LECTURE 25

QR ALGORITHM WITH SHIFTS

QR algorithm, when it converges, converges to an upper triangular matrix only q-linearly, i.e.

$$\underbrace{Q_k^* \cdots Q_1^* \underbrace{A}_{\text{Hessenberg}} Q_1 \cdots Q_k}_{A_k} \longrightarrow \underbrace{T}_{\substack{\text{upper} \\ \text{triangular}}} \quad \text{as } k \to \infty$$

and subdiagonal entries of A approach zero q-linearly. For instance typically

$$\lim_{k\to\infty} \frac{|a_{n,n-1}^{(k+1)} - 0|}{|a_{n,n-1}^{(k)} - 0|} = \frac{|\lambda_n|}{|\lambda_{n-1}|}$$

$a_{n,n-1}^{(k)}$: $(n,n-1)$ entry of A_k

λ_n : Smallest eigenvalue of A in modulus

λ_{n-1} : Second smallest eigenvalue of A in modulus

$$\begin{bmatrix} X & X & \cdots & X \\ X & X & \cdots & X \\ & X & \cdots & X \\ 0 & & \ddots & \vdots \\ & & X & X \end{bmatrix} \longrightarrow \begin{bmatrix} X & X & \cdots & X & X \\ X & X & \cdots & X & X \\ & X & \ddots & \vdots & \vdots \\ & & X & X & X \\ & & & \underbrace{\approx 0}_{a_{n,n-1}^{(k)}} & \underbrace{X}_{a_{n,n}^{(k)}} \end{bmatrix}$$

(almost an eigenvalue)

To speed up convergence shift A_k by $\mu_k = a_{n,n}^{(k)}$

Define the sequence $\{A_k\}$ such that

(i) $A_0 = A$

$\boxed{\text{FOR } k \geq 0}$ (ii) $(A_k - \mu_k I) = Q_{k+1} R_{k+1}$ and $A_{k+1} = R_{k+1} Q_{k+1} + \mu_k I$

ALGORITHM (QR Algorithm with Shifts)

* Given $A \in \mathbb{C}^{n \times n}$ in Hessenberg form.
* Produce a sequence $\{A_k\}$ such that typically $\lim_{k \to \infty} A_k$ is upper triangular.

$A_0 = A$

for $k = 0, 1, \ldots$

  Choose a shift μ_k

  Compute a QR factorization $A_k - \mu_k I = Q_{k+1} R_{k+1}$

  $A_{k+1} = R_{k+1} Q_{k+1} + \mu_k I$

end

COMMON SHIFT STRATEGIES

Rayleigh shift: $\mu_k = a_{n,n}^{(k)}$

Wilkinson shift: μ_k is the eigenvalue of $A_k(n-1:n, n-1:n)$ closest to $a_{n,n}^{(k)}$

EXAMPLE

$$A = \begin{bmatrix} 3 & 4 \\ 2 & 1 \end{bmatrix}$$ has eigenvalues $\lambda_1 = 5$, $\lambda_2 = -1$

Apply QR algorithm with fixed shifts $\mu = -0.8$

$$Q_1 R_1 = A + 0.8I$$

$$A_1 = R_1 Q_1 - 0.8I = \begin{bmatrix} 5.04 & -1.87 \\ 0.13 & -1.04 \end{bmatrix}$$

$$Q_2 R_2 = A_1 + 0.8I$$

$$A_2 = R_2 Q_2 - 0.8I = \begin{bmatrix} 5.00 & 2.00 \\ 0.00 & -1.00 \end{bmatrix}$$

REMARKS

(i) A_{k+1} and A_k are unitarily similar.

$$(A_k - \mu_k I) = Q_{k+1} R_{k+1} \text{ and } A_{k+1} = R_{k+1} Q_{k+1} + \mu_k I$$

$$\Longrightarrow$$

$$R_{k+1} = Q_{k+1}^* (A_k - \mu_k I) \text{ and } A_{k+1} = R_{k+1} Q_{k+1} + \mu_k I$$

$$\Longrightarrow$$

$$A_{k+1} = Q_{k+1}^* (A_k - \mu_k I) Q_{k+1} + \mu_k I$$

$$= Q_{k+1}^* A_k Q_{k+1}$$

(ii) Rate of convergence with Rayleigh or Wilkinson shifts is typically q-quadratic.

QR ALGORITHM WITH DOUBLE SHIFTS

In practice Wilkinson shift is the standard shift strategy used.

When the matrix A is real, ideally we like to remain in real arithmetic, i.e. ideally we like A_k to be real for all k.

