MATRIX THEORY - CHAPTER 2

FALL 2017

1. Unitary Matrix

Definition 1. *A matrix $U \in M_n(\mathbb{C})$ is said to be unitary if $U^*U = I$.*

Theorem 1. *If $U \in M_n$, the following are equivalent:*

(a) *$U^*U = I$, i.e. U is unitary.*
(b) *U is nonsingular and $U^{-1} = U^*$.*
(c) *$UU^* = I$, i.e. U^* is unitary.*
(d) *The column vectors of U form an orthonomal basis of $\mathbb{C}^n$.*
(e) *The row vectors of U form an orthonomal basis of $\mathbb{C}^n$.*
(f) *The map $x \to Ux$ is an isometry, i.e. for all $x \in \mathbb{C}^n$, $\|x\| = \|Ux\|$.*

Remark 1. *All nonsingular matrices form a group $GL(n, \mathbb{C})$; all unitary matrices form a group $U(n)$, which is a subgroup of $GL(n, \mathbb{C})$.*

All nonsingular real matrices form a group $GL(n, \mathbb{R})$; all unitary real matrices (also called real orthogonal matrices) form a group $O(n)$, which is a subgroup of $GL(n, \mathbb{R})$.

Example 1. $U = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$*: U defines a rotation on the plane.*

Example 2. *$U_w = I - 2(w^*w)^{-1}ww^*$, where $0 \neq w \in \mathbb{C}^n$. U_w defines a reflection w.r.t. the plane perpendicular to w, which is called the Householder transformation.*

Example 3. $U = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{3}} & \frac{1}{\sqrt{6}} \\ 0 & \frac{1}{\sqrt{3}} & \frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} & \frac{-1}{\sqrt{6}} \end{bmatrix}$

2. Unitary equivalence

Definition 2. *A matrix B is said to be unitarily equivalent to A if there exists some unitary matrix U such that $B = U^*AU$.*

Theorem 2. *Unitary equivalence is an equivalent relation, i.e. if denoting "B is unitarily equivalent to A" by $B \sim A$, then*

- *reflective: $A \sim A$;*
- *symmetric: $B \sim A$ implies $A \sim B$;*
- *transitive: $C \sim B$ and $B \sim A$ imply $C \sim A$.*

Theorem 3. *If $A = [a_{ij}]$ and $B = [b_{ij}]$ are unitary equivalent, then*

$$\sum_{ij} |a_{ij}|^2 = \sum_{ij} |b_{ij}|^2.$$

Remark 2. *Unitary equivalence is finer than similarity.*

Example 4. $A = \begin{bmatrix} 3 & 1 \\ -2 & 0 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix}$. *Since $\sigma(A) = \sigma(B) = \{1, 2\}$, there are similar. Since $3^2 + 1^2 + (-2)^2 + 0^2 \neq 1^2 + 1^2 + 0^2 + 2^2$, they are not unitarily equivalent.*

3. Schur's Lemma

Theorem 4 (Schur). *Given $A \in M_n(\mathbb{C})$ with eigenvalues $\lambda_1, ..., \lambda_n$ in any prescribed order, there is a unitary matrix $U \in M_n(\mathbb{C})$ such that*

$$U^*AU = T = [t_{ij}]$$

is upper triangular with diagonal entries $\lambda_1, ..., \lambda_n$.

Proof. Induction method. When $n = 1$, it is obviously true. Assume the theorem is true for $(n-1)\times(n-1)$ matrices, we prove it is also true $n \times n$ matrices. □

Remark 3. *Notice that*

- *In Schur's theorem, neither T nor U is unique.*
- *The diagonal entries of T are the eigenvalues of A. We can arrange them in any order by choosing suitable U.*

4. Implications of Schur's theorem

Theorem 5 (Cayley-Hamilton). *For any $A \in M_n(\mathbb{C})$, let $p_A(t) = \det(tI - A)$. Then $p_A(A) = 0$.*

Proof. By Schur's theorem, there exists some unitary matrix U and upper triangular matrix T such that $U^*AU = T$. Then $p_A(A) = U^*p_A(T)U$. So we only need to show $p_A(T) = (T - \lambda_1 I)\cdots(T - \lambda_n I) = 0$. This is done by direct computation. □

Corollary 1. *If $A \in M_n(\mathbb{C})$ is nonsingular, then there exists a polynomial $q(t)$ of order $\leq n-1$ such that $A^{-1} = q(A)$.*

Example 5. $A = \begin{bmatrix} 3 & 1 \\ -2 & 0 \end{bmatrix}$, *$p_A(t) = t^2 - 3t + 2$. Hence $A^2 - 3A + 2I = 0$. This implies*

$$A^2 = 3A - 2I, \quad A^3 = A(3A - 2I) = 7A - 6I, \quad \cdots$$

and

$$A^{-1} = \frac{-1}{2}(A - 3I), \quad A^{-2} = \frac{1}{4}(A - 3I)^2 = -\frac{13}{8}A + \frac{21}{8}I, \quad \cdots$$

