Recent Advance in Numerical Algorithms for Large/Sparse Eigenvalue Problems

Wen-Wei Lin

Department of Applied Mathematics
National Chiao Tung University, Taiwan

Outline

1. Power Method
2. Inverse Power Iteration

Eigenproblems

- Standard eigenproblems: $Ax = \lambda x$
- Generalized eigenproblems: $Ax = \lambda Bx$
- Higher order poly. eigenproblems:
  $(A_0 + \lambda A_1 + ... + \lambda^n A_n)x = 0$
- Eigenproblems of λ-matrices: $F(\lambda)x = 0$

What do we care ?

(i) In theory: eigenstructure, spectral decomposition, canonical form, . . . , etc.

(ii) In computation: eigenvalues, eigenvectors, invariant subspaces, . . . , etc.

Power Method

Given $A \in \mathbb{C}^{n \times n}$. Let A be diagonalizable and

$$Ax_i = \lambda_i x_i, \quad i = 1, \ldots, n$$

with

$$|\lambda_1| > |\lambda_2| \geq \cdots \geq |\lambda_n|.$$

Goal

Find the maximum eigenvalue λ_1 and the corresponding eigenvector x_1.

Let $u_0 \neq 0$ be a given vector. From the expansion

$$u_0 = \sum_{i=1}^{n} \alpha_i x_i, \quad \alpha_1 \neq 0,$$

follows that

$$\begin{aligned} A^k u_0 &= \sum_{i=1}^{n} \alpha_i \lambda_i^k x_i \\ &= \lambda_1^k \left\{ \alpha_1 x_1 + \sum_{i=2}^{n} \alpha_i \left(\frac{\lambda_i}{\lambda_1} \right)^k x_i \right\}. \qquad (1) \end{aligned}$$

Thus

$$\lambda_1^{-k} A^k u_0 \to \alpha_1 x_1 \qquad \text{as } k \to \infty.$$

Algorithm (Power Method with 2-norm)

Choose an initial $u \neq 0$ *with* $\|u\|_2 = 1$.
Iterate until convergence
Compute $v = Au;\ m = \|v\|_2;\ u := v/m$

Theorem

The sequence defined by Algorithm 1 is satisfied

$$\lim_{k\to\infty} m_k = |\lambda_1|,$$

$$\lim_{k\to\infty} \varepsilon^k u_k = \frac{x_1}{\|x_1\|}\frac{\alpha_1}{|\alpha_1|}, \quad \textit{where } \varepsilon = \frac{|\lambda_1|}{\lambda_1}.$$

Proof: It is obvious that

$$u_k = A^k u_0/\|A^k u_0\|, \quad m_k = \|A^k u_0\|/\|A^{k-1}u_0\|. \qquad (2)$$

This follows from ${\lambda_1}^{-k}A^k u_0 \to \alpha_1 x_1$ that

$$|\lambda_1|^{-k}\|A^k u_0\| \to |\alpha_1|\|x_1\|$$
$$|\lambda_1|^{-k+1}\|A^{k-1}u_0\| \to |\alpha_1|\|x_1\|$$

and then

$$|\lambda_1|^{-1}\|A^k u_0\|/\|A^{k-1}u_0\| = |\lambda_1|^{-1}m_k \to 1.$$

From (1) follows now for $k \to \infty$

$$\begin{aligned}\varepsilon^k u_k &= \varepsilon^k \frac{A^k u_0}{\|A^k u_0\|} = \frac{\alpha_1 x_1 + \sum_{i=2}^n \alpha_i \left(\frac{\lambda_i}{\lambda_1}\right)^k x_i}{\|\alpha_1 x_1 + \sum_{i=2}^n \alpha_i \left(\frac{\lambda_i}{\lambda_1}\right)^k x_i\|} \\ &\to \frac{\alpha_1 x_1}{\|\alpha_1 x_1\|} = \frac{x_1}{\|x_1\|}\frac{\alpha_1}{|\alpha_1|}.\end{aligned}$$

Algorithm (Power Method with Linear Function)

Choose an initial $u_0 \neq 0$.
Iterate until convergence
Compute $v_{k+1} = Au_k,\ m_{k+1} = \ell(v_{k+1}),\ u_{k+1} = v_{k+1}/m_{k+1}$
where $\ell(v_{k+1})$, *e.g.* $e_1(v_{k+1})$ *or* $e_n(v_{k+1})$, *is a linear functional.*

