QR-method lecture 3

SF2524 - Matrix Computations for Large-scale Systems

Agenda QR-method

1. Decompositions (lecture 1)
   - Jordan form
   - Schur decomposition
   - QR-factorization
2. Basic QR-method (lecture 1)
3. Improvement 1: Two-phase approach
   - Hessenberg reduction (lecture 1)
   - Hessenberg QR-method (lecture 2)
4. Improvement 2: Acceleration with shifts (lecture 2)
5. **Convergence theory (now)**

Agenda QR-method

1. Decompositions (lecture 1)
   - Jordan form
   - Schur decomposition
   - QR-factorization
2. Basic QR-method (lecture 1)
3. Improvement 1: Two-phase approach
   - Hessenberg reduction (lecture 1)
   - Hessenberg QR-method (lecture 2)
4. Improvement 2: Acceleration with shifts (lecture 2)
5. **Convergence theory (now)**

Reading instructions

Point 1: TB Lecture 24

Points 2-4: Lecture notes "QR-method" on course web page

Point 5: TB Chapter 28

(Extra reading: TB Chapter 25-26, 28-29)

Convergence theory - TB Chapter 28

Didactic simplification for convergence of QR-method: Assume $A = A^T$.

Convergence theory - TB Chapter 28

Didactic simplification for convergence of QR-method: Assume $A = A^T$.

Convergence characterization

(1) Artificial algorithm: USI - Unnormalized Simultaneous Iteration

Convergence theory - TB Chapter 28

Didactic simplification for convergence of QR-method: Assume $A = A^T$.

Convergence characterization

(1) Artificial algorithm: USI - Unnormalized Simultaneous Iteration

(2) Show convergence properties of USI

Convergence theory - TB Chapter 28

Didactic simplification for convergence of QR-method: Assume $A = A^T$.

Convergence characterization

(1) Artificial algorithm: USI - Unnormalized Simultaneous Iteration

(2) Show convergence properties of USI

(3) Artificial algorithm: NSI - Normalized Simultaneous Iteration

Convergence theory - TB Chapter 28

Didactic simplification for convergence of QR-method: Assume $A = A^T$.

Convergence characterization

(1) Artificial algorithm: USI - Unnormalized Simultaneous Iteration

(2) Show convergence properties of USI

(3) Artificial algorithm: NSI - Normalized Simultaneous Iteration

(4) Show: USI $\Leftrightarrow$ NSI $\Leftrightarrow$ QR-method

Definition: Unnormalized simultaneous iteration (USI)

A generalization of power method with n vectors "simultaneously"

Definition: Unnormalized simultaneous iteration (USI)

A generalization of power method with n vectors "simultaneously"

$$V^{(0)} = [v_1^{(0)}, \dots, v_n^{(0)}] \in \mathbb{R}^{m \times n}.$$

Define

$$V^{(k)} := A^k V^{(0)}.$$

Definition: Unnormalized simultaneous iteration (USI)

A generalization of power method with n vectors "simultaneously"

$$V^{(0)} = [v_1^{(0)}, \ldots, v_n^{(0)}] \in \mathbb{R}^{m \times n}.$$

Define

$$V^{(k)} := A^k V^{(0)}.$$

A QR-factorization generalizes the normalization step:

$$\hat{Q}^{(k)} \hat{R}^{(k)} = V^{(k)}.$$

Definition: Unnormalized simultaneous iteration (USI)

A generalization of power method with n vectors "simultaneously"

$$V^{(0)} = [v_1^{(0)}, \dots, v_n^{(0)}] \in \mathbb{R}^{m \times n}.$$

Define

$$V^{(k)} := A^k V^{(0)}.$$

A QR-factorization generalizes the normalization step:

$$\hat{Q}^{(k)} \hat{R}^{(k)} = V^{(k)}.$$

Convergence of USI

Convergence of USI

Assumptions:

- Let eigenvalues ordered and assume:

$$|\lambda_1| > |\lambda_2| > \cdots > |\lambda_{n+1}| \geq |\lambda_{n+2}| \geq \cdots \geq |\lambda_m|.$$

Convergence of USI

Assumptions:

- Let eigenvalues ordered and assume:

$$|\lambda_1| > |\lambda_2| > \cdots > |\lambda_{n+1}| \geq |\lambda_{n+2}| \geq \cdots \geq |\lambda_m|.$$

- Assume leading principal submatrices of $\hat{Q}^T V^{(0)}$ are nonsingular, where $\hat{Q} = (q_1, \ldots, q_n)$ are the eigenvectors.

