalues will appear in decreasing order of absolute value.

This theorem is a direct result of the following lemma.

Lemma 5

The $\mathbf{A}^{(k)}$ computed by “pure” QR algorithm is *identical* to the matrix $(\mathbf{Q}_n^{(k)})^* \mathbf{A} \mathbf{Q}_n^{(k)}$ implicitly computed by running subspace iteration on matrix $\mathbf{A} \in \mathbb{C}^{m \times m}$ with $n = m$ and $\mathbf{Q}_n^{(0)} = \mathbf{I}$. (We need an assumption about QR factorizations used in subspace iteration.)

Proof. We use induction. For $k = 1$, let $\mathbf{Q}_n^{(1)} := \mathbf{Q}^{(1)}$ and $\mathbf{R}_n^{(1)} := \mathbf{R}^{(1)}$. We have $\mathbf{A}^{(1)} = (\mathbf{Q}_n^{(1)})^* \mathbf{A} \mathbf{Q}_n^{(1)}$. Assume $\mathbf{A}^{(k-1)} = (\mathbf{Q}_n^{(k-1)})^* \mathbf{A} \mathbf{Q}_n^{(k-1)}$. Then from the “pure” QR algorithm and the induction hypothesis, we have

$$\mathbf{Q}^{(k)} \mathbf{R}^{(k)} = (\mathbf{Q}_n^{(k-1)})^* \mathbf{A} \mathbf{Q}_n^{(k-1)}, \quad \text{i.e., } \mathbf{A} \mathbf{Q}_n^{(k-1)} = \mathbf{Q}_n^{(k-1)} \mathbf{Q}^{(k)} \mathbf{R}^{(k)}.$$

Let $\mathbf{Q}_n^{(k)} := \mathbf{Q}_n^{(k-1)} \mathbf{Q}^{(k)}$ and $\mathbf{R}_n^{(k)} := \mathbf{R}^{(k)}$ be the matrices computed in step k of subspace iteration (QR factorization $\mathbf{Q}_n^{(k)} \mathbf{R}_n^{(k)} = \mathbf{A} \mathbf{Q}_n^{(k-1)}$). By $\mathbf{R}^{(k)} = \mathbf{R}_n^{(k)} = (\mathbf{Q}_n^{(k)})^* \mathbf{A} \mathbf{Q}_n^{(k-1)}$ and $\mathbf{Q}^{(k)} = (\mathbf{Q}_n^{(k-1)})^* \mathbf{Q}_n^{(k)}$, we have

$$\mathbf{A}^{(k)} = \mathbf{R}^{(k)} \mathbf{Q}^{(k)} = (\mathbf{Q}_n^{(k)})^* \mathbf{A} \mathbf{Q}_n^{(k-1)} (\mathbf{Q}_n^{(k-1)})^* \mathbf{Q}_n^{(k)} = (\mathbf{Q}_n^{(k)})^* \mathbf{A} \mathbf{Q}_n^{(k)}.$$

This completes the proof. □

- From the earlier analysis, we know that the convergence rate of “pure” QR algorithm depends on the absolute values of the ratios of eigenvalues. To speed convergence, we can use **shift and invert** techniques.

3. QR algorithm with shifts

Algorithm 3: QR algorithm with shifts

$$\mathbf{A}^{(0)} = \mathbf{A}$$

for $k = 1, 2, 3, \dots$,

Pick a shift $\mu^{(k)}$ near an eigenvalue of \mathbf{A}

$$\mathbf{Q}^{(k)} \mathbf{R}^{(k)} = \mathbf{A}^{(k-1)} - \mu^{(k)} \mathbf{I} \quad (\text{QR factorization})$$

$$\mathbf{A}^{(k)} = \mathbf{R}^{(k)} \mathbf{Q}^{(k)} + \mu^{(k)} \mathbf{I}$$

end

Proposition 6

We have $\mathbf{A}^{(k)} = (\mathbf{Q}^{(k)})^* \mathbf{A}^{(k-1)} \mathbf{Q}^{(k)} = (\underline{\mathbf{Q}}^{(k)})^* \mathbf{A} \underline{\mathbf{Q}}^{(k)}$, where $\underline{\mathbf{Q}}^{(k)} := \mathbf{Q}^{(1)} \mathbf{Q}^{(2)} \dots \mathbf{Q}^{(k)}$.

