Advanced Numerical Methods	November 23, and 25, 2011

Lecture Notes 7: Eigenvalue Problem

Lecturer: Che-Rung Lee *Scribe: Zhen-Yu Peng*

3 Eigenvalue Problem

3.1 Definition

- Given an $n \times n$ matrix A, an eigenvalue λ of A is a scalar such that $A\vec{x} = \lambda\vec{x}$ for a nonzero vector $\vec{x}$, $\vec{x}$ is call an eigenvector.

- If there exists n linearly independent eigenvectors $\vec{x}_1$, $\vec{x}_2$, ..., $\vec{x}_n$, we said A is diagonalizable.

$$X = \begin{bmatrix} \vec{x}_1, \vec{x}_2, \vec{x}_3, \ldots, \vec{x}_n \end{bmatrix}$$

$$\begin{aligned} AX &= \begin{bmatrix} A\vec{x}_1, A\vec{x}_2, A\vec{x}_3, \ldots, A\vec{x}_n \end{bmatrix} \\ &= \begin{bmatrix} \lambda_1\vec{x}_1, \lambda_2\vec{x}_2, \lambda_3\vec{x}_3, , \ldots, \lambda_n\vec{x}_n \end{bmatrix} \\ &= X \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix} \end{aligned}$$

then

$$A = X \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix} X^{-1}$$

$$\Rightarrow A = X\Lambda X^{-1}$$
$$\Rightarrow \Lambda = X^{-1}AX$$

- Given λ, its eigenvector = ?

$$\begin{aligned} & A\vec{x} = \lambda\vec{x} \\ \Leftrightarrow & A\vec{x} - \lambda\vec{x} = \vec{0} \\ \Leftrightarrow & (A - \lambda I)\vec{x} = \vec{0} \end{aligned}$$

then we can use kernel to find the eigenvectors correspond to λ.

$$ker(A - \lambda I) = \{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_k\}$$

- Given $\vec{x}$, its eigenvalue = ?

$$\begin{aligned} & A\vec{x} = \lambda\vec{x} \\ \Rightarrow & \vec{x}^T A\vec{x} = \vec{x}^T \lambda\vec{x} \\ \Rightarrow & \lambda = \frac{\vec{x}^T A\vec{x}}{\vec{x}^T\vec{x}} \qquad (Rayleigh \quad quotient) \end{aligned}$$

3.2 Power Method

- We can use this algorithm to find the eigenvector whose eigenvalue is the largest.
- Algorithm
  1. Choose a random vector $\vec{v}_0$, $\parallel \vec{v}_0 \parallel = 1$.
  2. For i = 1, 2, 3, ..., until converge.
  3. $\quad \vec{y}_i = A\vec{v}_{i-1}$
  4. $\quad \vec{v}_i = \frac{\vec{y}_i}{\|\vec{y}_i\|}$

     end for
  5. $\vec{x} = \vec{v}_i, \lambda = \frac{\vec{x}^T A\vec{x}}{\vec{x}^T\vec{x}}$
- How do we know the eigenvector is converged?

$$\begin{aligned} \vec{v}_1 &= \alpha_1 A\vec{v}_0 \\ \vec{v}_2 &= \alpha_2 A\vec{v}_1 = \alpha_2 A^2\vec{v}_0 \\ \vec{v}_3 &= \alpha_3 A\vec{v}_2 = \alpha_3 A^3\vec{v}_0 \\ &\vdots \\ \vec{v}_i &= \alpha_i A\vec{v}_{i-1} = \alpha_i A^i\vec{v}_0 \end{aligned}$$

