

# Lecture 7: Eigenvalue problem



School of Mathematical Sciences, Xiamen University

# 1. Eigenvalues

- The *eigenvalues* of a matrix  $\mathbf{A} \in \mathbb{C}^{m \times m}$  are the  $m$  roots of its *characteristic polynomial*

$$p(z) = \det(z\mathbf{I} - \mathbf{A}).$$

- We have

$$\det(\mathbf{A}) = \lambda_1 \lambda_2 \cdots \lambda_m, \quad \text{tr}(\mathbf{A}) = \lambda_1 + \lambda_2 + \cdots + \lambda_m.$$

## Theorem 1 (Gerschgorin's theorem)

*Every eigenvalue of  $\mathbf{A}$  lies in at least one of the  $m$  circular disks in the complex plane with centers  $a_{ii}$  and radii  $\sum_{j \neq i} |a_{ij}|$ . Moreover, if  $n$  of these disks form a connected domain that is disjoint from the other  $m - n$  disks, then there are precisely  $n$  eigenvalues of  $\mathbf{A}$  within this domain.*

The proof is left as an exercise.

## Theorem 2

*Eigenvalues of  $\mathbf{A}$  are continuous functions of entries of  $\mathbf{A}$ .*

## Proof.

See Demmel's book: Proposition 4.4, Page 149, **Applied numerical linear algebra**. □

## Remark 3

*Eigenvalues of  $\mathbf{A}$  are not necessarily differentiable everywhere.*

**Example:** Consider the  $m \times m$  matrix

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & \ddots & \ddots & \\ & & & 0 & 1 \\ \varepsilon & & & & 0 \end{bmatrix}. \quad \lambda_j(\varepsilon) = \varepsilon^{\frac{1}{m}} \exp\left(\frac{i2j\pi}{m}\right).$$

## 2. Eigenvectors

- A nonzero vector  $\mathbf{y} \in \mathbb{C}^m$  is called a *left eigenvector* of  $\mathbf{A} \in \mathbb{C}^{m \times m}$  corresponding to  $\lambda \in \Lambda(\mathbf{A})$  if  $\mathbf{y}^* \mathbf{A} = \lambda \mathbf{y}^*$ .
- A nonzero vector  $\mathbf{x} \in \mathbb{C}^m$  is called a (*right*) *eigenvector* of  $\mathbf{A} \in \mathbb{C}^{m \times m}$  corresponding to  $\lambda \in \Lambda(\mathbf{A})$  if  $\mathbf{A} \mathbf{x} = \lambda \mathbf{x}$ .

### Theorem 4

If  $\mathbf{A} \in \mathbb{C}^{m \times m}$  and if  $\lambda, \mu \in \Lambda(\mathbf{A})$ , with  $\lambda \neq \mu$ , then any left eigenvector of  $\mathbf{A}$  corresponding to  $\mu$  is orthogonal to any right eigenvector of  $\mathbf{A}$  corresponding to  $\lambda$ .

### Proof.

Let  $\mathbf{y}^* \mathbf{A} = \mu \mathbf{y}^*$  and  $\mathbf{A} \mathbf{x} = \lambda \mathbf{x}$ . We have

$$\mathbf{y}^* \mathbf{A} \mathbf{x} = \mathbf{y}^* (\lambda \mathbf{x}) = \lambda (\mathbf{y}^* \mathbf{x}), \quad \mathbf{y}^* \mathbf{A} \mathbf{x} = (\mu \mathbf{y}^*) \mathbf{x} = \mu (\mathbf{y}^* \mathbf{x}).$$

Then,  $\mathbf{y}^* \mathbf{x} = 0$  follows from  $\lambda \neq \mu$ . □

### 3. Geometric multiplicity and algebraic multiplicity

- The *geometric multiplicity* of an eigenvalue  $\lambda$  is the dimension of the null-space of  $\mathbf{A} - \lambda\mathbf{I}$ , which is an *eigenspace* corresponding to the eigenvalue  $\lambda$ .
- The *algebraic multiplicity* of an eigenvalue  $\lambda$  is its multiplicity as a root of the characteristic polynomial. The algebraic multiplicity of an eigenvalue is at least as great as its geometric multiplicity.
- An eigenvalue is *simple* if its algebraic multiplicity is 1. Otherwise, *multiple*.