Proposition 7

We have the factorization (for $k \geq 1$)

$$(\mathbf{A} - \mu^{(k)}\mathbf{I})(\mathbf{A} - \mu^{(k-1)}\mathbf{I}) \cdots (\mathbf{A} - \mu^{(1)}\mathbf{I}) = \underline{\mathbf{Q}}^{(k)}\underline{\mathbf{R}}^{(k)},$$

where $\underline{\mathbf{Q}}^{(k)} := \mathbf{Q}^{(1)}\mathbf{Q}^{(2)} \cdots \mathbf{Q}^{(k)}$, and $\underline{\mathbf{R}}^{(k)} := \mathbf{R}^{(k)}\mathbf{R}^{(k-1)} \cdots \mathbf{R}^{(1)}$.

Proof. We use induction. For $k = 1$, $\mathbf{A} - \mu^{(1)}\mathbf{I} = \mathbf{Q}^{(1)}\mathbf{R}^{(1)} = \underline{\mathbf{Q}}^{(1)}\underline{\mathbf{R}}^{(1)}$. Assume $(\mathbf{A} - \mu^{(k-1)}\mathbf{I})(\mathbf{A} - \mu^{(k-2)}\mathbf{I}) \cdots (\mathbf{A} - \mu^{(1)}\mathbf{I}) = \underline{\mathbf{Q}}^{(k-1)}\underline{\mathbf{R}}^{(k-1)}$. By Proposition 6, we have $\mathbf{A}^{(k-1)} = (\underline{\mathbf{Q}}^{(k-1)})^* \mathbf{A} \underline{\mathbf{Q}}^{(k-1)}$. Then

$$\begin{aligned} & (\mathbf{A} - \mu^{(k)}\mathbf{I})(\mathbf{A} - \mu^{(k-1)}\mathbf{I}) \cdots (\mathbf{A} - \mu^{(1)}\mathbf{I}) = (\mathbf{A} - \mu^{(k)}\mathbf{I})\underline{\mathbf{Q}}^{(k-1)}\underline{\mathbf{R}}^{(k-1)} \\ &= (\mathbf{A}\underline{\mathbf{Q}}^{(k-1)} - \mu^{(k)}\underline{\mathbf{Q}}^{(k-1)})\underline{\mathbf{R}}^{(k-1)} \\ &= (\underline{\mathbf{Q}}^{(k-1)}\mathbf{A}^{(k-1)} - \mu^{(k)}\underline{\mathbf{Q}}^{(k-1)})\underline{\mathbf{R}}^{(k-1)} \\ &= \underline{\mathbf{Q}}^{(k-1)}(\mathbf{A}^{(k-1)} - \mu^{(k)}\mathbf{I})\underline{\mathbf{R}}^{(k-1)} = \underline{\mathbf{Q}}^{(k-1)}\mathbf{Q}^{(k)}\mathbf{R}^{(k)}\underline{\mathbf{R}}^{(k-1)} = \underline{\mathbf{Q}}^{(k)}\underline{\mathbf{R}}^{(k)}. \end{aligned}$$

This completes the proof. □

- Connection with shifted power iteration: By

$$(\mathbf{A} - \mu^{(k)}\mathbf{I})(\mathbf{A} - \mu^{(k-1)}\mathbf{I}) \cdots (\mathbf{A} - \mu^{(1)}\mathbf{I}) = \underline{\mathbf{Q}}^{(k)}\underline{\mathbf{R}}^{(k)},$$

the first column of $\underline{\mathbf{Q}}^{(k)}$ is the result of applying k steps of shifted power iteration on the matrix \mathbf{A} using the starting vector \mathbf{e}_1 and the shifts $\mu^{(j)}$, $j = 1 : k$.

- Connection with shifted inverse iteration: By

$$\underline{\mathbf{Q}}^{(k)} = (\mathbf{A}^* - \overline{\mu^{(k)}}\mathbf{I})^{-1}(\mathbf{A}^* - \overline{\mu^{(k-1)}}\mathbf{I})^{-1} \cdots (\mathbf{A}^* - \overline{\mu^{(1)}}\mathbf{I})^{-1}(\underline{\mathbf{R}}^{(k)})^*,$$

the last column of $\underline{\mathbf{Q}}^{(k)}$ is the result of applying k steps of shifted inverse iteration on the matrix \mathbf{A}^* using the starting vector \mathbf{e}_m and the shifts $\overline{\mu^{(j)}}$, $j = 1 : k$.

If the shifts are good eigenvalue estimates, the last column of $\underline{\mathbf{Q}}^{(k)}$, i.e., $\underline{\mathbf{Q}}^{(k)}\mathbf{e}_m$, converges quickly to an eigenvector of \mathbf{A}^* .