Theorem

Suppose $\ell(x_1) \neq 0$ *and* $\ell(v_k) \neq 0, k = 1, 2, \ldots,$ *then*

$$\lim_{k\to\infty} m_k = \lambda_1$$

$$\lim_{k\to\infty} u_k = \frac{x_1}{\ell(x_1)}.$$

Proof: As above we show that

$$u_k = A^k u_0 / \ell(A^k u_0), \quad m_k = \ell(A^k u_0)/\ell(A^{k-1} u_0).$$

From (1) we get for $k \to \infty$

$${\lambda_1}^{-k} \ell(A^k u_0) \to \alpha_1 \ell(x_1),$$

$${\lambda_1}^{-k+1} \ell(A^{k-1} u_0) \to \alpha_1 \ell(x_1),$$

thus

$${\lambda_1}^{-1} m_k \to 1, \ k \to \infty.$$

Similarly for $k \to \infty$,

$$u_k = \frac{A^k u_0}{\ell(A^k u_0)} = \frac{\alpha_1 x_1 + \sum_{i=2}^n \alpha_i (\frac{\lambda_i}{\lambda_1})^k x_i}{\ell(\alpha_1 x_1 + \sum_{i=2}^n \alpha_i (\frac{\lambda_i}{\lambda_1})^k x_i)} \to \frac{\alpha_1 x_1}{\alpha_1 \ell(x_1)}$$

- Note that:

$$
\begin{aligned}
m_k &= \frac{\ell(A^k u_0)}{\ell(A^{k-1}u_0)} = \lambda_1 \frac{\alpha_1 \ell(x_1) + \sum_{i=2}^{n} \alpha_i (\frac{\lambda_i}{\lambda_1})^k \ell(x_i)}{\alpha_1 \ell(x_1) + \sum_{i=2}^{n} \alpha_i (\frac{\lambda_i}{\lambda_1})^{k-1} \ell(x_i)} \\
&= \lambda_1 + O\left(\left| \frac{\lambda_2}{\lambda_1} \right|^{k-1} \right).
\end{aligned}
$$

That is the convergent rate is $\left| \frac{\lambda_2}{\lambda_1} \right|$.

Inverse Power Iteration

Goal

Find the eigenvalue of A that is in a given region or closest to a certain scalar σ and the corresponding eigenvector.

- This is achieved by iterating with the matrix $(A - \sigma I)^{-1}$.
- Often, σ is referred to as the "Shift".

Algorithm (Inverse power method with a fixed shift)

Choose an initial $u_0 \neq 0$.
For $k = 0, 1, 2, \ldots$
  Compute $v_{k+1} = (A - \sigma I)^{-1} u_k$ *and* $m_{k+1} = \ell(v_{k+1})$.
  Set $u_{k+1} = v_{k+1}/m_{k+1}$

- The convergence of Algorithm 3 is $|\frac{\lambda_1-\sigma}{\lambda_2-\sigma}|$ whenever λ_1 and λ_2 are the closest and the second closest eigenvalues to σ.
- Algorithm 3 is linearly convergent.

Algorithm (Inverse power method with variant shifts)

Choose an initial $u_0 \neq 0$.
Given $\sigma_0 = \sigma$.
For $k = 0, 1, 2, \ldots$
  Compute $v_{k+1} = (A - \sigma_k I)^{-1} u_k$ *and* $m_{k+1} = \ell(v_{k+1})$.
  Set $u_{k+1} = v_{k+1}/m_{k+1}$ *and* $\sigma_{k+1} = \sigma_k + 1/m_{k+1}$.

- Above algorithm is locally quadratic convergent.

Connection with Newton method

Consider the nonlinear equations:

$$F\left(\begin{bmatrix} u \\ \lambda \end{bmatrix}\right) \equiv \begin{bmatrix} Au - \lambda u \\ \ell^T u - 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \qquad (3)$$

Newton method for (3): for $k = 0, 1, 2, \ldots$

$$\begin{bmatrix} u_{k+1} \\ \lambda_{k+1} \end{bmatrix} = \begin{bmatrix} u_k \\ \lambda_k \end{bmatrix} - \left[F'\left(\begin{bmatrix} u_k \\ \lambda_k \end{bmatrix}\right)\right]^{-1} F\left(\begin{bmatrix} u_k \\ \lambda_k \end{bmatrix}\right).$$

