CS475 / CM375
Lecture 15: Nov 1st, 2011

Rayleigh Quotient Iteration

and QR Algorithm

Reading: [TB] Chapters 27, 28

Inverse Iteration Algorithm

$v^{(0)}$ = some vector with $\left|\left|v^{(0)}\right|\right| = 1$

For $k = 1,2, \ldots$

  Solve $(A - \mu I)w = v^{(k-1)}$

  $v^{(k)} = w/\left|\left|w\right|\right|$

  $\lambda^{(k)} = \left(v^{(k)}\right)^T A v^{(k)}$

end

Inverse Iteration Algorithm

- Notes
  1. If $\mu = \lambda_j$, then $A - \mu I$ is singular
     If $\mu \approx \lambda_j$, then $A - \mu I$ is close to being singular
     It turns out to be OK if the linear system is solved stably
  2. Like power iteration, inverse iteration has linear convergence
  3. Unlike power iteration, we can choose which eigenvector to compute by choosing μ close to the corresponding λ_j

Inverse Iteration Algorithm

- Theorem: Suppose λ_j is closest to μ and λ_l is the second closest,

  i.e., $|\mu - \lambda_j| < |\mu - \lambda_l| \leq |\mu - \lambda_i| \quad \forall i \neq j.$

  Also suppose $q_j^T v^{(0)} \neq 0$. Then

$$\left|\left|v^{(k)} - (\pm q_j)\right|\right| = O\left(\left|\frac{\mu - \lambda_j}{\mu - \lambda_l}\right|^k\right)$$

and $\left|\lambda^{(k)} - \lambda_j\right| = O\left(\left|\frac{\mu-\lambda_j}{\mu-\lambda_l}\right|^{2k}\right)$ as $k \to \infty$

Rayleigh Quotient Iteration

- Rayleigh quotient gives an eigenvalue estimate from an eigenvector estimate
- Inverse iteration gives an eigenvector estimate from an eigenvalue estimate
- Idea: combine the two.

Rayleigh Quotient Iteration Algorithm

$v^{(0)}$ = some vector with $\left|\left|v^{(0)}\right|\right| = 1$

$\lambda^{(0)} = \left(v^{(0)}\right)^T A v^{(0)} = r\left(v^{(0)}\right)$

For $k = 1,2, \ldots$

Solve $\left(A - \lambda^{(k-1)} I\right) w = v^{(k-1)}$

$v^{(k)} = w / ||w||$

$\lambda^{(k)} = \left(v^{(k)}\right)^T A v^{(k)}$

end

Rayleigh Quotient Iteration

- Theorem: RQI converges for most starting vectors $v^{(0)}$. The convergence is cubic:

$$\left|\left|v^{(k+1)} - (\pm q_j)\right|\right| = O\left(\left|\left|v^{(k)} - (\pm q_j)\right|\right|^3\right)$$
$$\left|\lambda^{(k+1)} - \lambda_j\right| = O\left(\left|\lambda^{(k)} - \lambda_j\right|^3\right)$$

Rayleigh Quotient Iteration

- Proof sketch: Suppose $\left|\left|v^{(k)} - q_j\right|\right| \leq \epsilon$. Then

$$\left|r\left(v^{(k)}\right) - r(q_j)\right| = O\left(\left|\left|v^{(k)} - q_j\right|\right|\right)^2$$

i.e., $\left|\lambda^{(k)} - \lambda_j\right| = \epsilon^2$. It can be proven that

$$\left|\left|v^{(k+1)} - q_j\right|\right| = O\left(\left|\lambda^{(k)} - \lambda_j\right| \left|\left|v^{(k)} - q_j\right|\right|\right) = O(\epsilon^3)$$

Thus $\left|\left|v^{(k)} - q_j\right|\right| \leq \epsilon$

$\implies \left|\lambda^{(k)} - \lambda_j\right| \leq \epsilon^2$

$\implies \left|\lambda^{(k)} - \lambda_j\right| \left|\left|v^{(k)} - q_j\right|\right| \leq \epsilon^3$

Example

- Example of cubic convergence for Rayleigh quotient iteration

$$A = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 3 & 1 \\ 1 & 1 & 4 \end{bmatrix} \quad v^{(0)} = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

Second Example

- Comparison of RQI and Power iteration

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

Complexity

- Each step of power iteration involves $Av^{(k-1)}$, which takes $O(n^2)$ flops.
- Each step of inverse iteration solves $(A-\mu I)^{-1}w = v^{(k-1)}$, which takes $O(n^3)$ flops.
  One can pre-compute and store L, U factors of $A-\mu I$. Thus each step takes $O(n^2)$ flops for forward and backward solves.
- The matrix $A-\lambda^{(k-1)}I$ changes in each step of RQI. Hence it takes $O(n^3)$ flops in general.
- If A is tridiagonal, all 3 methods take $O(n)$ flops per iteration.

