CS475 / CS675
Lecture 14: June 16, 2016

Eigenvalue problems

Reading: [TB] Chapters 24, 25

Eigenvalue problem

- Example: $A = \begin{bmatrix} 1 & 0 & 0 \\ 2 & -1 & 2 \\ 4 & -4 & 5 \end{bmatrix}$

  $\det(\lambda I - A) =$

Example continued

- $\lambda = 1$:

- $\lambda = 3$:

Example Continued

- Thus $AX =$

  $X\Lambda =$

- Note: we never compute eigenvalues by finding the roots of the characteristic polynomial

Eigenvalues

- <u>Gershgorin Theorem:</u> Let A be any square matrix. The eigenvalues λ of A are located in the union of the n disks:

$$|\lambda - a_{ii}| \leq \sum_{j\neq i} |a_{ij}|$$

- <u>Proof:</u> Consider (λ, x) such that $Ax = \lambda x,\ x \neq 0.$

Scale x such that $||x||_\infty = 1 = x_i$ for some i

Then $\lambda x_i = (Ax)_i = \sum_{j=1}^{n} a_{ij} x_j = a_{ii} x_i + \sum_{j\neq i} a_{ij} x_j$

$$\Longrightarrow |\lambda - a_{ii}| = \left|\sum_{j\neq i} a_{ij} x_j\right| \leq \sum_{j\neq i} \left|a_{ij} x_j\right| \leq \sum_{j\neq i} |a_{ij}|$$

Example

- $A = \begin{bmatrix} 4 & -0.5 & 0 \\ 0.6 & 5 & -0.6 \\ 0 & 0.5 & 3 \end{bmatrix}$
- Picture:

Rayleigh quotient

- Assume A is real and symmetric. Thus A has real eigenvalues and a complete set of orthogonal eigenvectors

$$\{\lambda_1, \dots, \lambda_n\}, \{q_1, \dots, q_n\} \qquad \left|\left|q_j\right|\right| = 1$$

- Def: The Rayleigh quotient of a vector x is:

$$r(x) = \frac{x^T A x}{x^T x}$$

Rayleigh quotient

- Notes
  1. If x is an eigenvector, then $r(x)$ is an eigenvalue
  2. Given x, find α such that

     $$\min_{\alpha} \left\| \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \alpha - Ax \right\|_2 \quad (m \times 1 \text{ least squares})$$

     The normal equations: $(x^T x)\alpha = x^T (Ax)$

     $$\alpha = r(x)$$
  3. Theorem: Let q_j be an eigenvector and $x \approx q_j$

     then $r(x) - r(q_j) = O\left(\left\|x - q_j\right\|^2\right)$ as $x \to q_j$

Power Iteration

- Let $v^{(0)}$ = approximate eigenvector, $\left|\left|v^{(0)}\right|\right| = 1$ and $\{q_i\}$ = set of eigenvectors

- Then $v^{(0)} = c_1 q_1 + c_2 q_2 + \cdots + c_n q_n$
  $\Longrightarrow Av^{(0)} = c_1 \lambda_1 q_1 + c_2 \lambda_2 q_2 + \cdots + c_n \lambda_n q_n$

Power Iteration

- Similarly, $A^k v^{(0)} = c_1 \lambda_1^k q_1 + c_2 \lambda_2^k q_2 + \cdots + c_n \lambda_n^k q_n$

$$= \lambda_1^k \left(c_1 q_1 + c_2 \left(\frac{\lambda_2}{\lambda_1}\right)^k q_2 + \cdots + c_n \left(\frac{\lambda_n}{\lambda_1}\right)^k q_n \right)$$

- Suppose $|\lambda_1| > |\lambda_2| \geq \cdots \geq |\lambda_n|$.
  Then $\left|\frac{\lambda_i}{\lambda_1}\right|^k \to 0$ as $k \to \infty$

$$A^k v^{(0)} \sim c_1 \lambda_1^k q_1 \quad \text{for large } k$$

i.e., $q_1 \sim \frac{A^k v^{(0)}}{\left\|A^k v^{(0)}\right\|}$

Example

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

Power Iteration Algorithm

$v^{(0)} =$ initial guess, $\left|\left|v^{(0)}\right|\right| = 1$

for $k = 1,2, \ldots$

$$w = Av^{(k-1)}$$

$$v^{(k)} = \frac{w}{||w||}$$

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

end

Power Iteration Algorithm

- Notes
  1. We normalize $Av^{(k-1)}$ in each computation of $v^{(k)}$
  2. <u>Theorem:</u> Suppose $|\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq \cdots \geq |\lambda_n|$ and $q_1^T v^{(0)} \neq 0$. Then
     $$\left\| v^{(k)} - (\pm q_1) \right\| = O\left(\left| \frac{\lambda_2}{\lambda_1} \right|^k \right) \text{ and}$$
     $$\left| \lambda^{(k)} - \lambda_1 \right| = O\left(\left| \frac{\lambda_2}{\lambda_1} \right|^{2k} \right) \text{ as } k \to \infty$$
  3. It only computes q_1
  4. The convergence is linear, the convergence rate $= \left| \frac{\lambda_2}{\lambda_1} \right|$
  5. The convergence can be slow if $|\lambda_1| \approx |\lambda_2|$

Inverse Iteration

- <u>Idea 1:</u> Use A^{-1} to compute the smallest eigenvalue

  Note: $\Lambda(A^{-1}) = \left\{\frac{1}{\lambda_1}, \frac{1}{\lambda_2}, \ldots, \frac{1}{\lambda_n}\right\}$

- Thus: $v^{(0)} = c_1 q_1 + c_2 q_2 + \cdots + c_n q_n$

$$A^{-1} v^{(0)} = c_1 \frac{1}{\lambda_1} q_1 + \cdots + c_n \frac{1}{\lambda_n} q_n$$

$$\vdots$$

$$A^{-k} v^{(0)} = c_1 \left(\frac{1}{\lambda_1}\right)^k q_1 + \cdots + c_n \left(\frac{1}{\lambda_n}\right)^k q_n$$

$$= \left(\frac{1}{\lambda_n}\right)^k \left[c_1 \left(\frac{\lambda_n}{\lambda_1}\right)^k q_1 + \cdots + c_{n-1} \left(\frac{\lambda_n}{\lambda_{n-1}}\right)^k q_{n-1} + c_n q_n\right]$$

$$\therefore A^{-k} v^{(0)} \approx c_n \left(\frac{1}{\lambda_n}\right)^k q_n$$

Inverse Iteration

- Idea 2: Shifting. Consider $B = A - \mu I \quad \mu \notin \Lambda(A)$
  Then B has the same eigenvectors as A and
  its eigenvalues are $\{\lambda_j - \mu\}$, where $\lambda_j \in \Lambda(A)$.

- If μ is close to λ_j, then $\lambda_j - \mu$ would be the smallest eigenvalue of B.

- We can apply idea 1 to compute $\lambda_j - \mu$

Example

- $A = \begin{bmatrix} 21 & 7 & -1 \\ 5 & 7 & 7 \\ 4 & -4 & 20 \end{bmatrix}$ $\Lambda(A) = \{8,16,24\}, \quad \mu = 15$