Lecture 18
QR Algorithm and Simultaneous Iteration

Songting Luo

Department of Mathematics
Iowa State University

MATH 562 Numerical Analysis II

Outline

Outline

QR Algorithm

- Most basic version of QR algorithm is remarkably simple:

Algorithm: "Pure" QR Algorithm

$\mathbf{A}^{(0)} = \mathbf{A}$
for $k = 1, 2, \ldots$
 $\mathbf{Q}^{(k)}\mathbf{R}^{(k)} = \mathbf{A}^{(k-1)}$
 $\mathbf{A}^{(k)} = \mathbf{R}^{(k)}\mathbf{Q}^{(k)}$

- With some suitable assumptions, $\mathbf{A}^{(k)}$ converge to Schur form of $\mathbf{A}$ (diagonal if $\mathbf{A}$ is symmetric)
- Similarity transformation of $\mathbf{A}$:

$$\mathbf{A}^{(k)} = \mathbf{R}^{(k)}\mathbf{Q}^{(k)} = \left(\mathbf{Q}^{(k)}\right)^T \mathbf{A}^{(k-1)}\mathbf{Q}^{(k)}$$

- But why it works?

Unnormalized Simultaneous Iteration

- To understand QR algorithm, first consider simple algorithm
- Simultaneous iteration is power iteration applied to several vectors
- Start with linearly independent $\mathbf{v}_1^{(0)}, \cdots, \mathbf{v}_n^{(0)}$
- We know from power iteration that $\mathbf{A}^k\mathbf{v}_1^{(0)}$ converges to $\mathbf{q}_1$
- With some assumptions, the space $\langle \mathbf{A}^k\mathbf{v}_1^{(0)}, \cdots, \mathbf{A}^k\mathbf{v}_n^{(0)} \rangle$ converge to $\langle \mathbf{q}_1, \ldots, \mathbf{q}_n \rangle$
- Notation: Define initial matrix $\mathbf{V}^{(0)}$ and matrix $\mathbf{V}^{(k)}$ at step k:

$$\mathbf{V}^{(0)} = [\mathbf{v}_1^{(0)} | \cdots | \mathbf{v}_n^{(0)}], \quad \mathbf{V}^{(k)} = \mathbf{A}^k\mathbf{V}^{(0)} = [\mathbf{v}_1^{(k)} | \cdots | \mathbf{v}_n^{(k)}],$$

Unnormalized Simultaneous Iteration

- Define orthogonal basis for column space of $\mathbf{V}^{(k)}$ by reduced QR factorization $\hat{\mathbf{Q}}^{(k)}\hat{\mathbf{R}}^{(k)} = \mathbf{V}^{(k)}$
- We assume that
  - leading $n+1$ eigenvalues are distinct, and
  - all leading principal submatrices of $\hat{\mathbf{Q}}^T\mathbf{V}^{(0)}$ are nonsingular where $\hat{\mathbf{Q}} = [\mathbf{q}_1|\cdots|\mathbf{q}_n]$.
- We then have columns of $\hat{\mathbf{Q}}^{(k)}$ converge to eigenvectors of $\mathbf{A}$:

$$\|\mathbf{q}_j^{(k)} - (\pm\mathbf{q}_j)\| = O(c^k),$$

  where $c = \max_{1\leqslant k\leqslant n} |\lambda_{k+1}|/|\lambda_k|$
- Proof idea: Show that subspace of any leading j columns of $\mathbf{V}^{(k)} = \mathbf{A}^k\mathbf{V}^{(0)}$ converges to subspace of first j eigenvectors of $\mathbf{A}$, so does the subspace of any leading j columns of $\hat{\mathbf{Q}}^{(k)}$.

Simultaneous Iteration

- Matrices $\mathbf{V}^{(k)} = \mathbf{A}^k \mathbf{V}^{(0)}$ are highly ill-conditioned
- Orthonormalize at each step rather than at the end

Algorithm: Simultaneous Iteration

Pick $\hat{\mathbf{Q}}^{(0)} \in \mathbb{R}^{m \times n}$
for $k = 1, 2, \ldots$
  $\mathbf{Z} = \mathbf{A}\hat{\mathbf{Q}}^{(k-1)}$
  $\hat{\mathbf{Q}}^{(k)}\hat{\mathbf{R}}^{(k)} = \mathbf{Z}$

- Column spaces of $\hat{\mathbf{Q}}^{(k)}$ and $\mathbf{Z}^{(k)}$ are both equal to column space of $\mathbf{A}^k\hat{\mathbf{Q}}^{(0)}$, therefore same convergence as before

Simultaneous Iteration $\Leftrightarrow$ QR Algorithm

Algorithm: Simultaneous Iteration

Pick $\hat{\mathbf{Q}}^{(0)} \in \mathbb{R}^{m\times n}$
for $k = 1, 2, \ldots$
$\quad \mathbf{Z} = \mathbf{A}\hat{\mathbf{Q}}^{(k-1)}$
$\quad \hat{\mathbf{Q}}^{(k)}\hat{\mathbf{R}}^{(k)} = \mathbf{Z}$

