Power and inverse power methods

Tsung-Ming Huang

Department of Mathematics
National Taiwan Normal University, Taiwan

February 15, 2011

Outline

Definition 1

1. An eigenvalue whose geometric multiplicity is less than its algebraic multiplicity is defective.
2. The matrix $A \in \mathbb{C}^{n\times n}$ has a complete system of eigenvectors if it has n linearly independent eigenvectors.

Let A be a nondefective matrix and (λ_i, x_i) for $i = 1, \cdots, n$ be a complete set of eigenpairs of A. That is $\{x_1, \cdots, x_n\}$ is linearly independent. Hence, for any $u_0 \neq 0$, $\exists\ \alpha_1, \cdots, \alpha_n$ such that

$$u_0 = \alpha_1 x_1 + \cdots + \alpha_n x_n.$$

Now $A^k x_i = \lambda_i^k x_i$, so that

$$A^k u_0 = \alpha_1 \lambda_1^k x_1 + \cdots + \alpha_n \lambda_n^k x_n. \tag{1}$$

If $|\lambda_1| > |\lambda_i|$ for $i \geq 2$ and $\alpha_1 \neq 0$, then

$$\frac{1}{\lambda_1^k} A^k u_0 = \alpha_1 x_1 + (\frac{\lambda_2}{\lambda_1})^k \alpha_2 x_2 + \cdots + \alpha_n (\frac{\lambda_n}{\lambda_1})^k x_n \rightarrow \alpha_1 x_1 \text{ as } k \rightarrow \infty.$$

Algorithm 1 (Power Method with 2-norm)

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

Theorem 2

The sequence defined by Algorithm 1 is satisfied

$$\lim_{i\to\infty} k_i = |\lambda_1|$$

$$\lim_{i\to\infty} \varepsilon^i u_i = \frac{x_1}{\|x_1\|}\frac{\alpha_1}{|\alpha_1|}, \quad \text{where } \varepsilon = \frac{|\lambda_1|}{\lambda_1}$$

Proof: It is obvious that

$$u_s = A^s u_0 / \|A^s u_0\|, \quad k_s = \|A^s u_0\| / \|A^{s-1} u_0\|. \tag{2}$$

This follows from ${\lambda_1}^{-s} A^s u_0 \longrightarrow \alpha_1 x_1$ that

$$|\lambda_1|^{-s} \|A^s u_0\| \longrightarrow |\alpha_1| \|x_1\|$$

$$|\lambda_1|^{-s+1} \|A^{s-1} u_0\| \longrightarrow |\alpha_1| \|x_1\|$$

and then

$$|\lambda_1|^{-1} \|A^s u_0\| / \|A^{s-1} u_0\| = |\lambda_1|^{-1} k_s \longrightarrow 1.$$

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

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

Algorithm 2 (Power Method with Linear Function)

Choose an initial $u \neq 0$.
Iterate until convergence
Compute $v = Au; \; k = \ell(v); \; u := v/k$
where $\ell(v)$, *e.g.* $e_1(v)$ *or* $e_n(v)$, *is a linear functional.*

Theorem 3

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

$$\begin{aligned} \lim_{i\to\infty} k_i &= \lambda_1 \\ \lim_{i\to\infty} u_i &= \frac{x_1}{\ell(x_1)}. \end{aligned}$$

Proof: As above we show that

$$u_i = A^i u_0 / \ell(A^i u_0), \quad k_i = \ell(A^i u_0)/\ell(A^{i-1} u_0).$$

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

$$\lambda_1{}^{-i} \ell(A^i u_0) \longrightarrow \alpha_1 \ell(x_1),$$

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

thus

$$\lambda_1{}^{-1} k_i \longrightarrow 1.$$

Similarly for $i \longrightarrow \infty$,

$$u_i = \frac{A^i u_0}{\ell(A^i u_0)} = \frac{\alpha_1 x_1 + \sum_{j=2}^{n} \alpha_j (\frac{\lambda_j}{\lambda_1})^i x_j}{\ell(\alpha_1 x_1 + \sum_{j=2}^{n} \alpha_j (\frac{\lambda_j}{\lambda_1})^i x_j)} \longrightarrow \frac{\alpha_1 x_1}{\alpha_1 \ell(x_1)}$$

