CS475 / CS675
Lecture 17: June 28, 2016

QR Algorithm and

Reduction to Hessenberg

Reading: [TB] Chapt 28

Simultaneous iteration vs QR algorithm

- QR algorithm can be viewed as simultaneous iteration with $\hat{Q}^{(0)} = I$ and $p = n$.

- We can drop the hats on $\hat{Q}^{(k)}$, $\hat{R}^{(k)}$

- $\underline{Q}^{(k)} = Q$'s from simultaneous iteration,
  $\overline{Q}^{(k)} = Q$'s from QR algorithm

Simultaneous iteration revisited

- Simultaneous iteration can be written as:

$\underline{Q}^{(0)} = I$

For $k = 1,2, \ldots$

$Z^{(k)} \leftarrow A\underline{Q}^{(k-1)}$

$\underline{Q}^{(k)}R^{(k)} \leftarrow Z^{(k)}$

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

$\underline{R}^{(k)} = R^{(k)}R^{(k-1)} \ldots R^{(1)}$

} New matrices for proof purpose

end

QR algorithm revisited

- QR algorithm can be written as:

  $A^{(0)} = A$

  For $k = 1,2, \dots$

  $\quad Q^{(k)}R^{(k)} \leftarrow A^{(k-1)}$

  $\quad A^{(k)} \leftarrow R^{(k)}Q^{(k)}$

  $\quad \underline{Q}^{(k)} = Q^{(1)}Q^{(2)} \dots Q^{(k)}$

  $\quad \underline{R}^{(k)} = R^{(k)}R^{(k-1)} \dots R^{(1)}$

  } New matrices for proof purpose

  end

Equivalence

- <u>Theorem:</u> The two algorithms generate identical sequences of matrices $\underline{R}^{(k)}$, $\underline{Q}^{(k)}$ and $A^{(k)}$ and they are

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

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

Equivalence

- Proof: by induction. The case $k = 0$ is trivial since $A^0 = \underline{Q}^{(0)} = \underline{R}^{(0)} = I$ and $A^{(0)} = A$. Suppose it is true for $k - 1$.
  Simultaneous iteration:

Equivalence

- Proof continued…
  QR algorithm:

Convergence of the QR algorithm

(1) $\Longrightarrow$ QR algorithm effectively computes Q, R factors of A^k i.e., orthonormal basis for A^k

(2) $\Longrightarrow$ The diagonal of $A^{(k)}$ are Rayleigh quotients of column vectors of $\underline{Q}^{(k)}$

- As columns of $\underline{Q}^{(k)} \longrightarrow$ eigenvectors,
  the Rayleigh quotients $\longrightarrow$ eigenvalues

Convergence of the QR algorithm

- $A_{ij}^{(k)} = \left(\underline{q}_i^{(k)}\right)^T A(\underline{q}_j^{(k)})$
  - Here $\underline{q}_i^{(k)}, \underline{q}_j^{(k)}$ are columns i and j of $\underline{Q}^{(k)}$
  - Eventually $\underline{q}_j^{(k)} \rightarrow q_j, \underline{q}_i^{(k)} \rightarrow q_i, A\underline{q}_j^{(k)} \approx \lambda_j q_j$
  - Therefore $A_{ij}^{(k)} \approx \lambda_j q_i^T q_j = 0 \quad \forall i \neq j$
- $\therefore A^{(k)}$ converges to a diagonal matrix

Convergence of the QR algorithm

- Theorem: Assume $|\lambda_1| > |\lambda_2| > \cdots > |\lambda_n|$ and Q has all nonsingular leading principal minors. As $k \to \infty$, $A^{(k)}$ converges linearly to $diag(\lambda_1, \ldots, \lambda_n)$ and $\underline{Q}^{(k)}$ converges at the same rate to Q. The rate of convergence is

$$C = \max_k \left| \frac{\lambda_{k+1}}{\lambda_k} \right|$$

Example

- $A = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 3 & 1 \\ 1 & 1 & 4 \end{bmatrix} = A^{(0)}$

Example

- $A = \begin{bmatrix} 21 & 7 & -1 \\ 5 & 7 & 7 \\ 4 & -4 & 20 \end{bmatrix} = A^{(0)}$

Practical QR

- It is expensive to compute the QR factorization of a square matrix $\left(\frac{4}{3}n^3\ flops\right)$
- In practice, we first reduce A to a Hessenberg matrix if $A \neq A^T$ and to a tridiagonal matrix if $A = A^T$
- The resulting QR factorization would be $O(n^2)$ if $A \neq A^T$ and $O(n)$ if $A = A^T$

Reduction to Hessenberg or Tridiagonal

- The matrix can be nonsymmetric in general
- Why Hessenberg? Why not triangular?

Reduction to Hessenberg or Tridiagonal

- Be less ambitious and choose Q_1^T that leaves 1st row unchanged

Reduction to Hessenberg or Tridiagonal

- In general:
  $Q = Q_1 Q_2 \dots Q_{n-2}$ and $Q^T A Q =$ upper Hessenberg

- Complexity:
  - Flops(Reduction to Hessenberg) $\approx \frac{10}{3} n^3$
  - Flops(Reduction to tridiagonal) $\approx \frac{4}{3} n^3$