

Lecture VIII

- **Vector norms and the associated matrix norms**
- **Some max and minimax principles**
- **Location of eigenvalues**

Vector Norms

Any **regular vector-norm** satisfies:

$|x|$ denotes the (**generic**) norm of vector x .

- $|x| > 0$, if $x \neq 0$; $|x| = 0$, if $x = 0$.
- $|\lambda x| = |\lambda| \cdot |x|$, for any scalar λ .
- $|x + y| \leq |x| + |y|$, "**triangle inequality**"
- $|x|$ depends continuously on x .
- $\exists \alpha, \beta > 0$ such that

$$\alpha \max_k |x_k| \leq |x| \leq \beta \max_k |x_k|, \quad \forall \text{ vector } x.$$

Comment

It is of interest to note that the last 2 properties in the definition of norm follow from the first 3 properties.

Examples of Norm

(1) The **Euclidean** norm of x is:

$$|x| = \left(\sum_{k=1}^n |x_k|^2 \right)^{\frac{1}{2}} \doteq \|x\| \text{ (or sometimes, } |x|_2 \text{)}.$$

(2) **Riemannian metric**:

$$|x| = \langle Px, x \rangle^{\frac{1}{2}} = \left(\sum_j \sum_k p_{jk} x_k \bar{x}_j \right)^{\frac{1}{2}}$$

where P is positive definite.

Examples of Norm

(3) The "Manhattan", or l_1 norm of x is:

$$|x|_1 = \sum_{k=1}^n |x_k|.$$

(4) The " l_∞ " norm: $|x|_\infty = \max_{1 \leq i \leq n} |x_i|.$

(5) The " l_p " norm: $|x|_p = \left(\sum_{i=1}^n |x_i|^p \right)^{1/p}.$

Hardy, Littlewood, Polya, *Inequalities*, Cambridge Univ. Press, 1988

Exercise

All norms over any finite-dimensional space are equivalent.

Some useful inequalities:

$$|x|_1 \leq \sqrt{n} |x|_2, \quad |x|_1 \leq n |x|_\infty,$$

$$|x|_2 \leq |x|_1, \quad |x|_2 \leq \sqrt{n} |x|_\infty,$$

$$|x|_\infty \leq |x|_1, \quad |x|_\infty \leq |x|_2.$$

The Related Matrix Norm

For any given regular vector-norm $|x|$, we can define the related matrix-norm as

$$|A| = \max_{x \neq 0} \frac{|Ax|}{|x|}, \quad A : n \times m \text{ matrix.}$$

Remark: It reduces to the vector norm when $m = 1$.

Example of matrix-norm

From the Euclidean vector-norm, define the related *spectral* matrix-norm:

$$\begin{aligned}\|A\|_2 &= \max_{x \neq 0} \frac{\|Ax\|_2}{\|x\|_2} \\&= \max_{x \neq 0} \sqrt{\frac{\langle Ax, Ax \rangle}{\langle x, x \rangle}} = \max_{x \neq 0} \sqrt{\frac{\langle A^* Ax, x \rangle}{\langle x, x \rangle}} \\&= \sqrt{\lambda_{\max}(A^* A)}, \text{ using Rayleigh's principle below.}\end{aligned}$$

↘ (maximum) Singular Value of A

For example,

$$A = \begin{pmatrix} 0 & -2 \\ 1 & 0 \end{pmatrix} \Rightarrow A^* A = \begin{pmatrix} 1 & 0 \\ 0 & 4 \end{pmatrix}$$

$$\Rightarrow \|A\| = \sqrt{4} = 2.$$

Other Matrix Norms

- **Frobenius norm** (or, Euclidean norm, l_2 -norm, Schur norm, Hilbert-Schmidt norm):

$$|A|_F := |A|_2 = \left(\sum_{i,j=1}^n |a_{ij}|^2 \right)^{1/2}.$$

- **l_1 -norm**: $|A|_1 = \sum_{i,j=1}^n |a_{ij}| = \sum_{j=1}^n |a^j|_1.$

- **l_∞ -norm**: $|A|_\infty = \sum_{j=1}^n |a^j|_\infty$, where $A = \begin{bmatrix} a^1 & a^2 & \cdots & a^n \end{bmatrix}.$

Fundamental Theorem

For any matrix norm $\|\bullet\|$, then

$$\rho(A) \leq \|A\|, \quad \text{with } A \text{ a square matrix}$$

where $\rho(A)$ is the **spectral radius** of A , i.e.

$$\rho(A) = \max \{|\lambda| : \lambda \text{ is an eigenvalue of } A\}.$$

Remark: Provided an upper-bound for all eigenvalues of any given matrix.