Theorem 6 (Block diagonalization). *For $A \in M_n(\mathbb{C})$, let $p_A(t) = (t-\lambda_1)^{n_1}\cdots(t-\lambda_k)^{n_k}$ where $\lambda_1, ..., \lambda_k$ are distinct eigenvalues. Then A is similar to a block diagonal matrix*

$$T = \begin{bmatrix} T_1 & & & \\ & T_2 & & \\ & & \ddots & \\ & & & T_k \end{bmatrix}, \quad T_i = \begin{bmatrix} \lambda_i & * & * \\ & \ddots & * \\ & & \lambda_i \end{bmatrix}.$$

Proof. First, we can choose $U \in U(n)$ such that $U^*AU = T$, where $T = [t_{ij}]$ is upper triangular such that the diagonal entries are arranged as $\lambda_1, ..., \lambda_1, \lambda_2, ...\lambda_2, ..., \lambda_k, ..., \lambda_k$. Consider t_{rs} with $r < s$ and $t_{rr} \neq t_{ss}$. Then after the following similar transformation,

$$(I + \alpha E_{rs})T(I - \alpha E_{rs})$$

t_{rs} becomes $t_{rs} + \alpha t_{ss} - \alpha t_{rr}$. Hence by choosing $\alpha = \frac{t_{rs}}{t_{rr} - t_{ss}}$, the (r, s)-entry is *zero*. Finally, we deal with all such t_{rs} in the same way as above in the order $(n-1, n), (n-2, n-1), (n-2, n), (n-3, n-2), (n-3, n-1), (n-3, n), \cdots$. This can make all (r, s)-entry to be zero if $r < s, t_{rr} \neq t_{ss}$. □

Remark 4. *The elementary row/column operations can be represented by matrix product from the left/right.*

- *Type I.*
  exchange r-th row and s-th row of A ($s \neq r$): $(I - E_{rr} - E_{ss} + E_{rs} + E_{sr})A$;
  exchange r-th column and s-th column of A ($s \neq r$): $A(I - E_{rr} - E_{ss} + E_{rs} + E_{sr})$.
- *Type II.*
  multiply the r-th row of A by $\alpha \neq 0$: $(I + (\alpha - 1)E_{rr})A$;
  multiply the r-th column of A by $\alpha \neq 0$: $A(I + (\alpha - 1)E_{rr})$.

- *Type III.*
  multiply the s-th row of A by α *and add it to the r-th row:* $(I+\alpha E_{rs})A$;
  multiply the r-th column of A by α *and add it to the s-th row:* $A(I+\alpha E_{rs})$;

Example 6.

$$T = \begin{bmatrix} 2 & 1 & 3 & 4 \\ & 2 & 0 & 6 \\ & & 5 & 7 \\ & & & 5 \end{bmatrix}$$

We will make similar transformation to T *such that the small block* $\begin{bmatrix} 3 & 4 \\ 0 & 6 \end{bmatrix}$ *becomes zero. We will deal with the entries in the order* $(3,4),(2,3),(2,4),(1,2),(1,3),(1,4)$.

- $(3,4)$: *since* $t_{33} = t_{44} = 5$, *we can not make this entry to be zero and just leave it there.*
- $(2,3)$: *since* t_{23} *is already zero, we don't need to make any change.*
- $(2,4)$: *since* $t_{22} \neq t_{44}$, *we can make a similar transformation*

$$T' = (I+\alpha_1 E_{24})T(I-\alpha_1 E_{24}) = \begin{bmatrix} 2 & 1 & 3 & 4-\alpha_1 \\ & 2 & 0 & 6+5\alpha_1-2\alpha_1 \\ & & 5 & 7 \\ & & & 5 \end{bmatrix} = \begin{bmatrix} 2 & 1 & 3 & 6 \\ & 2 & 0 & 0 \\ & & 5 & 7 \\ & & & 5 \end{bmatrix} \quad (\alpha_1 = -2)$$

- $(1,2)$: *since* $t'_{11} = t'_{22} = 2$, *we can not make this entry to be zero and just leave it there.*
- $(1,3)$: *since* $t'_{11} \neq t'_{33}$, *we can make a similar transformation*

$$T'' = (I+\alpha_2 E_{13})T'(I-\alpha_2 E_{13}) = \begin{bmatrix} 2 & 1 & 3+5\alpha_2-2\alpha_2 & 6+7\alpha_2 \\ & 2 & 0 & 0 \\ & & 5 & 7 \\ & & & 5 \end{bmatrix} = \begin{bmatrix} 2 & 1 & 0 & -1 \\ & 2 & 0 & 0 \\ & & 5 & 2 \\ & & & 5 \end{bmatrix} \quad (\alpha_2 = -1)$$

