Advanced Linear Algebra Week 2 Day 1

2018/09/17 – Jonathan Hayase, updated by Prof. Weiqing Gu

$$\mathcal{V}/_{\mathbb{K}} \xrightarrow{f} \mathcal{W}/_{\mathbb{K}}$$

f is linear if:

i. $f(\boldsymbol{v}+\boldsymbol{w}) = f(\boldsymbol{v}) + f(\boldsymbol{w})$

ii. $f(c\boldsymbol{v}) = cf(v)$ for all $c \in \mathbb{K}$.

Key: V^n is finite dimensional, with $\{\boldsymbol{u}_1, \dots, \boldsymbol{u}_n\}$ basis of V^n and $\{\boldsymbol{w}_1, \dots, \boldsymbol{w}_m\}$ is a basis of W^m.

$$\forall v \in V \implies v = a_1\boldsymbol{u}_1 + \cdots + a_n\boldsymbol{u}_n$$

$$L(\boldsymbol{v}) = a_1 L(\boldsymbol{u}_1) + \cdots + a_n L(\boldsymbol{u}_n)$$

$$\begin{cases} L(\boldsymbol{u}_1) = \alpha_{11}\boldsymbol{w}_1 + \alpha_{21}\boldsymbol{w}_2 \\ L(\boldsymbol{u}_2) = \alpha_{12}\boldsymbol{w}_1 + \alpha_{22}\boldsymbol{w}_2 \end{cases}$$

$$(L(\boldsymbol{u}_1), L(\boldsymbol{u}_2)) = (\boldsymbol{w}_1, \boldsymbol{w}_2) \underbrace{\begin{bmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{bmatrix}}_{A}$$

Where A is the matric representation of L w.r.t. bases $\{\boldsymbol{u}_1, \boldsymbol{u}_2\}$ and $\{\boldsymbol{w}_1, \boldsymbol{w}_2\}$

Recall: If we have $B = P^{-1}AP$, then A and B have the same eigenvectors.

$$L \longleftrightarrow A$$

$$L : V^n \to V^n$$

for finite n	for infinite dimensions
If L is 1–1 then L is onto	does not hold
If L is onto then L is 1–1	does not hold If L is a bijection, then L is invertible
does not hold[a]	

This may not hold in the infinite dimensional case! For example, consider:

$$T\big((x_1, x_2, \dots)\big) = (0, x_1, \dots)$$

This is one to one, since the outputs uniquely determine the inputs, but the map is not onto since it does not cover any sequence not starting with 0.

Similarly,

$$T\big((x_1, x_2, \dots)\big) = (x_2, x_3, \dots)$$

1 Eigenvalues and Eigenvectors

On V^n for finite n. We seek $L(\boldsymbol{x}) = \lambda \boldsymbol{x}$ for $\boldsymbol{x} \neq 0$. Given a basis B we then write

$$A\boldsymbol{x} = \lambda \boldsymbol{x} \implies A\boldsymbol{x} - \lambda \boldsymbol{x} = 0 \implies (A - \lambda I) = 0$$

We want $x \neq 0$ i.e. the homogenerous system must have onzero solution $\iff$ $A - \lambda I$ is not invertible $\iff$ $\det(A - \lambda I) = 0$.

Solve the λ's first, then plug back $(A - \lambda I)\,\boldsymbol{x} = 0$ and solve for $\boldsymbol{x}$ to find eigenvectors.

If V is on infinite dimensional vector space, say take

$$(x_0, x_1, \dots) \overset{T}{\mapsto} (x_1, x_2, \dots)$$

(T is called a *linear operator*.) Then $Tx = \lambda x$ yields

$$\left(1, \lambda, \lambda^2, \dots, \lambda^n, \dots\right)$$

or

$$\left(\lambda, \lambda^2, \lambda^3, \dots, \lambda^{n+1}, \dots\right) = \lambda \left(1, \lambda, \lambda^2, \dots, \lambda^n, \dots\right)$$

True for $\forall \lambda \in R$, so T has uncountably many eigenvalues.