- Connection with Rayleigh quotient iteration: Choose

$$\begin{aligned}\mu^{(1)} &= r(\mathbf{e}_m) = \mathbf{e}_m^* \mathbf{A} \mathbf{e}_m, \\ \mu^{(k+1)} &= r(\underline{\mathbf{Q}}^{(k)} \mathbf{e}_m) = (\underline{\mathbf{Q}}^{(k)} \mathbf{e}_m)^* \mathbf{A} (\underline{\mathbf{Q}}^{(k)} \mathbf{e}_m), \quad k \geq 1,\end{aligned}$$

as the shifts. Then $\overline{\mu^{(k+1)}}$ and $\underline{\mathbf{Q}}^{(k)} \mathbf{e}_m$ are identical to those computed by the Rayleigh quotient iteration on \mathbf{A}^* starting with \mathbf{e}_m . Assume the algorithm converges. Then $\underline{\mathbf{Q}}^{(k)} \mathbf{e}_m$ converges quadratically or cubically to an eigenvector of \mathbf{A}^* .

- *Rayleigh quotient shift* $\mu^{(k+1)} = \mathbf{A}_{mm}^{(k)}$: In the QR algorithm, we have

$$\mathbf{A}_{mm}^{(k)} = \mathbf{e}_m^* \mathbf{A}^{(k)} \mathbf{e}_m = \mathbf{e}_m^* (\underline{\mathbf{Q}}^{(k)})^* \mathbf{A} \underline{\mathbf{Q}}^{(k)} \mathbf{e}_m = r(\underline{\mathbf{Q}}^{(k)} \mathbf{e}_m),$$

which means that the Rayleigh quotient $r(\underline{\mathbf{Q}}^{(k)} \mathbf{e}_m)$ appears as the (m, m) entry of $\mathbf{A}^{(k)}$. So it comes for free!

- Other issues: *Wilkinson shift ...*

4. Upper Hessenberg structure in QR algorithm

Proposition 8

Upper Hessenberg structure is preserved by QR algorithm.

Proof.

For the upper Hessenberg matrix $\mathbf{H}^{(k-1)} - \mu^{(k)}\mathbf{I}$, it is easy to show that there exists a QR factorization $\mathbf{Q}^{(k)}\mathbf{R}^{(k)} = \mathbf{H}^{(k-1)} - \mu^{(k)}\mathbf{I}$ such that $\mathbf{Q}^{(k)}$ is upper Hessenberg. Then it is easy to confirm that $\mathbf{R}^{(k)}\mathbf{Q}^{(k)}$ remains upper Hessenberg and adding $\mu^{(k)}\mathbf{I}$ does not change this. \square

Proposition 9

Hermitian tridiagonal structure is preserved by QR algorithm.

Hint: Hermitian + tridiagonal = Hermitian + upper Hessenberg. \square

For simplicity, in subsections 4.1 – 4.3, we only consider the real case.

4.1. Implicit Q theorem

Definition 10

An upper Hessenberg matrix \mathbf{H} is unreduced if all $(j+1, j)$ entries of \mathbf{H} are nonzero.

Theorem 11 (Implicit Q theorem)

Let $\mathbf{A} \in \mathbb{R}^{m \times m}$. Suppose that $\mathbf{Q}^\top \mathbf{A} \mathbf{Q} = \mathbf{H}$ is unreduced upper Hessenberg and \mathbf{Q} is orthogonal. Then columns 2 to m of \mathbf{Q} are determined uniquely (up to signs) by the first column of \mathbf{Q} .

Remark 12

*Implicit Q theorem implies that QR algorithm can be implemented cheaply on unreduced upper Hessenberg matrices. The implementation will be **implicit** in the sense that we do not explicitly compute the QR factorization of an unreduced upper Hessenberg matrix each iteration but rather construct \mathbf{Q} implicitly as a product of Givens rotations and other simple orthogonal matrices. See subsections 4.2 – 4.3 for details.*

Proof of implicit Q theorem.

Suppose that $\mathbf{Q}^\top \mathbf{A} \mathbf{Q} = \mathbf{H}$ and $\mathbf{V}^\top \mathbf{A} \mathbf{V} = \mathbf{G}$ are both unreduced upper Hessenberg, \mathbf{Q} and \mathbf{V} are orthogonal, and the first columns of \mathbf{Q} and \mathbf{V} are equal. Let $(\mathbf{X})_i$ denote the i th column of \mathbf{X} . Let $\mathbf{W} := \mathbf{V}^\top \mathbf{Q}$. By

$$\mathbf{G} \mathbf{W} = \mathbf{G} \mathbf{V}^\top \mathbf{Q} = \mathbf{V}^\top \mathbf{A} \mathbf{Q} = \mathbf{V}^\top \mathbf{Q} \mathbf{H} = \mathbf{W} \mathbf{H},$$

we have

$$\mathbf{G}(\mathbf{W})_i = \mathbf{W}(\mathbf{H})_i = \sum_{j=1}^{i+1} h_{ji}(\mathbf{W})_j.$$