Similarity Transformation

- <u>Def:</u> If $X \in \Re^{n\times n}$ is nonsingular, then $A \rightarrow X^{-1}AX$ is called a similarity transformation of A.
- <u>Def:</u> A and B are similar if $B = X^{-1}AX$ for some nonsingular X.
- <u>Theorem:</u> If A, B are similar, then they have the same characteristic polynomial and the same eigenvalues
- <u>Proof:</u>

QR Algorithm

- Idea: Apply a sequence of similarity transformations to A, which will converge to a diagonal matrix.
- Consider $A^{(k-1)}$. Compute its QR factorization.

  i.e., $R^{(k)} = \left(Q^{(k)}\right)^T A^{(k-1)}$
- Then $R^{(k)}Q^{(k)} = \underbrace{\left(Q^{(k)}\right)^T A^{(k-1)} Q^{(k)}}_{A^{(k)}}$
- Clearly $A^{(k-1)}$ and $A^{(k)}$ are similar

QR Algorithm

$A^{(0)} = A$

For $k = 1,2, \ldots$

$Q^{(k)}R^{(k)} = A^{(k-1)}$ (QR factorization of $A^{(k-1)}$)

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

End

How does it work?

Unnormalized Simultaneous Iteration

- Apply power iteration to several vectors at once and maintain linear independence among the vectors.

- Start with: $v_1^{(0)}, v_2^{(0)}, \dots, v_p^{(0)}$
  Then $A^k v_1^{(0)}$ converges to q_1 where $|\lambda_1|$ is largest.
  Thus span $\{A^k v_1^{(0)}, \dots, A^k v_p^{(0)}\}$ should converge to $\{q_1, \dots, q_p\}$ where $\lambda_1, \dots, \lambda_p$ are the p largest eigenvalues.

Unnormalized Simultaneous Iteration

- Write $V^{(0)} = \left[v_1^{(0)} \;\; v_2^{(0)} \;\; \dots \;\; v_p^{(0)}\right]$
- Define $V^{(k)} \equiv A^k V^{(0)} = \left[v_1^{(k)} \;\; v_2^{(k)} \;\; \dots \;\; v_p^{(k)}\right]$
- Compute a reduced QR factorization of $V^{(k)}$:
  $\hat{Q}^{(k)} \hat{R}^{(k)} = V^{(k)} \qquad \hat{Q}^{(k)} \in \Re^{n \times p}, \hat{R}^{(k)} \in \Re^{p \times p}$
  As $k \to \infty$, $\hat{Q}^{(k)}$ should $\to \hat{Q} \equiv [q_1, \; q_2, \; \dots \; q_p]$

Unnormalized Simultaneous Iteration

- Assumption 1: the leading $p + 1$ eigenvalues are distinct in absolute values
  $|\lambda_1| > |\lambda_2| > \cdots > |\lambda_p| > |\lambda_{p+1}| \geq |\lambda_{p+2}| \geq \cdots \geq |\lambda_n|$
- Assumption 2: All the leading principal minors of $\hat{Q}^T V^{(0)}$ are nonsingular
- Theorem: Suppose block power iteration is carried out and assumptions 1 & 2 hold. Then as $k \to \infty$

  $$\left\| q_j^{(k)} - (\pm q_j) \right\| = O(c^k) \quad j = 1,2,\dots,p$$

  Where $c = \max_{1 \leq k \leq p} \left| \frac{\lambda_{k+1}}{\lambda_k} \right| < 1$

Unnormalized Simultaneous Iteration

- As $k \to \infty$, the vectors $v_1^{(k)}, \dots, v_p^{(k)}$ all converge to multiples of the same dominant eigenvector q_1
- Idea: orthogonalize the vectors at each step

Simultaneous Iteration Algorithm

Pick $\hat{Q}^{(0)} \in \Re^{n \times p}$ with orthonormal columns

For $k = 1,2, \ldots$

$Z^{(k)} = A\hat{Q}^{(k-1)}$ power iteration

$\hat{Q}^{(k)}\hat{R}^{(k)} = Z^{(k)}$ reduced QR factorization

End

Note: The column space of $\hat{Q}^{(k)}$ and $Z^{(k)}$ are the same. They are both equal to that of $A^{(k)}\hat{Q}^{(0)}$.