2 Spectrum of a bounded linear operator

If V is finite, $L : V \to V$, then the spectrum of L is the span of its eigenvectors.

Definition: A operator number λ is said to be in the spectrum of a bounded linear operator T if $\lambda I - T$ is not invertible.

Spectrum Mapping Theorem

$$\begin{tikzcd}
A \arrow[r] \arrow[d, leftrightarrow] & a_2A^2 + a_1A + a_0I = q(A) \arrow[d, leftrightarrow] \\
\lambda \arrow[r] & a_2\lambda^2 + a_1\lambda + a_0
\end{tikzcd}$$

M is an eigenvalue of $q(A)$ iff $\epsilon\lambda$ (an eigenvalue of A) such that $M = q(\lambda)$

Remark: A and $q(A)$ share the same set of eigenvalues.

3 Cayley-Hamilton Theorem

Every matrix satisfies its own characteristic polynomial, so $P_A(A) = 0$.

You may think: $P_A(\lambda) = \det(A - \lambda I) = 0$. Then $P_A(A) = \det(A - AI) = 0$. But this doesn't really make sense, since we can't subtract A from a scalar!

Actual proof:

$$\begin{aligned}P_A(\lambda) &= \lambda^n - \underbrace{(a_{11}+a_{22}+\cdots+a_{nn})}_{a_{n-1}}\lambda^{n-1}+\cdots+\underbrace{(-1)^n\det A}_{a_0}\\ &= \lambda^n + a_{n-1}\lambda^{n-1} + + a_1\lambda + a_0\end{aligned}$$

e.g.

$$\begin{aligned}P_A(\lambda) &= \begin{bmatrix} a_1-\lambda & a_{12} \\ a_{21} & a_{22}-\lambda\end{bmatrix}\\ &= (a_{11}-\lambda)(a_{22}-\lambda) - a_{21}a_{22}\\ &= \lambda^2 - \underbrace{(a_{11}+a_{22})}_{\operatorname{tr} A} + \underbrace{a_{11}a_{22}-a_{12}a_{21}}_{\det A}\end{aligned}$$

$$\begin{gathered}P_A(\lambda) = \det(\lambda I - A)\\ \text{let } \lambda = 0 \implies \det -A = (-1)^n \det A\end{gathered}$$

$$\det\begin{vmatrix} a_{11} & a_{12} & a_{13}\\ a_{21} & a_{22} & a_{23}\\ a_{31} & a_{32} & a_{33}\end{vmatrix} = (-1)^{(1+1)}a_{11}\begin{vmatrix} a_{22} & a_{23}\\ a_{32} & a_{33}\end{vmatrix} + (-1)^{(1+2)}a_{12}\begin{vmatrix} a_{21} & a_{23}\\ a_{31} & a_{33}\end{vmatrix} + a_{13}\begin{vmatrix} a_{21} & a_{22}\\ a_{31} & a_{32}\end{vmatrix}$$

Consider the adjoin of A

$$\begin{bmatrix} c_{11} & c_{12} & \cdots\\ c_{21} & \ddots & \ddots\\ \vdots & \ddots & \ddots\end{bmatrix}.$$

Since each every of $B(A)$ is cofactor of $\Lambda I - A$, thus the degree is $\leq n-1$. Let's write $B(\lambda) = \lambda^{n-1}B_0 + \lambda^{n-2}B_1 + \cdots + \lambda B_{n-2} + B_{n-1}$ where B_i are scalar matrices.