Thus,

$$h_{i+1,i}(\mathbf{W})_{i+1} = \mathbf{G}(\mathbf{W})_i - \sum_{j=1}^i h_{ji}(\mathbf{W})_j.$$

Since $(\mathbf{W})_1 = \mathbf{e}_1$ and \mathbf{G} is upper Hessenberg, we can use induction on i to show that $(\mathbf{W})_i$ is nonzero in entries 1 to i only; i.e., \mathbf{W} is upper triangular. Since \mathbf{W} is also orthogonal, then \mathbf{W} is diagonal and

$$\mathbf{W} = \text{diag}\{1, \pm 1, \dots, \pm 1\},$$

which implies

$$\mathbf{V} \text{diag}\{1, \pm 1, \dots, \pm 1\} = \mathbf{Q}. \quad \square$$

4.2. Implicit single shift QR algorithm ($\mu^{(k)} \in \mathbb{R}$)

- To compute $\mathbf{H}^{(k)} = (\mathbf{Q}^{(k)})^\top \mathbf{H}^{(k-1)} \mathbf{Q}^{(k)}$ from $\mathbf{H}^{(k-1)}$ in the QR algorithm (assume that $\mathbf{H}^{(k)}$ is unreduced), we will need only to
 - (1) compute the first column of $\mathbf{Q}^{(k)}$ (which is parallel to the first column of $\mathbf{H}^{(k-1)} - \mu^{(k)} \mathbf{I}$ and so can be gotten just by normalizing this column vector).
 - (2) choose other columns of $\mathbf{Q}^{(k)}$ such that $\mathbf{Q}^{(k)}$ is orthogonal and $(\mathbf{Q}^{(k)})^\top \mathbf{H}^{(k-1)} \mathbf{Q}^{(k)}$ is unreduced upper Hessenberg.
- By the implicit Q theorem, we know that we will have computed $\mathbf{H}^{(k)}$ **correctly** because $\mathbf{Q}^{(k)}$ is unique up to signs, which do not matter. Signs do not matter because changing the signs of the columns of $\mathbf{Q}^{(k)}$ is the same as changing $\mathbf{Q}^{(k)} \mathbf{R}^{(k)} = \mathbf{H}^{(k-1)} - \mu^{(k)} \mathbf{I}$ to

$$(\mathbf{Q}^{(k)} \mathbf{S}^{(k)}) (\mathbf{S}^{(k)} \mathbf{R}^{(k)}) = \mathbf{H}^{(k-1)} - \mu^{(k)} \mathbf{I},$$

where $\mathbf{S}^{(k)} = \text{diag}\{1, \pm 1, \dots, \pm 1\}$.

- To see how to use the implicit Q theorem, we use a 5×5 example.

$$1. \mathbf{Q}_1^\top = \begin{bmatrix} c_1 & s_1 & & & \\ -s_1 & c_1 & & & \\ & & 1 & & \\ & & & 1 & \\ & & & & 1 \end{bmatrix}, \mathbf{H}_1 := \mathbf{Q}_1^\top \mathbf{H} \mathbf{Q}_1 = \begin{bmatrix} \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ + & \times & \times & \times & \times \\ 0 & 0 & \times & \times & \times \\ 0 & 0 & 0 & \times & \times \end{bmatrix}$$

Here c_1 and s_1 are unknown, and $\mathbf{H} = \mathbf{H}^{(k-1)}$.

$$2. \mathbf{Q}_2^\top = \begin{bmatrix} 1 & & & & \\ & c_2 & s_2 & & \\ & -s_2 & c_2 & & \\ & & & 1 & \\ & & & & 1 \end{bmatrix}, \mathbf{Q}_2^\top \mathbf{H}_1 = \begin{bmatrix} \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times \\ 0 & 0 & \times & \times & \times \\ 0 & 0 & 0 & \times & \times \end{bmatrix},$$

$$\mathbf{H}_2 := \mathbf{Q}_2^\top \mathbf{H}_1 \mathbf{Q}_2 = \begin{bmatrix} \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times \\ 0 & + & \times & \times & \times \\ 0 & 0 & 0 & \times & \times \end{bmatrix}$$

$$3. \mathbf{Q}_3^\top = \begin{bmatrix} 1 & & & & \\ & 1 & & & \\ & & c_3 & s_3 & \\ & & -s_3 & c_3 & \\ & & & & 1 \end{bmatrix}, \quad \mathbf{Q}_3^\top \mathbf{H}_2 = \begin{bmatrix} \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times \\ 0 & 0 & \times & \times & \times \\ 0 & 0 & 0 & \times & \times \end{bmatrix},$$

