

Unitary Similarities and Schur's Theorem

Wang F.¹ and Tan V.²

*Department of Mathematics, National University of Singapore,
Singapore 117543*

INTRODUCTION

Given a square matrix \mathbf{A} , if it has full rank of linearly independent eigenvectors, then there exists an invertible matrix \mathbf{P} such that, $\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{\Lambda}$, where $\mathbf{\Lambda}$ is a diagonal matrix. However, some square matrices do not have this beautiful property. That is, not all square matrices are diagonalizable. The purpose of this project is to remedy this deficiency by means of Schur's Theorem. Schur's Theorem is almost the same as the diagonalization theorem. It is very useful for analyzing the structure of matrices.

UNITARY MATRICES AND SCHUR'S THEOREM

The Hermitian Inner Product

This section aims to extend the definitions of Euclidean inner product and Euclidean norm to the complex vector space \mathbb{C}^n , so that we can talk about matrices containing complex entries. The Cauchy-Swards inequality and Triangle Inequality respected to the hermitian inner product are proven in this section.

Unitary Matrices

Unitary matrix is the extension of orthogonal matrix over the complex field. A square matrix \mathbf{U} is unitary if and only if $\langle \mathbf{U}\mathbf{x}, \mathbf{U}\mathbf{y} \rangle = \langle \mathbf{x}, \mathbf{y} \rangle$, if and only if $\| \mathbf{U}\mathbf{x} \| = \| \mathbf{x} \|$ for every vector \mathbf{x} and \mathbf{y} . Note that Gram-Schmidt process is a tool to construct an orthonormal basis of a vector space. We can use it to show that *every invertible matrix is a product of a unitary matrix and an upper triangular matrix*. It is known as the **QR** Decomposition.

Schur's Theorem

Issai Schur (1875-1941) first proved that *every square matrix \mathbf{A} is unitarily similar to an upper triangular matrix \mathbf{T}* . i.e. $\mathbf{T} = \mathbf{U}^H \mathbf{A} \mathbf{U}$ for some unitary matrix \mathbf{U} and upper triangular matrix \mathbf{T} . Note that the diagonal entries of \mathbf{T} are exactly the eigenvalues of \mathbf{A} . This decomposition represents the structure of matrices. Schur's Theorem is widely used to solve matrix problems related to

¹ Student

² Assistant Professor

eigenvalues. This beautiful theorem is proven in this section by mathematical induction. Moreover, we state and prove the Generalized Schur's Theorem for a pair of square matrices: *For any square matrices \mathbf{A} and \mathbf{B} of the same size, there exist unitary matrices \mathbf{U} and \mathbf{V}^H , and upper triangular matrices \mathbf{S} and \mathbf{T} , such that $\mathbf{V}^H \mathbf{A} \mathbf{U} = \mathbf{S}$ and $\mathbf{V}^H \mathbf{B} \mathbf{U} = \mathbf{T}$.*

Cayley-Hamilton Theorem

Arthur Cayley (1821-1895) asserted that *a square matrix is a root of its characteristic polynomial* in his "Memoir on the Theory of Matrices" in 1858, known as Cayley-Hamilton Theorem. This section offers a nice proof based on the Schur's Theorem.

APPLICATIONS OF SCHUR'S THEOREM

Hermitian Matrices

Two special types of matrices are introduced in this section, called hermitian matrices and skew-hermitian matrices. They are analogs of real symmetric matrices and skew-symmetric matrices respectively over the complex field. By Schur's Theorem, we can simply prove that *every hermitian matrix is a unitarily diagonalizable matrix with only real eigenvalues*. This theorem yields a wealth of conclusions of hermitian matrix.

Normal Matrices

A normal matrix is a square matrix which is commutative with its hermitian transpose. In fact, *a square matrix is normal if and only if it is unitarily diagonalizable*. Note that every unitary, hermitian or skew-hermitian matrix is a normal matrix. Moreover, let \mathbf{A} be a normal matrix. We can determine whether it is unitary, hermitian or skew-hermitian by checking the eigenvalues:

- a) \mathbf{A} is unitary $\Leftrightarrow |\lambda| = 1$;
- b) \mathbf{A} is hermitian $\Leftrightarrow \text{Im}(\lambda) = 0$;
- c) \mathbf{A} is skew-hermitian $\Leftrightarrow \text{Re}(\lambda) = 0$;

for every eigenvalue λ of \mathbf{A} .