(adjoint M) $M = (\det M)\, I$, where M an $n\times n$ matrix. In our case,$B(\lambda)(\lambda I - A) = \det(\lambda I - A)\, I$.

$$\text{RHS} = \left(\lambda^n + a_n\lambda^{n-1} + a_{n-1}\lambda^{n-2} + \cdots + a_1\lambda + a_0\right) I$$

$$\begin{aligned}\text{LHS} &= \left(\lambda^{n-1}B_0 + \lambda^{n-2}B_1 + \lambda^{n03}B_2 + \cdots + \lambda^2 B_{n-2} + \lambda B_{n-1}\right)\\ &= \lambda^n B_0 + \lambda^{n-1}B_1 + \lambda^{n-2}B_2 + \cdots + \lambda^2 B_{n-2} + \lambda B_{n-1}\\ &\quad - \lambda^{n-1}B_0A - \lambda^{n-2}B_aA - \cdots - \lambda^2 B_{n-3}A - B_{n-1}A.\end{aligned}$$

Compare the left and right hand sides, we have

$$\begin{aligned}B_0 &= I\\ B_1 - B_0A &= a_1 I\\ B_2 - B_1A &= a_2 I\\ &\vdots\\ B_{n-2} - B_{n-3}A &= a_{n-2}I\\ B_{n-1} - B_{n-2}A &= a_{n-1}I\\ -B_{n-1}A &= a_n\end{aligned}$$

We can then multiply the first row by powers of A, yielding

$$\begin{aligned}
\cancel{B_0 A^n} &= IA^n \\
\cancel{B_1 A^{n-1}} \cancel{-B_0 A^n} &= a_1 I A^{n-1} \\
\cancel{B_2 A^{n-2}} - \cancel{B_1 A^{n-1}} &= a_2 I A^{n-2} \\
&\vdots \\
\cancel{B_{n-2} A^2} - \cancel{B_{n-3} A^3} &= a_{n-2} I A^2 \\
\cancel{B_{n-1} A} - \cancel{B_{n-2} A^2} &= a_{n-1} I A \\
\cancel{-B_{n-1} A} &= a_n
\end{aligned}$$

Thus, adding yields

$$0 = P_A(A)$$

As desired!

4 Unitary Matrix

Definition a matrix $A \in M_n$ is called *unitary* if $UU^* = I = U^*U$. We have $U^* \triangleq U^T$ if U is real, and then U is called *orthogonal*.

Properties: U is invertible ($U^{-1} = U^*$) and $|\det U| = 1$.

Key theorem: U is unitary $\iff \|U\boldsymbol{x}\| = \|\boldsymbol{x}\| \iff$ There exists orthonormal system $\{U_1, \ldots, U_n\}$ such that $\langle U_i, U_j \rangle = \delta_{ij}$.

5 Schur's Theorem

Given $A \in M_n(\mathbb{C})$ with eigenvalues $\lambda_1, \ldots, \lambda_n$, which could be complex, counting multiplicities then, there exists a unitary matrix $U \in M_n(\mathbb{C})$ such that

$$A = U \begin{bmatrix} \lambda_1 & x & \cdots & x \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & x \\ 0 & \cdots & 0 & \lambda_n \end{bmatrix} U^*$$

Proof by induction: $n = 1$ true. Now, suppose for $n-1 \times n-1$ matrix $\tilde{A}$, there exists unitary $\tilde{W}$ s.t.

$$\tilde{W}^* A \tilde{W} = \begin{bmatrix} \lambda_2 & x & \cdots & x \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & x \\ 0 & \cdots & 0 & \lambda_n \end{bmatrix}$$

Now, for $n \times n$ matrix A, let $\lambda_1, \ldots, \lambda_n$ are eigenvalues with $\boldsymbol{v}_1$ corresponding to λ_1.