Indeed, we have

$$|\lambda||x| = |\lambda x| = |Ax| \leq \|A\| \bullet |x|$$

where x is an associated eigenvector.

So, $\rho(A) \leq \|A\|$.

Example

Verify this theorem on the following matrices:

$$1) A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix};$$

$$2) A = \begin{pmatrix} 0 & j \\ -j & 0 \end{pmatrix}.$$

You may consider the l_2 -norm only.

Question

What are the **tightest** bounds in

$$c_1 \|x\|^2 \leq \langle Hx, x \rangle \leq c_2 \|x\|^2$$

where H is an $n \times n$ Hermitian matrix.

Simplified Question

Given a Hermitian matrix H ,
what is the maximum of

$$\bar{u}^T H u := \langle H u, u \rangle, \quad \|u\| = 1$$

The Rayleigh Principle

Consider a Hermitian matrix H . Then,

$$\max_{\|u\|=1} \langle Hu, u \rangle = \lambda_1$$

where λ_1 is the largest eigenvalue of H .

Moreover, the equality is attained with u being a λ_1 -associated eigenvector.

Corollary

If H is Hermitian and λ_1 is its largest eigenvalue, then

$$\lambda_1 = \max_{x \neq 0} \frac{\langle Hx, x \rangle}{\langle x, x \rangle},$$

where $\frac{\langle Hx, x \rangle}{\langle x, x \rangle}$ is called Rayleigh quotient.

Comment

As a direct consequence, we obtain a useful inequality:

$$\left| x^T H x \right| \leq \lambda_{\max} (H) |x|^2, \quad \forall x$$

when H is also positive definite.

An Example

Consider the matrix $H = \begin{pmatrix} 4 & 3i \\ -3i & 2 \end{pmatrix}$.

- Is it Hermitian?
- Compute its eigenvalues.
- Verify the Rayleigh's Principle.

Answer

- Yes, it is Hermitian, and therefore, its eigenvalues must be real.
- $\lambda_{1,2} = 3 \pm \sqrt{10}$.
- $$\frac{\langle Hx, x \rangle}{\langle x, x \rangle} = \frac{4|x_1|^2 + 3i\bar{x}_1x_2 - 3i\bar{x}_2x_1 + 2|x_2|^2}{|x_1|^2 + |x_2|^2} \leq 3 + \sqrt{10} \quad (\text{using CFT. Do you know why/how?})$$

where the equality is attained when $\begin{bmatrix} x_1 & x_2 \end{bmatrix}^T = x$ is an eigenvector associated with $3 + \sqrt{10}$.

Comment

The Rayleigh principle *cannot* be applied to non-Hermitian matrix. Here is a simple counter-example:

$$A = \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix}, \quad b \neq 0.$$

Proof of the Rayleigh Principle

As shown previously with canonical diagonal form, a Hermitian matrix H only has real eigenvalues $\lambda_1 \geq \dots \geq \lambda_n$ and has mutually orthogonal eigenvectors

$$\{u^i\}_{i=1}^n, \text{ with } \|u^i\| = \sqrt{\langle u^i, u^i \rangle} = 1.$$

Notice that $\lambda_i = \langle Hu^i, u^i \rangle, \forall 1 \leq i \leq n.$

Proof (cont'd)

Since $\{u^i\}_{i=1}^n$ are mutually orthogonal, then, every **unit** vector u can be written as

$$u = c_1 u^1 + \cdots + c_n u^n,$$

$$\text{with } |c_1|^2 + \cdots + |c_n|^2 = 1.$$

On the other hand, $Hu = \sum c_i \lambda_i u^i$, implying

$$\langle Hu, u \rangle = \sum \lambda_i |c_i|^2 \leq \lambda_1, \text{ as wished.}$$

On Other Eigenvalues

Consider Hermitian H having real eigenvalues

$\lambda_1 \geq \dots \geq \lambda_n$. Let $\{u^k\}_{k=1}^{i-1}$ be mutually orthogonal unit eigenvectors associated with $\{\lambda_k\}_{k=1}^{i-1}$. Then,

$$\lambda_i = \max_{\substack{\|u\|=1 \\ \langle u, u^k \rangle = 0 \\ 1 \leq k \leq i-1}} \langle Hu, u \rangle$$

Benefit of quadratic forms

Courant's MinMax Theorem

To **independently** evaluate each eigenvalue:

Theorem: Let $\lambda_1 \geq \dots \geq \lambda_n$ be the eigenvalues of a Hermitian matrix $H \in \mathbb{C}^{n \times n}$.

