Lecture VI

Extensions to Complex Matrices, in particular Hermitian Matrices.

Key Notions:

- **Unitary matrices**
- **Unitary equivalence**
- **Schur's unitary triangularization**
- **QR factorization**
- **Congruence and simultaneous diagonalization**

Orthogonality Between Complex Vectors

Given any pair of (*complex*) vectors $x, y \in \mathbb{C}^n$, the inner product is defined as

$$\langle x,\ y \rangle \triangleq y^* x$$

$$= x_1 \bar{y}_1 + x_2 \bar{y}_2 + \cdots + x_n \bar{y}_n .$$

They are said to be orthogonal, if

$$\langle x,\ y \rangle = 0.$$

Facts about the Inner Product

It can be easily checked that the inner product enjoys the following properties:

- $\langle x,\ y+z\rangle = \langle x,\ y\rangle + \langle x,\ z\rangle,\ \forall x, y, z \in \mathbb{C}^n.$
- $\langle x,\ \alpha y\rangle = \bar{\alpha}\langle x,\ y\rangle,\ \forall \alpha \in \mathbb{C},$ scalar.
- $\langle x,\ x\rangle = \begin{cases} \geq 0, & \forall x \in \mathbb{C}^n; \\ = 0, & \text{if and only if } x = 0. \end{cases}$

Orthogonal & Orthonomral Sets of Vectors

- A set of vectors $x^i \in \mathbb{C}^n$ is said to be orthogonal, if $\left\langle x^i,\ x^j \right\rangle = 0,\ \forall 1 \le i,\ j \le k, i \ne j.$

- A set of vectors $x^i \in \mathbb{C}^n$ is said to be orthonormal if, additionally, $\left\| x^i \right\| := \sqrt{\left\langle x^i,\ x^i \right\rangle} = 1,\ \forall 1 \le i \le k.$

Remark

Any orthogonal set of nonzero vectors $\left\{y^i\right\}_{i=1}^{k}$ can be made an orthonormal set, by defining

$$x^i := \frac{1}{\sqrt{\left\langle y^i, y^i \right\rangle}} y^i, \ \forall 1 \le i \le k.$$

Fundamental Results

1) Any orthogonal set of nonzero vectors is linearly independent.

2) Any orthonormal set of vectors is linearly independent.

Unitary Matrix

A matrix $U \in \mathbb{C}^{n\times n}$ is said to be unitary if $U^*U = I$. (Recall that $U^* \triangleq \bar{U}^T$)

Of course, a real orthogonal matrix $O \in \mathbb{R}^{n\times n}$ is unitary, but the converse is not true. Can you find some examples?

Complex Orthogonal Matrix

A matrix $A \in \mathbb{C}^{n \times n}$ is said to be complex orthogonal, if:

$$A^T A = I.$$

<u>*Remark*</u>:
A complex orthogonal matrix is unitary <u>if and only if</u> it is real.

Equivalent Characterizations

The following are equivalent:

- U is unitary;
- U is nonsingular and $U^* = U^{-1}$;
- $UU^* = I$;
- U^* is unitary;
- The columns of U form an orthonormal set;
- The rows of U form an orthonormal set;
- For any $x \in \mathbb{C}^n$, $y = Ux$ satisfies $y^* y = x^* x$.

Exercise

Are the following statements true or false?

1) For any given real parameters θ_i, $1 \le i \le n$,

$U = diag\left\{e^{j\theta_k}\right\}$ is always unitary.

2) Any diagonal unitary matrix can always be put into the above form.

3) Any diagonalizable unitary matrix can be transformed to the above form.

Question

How to apply a unitary matrix, instead of a real orthogonal matrix, to transform a Hermitian matrix into a canonical diagonal form?

Review: Canonical Form of a Real Symmetrical Matrix

Let $A \in \mathbb{R}^{n \times n}$ be a real symmetric matrix. Then, it can be transformed into the diagonal form by using an orthogonal matrix O so that

$$O^T A O = \begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{pmatrix}$$

where $\{\lambda_i\}_{i=1}^n$ are the eigenvalues of A.