Since

$$F'\left(\begin{bmatrix} u \\ \lambda \end{bmatrix}\right) = \begin{bmatrix} A - \lambda I & -u \\ \ell^T & 0 \end{bmatrix},$$

the Newton method can be rewritten by component-wise

$$\begin{aligned} (A - \lambda_k)u_{k+1} &= (\lambda_{k+1} - \lambda_k)u_k && (4) \\ \ell^T u_{k+1} &= 1. && (5) \end{aligned}$$

Let

$$v_{k+1} = \frac{u_{k+1}}{\lambda_{k+1} - \lambda_k}.$$

Substituting v_{k+1} into (4), we get

$$(A - \lambda_k I)v_{k+1} = u_k.$$

By equation (5), we have

$$m_{k+1} = \ell(v_{k+1}) = \frac{\ell(u_{k+1})}{\lambda_{k+1} - \lambda_k} = \frac{1}{\lambda_{k+1} - \lambda_k}.$$

It follows that

$$\lambda_{k+1} = \lambda_k + \frac{1}{m_{k+1}}.$$

Hence the Newton's iterations (4) and (5) are identified with Algorithm 4.

Theorem

Let $u \neq 0$ and for any μ set $r_\mu = Au - \mu u$. Then $\|r_\mu\|_2$ is minimized when

$$\mu = \theta = u^* Au / u^* u.$$

In this case $r_\theta \perp u$.

Proof : W.L.O.G. assume $\|u\|_2 = 1$. Let $[u\ U]$ be unitary and set

$$\begin{bmatrix} u^* \\ U^* \end{bmatrix} A \begin{bmatrix} u & U \end{bmatrix} = \begin{bmatrix} u^* Au & u^* AU \\ U^* Au & U^* AU \end{bmatrix} \equiv \begin{bmatrix} \theta & h^* \\ g & B \end{bmatrix}.$$

Then

$$\begin{aligned} \begin{bmatrix} u^* \\ U^* \end{bmatrix} r_\mu &= \begin{bmatrix} u^* \\ U^* \end{bmatrix} Au - \mu \begin{bmatrix} u^* \\ U^* \end{bmatrix} u \\ &= \begin{bmatrix} u^* \\ U^* \end{bmatrix} A \begin{bmatrix} u & U \end{bmatrix} \begin{bmatrix} u^* \\ U^* \end{bmatrix} u - \mu \begin{bmatrix} u^* \\ U^* \end{bmatrix} u \\ &= \begin{bmatrix} \theta & h^* \\ g & B \end{bmatrix} \begin{bmatrix} u^* \\ U^* \end{bmatrix} u - \mu \begin{bmatrix} u^* \\ U^* \end{bmatrix} u \\ &= \begin{bmatrix} \theta & h^* \\ g & B \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} - \mu \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} \theta - \mu \\ g \end{bmatrix}. \end{aligned}$$

It implies that

$$\|r_\mu\|_2^2 = \| \begin{bmatrix} u^* \\ U^* \end{bmatrix} r_\mu \|_2^2 = \| \begin{bmatrix} \theta - \mu \\ g \end{bmatrix} \|_2^2 = |\theta - \mu|^2 + \|g\|_2^2.$$

Therefore

$$\min_\mu \|r_\mu\|_2 = \|g\|_2 = \|r_\theta\|_2.$$

That is

$$\mu = \theta = u^* A u,$$

and

$$u^* r_\theta = u^*(Au - \theta u) = u^* A u - \theta = 0.$$

Hence, $r_\theta \perp u$. $\square$

Definition

Let u be a nonzero vector. Then $u^* A u / u^* u$ is called a Rayleigh quotient.

If u is an eigenvector corresponding to an eigenvalue λ of A, then

$$\frac{u^* A u}{u^* u} = \frac{\lambda u^* u}{u^* u} = \lambda.$$

$\Rightarrow u_k^* A u_k / u_k^* u_k$ provide a sequence of approximation to λ in the power method.

Algorithm (Inverse power method with Rayleigh Quotient)

Choose an initial $u_0 \neq 0$ *with* $\|u_0\|_2 = 1$.
Compute $\sigma_0 = u_0^\top A u_0$.
For $k = 0, 1, 2, \ldots$
  Compute $v_{k+1} = (A - \sigma_k I)^{-1} u_k$.
  Set $u_{k+1} = v_{k+1} / \|v_{k+1}\|_2$ *and* $\sigma_{k+1} = u_{k+1}^\top A u_{k+1}$.

- For symmetric A, above algorithm is cubically convergent.