#### Remark 5

*Simple eigenvalue of  $\mathbf{A}$  is differential at  $\mathbf{A} \in \mathbb{C}^{m \times m}$ .*

#### Theorem 6

*An eigenvalue is multiple if and only if it has a pair of orthogonal left and right eigenvectors.*

The proof is left as an exercise.

## 4. Jordan form

### Theorem 7

*For any square matrix  $\mathbf{A}$  there exists a similar matrix  $\mathbf{J} = \mathbf{SAS}^{-1}$  such that*

$$\mathbf{J} = \text{diag}\{\mathbf{J}_1, \mathbf{J}_2, \dots, \mathbf{J}_k\}$$

*where each  $\mathbf{J}_i$  is a Jordan block:  $\mathbf{J}_i =$*

$$\begin{bmatrix} \lambda_i & 1 & & & \\ & \lambda_i & 1 & & \\ & & \ddots & \ddots & \\ & & & \lambda_i & 1 \\ & & & & \lambda_i \end{bmatrix}.$$

- Up to permuting the order of the  $\mathbf{J}_i$ , the Jordan form is unique.
- Up to a nonzero constant, there are only one left eigenvector and one right eigenvector per  $\mathbf{J}_i$ .
- **Discussion:** How to determine the rank of  $\mathbf{A}$  via its Jordan form?

- Jordan form is a discontinuous function of  $\mathbf{A}$ , so any rounding error can change it completely. Therefore, Jordan form is theoretically useful only.

**Example:** Consider the matrix

$$\mathbf{A}(\varepsilon) = \begin{bmatrix} \varepsilon & 1 & & \\ & 2\varepsilon & \ddots & \\ & & \ddots & 1 \\ & & & m\varepsilon \end{bmatrix}.$$

It is easy to show that

$$\lim_{\varepsilon \rightarrow 0} \mathbf{J}(\mathbf{A}(\varepsilon)) \neq \mathbf{J}(\mathbf{A}(0)) = \begin{bmatrix} 0 & 1 & & \\ & 0 & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{bmatrix}.$$

## 5. Schur form

### Theorem 8 (Schur factorization)

If  $\mathbf{A} \in \mathbb{C}^{m \times m}$ , then there exists a unitary matrix  $\mathbf{Q} \in \mathbb{C}^{m \times m}$  and an upper-triangular matrix  $\mathbf{T} \in \mathbb{C}^{m \times m}$  such that  $\mathbf{A} = \mathbf{Q}\mathbf{T}\mathbf{Q}^*$ .

Proof. By induction on the dimension  $m$  of  $\mathbf{A}$ .  $\square$

### Remark 9

See Demmel's book (Applied numerical linear algebra, Theorem 4.3, Page 147) for real Schur form of a real matrix  $\mathbf{A}$ .

**Exercise:** Let  $\lambda_1, \dots, \lambda_m$  be the  $m$  eigenvalues of  $\mathbf{A} \in \mathbb{C}^{m \times m}$ . Let

$$\mathbf{M} = \frac{\mathbf{A} + \mathbf{A}^*}{2}, \quad \mathbf{N} = \frac{\mathbf{A} - \mathbf{A}^*}{2}.$$

Prove that

$$\sum_{i=1}^m |\lambda_i|^2 \leq \|\mathbf{A}\|_{\text{F}}^2, \quad \sum_{i=1}^m |\operatorname{Re} \lambda_i|^2 \leq \|\mathbf{M}\|_{\text{F}}^2, \quad \sum_{i=1}^m |\operatorname{Im} \lambda_i|^2 \leq \|\mathbf{N}\|_{\text{F}}^2.$$



- Let  $\mathbf{A} = \mathbf{Q}\mathbf{T}\mathbf{Q}^*$  be a Schur factorization. If  $\{\lambda, \mathbf{x}\}$  is an eigenpair of  $\mathbf{T}$ , then  $\{\lambda, \mathbf{Q}\mathbf{x}\}$  is an eigenpair of  $\mathbf{A}$ .

