Linear Analysis
Lecture 17

The QR for Algorithm for the Shur Factorization

The QR algorithm is used to compute a specific Schur unitary triangularization of a matrix $A \in \mathbb{C}^{n \times n}$.
The algorithm is **iterative**. That is, we generate a sequence

$$A = A_0,\ A_1,\ A_2,\ \ldots$$

of matrices that are unitarily similar to A. The goal is to get the subdiagonal elements to converge to zero, since then the eigenvalues will appear on the diagonal.
If A is Hermitian, then so are $A_1,\ A_2,\ \ldots$, so if the subdiagonal elements $\to 0$, also the superdiagonal elements converge to 0, and (in the limit) we have diagonalized A.

Variations of the QR algorithm are the most commonly used methods for computing all the eigenvalues (and eigenvectors if wanted) of a matrix. It behaves well numerically since all the similarity transformations are unitary.

Upper-Hessenberg Form

When used in practice, a matrix is first reduced to **upper-Hessenberg form**

$$\begin{bmatrix} * & \cdots & * & * \\ * & \dots & * & * \\ 0 & \ddots & & \vdots \\ 0 & \dots & * & * \end{bmatrix}$$

($h_{ij} = 0$ for $i > j + 1$) using unitary similarity transformations built from Householder reflections (or Givens rotations), quite analogous to computing a QR factorization.

However, since similarity transformations are being performed, we require left and right multiplication by the Householder transformations — leading to an inability to zero out the first subdiagonal $(i = j + 1)$ in the process.

If A is Hermitian and upper-Hessenberg, A is tridiagonal.

This initial reduction decreases the computational cost of the QR algorithm. It is successful because the upper-Hessenberg form is preserved by the iterations: if A_k is upper Hessenberg, so is A_{k+1}.

There are many variants of the QR algorithm. We consider the basic algorithm over $\mathbb{C}$.

The Basic QR Algorithm

Given $A \in \mathbb{C}^{n\times n}$, let $A_0 = A$. For $k = 0, 1, 2, \ldots$, starting with A_k, do a QR factorization of A_k:

$$A_k = Q_k R_k.$$

The set

$$A_{k+1} = R_k Q_k \; .$$

Remark

$$R_k = Q_k^H A_k \quad \text{so} \quad A_{k+1} = Q_k^H A_k Q_k$$

is unitarily similar to A_k. In general,

$$A_{k+1} = Q_k^H Q_{k-1}^H \ldots Q_0^H A_0 Q_0 \ldots Q_{k-1} Q_k \; .$$

Thus

$$\|A_k\| = \|A_0\| \qquad \forall\ k = 1, 2, \ldots \; .$$

The algorithm uses the Q of the QR factorization of A_k to perform the next unitary similarity transformation.

Convergence of the QR Algorithm

We now show convergence. under the hypotheses that $A \in \mathbb{C}^{n\times n}$ has eigenvalues $\lambda_1, \ldots, \lambda_n$ with $|\lambda_1| > |\lambda_2| > \cdots > |\lambda_n| > 0$.

Lemma. Let $Q_j \in \mathbb{C}^{n\times n}$ $j = 1, 2, \ldots$ be unitary matrices, and let $R_j \in \mathbb{C}^{n\times n}$ $j = 1, 2, \ldots$ be upper triangular matrices with positive diagonal entries. Suppose

$$Q_j R_j \to I \quad \text{as} \quad j \to \infty \ .$$

Then $Q_j \to I$ and $R_j \to I$.