Convergence of USI

Assumptions:

- Let eigenvalues ordered and assume:

$$|\lambda_1| > |\lambda_2| > \cdots > |\lambda_{n+1}| \geq |\lambda_{n+2}| \geq \cdots \geq |\lambda_m|.$$

- Assume leading principal submatrices of $\hat{Q}^T V^{(0)}$ are nonsingular, where $\hat{Q} = (q_1, \ldots, q_n)$ are the eigenvectors.

Theorem (TB Theorem 28.1)

Suppose simultaneous iteration is started with $V^{(0)}$ and assumptions above are satisfied. Let q_j, $j = 1, \ldots, n$ be the first n eigenvectors of A. Then, as $k \to \infty$, the columns of the matrices $\hat{Q}^{(k)}$ convergence linearly to q_j

$$\|q_j^{(k)} - \pm q_j\| = \mathcal{O}(C^k), \;\; j = 1, \ldots, n,$$

Convergence of USI

Assumptions:

- Let eigenvalues ordered and assume:

$$|\lambda_1| > |\lambda_2| > \cdots > |\lambda_{n+1}| \geq |\lambda_{n+2}| \geq \cdots \geq |\lambda_m|.$$

- Assume leading principal submatrices of $\hat{Q}^T V^{(0)}$ are nonsingular, where $\hat{Q} = (q_1, \ldots, q_n)$ are the eigenvectors.

Theorem (TB Theorem 28.1)

Suppose simultaneous iteration is started with $V^{(0)}$ and assumptions above are satisfied. Let q_j, $j = 1, \ldots, n$ be the first n eigenvectors of A. Then, as $k \to \infty$, the columns of the matrices $\hat{Q}^{(k)}$ convergence linearly to q_j

$$\|q_j^{(k)} - \pm q_j\| = \mathcal{O}(C^k), \quad j = 1, \ldots, n,$$

where $C = \max_{1 \leq k \leq n} |\lambda_{k+1}|/|\lambda_k|$.

Convergence of USI

Assumptions:

- Let eigenvalues ordered and assume:

$$|\lambda_1| > |\lambda_2| > \cdots > |\lambda_{n+1}| \geq |\lambda_{n+2}| \geq \cdots \geq |\lambda_m|.$$

- Assume leading principal submatrices of $\hat{Q}^T V^{(0)}$ are nonsingular, where $\hat{Q} = (q_1, \ldots, q_n)$ are the eigenvectors.

Theorem (TB Theorem 28.1)

Suppose simultaneous iteration is started with $V^{(0)}$ and assumptions above are satisfied. Let q_j, $j = 1, \ldots, n$ be the first n eigenvectors of A. Then, as $k \to \infty$, the columns of the matrices $\hat{Q}^{(k)}$ convergence linearly to q_j

$$\|q_j^{(k)} - \pm q_j\| = \mathcal{O}(C^k), \;\; j = 1, \ldots, n,$$

where $C = \max_{1\leq k\leq n} |\lambda_{k+1}|/|\lambda_k|$.

* Show matlab demo on USI (video) *

Normalized Simultaneous Iteration (NSI)

Variants of the power method. Equivalent:

(i) $v_k = \frac{A^k v_0}{\|A^k v_0\|}$

Normalized Simultaneous Iteration (NSI)

Variants of the power method. Equivalent:

(i) $v_k = \frac{A^k v_0}{\|A^k v_0\|}$

(ii) $v_k = \frac{A v_{k-1}}{\|A v_{k-1}\|}$

Normalized Simultaneous Iteration (NSI)

Variants of the power method. Equivalent:

(i) $v_k = \frac{A^k v_0}{\|A^k v_0\|}$

(ii) $v_k = \frac{A v_{k-1}}{\|A v_{k-1}\|}$

USI is a generlization of (i).

Normalized Simultaneous Iteration (NSI)

Variants of the power method. Equivalent:

(i) $v_k = \frac{A^k v_0}{\|A^k v_0\|}$

(ii) $v_k = \frac{A v_{k-1}}{\|A v_{k-1}\|}$

USI is a generlization of (i).
NSI is a generalization of (ii).