Algorithm: "Pure" QR Algorithm

Let $\mathbf{A}^{(0)} = \mathbf{A}$
for $k = 1, 2, \ldots$
$\quad \mathbf{Q}^{(k)}\mathbf{R}^{(k)} = \mathbf{A}^{(k-1)}$
$\quad \mathbf{A}^{(k)} = \mathbf{R}^{(k)}\mathbf{Q}^{(k)}$

- QR algorithm is equivalent to simultaneous iteration with $\hat{\mathbf{Q}}^{(0)} = \mathbf{I}$
- Replace $\hat{\mathbf{R}}^{(k)}$ by $\mathbf{R}^{(k)}$ and $\hat{\mathbf{Q}}^{(k)}$ by $\underline{\mathbf{Q}}^{(k)}$, and introduce new statement $\mathbf{A}^{(k)} = \left(\underline{\mathbf{Q}}^{(k)}\right)^T \mathbf{A}\underline{\mathbf{Q}}^{(k)}$ in simultaneous iteration

Simultaneous Iteration $\Leftrightarrow$ QR Algorithm

Simultaneous Iteration

$\underline{\mathbf{Q}}^{(0)} = \mathbf{I}$
$\mathbf{Z} = \mathbf{A}\underline{\mathbf{Q}}^{(k-1)}$
$\underline{\mathbf{Q}}^{(k)}\mathbf{R}^{(k)} = \mathbf{Z}$
$\mathbf{A}^{(k)} = \left(\underline{\mathbf{Q}}^{(k)}\right)^T \mathbf{A}\underline{\mathbf{Q}}^{(k)}$

QR Algorithm

$\mathbf{A}^{(0)} = \mathbf{A}$
$\mathbf{Q}^{(k)}\mathbf{R}^{(k)} = \mathbf{A}^{(k-1)}$
$\mathbf{A}^{(k)} = \mathbf{R}^{(k)}\mathbf{Q}^{(k)}$
$\underline{\mathbf{Q}}^{(k)} = \mathbf{Q}^{(1)}\mathbf{Q}^{(2)}\dots\mathbf{Q}^{(k)}$

- $\underline{\mathbf{Q}}^{(k)} = \mathbf{Q}^{(1)}\mathbf{Q}^{(2)}\cdots\mathbf{Q}^{(k)}$. Let $\underline{\mathbf{R}}^{(k)} = \mathbf{R}^{(k)}\mathbf{R}^{(k-1)}\dots\mathbf{R}^{(1)}$
- Both schemes generate QR factorization $\mathbf{A}^k = \underline{\mathbf{Q}}^{(k)}\underline{\mathbf{R}}^{(k)}$ and projection $\mathbf{A}^{(k)} = \left(\underline{\mathbf{Q}}^{(k)}\right)^T \mathbf{A}\underline{\mathbf{Q}}^{(k)}$

Simultaneous Iteration $\Leftrightarrow$ QR Algorithm

Proof

Proof by induction. For $k = 0$ it is trivial for both algorithms.
For $k \geqslant 1$ with simultaneous iteration, $\mathbf{A}^{(k)}$ is given by definition, and

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

For $k \geqslant 1$ with QR algorithm,

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

and

$$\mathbf{A}^{(k)} = \left(\mathbf{Q}^{(k)}\right)^T \mathbf{A}^{(k-1)}\mathbf{Q}^{(k)} = \left(\underline{\mathbf{Q}}^{(k)}\right)^T \mathbf{A}\underline{\mathbf{Q}}^{(k)}$$

Convergence of QR Algorithm

- Since $\underline{\mathbf{Q}}^{(k)} = \hat{\mathbf{Q}}^{(k)}$ in simultaneous iteration, column vectors of $\underline{\mathbf{Q}}^{(k)}$ converge linearly to eigenvectors if $\mathbf{A}$ has distinct eigenvalues
- Fron $\mathbf{A}^{(k)} = \left(\underline{\mathbf{Q}}^{(k)}\right)^T \mathbf{A}\underline{\mathbf{Q}}^{(k)}$, diagonal entries of $\mathbf{Q}^{(k)}$ are Rayleigh quotients of column vectors of $\underline{\mathbf{Q}}^{(k)}$, so they converge linearly to eigenvalues of $\mathbf{A}$
- Off-diagonal entries of $\mathbf{A}^{(k)}$ converge to zeros, as they are generalized Rayleigh quotients involving approximations of distinct eigenvectors
- Overall, $\mathbf{A} = \underline{\mathbf{Q}}^{(k)}\mathbf{A}^{(k)}\left(\underline{\mathbf{Q}}^{(k)}\right)^T$. For a symmetric matrix, it converges to eigenvalue decomposition of $\mathbf{A}$