Proof Sketch. Let Q_{j_k} be any subsequence of Q_j. Since the set of unitary matrices in $\mathbb{C}^{n\times n}$ is compact, there exists a sub-subsequence $Q_{j_{k_l}}$ and a unitary Q such that $Q_{j_{k_l}} \to Q$. So

$$R_{j_{k_l}} = Q_{j_{k_l}}^H Q_{j_{k_l}} R_{j_{k_l}} \to Q^H \cdot I = Q^H.$$

So Q^H is unitary, upper triangular, with nonnegative diagonal elements, which implies easily that $Q^H = I$. Thus every subsequence of Q_j has in turn a sub-subsequence converging to I. Consequently, $Q_j \to I$, and thus $R_j = Q_j^H Q_j R_j \to I \cdot I = I$. □

Convergence of the QR Algorithm

Theorem: Suppose $A \in \mathbb{C}^{n \times n}$ has eigenvalues

$$\lambda_1, \ldots, \lambda_n \quad \text{with} \quad |\lambda_1| > |\lambda_2| > \cdots > |\lambda_n| > 0 \ .$$

Choose $X \in \mathbb{C}^{n \times n}$ such that

$$X^{-1} A X = \Lambda \equiv \operatorname{diag}(\lambda_1, \ldots, \lambda_n).$$

Suppose X^{-1} has an LU decomposition. Generate the sequence

$$A \ = \ A_0, \ A_1, \ A_2, \ \ldots$$

using the QR algorithm. Then

the subdiagonal entries of $A_k \to 0$ as $k \to \infty$,

and

the j^{th} diagonal entry $\to \lambda_j \quad 1 \leq j \leq n$.

Proof

Define

$$\widetilde{Q}_k = Q_0 Q_1 \cdots Q_k \quad \text{and} \quad \widetilde{R}_k = R_k \cdots R_0.$$

Then $A_{k+1} = \widetilde{Q}_k^H A \widetilde{Q}_k$.

Claim: $\widetilde{Q}_k \widetilde{R}_k = A^{k+1}$

Proof: Proceed by induction. Clearly the Claim holds for $k = 0$.
Suppose $\widetilde{Q}_{k-1} \widetilde{R}_{k-1} = A^k$. Then

$$R_k = A_{k+1} Q_k^H = \widetilde{Q}_k^H A \widetilde{Q}_k Q_k^H = \widetilde{Q}_k^H A \widetilde{Q}_{k-1},$$

so

$$\widetilde{R}_k = R_k \widetilde{R}_{k-1} = \widetilde{Q}_k^H A \widetilde{Q}_{k-1} \widetilde{R}_{k-1} = \widetilde{Q}_k^H A^{k+1},$$

so $\widetilde{Q}_k \widetilde{R}_k = A^{k+1}$. $\square$

Proof

Recall

$$X^{-1}AX = \Lambda \equiv \mathrm{diag}\,(\lambda_1, \ldots, \lambda_n).$$

Set

$X = QR,$		$X^{-1} = LU,$	
Q	unitary,	L	unit lower triangular,
R	nonsingular upper triangular,	U	nonsingular upper triangular.

Then

$$\begin{aligned} A^{k+1} &= X\Lambda^{k+1}X^{-1} \\ &= QR\Lambda^{k+1}LU \\ &= QR(\Lambda^{k+1}L\Lambda^{-(k+1)})\Lambda^{k+1}U. \end{aligned}$$

Let

$$E_{k+1} = \Lambda^{k+1}L\Lambda^{-(k+1)} - I \quad \text{and} \quad F_{k+1} = RE_{k+1}R^{-1}.$$

Proof

Claim: $E_{k+1} \to 0$ (and thus $F_{k+1} \to 0$) as $k \to \infty$.
Proof: Let ℓ_{ij} denote the elements of L. E_{k+1} is strictly lower triangular, and for $i > j$ its ij element is

$$\left(\frac{\lambda_i}{\lambda_j}\right)^{k+1} \ell_{ij} \to 0 \quad \text{since } |\lambda_i| < |\lambda_j|. \ \square$$