Normalized Simultaneous Iteration (NSI)

Variants of the power method. Equivalent:

(i) $v_k = \frac{A^k v_0}{||A^k v_0||}$

(ii) $v_k = \frac{A v_{k-1}}{||A v_{k-1}||}$

USI is a generlization of (i).

NSI is a generalization of (ii).

Algorithm: (Normalized) Simultaneous Iteration

- Input $\hat{Q}^{(0)} \in \mathbb{R}^{m \times n}$
- For $k = 1, \ldots,$
  - Set $Z = A\hat{Q}^{k-1}$
  - Compute QR-factorization $\hat{Q}^{(k)}\hat{R}^{(k)} = Z$

USI and NSI are equivalent.

USI and NSI are equivalent. More precisely:

Equivalence USI and NSI (TB Thm 28.2)

Suppose assumptions above are satisfied. If USI and NSI are started with the same vector they will generate the same sequence of matrices $\hat{Q}^k$ and $\hat{R}^k$.

Simultaneous iteration and QR-method

We will establish:

basic QR-method $\Leftrightarrow$ Simultaneous iteration with $\hat{Q}^{(0)} = I \in \mathbb{R}^{m\times m}$.

Simultaneous iteration and QR-method

We will establish:

$$\text{basic QR-method} \Leftrightarrow \text{Simultaneous iteration with } \hat{Q}^{(0)} = I \in \mathbb{R}^{m\times m}.$$

Simultaneous iteration satisfies

- $\underline{Q}^{(0)} = I$
- $Z_k = A\underline{Q}^{(k-1)}$
- $Z_k = \underline{Q}^{(k)} R^{(k)}$

Simultaneous iteration and QR-method

We will establish:

basic QR-method $\Leftrightarrow$ Simultaneous iteration with $\hat{Q}^{(0)} = I \in \mathbb{R}^{m\times m}$.

Simultaneous iteration satisfies

- $\underline{Q}^{(0)} = I$
- $Z_k = A\underline{Q}^{(k-1)}$
- $Z_k = \underline{Q}^{(k)}R^{(k)}$

QR-method satisfies

- $A^{(0)} = A$
- $A^{(k-1)} = Q^{(k)}R^{(k)}$
- $A^{(k)} = R^{(k)}Q^{(k)}$

Simultaneous iteration and QR-method

We will establish:

basic QR-method $\Leftrightarrow$ Simultaneous iteration with $\hat{Q}^{(0)} = I \in \mathbb{R}^{m\times m}$.

Simultaneous iteration satisfies

- $\underline{Q}^{(0)} = I$
- $Z_k = A\underline{Q}^{(k-1)}$
- $Z_k = \underline{Q}^{(k)}R^{(k)}$
- $A^{(k)} := (\underline{Q}^{(k)})^T A(\underline{Q}^{(k)})$

QR-method satisfies

- $A^{(0)} = A$
- $A^{(k-1)} = Q^{(k)}R^{(k)}$
- $A^{(k)} = R^{(k)}Q^{(k)}$

Simultaneous iteration and QR-method

We will establish:

basic QR-method $\Leftrightarrow$ Simultaneous iteration with $\hat{Q}^{(0)} = I \in \mathbb{R}^{m\times m}$.

Simultaneous iteration satisfies

- $\underline{Q}^{(0)} = I$
- $Z_k = A\underline{Q}^{(k-1)}$
- $Z_k = \underline{Q}^{(k)}R^{(k)}$
- $A^{(k)} := (\underline{Q}^{(k)})^T A(\underline{Q}^{(k)})$

QR-method satisfies

- $A^{(0)} = A$
- $A^{(k-1)} = Q^{(k)}R^{(k)}$
- $A^{(k)} = R^{(k)}Q^{(k)}$
- $\underline{Q}^{(k)} := Q^{(1)}\dots Q^{(k)}$

Simultaneous iteration and QR-method

We will establish:

basic QR-method $\Leftrightarrow$ Simultaneous iteration with $\hat{Q}^{(0)} = I \in \mathbb{R}^{m\times m}$.