$$\mathbf{H}_3 := \mathbf{Q}_3^\top \mathbf{H}_2 \mathbf{Q}_3 = \begin{bmatrix} \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times \\ 0 & 0 & \times & \times & \times \\ 0 & 0 & + & \times & \times \end{bmatrix}$$

$$4. \mathbf{Q}_4^\top = \begin{bmatrix} 1 & & & & \\ & 1 & & & \\ & & 1 & & \\ & & & c_4 & s_4 \\ & & & -s_4 & c_4 \end{bmatrix}, \quad \mathbf{Q}_4^\top \mathbf{H}_3 = \begin{bmatrix} \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times \\ 0 & 0 & \times & \times & \times \\ 0 & 0 & 0 & \times & \times \end{bmatrix},$$

$$\mathbf{H}_4 := \mathbf{Q}_4^\top \mathbf{H}_3 \mathbf{Q}_4 = \begin{bmatrix} \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times \\ 0 & 0 & \times & \times & \times \\ 0 & 0 & 0 & \times & \times \end{bmatrix}$$

Altogether $\mathbf{Q}^\top \mathbf{H}^{(k-1)} \mathbf{Q} = \mathbf{H}_4$ is upper Hessenberg, where

$$\mathbf{Q} = \mathbf{Q}_1 \mathbf{Q}_2 \mathbf{Q}_3 \mathbf{Q}_4 = \begin{bmatrix} c_1 & \times & \times & \times & \times \\ s_1 & \times & \times & \times & \times \\ & s_2 & \times & \times & \times \\ & & s_3 & \times & \times \\ & & & s_4 & c_4 \end{bmatrix}.$$

The first column of \mathbf{Q} is $[c_1 \ s_1 \ 0 \ \cdots \ 0]^\top$, which by the implicit Q theorem has uniquely determined the other columns of \mathbf{Q} (up to signs). We now choose the first column of \mathbf{Q} to be proportional to the first column of $\mathbf{H}^{(k-1)} - \mu^{(k)} \mathbf{I}$. Then we have $\mathbf{Q} = \mathbf{Q}^{(k)} \text{diag}\{1, \pm 1, \dots, \pm 1\}$, which means \mathbf{Q} is the Q-factor of a QR factorization of $\mathbf{H}^{(k-1)} - \mu^{(k)} \mathbf{I}$.

4.3. Implicit double shift QR algorithm ($\mu^{(k)} \in \mathbb{C}$)

- We describe how to maintain real arithmetic by shifting $\mu^{(k)}$ and $\overline{\mu^{(k)}}$ in succession:

$$\begin{aligned}\mathbf{Q}^{(k-1/2)} \mathbf{R}^{(k-1/2)} &= \mathbf{H}^{(k-1)} - \mu^{(k)} \mathbf{I} \\ \mathbf{H}^{(k-1/2)} &= \mathbf{R}^{(k-1/2)} \mathbf{Q}^{(k-1/2)} + \mu^{(k)} \mathbf{I} \\ &= (\mathbf{Q}^{(k-1/2)})^* \mathbf{H}^{(k-1)} \mathbf{Q}^{(k-1/2)}\end{aligned}$$

$$\begin{aligned}\mathbf{Q}^{(k)} \mathbf{R}^{(k)} &= \mathbf{H}^{(k-1/2)} - \overline{\mu^{(k)}} \mathbf{I} \\ \mathbf{H}^{(k)} &= \mathbf{R}^{(k)} \mathbf{Q}^{(k)} + \overline{\mu^{(k)}} \mathbf{I} \\ &= (\mathbf{Q}^{(k)})^* \mathbf{H}^{(k-1/2)} \mathbf{Q}^{(k)} \\ &= (\mathbf{Q}^{(k-1/2)} \mathbf{Q}^{(k)})^* \mathbf{H}^{(k-1)} \mathbf{Q}^{(k-1/2)} \mathbf{Q}^{(k)}\end{aligned}$$

Lemma 13

Assume $\mathbf{H}^{(0)} = \mathbf{H}$ is real. We can choose $\mathbf{Q}^{(k-1/2)}$ and $\mathbf{Q}^{(k)}$ such that

- (1) $\mathbf{Q}^{(k-1/2)}\mathbf{Q}^{(k)}$ is real,
- (2) $\mathbf{H}^{(k)}$ is therefore real,
- (3) the first column of $\mathbf{Q}^{(k-1/2)}\mathbf{Q}^{(k)}$ is easy to compute.