Extend $v_1 \neq 0$ into an orthonormal basis $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$.

$$
\begin{aligned}
A\boldsymbol{v}_1 &= \lambda_1 \boldsymbol{v}_1 \\
A\boldsymbol{v}_2 &= b_{11}\boldsymbol{v}_1 + b_{12}\boldsymbol{v}_2 + \cdots + b_{1n}\boldsymbol{v}_n \\
&\ \ \vdots
\end{aligned}
$$

$$
(A\boldsymbol{v}_1, \ldots, A\boldsymbol{v}_n) = (v_1, \ldots, v_n) \left[\begin{array}{c|ccc} \lambda_1 & x & \cdots & x \\ \hline 0 & & & \\ \vdots & & \tilde{A} & \\ 0 & & & \end{array}\right]
$$

Therefore,

$$
A = V \left[\begin{array}{c|ccc} \lambda_1 & x & \cdots & x \\ \hline 0 & & & \\ \vdots & & \tilde{A} & \\ 0 & & & \end{array}\right] V^{-1}
$$

Let

$$
w = \left[\begin{array}{c|ccc} 1 & x & \cdots & x \\ \hline 0 & & & \\ \vdots & & \tilde{W} & \\ 0 & & & \end{array}\right]
$$

then

$$
W^*W = \left[\begin{array}{c|ccc} 1 & x & \cdots & x \\ \hline 0 & & & \\ \vdots & & \tilde{W}^* & \\ 0 & & & \end{array}\right] \left[\begin{array}{c|ccc} 1 & x & \cdots & x \\ \hline 0 & & & \\ \vdots & & \tilde{W} & \\ 0 & & & \end{array}\right] = \left[\begin{array}{c|ccc} 1 & x & \cdots & x \\ \hline 0 & & & \\ \vdots & & \tilde{W}^*\tilde{W} & \\ 0 & & & \end{array}\right] = \left[\begin{array}{c|ccc} 1 & x & \cdots & x \\ \hline 0 & & & \\ \vdots & & I_{n-1} & \\ 0 & & & \end{array}\right] = I_n
$$

$$
\begin{aligned}
W^*AW &= \left[\begin{array}{c|ccc} 1 & x & \cdots & x \\ \hline 0 & & & \\ \vdots & & \tilde{W}^* & \\ 0 & & & \end{array}\right] \left[\begin{array}{c|ccc} \lambda_1 & x & \cdots & x \\ \hline 0 & & & \\ \vdots & & \tilde{A} & \\ 0 & & & \end{array}\right] \left[\begin{array}{c|ccc} 1 & x & \cdots & x \\ \hline 0 & & & \\ \vdots & & \tilde{W} & \\ 0 & & & \end{array}\right] \\
&= \left[\begin{array}{c|ccc} \lambda_1 & x & \cdots & x \\ \hline 0 & & & \\ \vdots & & \tilde{W} & \\ 0 & & & \end{array}\right] \left[\begin{array}{c|ccc} 1 & x & \cdots & x \\ \hline 0 & & & \\ \vdots & & A\tilde{W} & \\ 0 & & & \end{array}\right] \\
&= \left[\begin{array}{c|ccc} \lambda_1 & x & \cdots & x \\ \hline 0 & & & \\ \vdots & & \tilde{W}^*A\tilde{W} & \\ 0 & & & \end{array}\right] \\
&= \left[\begin{array}{c|ccc} \lambda_1 & x & \cdots & x \\ \hline 0 & \lambda_2 & & \\ \vdots & & \ddots & \\ 0 & & & \lambda_n \end{array}\right]
\end{aligned}
$$

So what?

Theorem: Given $A \in M_n(\mathbb{C})$ and $\epsilon > 0$, there exists a diagonalizable matrix $\tilde{A} \in M_n(\mathbb{C})$ s.t.

$$\sum_{1 \le i,j \le n} \left| a_{ij} - \tilde{a}_{ij} \right|^2 < \epsilon$$

Can be used to derive the Jordan Canonical form.

6 Dual Space

The dual space of $\mathcal{V}$ is the space of linear functionals.

$$f : {}^{W}/_{\mathbb{F}} \xrightarrow{\ell} \mathbb{F}$$

$\ell \to T$ also linear. Conside the set of $\ell : W \to \mathbb{F}$. $L(W, \mathbb{F})$ is called the dual space of W, denoted by W^*.