Simultaneous iteration satisfies

- $\underline{Q}^{(0)} = I$
- $Z_k = A\underline{Q}^{(k-1)}$
- $Z_k = \underline{Q}^{(k)} R^{(k)}$
- $A^{(k)} := (\underline{Q}^{(k)})^T A(\underline{Q}^{(k)})$

QR-method satisfies

- $A^{(0)} = A$
- $A^{(k-1)} = Q^{(k)} R^{(k)}$
- $A^{(k)} = R^{(k)} Q^{(k)}$
- $\underline{Q}^{(k)} := Q^{(1)} \cdots Q^{(k)}$

Define: $\underline{R}^{(k)} := R^{(k)} \cdots R^{(1)}$

Simultaneous iteration and QR-method

We will establish:

basic QR-method $\Leftrightarrow$ Simultaneous iteration with $\hat{Q}^{(0)} = I \in \mathbb{R}^{m\times m}$.

Simultaneous iteration satisfies

- $\underline{Q}^{(0)} = I$
- $Z_k = A\underline{Q}^{(k-1)}$
- $Z_k = \underline{Q}^{(k)}R^{(k)}$
- $A^{(k)} := (\underline{Q}^{(k)})^T A(\underline{Q}^{(k)})$

QR-method satisfies

- $A^{(0)} = A$
- $A^{(k-1)} = Q^{(k)}R^{(k)}$
- $A^{(k)} = R^{(k)}Q^{(k)}$
- $\underline{Q}^{(k)} := Q^{(1)}\cdots Q^{(k)}$

Define: $\underline{R}^{(k)} := R^{(k)}\cdots R^{(1)}$

Essentially: The above equations generate the same sequence of matrices

Simultaneous iteration and QR-method

We will establish:

basic QR-method $\Leftrightarrow$ Simultaneous iteration with $\hat{Q}^{(0)} = I \in \mathbb{R}^{m\times m}$.

Simultaneous iteration satisfies

- $\underline{Q}^{(0)} = I$
- $Z_k = A\underline{Q}^{(k-1)}$
- $Z_k = \underline{Q}^{(k)} R^{(k)}$
- $A^{(k)} := (\underline{Q}^{(k)})^T A(\underline{Q}^{(k)})$

QR-method satisfies

- $A^{(0)} = A$
- $A^{(k-1)} = Q^{(k)} R^{(k)}$
- $A^{(k)} = R^{(k)} Q^{(k)}$
- $\underline{Q}^{(k)} := Q^{(1)} \cdots Q^{(k)}$

Define: $\underline{R}^{(k)} := R^{(k)} \cdots R^{(1)}$

Essentially: The above equations generate the same sequence of matrices

More precisely ...

TB Theorem 28.3:

Theorem (Equivalence simultaneous iteration and QR-method)

The above processes generate identical sequences of vectors. In particular,

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

and

$$A^{(k)} = (\underline{Q}^{(k)})^T A(\underline{Q}^{(k)}).$$

TB Theorem 28.3:

Theorem (Equivalence simultaneous iteration and QR-method)

The above processes generate identical sequences of vectors. In particular,

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

and

$$A^{(k)} = (\underline{Q}^{(k)})^T A(\underline{Q}^{(k)}).$$

Beware: QR-factorization is not unique and equivalence only holds with one QR-factorization.

Important property:

$$A^{(k)} = (\underline{Q}^{(k)})^T A \underline{Q}^{(k)}$$

Consequences

- Recall from USI-NSI equivalence and USI convergence. The columns in $\hat{Q}^{(k)}$ satisfy
$$q_i^{(k)} = \pm q_i + O(C^k).$$
where $C = \max_{1<i<n} |\lambda_{i+1}|/|\lambda_i|$.

Important property:

$$A^{(k)} = (\underline{Q}^{(k)})^T A \underline{Q}^{(k)}$$

Consequences

- Recall from USI-NSI equivalence and USI convergence. The columns in $\hat{Q}^{(k)}$ satisfy

$$q_i^{(k)} = \pm q_i + O(C^k).$$

  where $C = \max_{1<i<n} |\lambda_{i+1}|/|\lambda_i|$.
- $(A^k)_{i,j} = (q_i^{(k)})^T A q_j^{(k)}$
  - Diagonal $i = j$: $(A^{(k)})_{i,i} = (q_i^{(k)})^T A q_i^{(k)}$