Proof. Since

$$\mathbf{Q}^{(k)}\mathbf{R}^{(k)} = \mathbf{H}^{(k-1/2)} - \overline{\mu^{(k)}}\mathbf{I} = \mathbf{R}^{(k-1/2)}\mathbf{Q}^{(k-1/2)} + (\mu^{(k)} - \overline{\mu^{(k)}})\mathbf{I},$$

we get

$$\begin{aligned} & \mathbf{Q}^{(k-1/2)}\mathbf{Q}^{(k)}\mathbf{R}^{(k)}\mathbf{R}^{(k-1/2)} \\ &= \mathbf{Q}^{(k-1/2)}(\mathbf{R}^{(k-1/2)}\mathbf{Q}^{(k-1/2)} + (\mu^{(k)} - \overline{\mu^{(k)}})\mathbf{I})\mathbf{R}^{(k-1/2)} \\ &= \mathbf{Q}^{(k-1/2)}\mathbf{R}^{(k-1/2)}\mathbf{Q}^{(k-1/2)}\mathbf{R}^{(k-1/2)} + (\mu^{(k)} - \overline{\mu^{(k)}})\mathbf{Q}^{(k-1/2)}\mathbf{R}^{(k-1/2)} \\ &= (\mathbf{H}^{(k-1)} - \mu^{(k)}\mathbf{I})^2 + (\mu^{(k)} - \overline{\mu^{(k)}})(\mathbf{H}^{(k-1)} - \mu^{(k)}\mathbf{I}) \\ &= (\mathbf{H}^{(k-1)})^2 - 2\operatorname{Re}(\mu^{(k)})\mathbf{H}^{(k-1)} + |\mu^{(k)}|^2\mathbf{I} =: \mathbf{M}. \end{aligned}$$

Note that

$$\mathbf{Q}^{(k-1/2)} \mathbf{Q}^{(k)} \mathbf{R}^{(k)} \mathbf{R}^{(k-1/2)} = \mathbf{M}$$

is a QR factorization of the real matrix \mathbf{M} . Therefore, $\mathbf{Q}^{(k-1/2)} \mathbf{Q}^{(k)}$ and $\mathbf{R}^{(k)} \mathbf{R}^{(k-1/2)}$ can be chosen real. This means that

$$\mathbf{H}^{(k)} = (\mathbf{Q}^{(k-1/2)} \mathbf{Q}^{(k)})^* \mathbf{H}^{(k-1)} \mathbf{Q}^{(k-1/2)} \mathbf{Q}^{(k)}$$

also is real if $\mathbf{H}^{(k-1)}$ is real. The first column of $\mathbf{Q}^{(k-1/2)} \mathbf{Q}^{(k)}$ is proportional to the first column of

$$(\mathbf{H}^{(k-1)})^2 - 2\operatorname{Re}(\mu^{(k)})\mathbf{H}^{(k-1)} + |\mu^{(k)}|^2 \mathbf{I},$$

whose sparsity pattern is $[\times \quad \times \quad \times \quad 0 \quad \cdots \quad 0]^\top$. Obviously, the first column of $\mathbf{Q}^{(k-1/2)} \mathbf{Q}^{(k)}$ is easy to compute since $\mathbf{H}^{(k-1)}$ is upper Hessenberg. □

The implicit Q theorem and the last lemma can be used to compute $\mathbf{H}^{(k)}$ from $\mathbf{H}^{(k-1)}$.

- We provide a 6×6 example. Let $\mathbf{H} = \mathbf{H}^{(k-1)}$.

1. Choose an orthogonal matrix

$$\mathbf{Q}_1^\top = \begin{bmatrix} \tilde{\mathbf{Q}}^\top & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix}, \quad \tilde{\mathbf{Q}}^\top \tilde{\mathbf{Q}} = \mathbf{I}_3,$$

where the first column of \mathbf{Q}_1 is proportional to the first column of

$$\mathbf{H}^2 - 2\operatorname{Re}(\mu^{(k)})\mathbf{H} + |\mu^{(k)}|^2\mathbf{I},$$

so

$$\mathbf{Q}_1^\top \mathbf{H} = \begin{bmatrix} \times & \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times & \times \\ + & \times & \times & \times & \times & \times \\ 0 & 0 & \times & \times & \times & \times \\ 0 & 0 & 0 & \times & \times & \times \\ 0 & 0 & 0 & 0 & \times & \times \end{bmatrix}, \quad \mathbf{Q}_1^\top \mathbf{H} \mathbf{Q}_1 = \begin{bmatrix} \times & \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times & \times \\ + & \times & \times & \times & \times & \times \\ + & + & \times & \times & \times & \times \\ 0 & 0 & 0 & \times & \times & \times \\ 0 & 0 & 0 & 0 & \times & \times \end{bmatrix}$$

