June 27, 2014

Rayleigh quotient

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

$$\{ \lambda_1, \ldots, \lambda_n \}, \ \{ q_1, \ldots, q_n \} \qquad \|q_j\| = 1$$

Def: The Rayleigh quotient of a vector x is:

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

Notes

1) If x is an eigenvector, then $r(x)$ is an eigenvalue.

2) Given x, find α such that

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

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

$$r(x) - r(q_J) = O(\|x - q_J\|^2) \qquad \text{as} \quad x \to q_J$$

Power iteration

Let $v^{(0)}$ = approx. eigenvector, $||v^{(0)}|| = 1$, and $\{ q_i \}$ = set of eigenvectors.

Then $\quad v^{(0)} = c_1 q_1 + c_2 q_2 + \ldots + c_n q_n$

$$A v^{(0)} = c_1 \lambda_1 q_1 + c_2 \lambda_2 q_2 + \ldots + c_n \lambda_n q_n$$

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

$$= \lambda_1^k (c_1 q_1 + c_2 (\lambda_2/\lambda_1)^k q_2 + \ldots + c_n (\lambda_n/\lambda_1)^k q_n)$$

Suppose $|\lambda_1| > |\lambda_2| \geq \ldots \geq |\lambda_n|$. Then $|\lambda_i / \lambda_1|^k \to 0$ as $k \to \infty$.

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

i.e.

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

Example

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

$w = A\, v^{(0)} = (27, 19, 20)^T$

$v^{(1)} = w/||w|| = (0.70, 0.49, 0.52)^T$

$\lambda^{(1)} = r(v^{(1)}) = 23.3235$

$w = A\, v^{(1)} = (17.62, 10.57, 11.19)^T$

$v^{(2)} = w/||w|| = (0.75, 0.45, 0.48)^T$

$\lambda^{(2)} = r(v^{(2)}) = 23.7250$

$w = A\, v^{(2)} = (18.50, 10.28, 10.71)^T$

$v^{(3)} = w/||w|| = (0.78, 0.43, 0.45)^T$

$\lambda^{(3)} = r(v^{(3)}) = 23.8670$

:

:

$q_1 = (0.8165, 0.4082, 0.4082)^T, \quad \lambda_1 = 24.$

Algorithm

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

for k = 1, 2, . . .

$w = A\, v^{(k-1)}$

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

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

end

Notes

1) We normalize $A\, v^{(k-1)}$ in each computation of $v^{(k)}$.

2) Theorem: Suppose $|\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq \ . . . \ \geq |\lambda_n|$, $q_1^T v^{(0)} \neq 0$. Then

$$\left\| v^{(k)} - (\pm q_1) \right\| = O\left(\left| \frac{\lambda_2}{\lambda_1} \right|^k \right), \quad \left| \lambda^{(k)} - \lambda_1 \right| = O\left(\left| \frac{\lambda_2}{\lambda_1} \right|^{2k} \right)$$

as $k \to \infty$

3) It only computes q_1.

4) The convergence is linear, the convergence rate = $|\lambda_2|/|\lambda_1|$.

5) The convergence can be slow if $|\lambda_1| \sim |\lambda_2|$.

Inverse iteration

Idea 1: Use A^{-1} to compute the smallest eigenvalue.

(Note: $\Lambda(A^{-1}) = \{ 1/\lambda_1, 1/\lambda_2, \ldots, 1/\lambda_n \}$.)

Thus $\quad v^{(0)} = c_1 q_1 + c_2 q_2 + \ldots + c_n q_n$

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

$$\vdots$$

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

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

$\therefore \quad A^{-k} v^{(0)} \sim c_n (1/\lambda_n)^k q_n \quad$ for large k

Idea 2: Shifting.

Consider $B = A - \mu I$, μ is not an eigenvalue of A. Then B has the same eigenvectors of A and its eigenvalues are $\{ \lambda_j - \mu \}$, $\lambda_j \in \Lambda(A)$.

If μ is close to λ_J, $\lambda_J - \mu$ would be the smallest eigenvalue of B.

We can apply idea 1 to compute $\lambda_J - \mu$.

Example

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

$v^{(0)} = (1, 1, 1)^T$

$w = (A-\mu I)^{-1} v^{(0)} = (0.032, 0.16, 0.30)^T$

$v^{(1)} = w/||w|| = (0.093, 0.46, 0.88)^T$

$\lambda^{(1)} = r(v^{(1)}) = 19.2000$

$w = (A-\mu I)^{-1} v^{(1)} = (-0.33, 0.40, 0.76)^T$

$v^{(2)} = w/||w|| = (-0.36, 0.44, 0.83)^T$

$\lambda^{(2)} = r(v^{(2)}) = 15.9749$

$w = (A-\mu I)^{-1} v^{(2)} = (-0.39, 0.40, 0.79)^T$

$v^{(3)} = w/||w|| = (-0.40, 0.41, 0.82)^T$

$\lambda^{(3)} = r(v^{(3)}) = 16.0290$

:

:

$q_2 = (-0.4082, 0.4082, 0.8165)^T, \quad \lambda_2 = 16.$