Then, for $1 \leq i < n$,

$$\lambda_{i+1} = \min_{v^1, \dots, v^i} \max_{\substack{\|u\|=1 \\ \langle u, v^k \rangle = 0 \\ k=1, \dots, i}} \langle Hu, u \rangle.$$

An Example

Consider the quadratic form

$$\langle Hu, u \rangle = 3u_1^2 + 2u_2^2 + u_3^2 \Rightarrow H = \text{diag}(3, 2, 1).$$

Thus, $\lambda_1 = 3$, $\lambda_2 = 2$, $\lambda_3 = 1$, and

its associated eigenvectors are:

$$u^1 = e^1 = \text{col}(1, 0, 0), \quad u^2 = e^2 = \text{col}(0, 1, 0),$$

$$u^3 = e^3 = \text{col}(0, 0, 1).$$

An Example (cont'd)

By Rayleigh's principle,

$$\begin{aligned}\lambda_2 &= \max_{\substack{\|u\|=1 \\ \langle u, e^1 \rangle = 0}} \left(3u_1^2 + 2u_2^2 + u_3^2 \right) \\ &= \max_{u_2^2 + u_3^2 = 1} \left(2u_2^2 + u_3^2 \right) = 2, \text{ as expected.}\end{aligned}$$

By Courant's principle,

$$\lambda_2 = \min_{v^1} \max_{\substack{\|u\|=1 \\ \langle u, v^1 \rangle = 0}} \left(3u_1^2 + 2u_2^2 + u_3^2 \right) := \min_{v^1} \phi(v^1).$$

An Example (cont'd)

Clearly, $\phi(0) = \lambda_1 = 3$ (Rayleigh's principle).

When $v^1 \neq 0$, to compute λ_2 is equivalent to solving two (constrained) nonlinear optimization problems.

Here, we apply a graphical proof to yield:

$$\lambda_2 = 2 = \phi(e^1) = \min \phi(v^1),$$

because $3u_1^2$ is the dominating term.

A Useful Test

For any $n \times n$ Hermitian matrix $H = (h_{ij})$, it is positive definite **if and only if** all its leading principal minors are positive:

$$h_{11} > 0, \det \begin{pmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{pmatrix} > 0, \dots,$$

$$\det \begin{pmatrix} h_{11} & \dots & h_{1n} \\ \vdots & \ddots & \vdots \\ h_{n1} & \dots & h_{nn} \end{pmatrix} > 0.$$

Recursive Proof

\Rightarrow : If H is positive definite, then all its eigenvalues are positive. Then, $\det H > 0$.

Let H_k denote the submatrix of H formed of the first k rows and columns of H .

Then, H_k must be positive definite, because

$$\langle H_k x^k, x^k \rangle = \langle Hx, x \rangle, \text{ with } x = \begin{pmatrix} x^k \\ 0 \end{pmatrix}.$$

That is: $\det H_k > 0$, as wished.

Recursive Proof

←: Now, $\det H_k > 0, \forall k$. In order to prove the positive definiteness of H , we need the following "Inclusion Principle".

The Inclusion Principle

From a Hermitian matrix $A = (a_{ij})_{n \times n}$, form an $(n-1) \times (n-1)$ matrix B by deleting the last row and column of A . Then, the eigenvalues $\alpha_1 \geq \dots \geq \alpha_n$ of A , and $\beta_1 \geq \dots \geq \beta_{n-1}$ of B satisfy:

$$\alpha_1 \geq \beta_1 \geq \alpha_2 \geq \dots \geq \alpha_{n-1} \geq \beta_{n-1} \geq \alpha_n.$$

Proof: using Courant's MinMax theorem.

Recursive Proof (cont'd)

\Leftarrow : Now, $\det H_k > 0, \forall k$. Then, $H_1 = (h_{11})$ is positive definite. By induction, assume that $H_k, k < n$, is positive definite. We then need to prove that H_{k+1} is positive definite.

Let $\alpha_1 \geq \dots \geq \alpha_{k+1}$ be the eigenvalues of H_{k+1} , and $\beta_1 \geq \dots \geq \beta_k$ be the eigenvalues of H_k .