Extension

It is possible to generalize this important result to (possibly complex) Hermitian matrices H, *i.e.*,

$$H^* = H.$$

In this case, we use unitary matrices U, instead of orthogonal matrices, *i.e.*,

$$U^* U = I.$$

Examples

- The matrix $\begin{pmatrix} 1 & 2+i \\ 2-i & -3 \end{pmatrix}$ is Hermitian.

- The matrix $\begin{pmatrix} 1 & 2+i \\ 2+i & -3 \end{pmatrix}$ is not Hermitian,

but is a complex symmetrical matrix.

Eigenvalues of Hermitian Matrices

The eigenvalues of a Hermitian matrix are real, and eigenvectors associated with distinct eigenvalues are orthogonal.

Canonical Transformation

If H is a Hermitian matrix, there exists a unitary matrix U such that

$$U^* H U = \begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{pmatrix}.$$

In particular, U becomes a real orthogonal matrix when H is a real symmetric matrix.

Idea of Proof

As in the case of real symmetric matrices, we use the Gram-Schmidt Orthogonalization Process, noting the following:

For complex vectors $x,\ y \in \mathbb{C}^n$, the inner product is defined as follows:

$$\langle x, y \rangle \triangleq \bar{y}^T x \triangleq \sum_{i=1}^{n} x_i \bar{y}_i .$$

Exercise

Compute the eigenvalues λ_1, λ_2 of

$$H = \begin{pmatrix} 1 & 2+i \\ 2-i & -3 \end{pmatrix}$$

and find a unitary matrix U that

reduces H to the diagonal form $\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$.

(Hint: use $\langle x, y \rangle = \sum_{i=1}^{n} x_i \overline{y}_i$ for *complex* vectors

x, y in the orthogonalization process.)

Schur's Unitary Triangularization

For *any* square, not necessarily Hermitian, $n \times n$ matrix A, there is a unitary matrix U for which

$$U^* A U = \begin{pmatrix} \lambda_1 & * & \cdots & * \\ 0 & & \ddots & \vdots \\ \vdots & & & * \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix}$$

with * being zero or nonzero scalars.

Algorithm

Step 1: Take a normalized eigenvector x^1 of A associated with an eigenvalue λ_1, and find $(n-1)$ vectors $\{y^2, \cdots, y^n\}$ so that $x^1, y^2, \cdots, y^n$ are linearly independent.

Algorithm

Step 2 : Apply the Gram-Schimidt orthonormalization procedure to $x^1,\ y^2,\cdots,y^n$ to produce an orthonormal set $x^1,\ z^2,\cdots,z^n$.

Define $U_1=\left[x^1,\ z^2,\cdots,z^n\right]$ which, clearly, is a unitary matrix.

Algorithm

Step 2 (cont'd) : Under $U_1 = \left[x^1,\ z^2, \cdots, z^n \right]$,

$$U_1^* A U_1 = \begin{pmatrix} \lambda_1 & * \\ 0 & A_1 \end{pmatrix}, \text{ with } A_1 \in \mathbb{C}^{(n-1)\times(n-1)}.$$

Of course, A_1 has eigenvalues $\lambda_2, \cdots, \lambda_n$.

Algorithm

Step 3: For $A_1 \in \mathbb{C}^{(n-1)\times(n-1)}$, apply Steps 1-2 to arrive at an orthonormal set $x^2,\ z_1^3,\ \ldots,\ z_1^n \in \mathbb{C}^{n-1}$ and a unitary matrix

$$U_2 = \left[x^2,\ z_1^3,\ \ldots,\ z_1^n \right] \in \mathbb{C}^{(n-1)\times(n-1)}$$

so that

$$U_2^* A_1 U_2 = \begin{pmatrix} \lambda_2 & * \\ 0 & A_2 \end{pmatrix}, \text{ with } A_2 \in \mathbb{C}^{(n-2)\times(n-2)}$$