Invariant Subspaces

A vector subspace φ of \mathbb{C}^n is an invariant subspace of a square matrix \mathbf{A} if the linear transformation $\mathbf{T}(\mathbf{u}) = \mathbf{A}\mathbf{u}$ maps φ to itself. In fact, *every non-trivial invariant subspace of a square matrix \mathbf{A} contains at least one eigenvalue of \mathbf{A}* . Moreover, we conjecture and successfully prove a nice property of invariant subspace in this section: *let \mathbf{A} be a square matrix. Then every non-trivial invariant subspace of \mathbf{A} can be spanned by a set of linearly independent eigenvectors of \mathbf{A} if and only if \mathbf{A} is diagonalizable*.

Simultaneous Diagonalization

Two matrices are said to be simultaneously diagonalizable if they are diagonalized by a same invertible matrix. That is, they share full rank of linearly independent eigenvectors. A necessary and sufficient condition of simultaneous diagonalization is proven in this section. *Two square matrices of the same size are simultaneously diagonalizable if and only if they are diagonalizable and commutative.*

The Frobenius Norm

Frobenius norm, stated by Ferdinand Georg Frobenius (1849-1917), is the most frequently used matrix norm in linear algebra. We prove the Schur's Inequalities respected to the Frobenius norm in this section: *Let $\lambda_1, \lambda_2, \dots, \lambda_n$ be the eigenvalues of \mathbf{A} . Denote $\mathbf{A}_H = \frac{\mathbf{A} + \mathbf{A}^H}{2}$ and $\mathbf{A}_S = \frac{\mathbf{A} - \mathbf{A}^H}{2}$.*

Then $\sum_{i=1}^n \lambda_i^2 \leq \|\mathbf{A}\|^2$, $\sum_{i=1}^n \text{Re}^2(\lambda_i) \leq \|\mathbf{A}_H\|^2$ and $\sum_{i=1}^n \text{Im}^2(\lambda_i) \leq \|\mathbf{A}_S\|^2$. These equalities hold if and only if \mathbf{A} is a norm matrix.

THE LIMIT OF \mathbf{A}^m

Gerschgorin's Theorem

The Gerschgorin's theorem provides an estimate for the location of the eigenvalues: *the spectrum of \mathbf{A} lies within the union of the Gerschgorin's disks. i.e. If λ_0 is an eigenvalue of \mathbf{A} , then $|\lambda_0 - a_{ii}| \leq c_i$ for some $1 \leq i \leq n$.* We give the definition of stochastic matrix in this section. By using the Gerschgorin's Theorem, we show that *if the diagonal entries of a stochastic matrix are all positive, then 1 is the only eigenvalue of the matrix with absolute value 1.*

The Limit of \mathbf{A}^m For $\rho(\mathbf{A}) \neq 1$

The convergence and limit of matrix sequence are introduced in this section. We prove that $\lim_{m \rightarrow \infty} \mathbf{A}^m$ diverges if $\rho(\mathbf{A}) > 1$, and $\lim_{m \rightarrow \infty} \mathbf{A}^m = \mathbf{O}$ if $\rho(\mathbf{A}) < 1$. As for complex numbers, we can also

talk about the series of matrices. In particular, *the geometric series $\sum_{m=0}^{\infty} \mathbf{A}^m$ converges if and only if*

$\rho(\mathbf{A}) < 1$. Moreover, $\sum_{m=0}^{\infty} \mathbf{A}^m = (\mathbf{I} - \mathbf{A})^{-1}$ when $\rho(\mathbf{A}) < 1$.

The Limit of \mathbf{A}^m For $\rho(\mathbf{A}) = 1$

We continue the work in the previous section to find the limit of \mathbf{A}^m when $\rho(\mathbf{A}) = 1$. *Let \mathbf{A} be a square matrix with $\rho(\mathbf{A}) = 1$. Then $\lim_{m \rightarrow \infty} \mathbf{A}^m$ exists if and only if*

- a) $\lambda = 1$ is the only eigenvalue with absolute value 1, and
- b) $a(1) = g(1)$.

Probability Vectors and Transition Matrices

Probability vectors and transition matrices are widely used in the mathematical models of time series. Suppose $\mathbf{A} > \mathbf{O}$ is a transition matrix. Then $\mathbf{A}^\infty = \lim_{m \rightarrow \infty} \mathbf{A}^m$ exists. Moreover, $\mathbf{A}^\infty = \mathbf{p}\mathbf{1}^\top$, where \mathbf{p} is the unique stationary vector of \mathbf{A} .

REFERENCES

- [1] Anton, H. (2000), "Elementary Linear Algebra", *John Wiley & Sons, Inc.*, New York.
- [2] Goldberg, J. L. (1991), "Matrix theory with applications", *McGraw-Hill, Inc.*, Columbus.
- [3] Lipschutz, S. (1987), "Theory and problems of linear algebra", *McGraw-Hill, Inc.*, Columbus.