## 6. Unitary diagonalization

- A matrix  $\mathbf{A}$  is called *unitarily diagonalizable* if there exists a unitary matrix  $\mathbf{Q}$  and a diagonal matrix  $\mathbf{\Lambda}$  such that  $\mathbf{A} = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^*$ .

**Examples:** Hermitian, skew-Hermitian, ...

- A matrix  $\mathbf{A}$  is called *normal* if  $\mathbf{A}^*\mathbf{A} = \mathbf{A}\mathbf{A}^*$ .

**Examples:** Hermitian, skew-Hermitian, ...

### Theorem 10

*A matrix is unitarily diagonalizable if and only if it is normal.*

### Proof.

“ $\Rightarrow$ ”: Easy. “ $\Leftarrow$ ” By Schur factorization of  $\mathbf{A}$ . □

## 7. Eigenvalue perturbation theory

### Theorem 11 (Bauer–Fike)

Suppose  $\mathbf{A} \in \mathbb{C}^{m \times m}$  is diagonalizable with  $\mathbf{A} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^{-1}$ , and let  $\mathbf{\Delta} \in \mathbb{C}^{m \times m}$  be arbitrary. For each eigenvalue  $\hat{\lambda}$  of  $\mathbf{A} + \mathbf{\Delta}$ , there exists an eigenvalue  $\lambda$  of  $\mathbf{A}$  such that

$$|\hat{\lambda} - \lambda| \leq \|\mathbf{V}\|_2 \|\mathbf{V}^{-1}\|_2 \|\mathbf{\Delta}\|_2.$$

Proof. Assume that  $\{\hat{\lambda}, \mathbf{V}\mathbf{y}\}$  is an eigenpair of  $\mathbf{A} + \mathbf{\Delta}$ . Then we have

$$(\hat{\lambda}\mathbf{I} - \mathbf{A})\mathbf{y} = \mathbf{V}^{-1}\mathbf{\Delta}\mathbf{V}\mathbf{y}.$$

Thus,  $\min_{\lambda \in \Lambda(\mathbf{A})} |\hat{\lambda} - \lambda| \leq \frac{\|(\hat{\lambda}\mathbf{I} - \mathbf{A})\mathbf{y}\|_2}{\|\mathbf{y}\|_2} \leq \|\mathbf{V}\|_2 \|\mathbf{V}^{-1}\|_2 \|\mathbf{\Delta}\|_2. \quad \square$

### Corollary 12

If  $\mathbf{A}$  is normal, i.e.,  $\mathbf{A}\mathbf{A}^* = \mathbf{A}^*\mathbf{A}$ , then for each eigenvalue  $\hat{\lambda}$  of  $\mathbf{A} + \mathbf{\Delta}$ , there is an eigenvalue  $\lambda$  of  $\mathbf{A}$  such that  $|\hat{\lambda} - \lambda| \leq \|\mathbf{\Delta}\|_2$ .

## 8. Hermitian matrix eigenvalues

### Theorem 13 (Courant–Fisher)

If  $\mathbf{A} \in \mathbb{C}^{m \times m}$  is Hermitian, then the eigenvalues  $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_m$  satisfy

$$\begin{aligned}\lambda_k &= \max_{\mathcal{S} \subseteq \mathbb{C}^m, \dim(\mathcal{S})=k} \min_{\mathbf{0} \neq \mathbf{y} \in \mathcal{S}} \frac{\mathbf{y}^* \mathbf{A} \mathbf{y}}{\mathbf{y}^* \mathbf{y}} \\ &= \min_{\mathcal{S} \subseteq \mathbb{C}^m, \dim(\mathcal{S})=m-k+1} \max_{\mathbf{0} \neq \mathbf{y} \in \mathcal{S}} \frac{\mathbf{y}^* \mathbf{A} \mathbf{y}}{\mathbf{y}^* \mathbf{y}},\end{aligned}$$

for  $k = 1, 2, \dots, m$ .