Now $\quad A^{k+1} = QR(I + E_{k+1})\Lambda^{k+1}U = Q(I + F_{k+1})R\Lambda^{k+1}U.$
Choose a QR factorization of $I + F_{k+1}$ (which is invertible)

$$I + F_{k+1} = \widehat{Q}_{k+1}\widehat{R}_{k+1} \to I,$$

where $\widehat{R}_{k+1}$ has positive diagonal entries. By the Lemma,

$$\widehat{Q}_{k+1} \to I \quad \text{and} \quad \widehat{R}_{k+1} \to I.$$

Since

$$A^{k+1} = Q(\widehat{Q}_{k+1}\widehat{R}_{k+1})R\Lambda^{k+1}U = (Q\widehat{Q}_{k+1})(\widehat{R}_{k+1}R\Lambda^{k+1}U)$$

and $\quad A^{k+1} = \widetilde{Q}_k\widetilde{R}_k$, the essential uniqueness of QR factorizations of invertible matrices implies there exists a unitary diagonal matrix D_k for which

$$Q\widehat{Q}_{k+1}D_k^H = \widetilde{Q}_k \quad \text{and} \quad D_k\widehat{R}_{k+1}\Lambda^{k+1}U = \widetilde{R}_k.$$

Proof

So $\widetilde{Q}_k D_k = Q\widehat{Q}_{k+1} \to Q$, and thus

$$D_k^H A_{k+1} D_k = D_k^H \widetilde{Q}_k^H A \widetilde{Q}_k D_k \to Q^H AQ.$$

But

$$\begin{aligned} Q^H AQ &= Q^H(X\Lambda X^{-1})Q \\ &= Q^H(QR\Lambda X^{-1})QRR^{-1} \\ &= R\Lambda R^{-1} \end{aligned}$$

is upper triangular with diagonal entries $\lambda_1, \ldots, \lambda_n$ in that order. Since D_k is unitary and diagonal, the diagonal and lower triangular entries of $R\Lambda R^{-1}$ and of $D_k R\Lambda R^{-1} D_k^H$ are the same, namely $\operatorname{diag}[\lambda_1, \ldots, \lambda_n]$ and

$$\|A_{k+1} - D_k R\Lambda R^{-1} D_k^H\| = \|D_k^H A_{k+1} D_k - R\Lambda R^{-1}\| \to 0.$$

The Theorem follows. □

Comments on the Proof

Note that the proof shows that there is a sequence $\{D_k\}$ of unitary diagonal matrices for which $D_k^H A_{k+1} D_k \to R\Lambda R^{-1}$. So although the superdiagonal $(i < j)$ elements of A_{k+1} may not converge, the magnitude of each superdiagonal element converges.

As a partial explanation for why the QR algorithm works, we show how the convergence of the first column of A_k to $[\lambda_1, 0, \ldots, 0]$ follows from the power method.

Suppose $A \in \mathbb{C}^{n\times n}$ is diagonalizable and has a unique e-value $\lambda_1 > 0$ of maximum modulus with unit e-vec x_1. Then if $x \in \mathbb{C}^n$ has $\langle x, x_1\rangle \neq 0$, then $A^k x/\|A^k x\| \to x_1/\|x_1\|$.

If X^{-1} has an LU factorization, the $(1,1)$ entry of X^{-1} is nonzero. Thus when e_1 is expanded in terms of the eigenvectors $x_1, \ldots, x_n$ (cols. of X), the x_1-coefficient is nonzero. So

$$A^{k+1}e_1/\|A^{k+1}e_1\| \to \alpha x_1 \quad \text{for some } \alpha \in \mathbb{C} \quad \text{with } |\alpha| = 1.$$

Let $(\widetilde{q}_k)_1$ denote the first column of $\widetilde{Q}_k$ and $(\widetilde{r}_k)_{11}$ denote the $(1,1)$-entry of $\widetilde{R}_k$. Then

$$A^{k+1}e_1 = \widetilde{Q}_k\widetilde{R}_k e_1 = (\widetilde{r}_k)_{11}\widetilde{Q}_k e_1 = (\widetilde{r}_k)_{11}(\widetilde{q}_k)_1,$$

so $(\widetilde{q}_k)_1 \to \alpha x_1$. Since $A_{k+1} = \widetilde{Q}_k^H A \widetilde{Q}_k$, the first column of A_{k+1} converges to $[\lambda_1, 0, \ldots, 0]$.