- $(1,4)$: *since* $t'_{11} \neq t'_{44}$, *we can make a similar transformation*

$$T''' = (I+\alpha_3 E_{14})T''(I-\alpha_3 E_{14}) = \begin{bmatrix} 2 & 1 & 0 & -1+5\alpha_3-2\alpha_3 \\ & 2 & 0 & 0 \\ & & 5 & 7 \\ & & & 5 \end{bmatrix} = \begin{bmatrix} 2 & 1 & 0 & 0 \\ & 2 & 0 & 0 \\ & & 5 & 2 \\ & & & 5 \end{bmatrix} \quad (\alpha_3 = \frac{1}{3})$$

Denote $S = (I+\alpha_3 E_{14})(I+\alpha_2 E_{13})(I+\alpha_1 E_{24})$. *Then* $STS^{-1} = T'''$ *is a block diagonal matrix.*

5. Normal matrix

Definition 3. *A matrix* $A \in M_n(\mathbb{C})$ *is said to be normal if* $A^*A = AA^*$.

Example 7. *The following are normal matrices:*

- *Unitary matrix:* $U^*U = I$
- *Hermitian matrix:* $A^* = A$
- *Skew-Hermitian matrix:* $A^* = -A$

Theorem 7. *If* A *is normal and* $B = U^*AU$ *for some unitary matrix* U, *then* B *is also normal.*

Theorem 8. *For* $A \in M_n(\mathbb{C})$, *the following are equivalent*

(a) *A is normal.*
(b) *A is unitary diagonalisable.*
(c) $\sum_{i,j=1}^n |a_{ij}|^2 = \sum_i^n |\lambda_i|^2$, *where* $\lambda_1, ..., \lambda_n$ *are the eigenvalues of* A.
(d) *There is an orthonormal set of* n *eigenvectors of* A.

Theorem 9. *If* $A \in M_n(\mathbb{C})$ *is Hermitian, then*

(a) *all eigenvalues of A is real.*
(b) *A is unitary diagonalisable.*

If $A \in M_n(\mathbb{R})$ is symmetric, then A is real orthogonally diagonalisable.

Example 8. $A = \begin{bmatrix} 0 & i \\ -i & 0 \end{bmatrix}$. *A is Hermitian and hence normal. The characteristic polynomial of A is* $\det(tI - A) = (t-1)(t+1)$.

$$\lambda_1 = 1, \alpha_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} i \\ 1 \end{pmatrix}, A\alpha_1 = \lambda_1 \alpha_1,$$
$$\lambda_2 = -1, \alpha_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -i \end{pmatrix}, A\alpha_2 = \lambda_2 \alpha_2.$$

Here α_1, α_2 form an orthonormal basis of $\mathbb{C}^2$. Define

$$U = [\alpha_1, \alpha_2] = \begin{bmatrix} \frac{i}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{-i}{\sqrt{2}} \end{bmatrix}.$$

Then U is unitary and

$$U^*AU = \begin{bmatrix} 1 & \\ & -1 \end{bmatrix}.$$

Example 9. $A = \begin{bmatrix} 8 & 4 & -1 \\ 4 & -7 & 4 \\ -1 & 4 & 8 \end{bmatrix}$. *A is real symmetric and hence normal. The characteristic polynomial of A is* $\det(tI - A) = (t-9)^2(t+9)$.

$$\lambda_1 = 9, \alpha_1 = \frac{1}{3} \begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix}, A\alpha_1 = \lambda_1 \alpha_1,$$
$$\lambda_2 = 9, \alpha_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, A\alpha_2 = \lambda_2 \alpha_2,$$
$$\lambda_3 = -9, \alpha_3 = \frac{1}{3\sqrt{2}} \begin{pmatrix} 1 \\ -4 \\ 1 \end{pmatrix}, A\alpha_3 = \lambda_3 \alpha_3,$$

Here $\alpha_1, \alpha_2, \alpha_3$ form orthonormal basis of $\mathbb{R}^3$. Define

$$Q = [\alpha_1, \alpha_2, \alpha_3] = \begin{bmatrix} \frac{2}{3} & \frac{1}{\sqrt{2}} & \frac{1}{3\sqrt{2}} \\ \frac{1}{3} & 0 & \frac{-4}{3\sqrt{2}} \\ \frac{2}{3} & \frac{-1}{\sqrt{2}} & \frac{1}{3\sqrt{2}} \end{bmatrix}.$$

Then Q is real orthogonal and

$$Q^TAQ = \begin{bmatrix} 9 & & \\ & 9 & \\ & & -9 \end{bmatrix}.$$