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.

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

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

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:

$$|x| = \langle Px, x \rangle^{\frac{1}{2}} = \left(\sum_{j} \sum_{k} p_{jk} x_k \overline{x}_j \right)^{\frac{1}{2}}$$

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: $|x|_\infty = \max_{1\le i\le n} |x_i|$.

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

Can you verify that these are indeed norms??

Exercise

All norms over any finite-dimensional space are equivalent.

Some useful inequalities:

$$|x|_1 \le \sqrt{n}\,|x|_2, \qquad |x|_1 \le n|x|_\infty,$$

$$|x|_2 \le |x|_1, \qquad |x|_2 \le \sqrt{n}\,|x|_\infty,$$

$$|x|_\infty \le |x|_1, \qquad |x|_\infty \le |x|_2.$$

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:

$$\|A\|_2 = \max_{x\neq 0} \frac{|Ax|_2}{|x|_2}$$

$$= \max_{x\neq 0} \sqrt{\frac{\langle Ax, Ax\rangle}{\langle x, x\rangle}} = \max_{x\neq 0} \sqrt{\frac{\langle A^* Ax, x\rangle}{\langle x, x\rangle}}$$

$$= \sqrt{\lambda_{\max}\left(A^* A\right)},$$ using Rayleigh's principle below.

(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 : $|A|_1 = \sum_{i,j=1}^{n} |a_{ij}|.$

- l_∞-norm: $|A|_\infty = \sum_{j=1}^{n} |a^j|_\infty$, where $A = \left[a^1 \; a^2 \; \cdots \; a^n \right].$

Fundamental Theorem

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

$$\rho(A) \le \|A\|, \qquad \text{with } A \text{ a square matrix}$$

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

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

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.

Morever, 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 $\dfrac{\langle Hx, x \rangle}{\langle x, x \rangle}$ is called Rayleigh quotient.

Comment

As a direct consequence, we obtain a useful inequality:

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

for H 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.
- $\lambda_{1,2} = 3 \pm \sqrt{10}.$
- $$\frac{\langle Hx, x\rangle}{\langle x, x\rangle} = \frac{4|x_1|^2 + 3i\bar{x}_1 x_2 - 3i\bar{x}_2 x_1 + 2|x_2|^2}{|x_1|^2 + |x_2|^2}$$

  $\leq 3 + \sqrt{10}$ (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 $\lambda_1 \geq \cdots \geq \lambda_n$ and has mutually orthogonal eigenvectors $\left\{u^i\right\}_{i=1}^n$, with $\left\|u^i\right\| = \sqrt{\left\langle u^i, u^i \right\rangle} = 1.$

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

Proof (cont'd)

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

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

with $\left|c_1\right|^2 + \cdots + \left|c_n\right|^2 = 1.$

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

$\left\langle Hu, u \right\rangle = \sum \lambda_i \left|c_i\right|^2 \le \lambda_1$, as wished.

On Other Eigenvalues

Consider Hermitian H having real eigenvalues $\lambda_1 \geq \cdots \geq \lambda_n$. Let $\left\{u^k\right\}_{k=1}^{i-1}$ be mutually orthogonal unit eigenvectors associated with $\left\{\lambda_k\right\}_{k=1}^{i-1}$. 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 $\lambda_1 \geq \cdots \geq \lambda_n$ be the eigenvalues of a Hermitian matrix $H \in \mathbb{C}^{n\times n}$.

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

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

An Example

Consider the quadratic form

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

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

its associated eigenvectors are:

$$u^1 = e^1 = col(1,\ 0,\ 0),\ u^2 = e^2 = col(0,\ 1,\ 0),$$

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

An Example (cont'd)

By Rayleigh's principle,

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

$$= \max_{u_2^2 + u_3^2 = 1} \left(2u_2^2 + u_3^2 \right) = 2, \text{ as expected.}$$

By Courant's principle,

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

An Example (cont'd)

Clearly, $\phi(0) = \lambda_1 = 3$ $\left(\text{Rayleigh's principle}\right)$.

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:

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

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

A Useful Test

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

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

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

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 $\left\langle H_k x^k, x^k \right\rangle = \langle Hx, x \rangle$, with $x = \left(x^k, 0\right)$.

