

# Lecture 8: Power/Inverse iteration, Rayleigh quotient iteration



School of Mathematical Sciences, Xiamen University

## 1. Eigenvalue problem and polynomial rootfinding problem

- The eigenvalues of a matrix  $\mathbf{A} \in \mathbb{C}^{m \times m}$  are the  $m$  roots of its characteristic polynomial  $p(z) = \det(z\mathbf{I} - \mathbf{A})$ .
- Suppose we have the monic polynomial

$$p(z) = z^m + a_{m-1}z^{m-1} + \cdots + a_1z + a_0.$$

It is not hard to verify that  $p(z) = \det(z\mathbf{I} - \mathbf{A})$ , where the  $m \times m$  matrix  $\mathbf{A}$  is

$$\mathbf{A} = \begin{bmatrix} 0 & & & & -a_0 \\ 1 & 0 & & & -a_1 \\ & 1 & 0 & & -a_2 \\ & & 1 & \ddots & \vdots \\ & & & \ddots & 0 & -a_{m-2} \\ & & & & 1 & -a_{m-1} \end{bmatrix}.$$

The matrix  $\mathbf{A}$  is called a *companion matrix* corresponding to  $p(z)$ .

- **Any eigenvalue solver must be iterative** because no explicit root expressing formula exists for polynomial of degree  $\geq 5$ . The goal of an eigenvalue solver is to produce sequences of numbers that converge rapidly towards eigenvalues.

- Convergence rate

Let  $e_1, e_2, e_3, \dots$  be a sequence of nonnegative numbers representing errors in some iterative process that converge to zero, and suppose there are a positive constant  $C$  and an exponent  $\alpha$  such that for all sufficiently large  $k$ ,  $e_{k+1} \leq C(e_k)^\alpha$ . Then,

- (1)  $\alpha = 1$  and  $C < 1$ , *linear convergence* or *geometric convergence*;
- (2)  $\alpha = 2$ , *quadratic convergence*;
- (3)  $\alpha = 3$ , *cubic convergence*; .....

## 2. Rayleigh quotient

- The *Rayleigh quotient* of a nonzero vector  $\mathbf{x} \in \mathbb{C}^m$  with respect to  $\mathbf{A}$  is the scalar

$$r(\mathbf{x}) = \frac{\mathbf{x}^* \mathbf{A} \mathbf{x}}{\mathbf{x}^* \mathbf{x}}.$$

## Theorem 1

Let  $\{\lambda, \mathbf{v}\}$  be an eigenpair of the matrix  $\mathbf{A}$ , i.e.,  $\mathbf{A}\mathbf{v} = \lambda\mathbf{v}$ . We further assume that  $\|\mathbf{v}\|_2 = 1$ .

- (i) If  $\mathbf{A}$  is non-normal, then the Rayleigh quotient  $r(\mathbf{x})$  is *generally* a linearly accurate estimate of the eigenvalue  $\lambda$ , i.e.,

$$|r(\mathbf{x}) - \lambda| = \mathcal{O}(\|\mathbf{x} - \mathbf{v}\|_2), \quad \text{as } \mathbf{x} \rightarrow \mathbf{v}.$$

- (ii) If  $\mathbf{A}$  is normal, then the Rayleigh quotient  $r(\mathbf{x})$  is a quadratically accurate estimate of the eigenvalue  $\lambda$ , i.e.,

$$|r(\mathbf{x}) - \lambda| = \mathcal{O}(\|\mathbf{x} - \mathbf{v}\|_2^2), \quad \text{as } \mathbf{x} \rightarrow \mathbf{v}.$$

## Proof.

Consider Schur form  $\mathbf{T} = \mathbf{Q}^* \mathbf{A} \mathbf{Q}$  with  $t_{11} = \lambda$  and  $\mathbf{Q}\mathbf{e}_1 = \mathbf{v}$ . □

### 3. Power iteration

---

**Algorithm 1:** Power iteration

---

$\mathbf{v}^{(0)}$  = some vector with  $\|\mathbf{v}^{(0)}\|_2 = 1$

**for**  $k = 1, 2, 3, \dots$ ,

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

$\mathbf{v}^{(k)} = \mathbf{w} / \|\mathbf{w}\|_2$

$\lambda^{(k)} = (\mathbf{v}^{(k)})^* \mathbf{A} \mathbf{v}^{(k)}$

**end**

---

- One application: Google's Pagerank.
- Power iteration can find only the eigenvector corresponding to the eigenvalue with the largest magnitude. The convergence is *linear*, which is very slow if the largest two eigenvalues are close in magnitude.