6.1 Basis Theorem

$$\dim W^* = \dim W$$

Proof: method 1: $L(W, B) \cong \{A_{m\times n}\} \cong M_{\mathbb{F}}(n, m)$. In our case, $M_{\mathbb{F}}(n, 1)$.

$$\dim M_{\mathbb{F}}(n, 1) = n$$

Method 2: Recall $\{\boldsymbol{u}_1, \dots, \boldsymbol{u}_n\}$ orthonormal, then

$$v = a_1\boldsymbol{u}_1 + \cdots a_n\boldsymbol{u}_n$$

With $a_i = \boldsymbol{v} \cdot \boldsymbol{u}_i$. We want to construct a basis for W^* which contains n basis vectors $\{\ell_1, \dots, \ell_n\}$.

$$\begin{array}{cccc}
\ell_1(\boldsymbol{v}_1) = 1 & \ell_2(\boldsymbol{v}_1) = 0 & \cdots & \ell_n(\boldsymbol{v}_1) = 0 \\
\ell_1(\boldsymbol{v}_2) = 0 & \ell_2(\boldsymbol{v}_2) = 1 & \cdots & \ell_n(\boldsymbol{v}_2) = 0 \\
\vdots & \dots & \ddots & \vdots \\
\ell_1(\boldsymbol{v}_n) = 0 & \ell_2(\boldsymbol{v}_n) = 0 & \cdots & \ell_n(\boldsymbol{v}_n) = 1
\end{array}$$

Here $\{\boldsymbol{v}_1, \dots, \boldsymbol{v}_n\}$ basis of W.

Let's prove $\{\ell_1, \dots, \ell_n\}$ are linearly independent.

Form $a_1\ell_1 + \cdots + a_n\ell_n = 0$ w.t.s $a_1 = \cdots = a_n = 0$.

$$(a_1\ell_1 + \cdots + a_i\ell_i + \cdots + a_n\ell_n)\,\boldsymbol{v}_i = 0(\boldsymbol{v}_i)$$

$$\implies a_1 \cancelto{0}{\ell(\boldsymbol{v}_1)} + \cdots + a_i \cancelto{1}{\ell_i(\boldsymbol{v}_i)} + \cdots + a_n \cancelto{0}{\ell_n(\boldsymbol{v}_i)} = 0$$

(from method 1, $\dim W^* = n$). Thus, $\{\ell_1, \dots, \ell_n\}$ forms a basis of W^*.

7 Double Dual

Let $^V/_{\mathbb{K}}$. Let's consider the dual space of V^* dual of W. What is V^{**}[1], the double dual!

$$V'' = L(V', \mathbb{K})$$

Theorem: $V'' \cong V$ i.e. the two are naturally isomorphic.

Proof:

$$V \to V''$$
$$\phi : x \mapsto L_x \begin{bmatrix} L_x : v' \to \mathbb{K} \\ L_x(\ell) = \ell(x) \\ (\ell : V \to \mathbb{K}) \end{bmatrix}$$

Want to show:

1. ϕ is bijection (1–1 and onto)
2. ϕ preserves linear structure (i.e. $\phi(cx + y) = c\phi(x) + \phi(y)$)

Let's show 2 first:

$$\phi(cx + y)(\ell) = L_{cx+y}(\ell) = \ell(cx + y) = c\ell(x) + \ell(y) = cL_x(\ell) + L_y(\ell) = c\phi(x) + \phi(y)$$

[1] or $(V')'$

Now let's show 1. Note that $\dim V'' = \dim V' = \dim V$. So all we need is to show is that ϕ is 1-to-1, so $\ker\phi = \{0\}$. Suppose $\phi(x) = 0$, want to show $x = 0$.

$$\phi_x = L_x = 0, L_x(\ell) = \ell(x) = 0$$

for all $\ell \in V' \implies x = 0$. Thus, ϕ is an isomorphism, so $V'' \cong V$.

e.g. start from inner product

$$\langle \underbrace{v_0}_{\text{fix}}, w\rangle = \ell(w)$$