### Theorem 14 (Interlacing property)

If  $\mathbf{A} \in \mathbb{C}^{m \times m}$  is Hermitian and  $\mathbf{A}_k = \mathbf{A}(1:k, 1:k)$ , then

$$\begin{aligned}\lambda_{k+1}(\mathbf{A}_{k+1}) &\leq \lambda_k(\mathbf{A}_k) \leq \lambda_k(\mathbf{A}_{k+1}) \leq \\ &\cdots \leq \lambda_2(\mathbf{A}_{k+1}) \leq \lambda_1(\mathbf{A}_k) \leq \lambda_1(\mathbf{A}_{k+1})\end{aligned}$$

for  $k = 1 : m - 1$ .

## Theorem 15 (Weyl)

Let  $\mathbf{A} \in \mathbb{C}^{m \times m}$  and  $\mathbf{B} \in \mathbb{C}^{m \times m}$  be Hermitian. Let  $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_m$  be eigenvalues. Then

$$|\lambda_k(\mathbf{A}) - \lambda_k(\mathbf{B})| \leq \|\mathbf{A} - \mathbf{B}\|_2, \quad k = 1, 2, \dots, m.$$

## Corollary 16

Let  $\mathbf{A} \in \mathbb{C}^{m \times n}$  and  $\mathbf{B} \in \mathbb{C}^{m \times n}$  be arbitrary. Let  $p = \min\{m, n\}$  and  $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_p$  be singular values. Then

$$|\sigma_k(\mathbf{A}) - \sigma_k(\mathbf{B})| \leq \|\mathbf{A} - \mathbf{B}\|_2, \quad k = 1, 2, \dots, p.$$

## Theorem 17

Let  $\mathbf{A} \in \mathbb{C}^{l \times m}$  and  $\mathbf{B} \in \mathbb{C}^{m \times n}$  be arbitrary. Let  $p = \min\{l, m, n\}$  and  $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_p$  be singular values. Then

$$\sigma_k(\mathbf{AB}) \leq \sigma_1(\mathbf{A})\sigma_k(\mathbf{B}), \quad k = 1, 2, \dots, p.$$

## 9. Generalized eigenvalue problem

- For  $\mathbf{A}, \mathbf{B} \in \mathbb{C}^{m \times m}$ ,  $\{\lambda, \mathbf{x}\}$  is called an eigenpair if  $\{\lambda, \mathbf{x}\}$  satisfies  $\mathbf{x} \neq \mathbf{0}$  and  $\mathbf{A}\mathbf{x} = \lambda\mathbf{B}\mathbf{x}$ .
- If  $\det(z\mathbf{B} - \mathbf{A})$  is not identically zero, the pencil  $z\mathbf{B} - \mathbf{A}$  is called *regular*. Otherwise it is called *singular*.
- When  $z\mathbf{B} - \mathbf{A}$  is regular,  $p(z) = \det(z\mathbf{B} - \mathbf{A})$  is called the *characteristic polynomial* of the pencil  $z\mathbf{B} - \mathbf{A}$  and the eigenvalues of the pencil  $z\mathbf{B} - \mathbf{A}$  are defined to be
  - (1) the roots of  $p(z) = 0$ ,
  - (2)  $\infty$  (with multiplicity  $m - \deg(p)$ ) if  $\deg(p) < m$ .
- Generalized Schur factorization for  $(\mathbf{A}, \mathbf{B})$ :

$$\mathbf{Q}^* \mathbf{A} \mathbf{Z} = \mathbf{S}, \quad \mathbf{Q}^* \mathbf{B} \mathbf{Z} = \mathbf{T},$$

where  $\mathbf{Q}$  and  $\mathbf{Z}$  are unitary, and  $\mathbf{S}$  and  $\mathbf{T}$  are upper-triangular.

- Real generalized Schur forms for real matrices  $\mathbf{A}$  and  $\mathbf{B}$ .
- QZ algorithm for generalized eigenvalue problem.