2. Choose a Householder reflector \mathbf{Q}_2^\top , which affects only rows 2,3, and 4 of $\mathbf{H}_1 := \mathbf{Q}_1^\top \mathbf{H} \mathbf{Q}_1$, zeroing out entries (3,1) and (4,1) of \mathbf{H}_1 (this means that \mathbf{Q}_2^\top is the identity matrix outside rows and columns 2 through 4):

$$\mathbf{Q}_2^\top = \begin{bmatrix} 1 & & & & & \\ & \times & \times & \times & & \\ & \times & \times & \times & & \\ & \times & \times & \times & & \\ & & & & 1 & \\ & & & & & 1 \end{bmatrix}, \quad \mathbf{Q}_2^\top \mathbf{H}_1 = \begin{bmatrix} \times & \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times & \times \\ 0 & + & \times & \times & \times & \times \\ 0 & 0 & 0 & \times & \times & \times \\ 0 & 0 & 0 & 0 & \times & \times \end{bmatrix},$$

$$\mathbf{H}_2 := \mathbf{Q}_2^\top \mathbf{H}_1 \mathbf{Q}_2 = \begin{bmatrix} \times & \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times & \times \\ 0 & + & \times & \times & \times & \times \\ 0 & + & + & \times & \times & \times \\ 0 & 0 & 0 & 0 & \times & \times \end{bmatrix}$$

3. Choose a Householder reflector \mathbf{Q}_3^\top , which affects only rows 3,4, and 5 of \mathbf{H}_2 , zeroing out entries (4,2) and (5,2) of \mathbf{H}_2 (this means that \mathbf{Q}_3^\top is the identity matrix outside rows and columns 3 through 5):

$$\mathbf{Q}_3^\top = \begin{bmatrix} 1 & & & & \\ & 1 & & & \\ & & \times & \times & \times \\ & & \times & \times & \times \\ & & \times & \times & \times \\ & & & & 1 \end{bmatrix}, \quad \mathbf{Q}_3^\top \mathbf{H}_2 = \begin{bmatrix} \times & \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times & \times \\ 0 & 0 & \times & \times & \times & \times \\ 0 & 0 & + & \times & \times & \times \\ 0 & 0 & 0 & 0 & \times & \times \end{bmatrix},$$

$$\mathbf{H}_3 := \mathbf{Q}_3^\top \mathbf{H}_2 \mathbf{Q}_3 = \begin{bmatrix} \times & \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times & \times \\ 0 & 0 & \times & \times & \times & \times \\ 0 & 0 & + & \times & \times & \times \\ 0 & 0 & + & + & \times & \times \end{bmatrix}$$

4. Choose a Householder reflector \mathbf{Q}_4^\top , which affects only rows 4,5, and 6 of \mathbf{H}_3 , zeroing out entries (5,3) and (6,3) of \mathbf{H}_3 (this means that \mathbf{Q}_4^\top is the identity matrix outside rows and columns 4 through 6):

$$\mathbf{Q}_4^\top = \begin{bmatrix} 1 & & & & & \\ & 1 & & & & \\ & & 1 & & & \\ & & & \times & \times & \times \\ & & & \times & \times & \times \\ & & & \times & \times & \times \end{bmatrix}, \quad \mathbf{Q}_4^\top \mathbf{H}_3 = \begin{bmatrix} \times & \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times & \times \\ 0 & 0 & \times & \times & \times & \times \\ 0 & 0 & 0 & \times & \times & \times \\ 0 & 0 & 0 & + & \times & \times \end{bmatrix},$$

$$\mathbf{H}_4 := \mathbf{Q}_4^\top \mathbf{H}_3 \mathbf{Q}_4 = \begin{bmatrix} \times & \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times & \times \\ 0 & 0 & \times & \times & \times & \times \\ 0 & 0 & 0 & \times & \times & \times \\ 0 & 0 & 0 & \times & \times & \times \end{bmatrix}$$

5. Choose a Givens rotation \mathbf{Q}_5^\top

$$\mathbf{Q}_5^\top = \begin{bmatrix} 1 & & & & & \\ & 1 & & & & \\ & & 1 & & & \\ & & & 1 & & \\ & & & & c & s \\ & & & & -s & c \end{bmatrix}, \quad \mathbf{Q}_5^\top \mathbf{H}_4 = \begin{bmatrix} \times & \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times & \times \\ 0 & 0 & \times & \times & \times & \times \\ 0 & 0 & 0 & \times & \times & \times \\ 0 & 0 & 0 & 0 & \times & \times \end{bmatrix},$$

$$\mathbf{H}_5 = \mathbf{Q}_5^\top \mathbf{H}_4 \mathbf{Q}_5 = \begin{bmatrix} \times & \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times & \times \\ 0 & 0 & \times & \times & \times & \times \\ 0 & 0 & 0 & \times & \times & \times \\ 0 & 0 & 0 & 0 & \times & \times \end{bmatrix}.$$