Important property:

$$A^{(k)} = (\underline{Q}^{(k)})^T A \underline{Q}^{(k)}$$

Consequences

- Recall from USI-NSI equivalence and USI convergence. The columns in $\hat{Q}^{(k)}$ satisfy
$$q_i^{(k)} = \pm q_i + O(C^k).$$
where $C = \max_{1<i<n} |\lambda_{i+1}|/|\lambda_i|$.
- $(A^k)_{i,j} = (q_i^{(k)})^T A q_j^{(k)}$
  - Diagonal $i = j$: $(A^{(k)})_{i,i} = (q_i^{(k)})^T A q_i^{(k)} = r(q_i^{(k)})$ =Rayleigh quotient

Important property:

$$A^{(k)} = (\underline{Q}^{(k)})^T A \underline{Q}^{(k)}$$

Consequences

- Recall from USI-NSI equivalence and USI convergence. The columns in $\hat{Q}^{(k)}$ satisfy
$$q_i^{(k)} = \pm q_i + O(C^k).$$
where $C = \max_{1<i<n} |\lambda_{i+1}|/|\lambda_i|$.
- $(A^k)_{i,j} = (q_i^{(k)})^T A q_j^{(k)}$
  - Diagonal $i = j$: $(A^{(k)})_{i,i} = (q_i^{(k)})^T A q_i^{(k)} = r(q_i^{(k)})$ =Rayleigh quotient $\Rightarrow (A^{(k)})_{i,i} = \lambda_i + O(C^{2k})$

Important property:

$$A^{(k)} = (\underline{Q}^{(k)})^T A \underline{Q}^{(k)}$$

Consequences

- Recall from USI-NSI equivalence and USI convergence. The columns in $\hat{Q}^{(k)}$ satisfy

$$q_i^{(k)} = \pm q_i + O(C^k).$$

where $C = \max_{1<i<n} |\lambda_{i+1}|/|\lambda_i|$.

- $(A^k)_{i,j} = (q_i^{(k)})^T A q_j^{(k)}$
  - Diagonal $i = j$: $(A^{(k)})_{i,i} = (q_i^{(k)})^T A q_i^{(k)} = r(q_i^{(k)})$ =Rayleigh quotient $\Rightarrow (A^{(k)})_{i,i} = \lambda_i + O(C^{2k})$
  - Off-diagonal $i \neq j$: $(A^{(k)})_{i,j}$

Important property:

$$A^{(k)} = (\underline{Q}^{(k)})^T A \underline{Q}^{(k)}$$

Consequences

- Recall from USI-NSI equivalence and USI convergence. The columns in $\hat{Q}^{(k)}$ satisfy

  $$q_i^{(k)} = \pm q_i + O(C^k).$$

  where $C = \max_{1<i<n} |\lambda_{i+1}|/|\lambda_i|$.
- $(A^k)_{i,j} = (q_i^{(k)})^T A q_j^{(k)}$
  - Diagonal $i = j$: $(A^{(k)})_{i,i} = (q_i^{(k)})^T A q_i^{(k)} = r(q_i^{(k)})$ =Rayleigh quotient $\Rightarrow (A^{(k)})_{i,i} = \lambda_i + O(C^{2k})$
  - Off-diagonal $i \neq j$: $(A^{(k)})_{i,j} = (q_i^{(k)})^T A q_j^{(k)} = O(C^k)$

Important property:

$$A^{(k)} = (\underline{Q}^{(k)})^T A \underline{Q}^{(k)}$$

Consequences

- Recall from USI-NSI equivalence and USI convergence. The columns in $\hat{Q}^{(k)}$ satisfy
$$q_i^{(k)} = \pm q_i + O(C^k).$$
where $C = \max_{1<i<n} |\lambda_{i+1}|/|\lambda_i|$.
- $(A^k)_{i,j} = (q_i^{(k)})^T A q_j^{(k)}$
  - Diagonal $i = j$: $(A^{(k)})_{i,i} = (q_i^{(k)})^T A q_i^{(k)} = r(q_i^{(k)}) =$Rayleigh quotient $\Rightarrow (A^{(k)})_{i,i} = \lambda_i + O(C^{2k})$
  - Off-diagonal $i \neq j$: $(A^{(k)})_{i,j} = (q_i^{(k)})^T A q_j^{(k)} = O(C^k)$

Hence, $A^{(k)}$ will approach a triangular matrix

* Matlab demo *