It follows from the inclusion principle that

$$\alpha_1 \geq \beta_1 \geq \dots \geq \beta_k \geq \alpha_{k+1}$$

Recursive Proof (cont'd)

\Leftarrow : $\alpha_1 \geq \beta_1 \geq \dots \geq \beta_k \geq \alpha_{k+1}$ implies
 $\alpha_1 > 0, \dots, \alpha_k > 0$.

It remains to prove $\alpha_{k+1} > 0$ to conclude the positive definiteness of H_{k+1} .

Using $\alpha_1 \cdots \alpha_k \bullet \alpha_{k+1} = \det H_{k+1} > 0$, we have
 $\alpha_{k+1} > 0$.

Question

How to provide a fine characterization for the location of the eigenvalues of a matrix?

Comment

The location of eigenvalues of an LTI system determines the stability nature of the system. See **Lecture XII**.

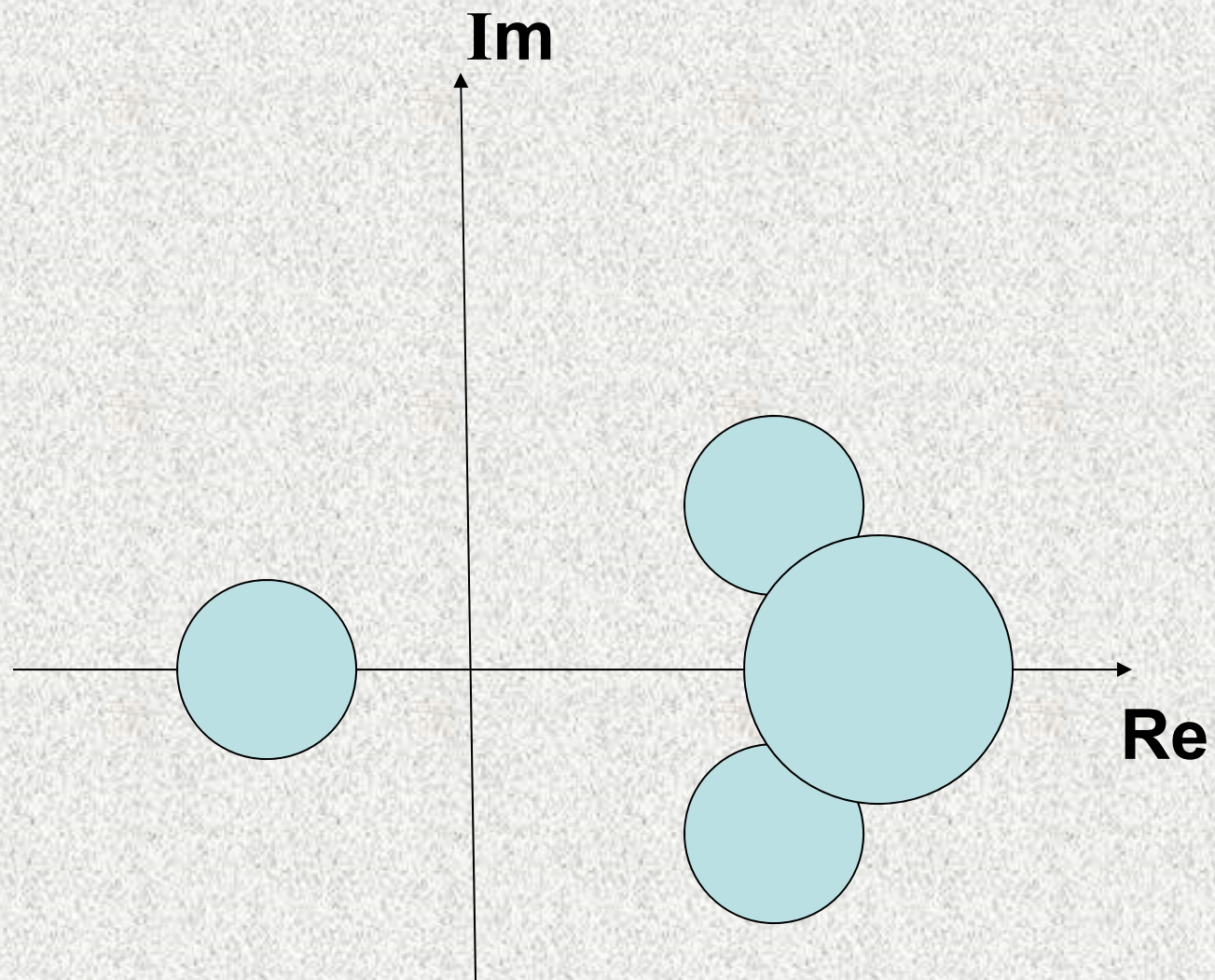
Gersgorin Disc Theorem

1) All eigenvalues of $A = [a_{ij}]_{n \times n}$ are located in the union

$$\text{of } n \text{ discs } \bigcup_{i=1}^n \left\{ z \in \mathbb{C} : |z - a_{ii}| \leq \sum_{j=1, j \neq i}^n |a_{ij}| \right\} \triangleq G(A).$$

2) If a union of k of these discs forms a connected region disjoint from the remaining $n - k$ discs, then there are exactly k eigenvalues in this region.

Proof: See the textbook (2nd Ed., 2013, pp. 387-389).



Essential Idea

Consider

$A = D + \varepsilon B$, ε sufficiently small
with $D = \text{diag}(a_{ii})$.

Observation:

The eigenvalues of perturbed matrix A should be "close" to those of the unperturbed matrix D .

Sketch of Proof

Take an eigenvector u associated with λ ,
and let $|u_m| = \max |u_j| \neq 0$. Using $(\lambda I - A)u = 0$,

$$(\lambda - a_{mm})u_m + \sum_{j \neq m} (-a_{mj})u_j = 0$$

$$\Rightarrow |\lambda - a_{mm}| \cdot |u_m| \leq \left| \sum_{j \neq m} a_{mj} u_j \right| \leq \sum_{j \neq m} |a_{mj}| \cdot |u_m|$$

$$\Rightarrow |\lambda - a_{mm}| \leq \sum_{j \neq m} |a_{mj}|, \text{ as wished.}$$

Comment