Let $A = X\Lambda X^{-1}$, $\Lambda = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix}$, $\mid \lambda_1 \mid > \mid \lambda_2 \mid > \cdots > \mid \lambda_n \mid$

$$
\begin{aligned}
A^k &= \underbrace{(X\Lambda X^{-1})(X\Lambda X^{-1})\cdots(X\Lambda X^{-1})}_{k} \\
&= X\Lambda^k X^{-1} \\
&= X \begin{bmatrix} \lambda_1^k & & & \\ & \lambda_2^k & & \\ & & \ddots & \\ & & & \lambda_n^k \end{bmatrix} X^{-1}
\end{aligned}
$$

$$
\begin{aligned}
\Rightarrow \vec{v}_k &= \alpha_k A^k \vec{v}_0 \\
&= \alpha_k X\Lambda^k X^{-1}\vec{v}_0, \text{ Let } \vec{z} = X^{-1}\vec{v}_0 \\
&= \alpha_k X\Lambda^k \vec{z} \\
&= \alpha_k \lambda_1^k X \begin{bmatrix} 1 & & & \\ & (\frac{\lambda_2}{\lambda_1})^k & & \\ & & \ddots & \\ & & & (\frac{\lambda_n}{\lambda_1})^k \end{bmatrix} \vec{z}, \text{ because } \lim_{k\to\infty}(\frac{\lambda_i}{\lambda_1})^k = 0 \\
&= \alpha_k \lambda_1^k X \begin{bmatrix} 1 & & & \\ & 0 & & \\ & & \ddots & \\ & & & 0 \end{bmatrix} \vec{z} \\
&= \alpha_k \lambda_1^k X \begin{bmatrix} \vec{z}(1) \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \vec{z}(1) \text{ is the first element in } \vec{z} \\
&= \alpha_k \lambda_1^k \begin{bmatrix} \vec{x}_1, \vec{x}_2, \vec{x}_3, \ldots, \vec{x}_n \end{bmatrix} \begin{bmatrix} \vec{z}(1) \\ 0 \\ \vdots \\ 0 \end{bmatrix} \\
&= \alpha_k \lambda_1^k \vec{z}(1)\vec{x}_1, \text{ because } \alpha_k \lambda_1^k \vec{z}(1) = 1 \\
&= \vec{x}_1
\end{aligned}
$$

- Shift Power Method
  We can shift the eigenvalue to speed up.

$$B = A - \mu I = X\Lambda X^{-1} - \mu(XX^{-1}) = X(\Lambda - \mu I)X^{-1}$$
$$= X \begin{bmatrix} \lambda_1 - \mu & & & \\ & \lambda_2 - \mu & & \\ & & \ddots & \\ & & & \lambda_n - \mu \end{bmatrix} X^{-1}$$
$$\vec{r}_k = A\vec{v}_k - \frac{\vec{v}_k^T A \vec{v}_k}{\vec{v}_k^T \vec{v}_k}\vec{v}_k$$

  Converge rate: $\lim_{k\to\infty} \frac{\|r_k\|}{\|r_{k-1}\|} = \frac{|\lambda_2|}{|\lambda_1|}$
  If the rate is very small, the speed of convergence will be fast. Choose μ which makes $\frac{|\lambda_2|}{|\lambda_1|}$ as small as possible.

- Invert Power Method
  We can find the eigenvector whose eigenvaule is the smallest.

$$B = A^{-1} = (X\Lambda X^{-1})^{-1} = X\Lambda^{-1}X^{-1}$$
$$= X \begin{bmatrix} \frac{1}{\lambda_1} & & & \\ & \frac{1}{\lambda_2} & & \\ & & \ddots & \\ & & & \frac{1}{\lambda_n} \end{bmatrix} X^{-1}$$

- Shift and Invert Power Method

$$B = (A - \mu I)^{-1} = X(\Lambda - \mu I)^{-1}X^{-1}$$
$$= X \begin{bmatrix} \frac{1}{\lambda_1 - \mu} & & & \\ & \frac{1}{\lambda_2 - \mu} & & \\ & & \ddots & \\ & & & \frac{1}{\lambda_n - \mu} \end{bmatrix} X^{-1}$$