The Wilkinson shift μ_k can be complex, but the matrix

$$(A_k - \mu_k I)(A_k - \overline{\mu}_k I)$$

$$=$$

$$A_k^2 - (\mu_k + \overline{\mu}_k) A_k + |\mu_k|^2 I$$

is real provided A_k is real.

The standard double-shift strategy with Wilkinson shifts μ_k generates a sequence $\{A_k\}$ such that

(i) $A_0 = A$

FOR EVEN $k \geq 0$ $\begin{cases} \text{(ii)} & (A_k - \mu_k I) = Q_{k+1} R_{k+1} \text{ and } A_{k+1} = R_{k+1} Q_{k+1} + \mu_k I \\ \text{(iii)} & (A_{k+1} - \overline{\mu}_k I) = Q_{k+2} R_{k+2} \text{ and } A_{k+2} = R_{k+2} Q_{k+2} + \overline{\mu}_k I \end{cases}$

REMARK

$$A_{k+2} = Q_{k+2}^{*} \; Q_{k+1}^{*} \, A_k \, Q_{k+1} \, Q_{k+2} \quad \text{is real.}$$

since

$$(A_k - \bar{\mu}_k I)(A_k - \mu_k I)$$

$$= (A_k - \bar{\mu}_k I) \, Q_{k+1} \, R_{k+1}$$

$$= Q_{k+1} \, (A_{k+1} - \bar{\mu}_k I) \, R_{k+1}$$

$$= \underbrace{Q_{k+1} \, Q_{k+2} \, R_{k+2} \, R_{k+1}}_{\text{QR factorization}}$$

of $(A_k - \bar{\mu}_k I)(A_k - \mu_k I)$

NOTE

$A_{k+1} = Q_{k+1}^{*} A_k Q_{k+1}$

$\Longrightarrow$

$(A_{k+1} - \bar{\mu}_k I) = Q_{k+1}^{*} (A_k - \bar{\mu}_k I) Q_{k+1}$

$\Longrightarrow$

$\underline{Q_{k+1} (A_{k+1} - \bar{\mu}_k I) = (A_k - \bar{\mu}_k I) Q_{k+1}}$

implying $Q_{k+1} Q_{k+2}$ is real.

SIMULTANEOUS POWER ITERATION

Power Iteration	Simultaneous Power Iteration
$z_k = A^k q_0$	$\underset{n\times n}{Z_k} = A^k \underset{n\times n}{Q_0}$
$q_k = z_k / \lVert z_k \rVert$	Q_k satisfies $Z_k = \underbrace{Q_k R_k}_{\text{QR factor}}$
$\hat{\lambda}_k = q_k^* A q_k$	$\hat{\Lambda}_k = Q_k^* A Q_k$

For simplicity assume $A \in \mathbb{C}^{n\times n}$ is Hermitian.

Power Iteration		Simultaneous Power Iteration
(1) q_k converges to a dominant eigenvector	WILL BE DISCUSSED	$Q_k \rightarrow V = [v_1 \ \dots \ v_n]$ as $k \to \infty$ where $v_1, \dots, v_n$ are orthonormal full set of eigenvectors
(2) $\hat{\lambda}_k$ converges to the eigenvalue with largest modulus		$\hat{\Lambda}_k$ converges to the full set of eigenvalues

As will be shown the <u>QR algorithm</u> is equivalent to <u>simultaneous power iteration</u>. (with $Q_0 = I$)

In particular $\boxed{\underbrace{A_k}_{\text{kth estimate by QR algorithm}} = \hat{\Lambda}_k \ (= Q_k^* A Q_k)}$

EXAMPLE

$$A = \begin{bmatrix} 3 & 4 \\ 2 & 1 \end{bmatrix}$$ with eigenvalues $\lambda_1 = 5$, $\lambda_2 = -1$

Recall when the QR algorithm is applied

$$A_3 = R_3 Q_3 = \begin{bmatrix} 5.01 & -1.97 \\ 0.03 & -1.01 \end{bmatrix}$$

Apply the simultaneous power iteration with $Q_0 = I$

$$* \; Z_3 = A^3(I) = \begin{bmatrix} 83 & 84 \\ 42 & 41 \end{bmatrix}$$

$$* \; Q_3 = \begin{bmatrix} -0.89 & -0.45 \\ -0.45 & 0.89 \end{bmatrix}$$ satisfies

$$Z_3 = \underbrace{Q_3 R_3}_{\text{QR Factorization}}$$

$$* \; \Lambda_3 = Q_3^* A Q_3 = \begin{bmatrix} 5.01 & -1.97 \\ 0.03 & -1.01 \end{bmatrix}$$

SUBSPACE ITERATION

Denote

* the eigenvalues of A by $\lambda_1, \dots, \lambda_n$ s.t.

$$|\lambda_1| \geq |\lambda_2| \geq \dots \geq |\lambda_n|,$$

* the associated ([redacted] orthonormal) eigenvectors by $v_1, v_2, \dots, v_n$.