- Note that:

$$
\begin{aligned}
k_i &= \frac{\ell(A^i u_0)}{\ell(A^{i-1}u_0)} = \lambda_1 \frac{\alpha_1 \ell(x_1) + \sum_{j=2}^{n} \alpha_j (\frac{\lambda_j}{\lambda_1})^i \ell(x_j)}{\alpha_1 \ell(x_1) + \sum_{j=2}^{n} \alpha_j (\frac{\lambda_j}{\lambda_1})^{i-1} \ell(x_j)} \\
&= \lambda_1 + O\left(| \frac{\lambda_2}{\lambda_1} |^{i-1} \right).
\end{aligned}
$$

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

Theorem 4

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

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

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

Proof: W.L.O.G. assume $\|u\|_2 = 1$. Let $\begin{pmatrix} u & U \end{pmatrix}$ be unitary and set

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

Then

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

It follows that

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

Hence

$$\min_{\mu} \|r_\mu\|_2 = \|g\|_2 = \|r_\nu\|_2.$$

That is $\mu = \nu = u^* A u$. On the other hand, since

$$u^* r_\mu = u^*(Au - \mu u) = u^* A u - \mu = 0,$$

it implies that $r_\mu \perp u$. ∎

Definition 5 (Rayleigh quotient)

Let u and v be vectors with $v^* u \neq 0$. Then $v^* A u / v^* u$ is called a Rayleigh quotient.

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

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

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

Inverse power method

Goal

Find the eigenpair (λ, x) of A where λ is belonged to a given region or closest to a certain scalar σ.

Let $\lambda_1, \cdots, \lambda_n$ be the eigenvalues of A. Suppose λ_1 is simple and $\sigma \approx \lambda_1$. Then

$$\mu_1 = \frac{1}{\lambda_1 - \sigma}, \mu_2 = \frac{1}{\lambda_2 - \sigma}, \cdots, \mu_n = \frac{1}{\lambda_n - \sigma}$$

are eigenvalues of $(A - \sigma I)^{-1}$ and $\mu_1 \to \infty$ as $\sigma \to \lambda_1$. Thus we transform λ_1 into a dominant eigenvalue μ_1.

The inverse power method is simply the power method applied to $(A - \sigma I)^{-1}$.

Algorithm 3 (Inverse power method with a fixed shift)

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

- 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.

Let (λ, x) be an eigenpair of A, i.e.,

$$Ax = \lambda x \Rightarrow (A - \sigma I)x = (\lambda - \sigma)x \Rightarrow (A - \sigma I)^{-1}x = \frac{1}{\lambda - \sigma}x \equiv \mu x.$$

It implies that

$$\lambda = \sigma + \mu^{-1}.$$

Algorithm 4 (Inverse power method with variant shifts)

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

- Above algorithm is locally quadratic convergent.

Connection with Newton method

Consider the nonlinear equations:

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

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

$$\begin{bmatrix} u_{i+1} \\ \lambda_{i+1} \end{bmatrix} = \begin{bmatrix} u_i \\ \lambda_i \end{bmatrix} - \left[F'\left(\begin{bmatrix} u_i \\ \lambda_i \end{bmatrix}\right)\right]^{-1} F\left(\begin{bmatrix} u_i \\ \lambda_i \end{bmatrix}\right).$$

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

$$(A - \lambda_i I)u_{i+1} = (\lambda_{i+1} - \lambda_i)u_i \tag{4}$$

$$\ell^T u_{i+1} = 1. \tag{5}$$

Let

$$v_{i+1} = \frac{u_{i+1}}{\lambda_{i+1} - \lambda_i}.$$

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

$$(A - \lambda_i I)v_{i+1} = u_i.$$

By equation (5), we have

$$k_{i+1} = \ell(v_{i+1}) = \frac{\ell(u_{i+1})}{\lambda_{i+1} - \lambda_i} = \frac{1}{\lambda_{i+1} - \lambda_i}.$$

It follows that

$$\lambda_{i+1} = \lambda_i + \frac{1}{k_{i+1}}.$$

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

Algorithm 5 (Inverse power method with Rayleigh Quotient)

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

- For symmetric A, Algorithm 5 is cubically convergent.