- Assume that  $\mathbf{A} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^{-1}$  is diagonalizable with  $\|\mathbf{v}_1\|_2 = 1$  and

$$\mathbf{\Lambda} = \text{diag}\{\lambda_1, \lambda_2, \dots, \lambda_m\}, \quad |\lambda_1| > |\lambda_2| \geq \dots \geq |\lambda_m|.$$

Let  $\begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_m \end{bmatrix} := \mathbf{V}^{-1}\mathbf{v}^{(0)}$ . If  $\alpha_1 \neq 0$ , then we have

$$\mathbf{A}^k \mathbf{v}^{(0)} = \mathbf{V} \mathbf{\Lambda}^k \mathbf{V}^{-1} \mathbf{v}^{(0)} = \mathbf{V} \begin{bmatrix} \alpha_1 \lambda_1^k \\ \alpha_2 \lambda_2^k \\ \vdots \\ \alpha_m \lambda_m^k \end{bmatrix} = \alpha_1 \lambda_1^k \mathbf{V} \begin{bmatrix} 1 \\ \frac{\alpha_2}{\alpha_1} \frac{\lambda_2^k}{\lambda_1^k} \\ \vdots \\ \frac{\alpha_m}{\alpha_1} \frac{\lambda_m^k}{\lambda_1^k} \end{bmatrix}.$$

Then  $\mathbf{v}^{(k)} = \frac{\mathbf{A}^k \mathbf{v}^{(0)}}{\|\mathbf{A}^k \mathbf{v}^{(0)}\|_2} \rightarrow e^{i\theta_k} \mathbf{v}_1$  and  $\lambda^{(k)} \rightarrow \lambda_1$ , where

$$\theta_k = k\theta + \theta_0, \quad e^{i\theta} = \lambda_1/|\lambda_1|, \quad e^{i\theta_0} = \alpha_1/|\alpha_1|.$$

## Theorem 2

Suppose that  $\mathbf{A}$  is diagonalizable, i.e.,  $\mathbf{A} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^{-1}$  with

$$\mathbf{\Lambda} = \text{diag}\{\lambda_1, \lambda_2, \dots, \lambda_m\}.$$

Furthermore, suppose  $\mathbf{e}_1^* \mathbf{V}^{-1} \mathbf{v}^{(0)} \neq 0$ ,  $\|\mathbf{v}_1\|_2 = 1$ , and

$$|\lambda_1| > |\lambda_2| \geq \dots \geq |\lambda_m|.$$

Then the iterates of power iteration satisfy, as  $k \rightarrow \infty$ ,

$$\|\mathbf{v}^{(k)} - e^{i\theta_k} \mathbf{v}_1\|_2 = \mathcal{O}\left(\left|\frac{\lambda_2}{\lambda_1}\right|^k\right),$$

and

$$|\lambda^{(k)} - \lambda_1| = \mathcal{O}\left(\left|\frac{\lambda_2}{\lambda_1}\right|^k\right) \quad \text{or} \quad \mathcal{O}\left(\left|\frac{\lambda_2}{\lambda_1}\right|^{2k}\right).$$

## 4. Shifted inverse iteration

### Proposition 3

*For any  $\mu$  that is not an eigenvalue, the eigenvectors of  $(\mathbf{A} - \mu\mathbf{I})^{-1}$  are the same as the eigenvectors of  $\mathbf{A}$ , and the corresponding eigenvalues are  $\{(\lambda_j - \mu)^{-1}\}$ , where  $\{\lambda_j\}$  are the eigenvalues of  $\mathbf{A}$ .*

---

#### Algorithm 2: Shifted inverse iteration

---

$\mathbf{v}^{(0)}$  = some vector with  $\|\mathbf{v}^{(0)}\|_2 = 1$

**for**  $k = 1, 2, 3, \dots$ ,

    Solve  $(\mathbf{A} - \mu\mathbf{I})\mathbf{w} = \mathbf{v}^{(k-1)}$  for  $\mathbf{w}$

$\mathbf{v}^{(k)} = \mathbf{w} / \|\mathbf{w}\|_2$

$\lambda^{(k)} = (\mathbf{v}^{(k)})^* \mathbf{A} \mathbf{v}^{(k)}$

**end**

---

- Like power iteration, shifted inverse iteration exhibits only *linear* convergence. **Other important issue: Exercise 27.5**