Algorithm

Step 4 : It is easy to check that,

$$V_2 = \begin{pmatrix} 1 & 0 \\ 0 & U_2 \end{pmatrix} \text{ and } U_1 V_2 \in \mathbb{C}^{n\times n}$$

are both unitary. In addition,

$$(U_1 V_2)^* A (U_1 V_2) = \left(\begin{array}{cccc} \lambda_1 & * & * & * \\ 0 & \lambda_2 & * & * \\ \hline \multicolumn{2}{c}{O_{(n-2)\times 2}} & \multicolumn{2}{c}{A_2} \end{array}\right)$$

Algorithm

Last Step : Continuing these steps to arrive at the last step, where we have produced unitary matrices $U_i \in \mathbb{C}^{(n-i+1)(n-i+1)}$, and $V_i \in \mathbb{C}^{n\times n}$, $i = 2, 3, \ldots, n-1$

so that

- $U = U_1 V_2 \cdots V_{n-1}$, and
- $U^* A U = \begin{pmatrix} \lambda_1 & * & \cdots & * \\ \vdots & \ddots & \cdots & \vdots \\ 0 & & \cdots & \lambda_n \end{pmatrix}.$

Some Applications of Schur's Theorem

- **Useful for solving algebraic, differential or difference linear equations.**

Do you know why?

Applications of Schur's Theorem

- **Cayley-Hamilton Theorem**

Let $p_A(\lambda)$ be the characteristic polynomial of A,

that is, $p_A(\lambda) = \det(\lambda I - A) = \lambda^n + \alpha_1 \lambda^{n-1} + \cdots + \alpha_n$.

Then, $p_A(A) := A^n + \alpha_1 A^{n-1} + \cdots + \alpha_n I = 0$.

See the textbook of Horn & Johnson (2nd ed., 2013), pp. 109~110.

Comment

Cayley-Hamilton Theorem is extremely important in linear systems theory.

Technical Remark

For any square $n \times n$ matrix A, for any integer $i \geq n$, there exist constants $c_{i1}, \ldots, c_{in}$ such that

$$A^i = c_{i1} A^{n-1} + \cdots + c_{in-1} A + c_{in} I, \ \forall i \geq n.$$

Exercise

Consider the matrix $A = \begin{pmatrix} 3 & 2 \\ -1 & 0 \end{pmatrix}$.

- Use Cayley-Hamilton Theorem to express $A^2,\ A^3,\ A^4$ as linear combinations of $A,\ I$.
- Use Cayley-Hamilton Theorem to find the inverse A^{-1}.

QR Factorization

For any (possibly nonsquare) matrix $A \in \mathbb{C}^{n \times m}$, *with* $n \geq m$, $\exists Q \in \mathbb{C}^{n \times m}$, $R \in \mathbb{C}^{m \times m}$ such that

- The columns of Q form an orthonormal set, and R is an upper triangular matrix;
- $A = QR$.

If, in addition, A is nonsingular, then the diagonal entries of R are positive. Moreover, in this case, Q and R are unique.

Remark

The factors Q and R may be taken real, if A is a real matrix.

Proof: See the textbook, pp.89~90, for the constructive procedure closely tied to the Gram-Schmidt (G-S) algorithm.

An Example

What is the QR factorization of

$$A = \begin{pmatrix} 1 & 0 \\ 2 & 3 \end{pmatrix}$$

Solution

For simplicity, denote $A = \begin{pmatrix} 1 & 0 \\ 2 & 3 \end{pmatrix} := \begin{pmatrix} a^1 & a^2 \end{pmatrix}$.

Then, let $q^1 = a^1 / \left\| a^1 \right\| = \begin{pmatrix} \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{pmatrix}^T$ and,

like in the G-S process, compute

$$y^2 = a^2 - \left(q^{1*} a^2 \right) q^1 = \begin{pmatrix} -\frac{6}{5} & \frac{3}{5} \end{pmatrix}^T$$

Solution (cont'd)

Now, let $q^2 = y^2 / \left\| y^2 \right\| = \left(\dfrac{-2}{\sqrt{5}} \quad \dfrac{1}{\sqrt{5}} \right)^T$.