Altogether $\mathbf{Q}^\top \mathbf{H}^{(k-1)} \mathbf{Q} = \mathbf{H}_5$ is upper Hessenberg, where

$$\mathbf{Q} = \mathbf{Q}_1 \mathbf{Q}_2 \mathbf{Q}_3 \mathbf{Q}_4 \mathbf{Q}_5 \quad \text{with} \quad \mathbf{Q} \mathbf{e}_1 = \mathbf{Q}_1 \mathbf{e}_1.$$

4.4. Two phases of QR algorithm

- First phase: reduce to an upper Hessenberg matrix

$$\begin{bmatrix} \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \end{bmatrix} \xrightarrow{\mathbf{Q}_1^*} \begin{bmatrix} \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times \end{bmatrix} \xrightarrow{\cdot \mathbf{Q}_1} \begin{bmatrix} \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \end{bmatrix}$$

$\mathbf{A} \qquad \qquad \mathbf{Q}_1^* \mathbf{A} \qquad \qquad \mathbf{Q}_1^* \mathbf{A} \mathbf{Q}_1$

$$\begin{bmatrix} \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \end{bmatrix} \xrightarrow{\mathbf{Q}_2^*} \begin{bmatrix} \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times \end{bmatrix} \xrightarrow{\cdot \mathbf{Q}_2} \begin{bmatrix} \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ & \times & \times & \times & \times \\ & \times & \times & \times & \times \\ & \times & \times & \times & \times \end{bmatrix}$$

$\mathbf{Q}_1^* \mathbf{A} \mathbf{Q}_1 \qquad \qquad \mathbf{Q}_2^* \mathbf{Q}_1^* \mathbf{A} \mathbf{Q}_1 \qquad \qquad \mathbf{Q}_2^* \mathbf{Q}_1^* \mathbf{A} \mathbf{Q}_1 \mathbf{Q}_2$

$$\begin{bmatrix} \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ & \times & \times & \times & \times \\ & & \times & \times & \times \\ & & & \times & \times \end{bmatrix} \qquad \underbrace{\mathbf{Q}_{m-2}^* \cdots \mathbf{Q}_2^* \mathbf{Q}_1^*}_{\mathbf{Q}^*} \mathbf{A} \underbrace{\mathbf{Q}_1 \mathbf{Q}_2 \cdots \mathbf{Q}_{m-2}}_{\mathbf{Q}} = \mathbf{H}.$$

- Second phase: generate a sequence of upper Hessenberg (or tridiagonal) matrices that converge to an upper triangular (or diagonal) matrix.

$$\begin{array}{ccc}
 \begin{bmatrix} \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \end{bmatrix} & \xrightarrow{\text{Phase 1}} & \begin{bmatrix} \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ & \times & \times & \times & \times \\ & & \times & \times & \times \\ & & & \times & \times \end{bmatrix} & \xrightarrow{\text{Phase 2}} & \begin{bmatrix} \times & \times & \times & \times & \times \\ & \times & \times & \times & \times \\ & & \times & \times & \times \\ & & & \times & \times \\ & & & & \times \end{bmatrix} \\
 \mathbf{A} \neq \mathbf{A}^* & & \mathbf{H} & & \mathbf{T}
 \end{array}$$

$$\begin{array}{ccc}
 \begin{bmatrix} \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \end{bmatrix} & \xrightarrow{\text{Phase 1}} & \begin{bmatrix} \times & \times & & & \\ \times & \times & \times & & \\ & \times & \times & \times & \\ & & \times & \times & \times \\ & & & \times & \times \end{bmatrix} & \xrightarrow{\text{Phase 2}} & \begin{bmatrix} \times & & & & \\ & \times & & & \\ & & \times & & \\ & & & \times & \\ & & & & \times \end{bmatrix} \\
 \mathbf{A} = \mathbf{A}^* & & \mathbf{H} & & \mathbf{D}
 \end{array}$$

5. Further reading

- J. L. Aurentz, T. Mach, L. Robol, R. Vandebril, and D. S. Watkins
Core-Chasing Algorithms for the Eigenvalue Problem
Society for Industrial and Applied Mathematics (SIAM), 2018
- A. Casulli and L. Robol
Rank-structured QR for Chebyshev rootfinding
SIAM J. Matrix Anal. Appl., 42, 1148–1171, 2021.
- D. S. Watkins
Understanding the QR algorithm
SIAM Review, 24, 427–440, 1982.
The QR algorithm revisited
SIAM Review, 50, 133–145, 2008.