## Theorem 4

Suppose that  $\mathbf{A}$  is diagonalizable, i.e.,  $\mathbf{A} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^{-1}$  with

$$\mathbf{\Lambda} = \text{diag}\{\lambda_1, \lambda_2, \dots, \lambda_m\}.$$

Suppose  $\lambda_j$  is the closest eigenvalue to  $\mu$  and  $\lambda_l$  is the second closest, that is,

$$|\lambda_j - \mu| < |\lambda_l - \mu| \leq |\lambda_i - \mu|$$

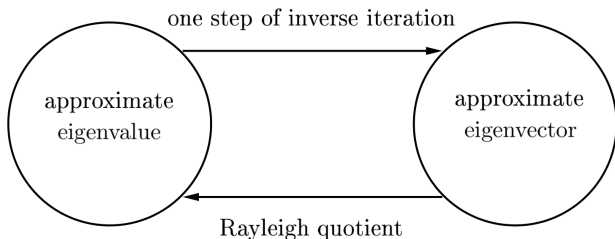
for each  $i \neq j$ . Furthermore, suppose  $\mathbf{e}_j^* \mathbf{V}^{-1} \mathbf{v}^{(0)} \neq 0$  and  $\|\mathbf{v}_j\|_2 = 1$ . Then the iterates of shifted inverse iteration satisfy, as  $k \rightarrow \infty$ ,

$$\|\mathbf{v}^{(k)} - e^{i\theta_k} \mathbf{v}_j\|_2 = \mathcal{O} \left( \left| \frac{\lambda_j - \mu}{\lambda_l - \mu} \right|^k \right), \quad (\text{Exercise : } \theta_k = ?)$$

and

$$|\lambda^{(k)} - \lambda_j| = \mathcal{O} \left( \left| \frac{\lambda_j - \mu}{\lambda_l - \mu} \right|^k \right) \quad \text{or} \quad \mathcal{O} \left( \left| \frac{\lambda_j - \mu}{\lambda_l - \mu} \right|^{2k} \right).$$

## 5. Rayleigh quotient iteration



---

### Algorithm 3: Rayleigh quotient iteration

---

$\mathbf{v}^{(0)}$  = some vector with  $\|\mathbf{v}^{(0)}\|_2 = 1$

$\lambda^{(0)} = (\mathbf{v}^{(0)})^* \mathbf{A} \mathbf{v}^{(0)}$

**for**  $k = 1, 2, 3, \dots$ ,

    Solve  $(\mathbf{A} - \lambda^{(k-1)} \mathbf{I}) \mathbf{w} = \mathbf{v}^{(k-1)}$  for  $\mathbf{w}$

$\mathbf{v}^{(k)} = \mathbf{w} / \|\mathbf{w}\|_2$

$\lambda^{(k)} = (\mathbf{v}^{(k)})^* \mathbf{A} \mathbf{v}^{(k)}$

**end**

---

## Theorem 5

*Rayleigh quotient iteration converges to an eigenpair*

$$\{\lambda, \mathbf{v}\}, \quad \|\mathbf{v}\|_2 = 1,$$

*for all except a set of measure zero of starting vectors  $\mathbf{v}^{(0)}$ . When it converges, the convergence is ultimately quadratic ( $\alpha = 2$ ) for non-normal case or cubic ( $\alpha = 3$ ) for normal case in the sense that if  $\mathbf{v}^{(k)}$  is sufficiently close to the eigenvector  $e^{i\theta_k} \mathbf{v}$ , then*

$$\|\mathbf{v}^{(k+1)} - e^{i\theta_{k+1}} \mathbf{v}\|_2 = \mathcal{O}(\|\mathbf{v}^{(k)} - e^{i\theta_k} \mathbf{v}\|_2^\alpha)$$

*and*

$$|\lambda^{(k+1)} - \lambda| = \mathcal{O}(|\lambda^{(k)} - \lambda|^\alpha)$$

*as  $k \rightarrow \infty$ .*

Example:

$$\mathbf{A} = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 3 & 1 \\ 1 & 1 & 4 \end{bmatrix}, \quad \mathbf{v}^{(0)} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} / \sqrt{3}.$$

The eigenvalue  $\lambda = 5.214319743377$

Power iteration:  $\lambda^{(0)} = 5.1818 \dots$

$$\lambda^{(1)} = 5.2081 \dots$$

$$\lambda^{(2)} = 5.2130 \dots$$

Rayleigh quotient iteration:  $\lambda^{(0)} = 5$

$$\lambda^{(1)} = 5.2131 \dots$$

$$\lambda^{(2)} = 5.214319743184 \dots$$

The convergence of Rayleigh quotient iteration is spectacular:  
each iteration triples the number of digits of accuracy.