Since A and A^T have the same eigenvalues, all eigenvalues of $A = [a_{ij}]_{n \times n}$ are also located in the union

$$\text{of } n \text{ discs } \bigcup_{j=1}^n \left\{ z \in \mathbb{C} : |z - a_{jj}| \leq \sum_{i=1, i \neq j}^n |a_{ij}| \right\} \triangleq G(A^T).$$

Clearly, all eigenvalues of any matrix A are inside $G(A) \cap G(A^T)$.

Example

By means of this theorem, we can give an estimate of all eigenvalues (when the exact values are **not** easy to obtain). For example,

$$A = \begin{pmatrix} 0 & 1 & -1 \\ 1.3 & 2 & -0.7 \\ 0.5 & 0.5i & 4i \end{pmatrix}.$$

Exercise 1

Let $A = (a_{ij})_{n \times n}$. Show that

$$|\det A|^2 \leq \prod_{i=1}^n \left(\sum_{j=1}^n |a_{ij}| \right)^2.$$

When does the equality hold?

Exercise 2

Assume $B = (b_{ij})_{n \times n}$ satisfies

$$|b_{ii}| > \sum_{j=1, j \neq i}^n |b_{ij}|, \text{ for all } i = 1, \dots, n.$$

Show that $\det B \neq 0$.

Exercise 3

Consider the matrix

$$A = \begin{pmatrix} 7 & -16 & 8 \\ -16 & 7 & -8 \\ 8 & -8 & -5 \end{pmatrix}$$

1) Use the Geršgorin theorem to say as much as you can about the location of the eigenvalues of A and its spectral radius.

2) Then, consider $D^{-1}AD$, with $D = \text{diag}(p_1, p_2, p_3) > 0$.

Can you obtain any improvement in your location of the eigenvalues via appropriate choice of parameters p_i .

Homework 8

1. If λ_n is the least eigenvalue of a Hermitian

matrix H , show that $\lambda_n = \min_{x \neq 0} \frac{\langle Hx, x \rangle}{\langle x, x \rangle}$.

2. Find all possible values of μ guaranteeing the positive-definiteness of

$$H = \begin{pmatrix} 1 & 2 & 3 \\ 2 & \mu & 4 \\ 3 & 4 & 5 \end{pmatrix}.$$

Homework 8

3. Show that $|x| = \max_k |x_k|$, denoted as $|x|_\infty$,

and $|x| = \sum_k |x_k|$, denoted as $|x|_1$,

are both norms.

What are their associated matrix norms?