Set $Q = \left(q^1 \ \ q^2 \right)$ which, by construction,

is orthonormal. Then, $R = \left(r_{ij} \right)$, (with r_{kj}=0$\forall k > j$)

can be determined according to the general formula:

$$a^j = \sum_{k=1}^{j} r_{kj} q^k, \ j = 1, 2, \ldots, m$$

$\boxed{\textbf{m = 2, here}}$

Solution (end)

$$\text{So, } r_{11} = \sqrt{5},\ r_{21} = 0,\ r_{12} = \frac{6}{\sqrt{5}}, r_{22} = \frac{3}{\sqrt{5}}.$$

$$\text{That is:} \quad R = \begin{pmatrix} \sqrt{5} & \frac{6}{\sqrt{5}} \\ 0 & \frac{3}{\sqrt{5}} \end{pmatrix}$$

It is directly verified that $A = QR$.

Application to Cholesky factorization

By means of QR factorization, any matrix $B \in \mathbb{C}^{n \times n}$ taking the form $B = A^* A$, with $A \in \mathbb{C}^{n \times n}$, can be written as:

$$B = LL^*, \text{ with } L \in \mathbb{C}^{n \times n} \text{ lower triangular.}$$

Moreover, this factorization is unique, if A is nonsingular.

Indeed, it suffices to write $A = QR$.

QR Numerical Algorithm

This is a powerful tool for computing the eigenvalues of a matrix.

QR Numerical Algorithm

Step 1: For any given $A_0 \in \mathbb{C}^{n\times n}$, factorize

$$A_0 = Q_0 R_0$$

Step 2: Define $A_1 = R_0 Q_0$, and factorize

$$A_1 = Q_1 R_1$$

Continuing this process, we have

$$\forall k \geq 1, \quad \begin{cases} A_k = Q_k R_k \\ A_{k+1} = R_k Q_k \end{cases}$$

Proposition

- Each A_k is unitarily equivalent to A_0, and thus they have the same eigenvalues.

- If A_0 has distinct eigenvalues, then A_k converges to an upper triangular matrix.

A Numerical Exercise

Use MATLAB simulation to validate the *QR* algorithm for the matrix

$$A = \begin{pmatrix} 1 & 0 \\ 2 & 3 \end{pmatrix}.$$

Congruence

Consider two matrices $A,\ B\ \in \mathbb{C}^{n\times n}$.

(1) B is said to be $*$*congruent to* A, if $B = SAS^*$ for some nonsingular matrix S.

(2) B is said to be *congruent*, or ^{T}*congruent to* A, if $B = SAS^T$ for some nonsingular matrix S.

Notice that both congruence are equivalence relations.
(Horn-Johnson, 2nd ed., 2013; p. 281)

Inertia

Consider a Hermitian matrix $A \in \mathbb{C}^{n \times n}$.

Its inertia is defined as the ordered triple:

$$i(A) = \left(i_+(A),\ i_-(A),\ i_0(A)\right)$$

where

$i_+(A) =$ the number of positive eigenvalues of A;

$i_-(A) =$ the number of negative eigenvalues of A;

$i_0(A) =$ the number of zero eigenvalues of A.

Sylvester's Law of Inertia

Hermitian matrices $A, B \in \mathbb{C}^{n\times n}$ are *congruent if and only if they have the same inertia, i.e., the same number of positive eigenvalues and the same number of negative eigenvalues.

For the proof, see (Horn-Johnson, 2nd Ed., 2013, p. 282)

Simultaneous Diagonalization

Consider two Hermitian matrices $A, B \in \mathbb{C}^{n\times n}$.

There is a unitary matrix $U \in \mathbb{C}^{n\times n}$ and real diagonal matrices $\Lambda,\ M$ such that $A = \mathrm{U}\Lambda\mathrm{U}^*,\ B = \mathrm{UMU}^*$

iff AB is Hermitian, that is, $AB = BA$.

See (Horn-Johnson, 2nd Edition, 2013, page 286.)

Homework VI

1. Transform the following Hermitian matrix

$$H = \begin{pmatrix} 0 & 2 & -1 \\ 2 & 5 & -6 \\ -1 & -6 & 8 \end{pmatrix}$$

into a diagonal form.

2. If a (real) Hermitian matrix H is positive definite, prove that $H = P^2$, for a positive definite matrix P.

Lecture VII

- The Jordan Canonical Form
- Examples and Applications