Given $q_1, \dots, q_m \in \mathbb{C}^n$ with $m \leq n$. Let

$$S = \text{span}\{q_1, \dots, q_m\},$$

and

$$AS = \{Ax : x \in S\}$$ (the image of S under the transformation $x \longrightarrow Ax$)

REMARK

AS is a subspace of $\mathbb{R}^n$.

THM (Subspace Iteration)

Let $U_m = \text{span}\{v_{m+1}, \dots, v_n\}$ and $S_k = A^k S$. If $S \cap U_m = \{0\}$ and $|\lambda_m| > |\lambda_{m+1}|$, then

$$S_k \longrightarrow T_m \quad \text{as } k \longrightarrow \infty$$

where $T_m = \text{span}\{v_1, \dots, v_m\}$.

CONVERGENCE OF SIMULTANEOUS ITERATION

Letting $S^{(m)} = \operatorname{span}\{e_1, \dots, e_m\} \subseteq$ [redacted] $\mathbb{C}^n$
and denoting the image of $S^{(m)}$ under
the transformation $x \longrightarrow A^k x$ by

$$S_k^{(m)} = \{ A^k x \; : \; x \in S^{(m)} \}$$

we deduce (by thm - subspace iteration)

$$S_k^{(m)} \longrightarrow T_m \qquad \text{as} \quad k \to \infty .$$

(Assuming A has distinct eigenvalues
and $S^{(m)} \cap U_m = \{0\}$.)

Consequently

$$\underbrace{\operatorname{Range}(A^k(:, 1{:}m))}_{S_k^{(m)}} \longrightarrow \underbrace{\operatorname{span}\{v_1, \dots, v_m\}}_{T_m} \quad \text{as } k \to \infty$$

QR factorization of $A^k(:, 1{:}m)$

$$A^k(:, 1{:}m) = \underset{n \times m}{Q} \underset{m \times m}{R}$$

$$= \left[q_1^{(k)} \dots q_m^{(k)} \right] R$$

Note

Recall that $\{q_1^{(k)}, \dots, q_\ell^{(k)}\}$ is an orthonormal basis for $\text{range}(A^k(:,1:\ell))$ for $\ell = 1, \dots, m$

(i) $\text{Range}(A^k(:,1)) \longrightarrow \text{span}\{v_1\}$ as $k \to \infty$.

Furthermore

$$\{q_1^{(k)}\} \text{ is a basis for } \text{Range}(A^k(:,1))$$

$$\Longrightarrow$$

$$q_1^{(k)} \longrightarrow \underbrace{c_1}_{\text{scalar}} v_1 \quad \text{as } k \to \infty$$

(ii) $\text{Range}(A^k(:,1:2)) \longrightarrow \text{span}\{v_1, v_2\}$ as $k \to \infty$.

Furthermore

$$\{q_1^{(k)}, q_2^{(k)}\} \text{ is an orthonormal basis for } \text{Range}(A^k(:,1:$$

$$\Longrightarrow (q_2^{(k)} \perp q_1^{(k)} \text{ and } q_1^{(k)} = c_1 v_1$$

$$q_2^{(k)} \longrightarrow \underbrace{c_2}_{\text{scalar}} v_2 \quad \text{as } k \to \infty$$

(iii) $\text{Range}(A^k(:,1:\ell)) \longrightarrow \text{span}\{v_1, \dots, v_\ell\}$ as $k \to \infty$

for $\ell = 1, \dots, m$.

$$\{q_1^{(k)}, \dots, q_\ell^{(k)}\} \text{ is an orthonormal basis for } \text{Range}(A^k(:,1:\ell),$$

$$\Longrightarrow (q_\ell^{(k)} \perp q_j^{(k)} \text{ and } q_j^{(k)} = c_j v_j \text{ for } j < \ell,$$

$$q_\ell^{(k)} \longrightarrow c_\ell v_\ell \quad \text{as } k \to \infty$$

THM (Simultaneous Iteration)

Let $A \in \mathbb{C}^{n\times n}$ be a Hermitian matrix with distinct eigenvalues and

$$A^k = Q_k R_k$$

be a QR factorization. Then generically the columns of Q_k approach the eigenvectors $v_1, v_2, \ldots, v_n$ of A as $k \to \infty$.