## 10. Matrix polynomial eigenvalue problem

- We consider the matrix polynomial

$$\mathcal{A}(z) := \sum_{i=0}^d z^i \mathbf{A}_i = z^d \mathbf{A}_d + z^{d-1} \mathbf{A}_{d-1} + \cdots + z \mathbf{A}_1 + \mathbf{A}_0,$$

where  $\mathbf{A}_i \in \mathbb{C}^{m \times m}$ . When  $d = 1$ , it is a generalized eigenvalue problem; and when  $d \geq 2$ , it is a nonlinear eigenvalue problem.

- The characteristic polynomial of the matrix polynomial  $\mathcal{A}(z)$  is

$$p(z) = \det(\mathcal{A}(z)).$$

The roots of  $p(z) = 0$  and  $\infty$  (with multiplicity  $md - \deg(p)$  if  $\deg(p) < md$ ) are defined to be the eigenvalues.

- Suppose that  $\lambda$  is an eigenvalue. A nonzero vector  $\mathbf{x}$  satisfying  $\mathcal{A}(\lambda)\mathbf{x} = \mathbf{0}$  is a right eigenvector for  $\lambda$ . A left eigenvector  $\mathbf{y}$  is defined analogously by  $\mathbf{y}^* \mathcal{A}(\lambda) = \mathbf{0}$ .

## 10.1. The quadratic eigenvalue problem

- F. Tisseur and K. Meerbergen, SIAM Review, 43: 235–286, 2001.

Consider the ODE system

$$\mathbf{M}\ddot{\mathbf{x}}(t) + \mathbf{B}\dot{\mathbf{x}}(t) + \mathbf{K}\mathbf{x}(t) = \mathbf{0},$$

where  $\mathbf{M}, \mathbf{B}, \mathbf{K} \in \mathbb{C}^{m \times m}$ . If we seek solutions of the form  $\mathbf{x}(t) = e^{\lambda t} \mathbf{x}(0)$ , we get

$$e^{\lambda t} (\lambda^2 \mathbf{M} \mathbf{x}(0) + \lambda \mathbf{B} \mathbf{x}(0) + \mathbf{K} \mathbf{x}(0)) = \mathbf{0},$$

i.e.,

$$\lambda^2 \mathbf{M} \mathbf{x}(0) + \lambda \mathbf{B} \mathbf{x}(0) + \mathbf{K} \mathbf{x}(0) = \mathbf{0}.$$

Thus  $\lambda$  is an eigenvalue and  $\mathbf{x}(0)$  is an eigenvector of the matrix polynomial

$$z^2 \mathbf{M} + z \mathbf{B} + \mathbf{K}.$$

- Tacoma Narrows Bridge, London Millennium Bridge, Humen Bridge

## 10.2. Linearization of matrix polynomial eigenvalue problem

- The generalized eigenvalue problem:

$$z \begin{bmatrix} \mathbf{A}_d & & & & \\ & \mathbf{I} & & & \\ & & \mathbf{I} & & \\ & & & \ddots & \\ & & & & \mathbf{I} \end{bmatrix} - \begin{bmatrix} -\mathbf{A}_{d-1} & -\mathbf{A}_{d-2} & \cdots & \cdots & -\mathbf{A}_0 \\ & \mathbf{I} & & & \\ & & \mathbf{I} & & \\ & & & \ddots & \\ & & & & \mathbf{I} \end{bmatrix}$$

- The standard eigenvalue problem if  $\mathbf{A}_d$  is nonsingular:

$$z\mathbf{I} - \begin{bmatrix} -\mathbf{A}_d^{-1}\mathbf{A}_{d-1} & -\mathbf{A}_d^{-1}\mathbf{A}_{d-2} & \cdots & \cdots & -\mathbf{A}_d^{-1}\mathbf{A}_0 \\ & \mathbf{I} & & & \\ & & \mathbf{I} & & \\ & & & \ddots & \\ & & & & \mathbf{I} \end{bmatrix}$$