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

Recursive Proof

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

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

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

It follows from the inclusion principle that

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

Recursive Proof (cont'd)

$\Leftarrow$: $\alpha_1 \geq \beta_1 \geq \cdots \geq \beta_k \geq \alpha_{k+1}$ implies

$\alpha_1 > 0, \ \ldots, \ \alpha_k > 0.$

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

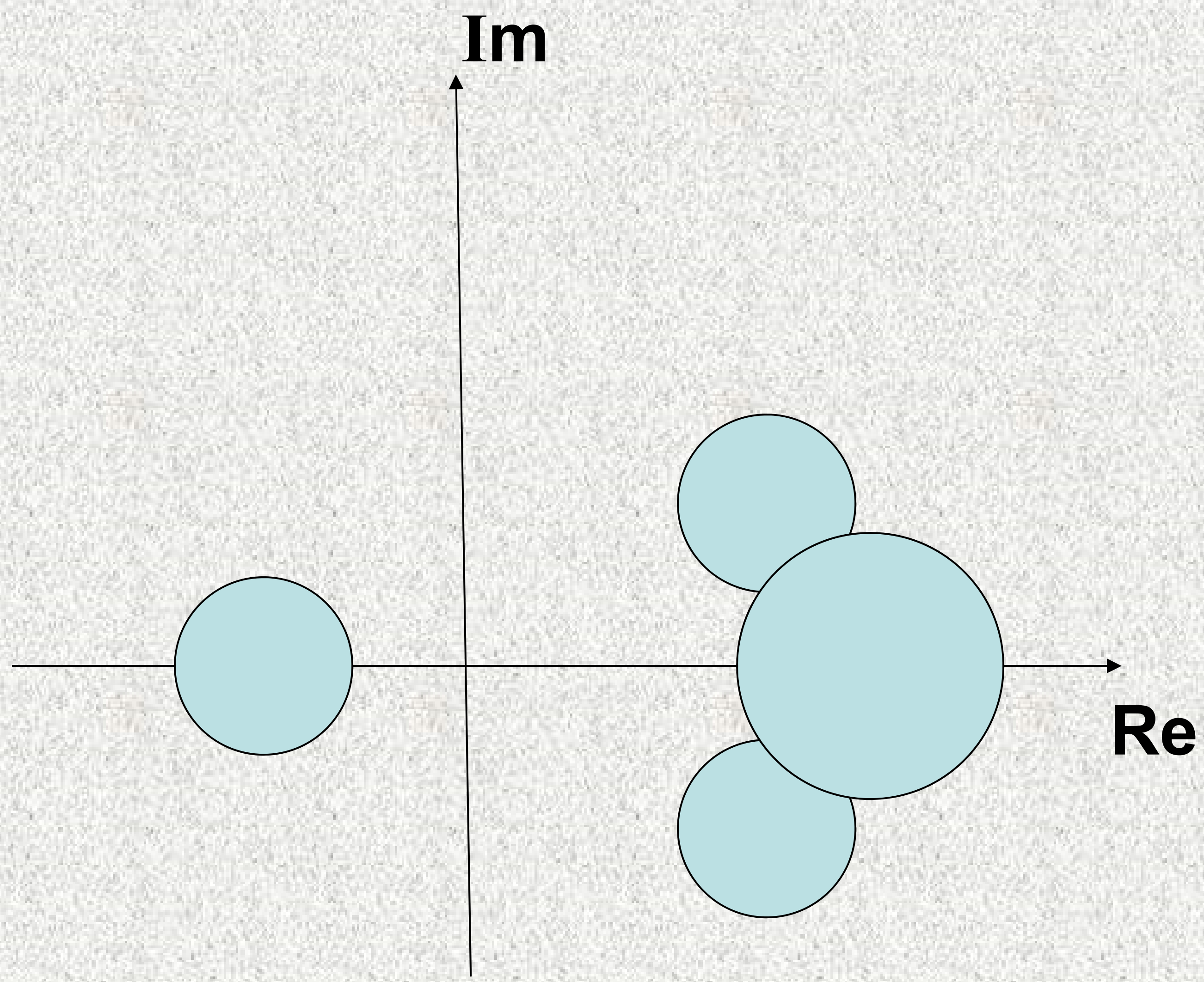
The location of eigenvalues of an LTI system determines the stability nature of the system. See **Lecture XII.**

Gersgorin Disc Theorem

1) All eigenvalues of $A = \left[a_{ij} \right]_{n \times n}$ are located in the union of n discs $\bigcup_{i=1}^{n} \left\{ z \in \mathbb{C} : \left| z - a_{ii} \right| \le \sum_{j=1,\, j \ne i}^{n} \left| a_{ij} \right| \right\} \triangleq G(A).$

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.


Im
Re

Essential Idea

Consider

$$A = D + \varepsilon B, \ \varepsilon \text{ sufficiently small}$$

with $D = diag\left(a_{ii}\right)$.

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 $(\lambda I - A)u = 0$,

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

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

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

Comment

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

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

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

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

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

When does the equality hold?

Exercise 2

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

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

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 $D = diag(p_1, p_2, p_3) > 0$.

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 $\lambda_n = \min_{x \neq 0} \dfrac{\langle Hx, x \rangle}{\langle x, x \rangle}$.

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?