vv214: Cayley-Hamilton Theorem. Adjoint, self-adjoint, normal operators. Symmetric matrices. Orthogonal diagonalization.

Dr.Olga Danilkina

UM-SJTU Joint Institute

July 16, 2020

1. Cayley-Hamilton Theorem and its applications.
2. Orthogonaly diagonalizable matrices.
3. Adjoint, self-adjoint, normal operators.

Cayley-Hamilton Theorem: Intro

1. Let $g(t) = a_0 + a_1 t + \ldots + a_k t^k$ and $D = diag(d_1, \ldots, d_n)$

$$\Rightarrow g(D) = diag(g(d_1), \ldots, g(d_n))$$

Let $f_D(\lambda)$ be a char poly $\Rightarrow f_D(D) = \begin{pmatrix} f_D(d_1) & & \\ & \ddots & \\ & & f_D(d_n) \end{pmatrix} =$

$$= \begin{pmatrix} \prod_{i=1}^{n}(d_1 - d_i) & & \\ & \ddots & \\ & & \prod_{i=1}^{n}(d_n - d_i) \end{pmatrix} = \underbrace{0}_{\text{zero matrix}}$$

$$\boxed{f_D(D) = 0}$$

Cayley-Hamilton Theorem: Intro

2. Let A be similar to $D \Rightarrow D = S^{-1}AS$ (A is diagonalizable)

$$g(D) = a_0 I_n + a_1 D + \ldots + a_k D^k =$$

$$= a_0 S^{-1} S + a_1 S^{-1} A S + \ldots + a_k S^{-1} A^k S = S^{-1}(a_0 I_n + a_1 A + \ldots + A^k) S$$

$$\Rightarrow g(D) = S^{-1} g(A) S \Rightarrow g(D) \sim g(A)$$

$$A \sim D \Rightarrow f_A(\lambda) = f_D(\lambda) \Rightarrow f_A(D) = f_D(D) = 0$$

$$f_A(A) = S \underbrace{f_A(D)}_{0} S^{-1} = 0$$

$$\boxed{f_A(A) = 0}$$

Cayley-Hamilton Theorem: Intro

1. Is $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ diagonalizable?

If $D = S^{-1}AS$ then $f_A(\lambda) = (\lambda - 1)^2 = f_D(\lambda) \Rightarrow D = I_2$

$$A = SDS^{-1} = SIS^{-1} = I \Rightarrow \text{contradiction}$$

2. Find a sequence of diagonalizable matrices that converges to A.

$$\{B_m\},\ B_m = \begin{pmatrix} 1 & 1 \\ 0 & 1 + \frac{1}{m} \end{pmatrix}$$

B_m has distinct eigenvalues and hence, diagonalizable

$$\Rightarrow f_{B_m}(B_m) = 0$$

$$B_m \to A \Rightarrow f_{B_m} \to f_A \text{ and determinant is a continuous function}$$

$$f_A(A) = \lim_{m \to \infty} f_{B_m}(B_m) = 0$$

Cayley-Hamilton Theorem

Theorem: Any $A \in M_{n\times n}(\mathbb{K})$ satisfies its own characteristic equation, i.e.

$$f_A(A) = (-A)^n + (tr\,A)(-A)^{n-1} + \ldots + (\det A)I_n = 0$$

Examples:

1. $A = \begin{pmatrix} 1 & -1 \\ 2 & 3 \end{pmatrix} \Rightarrow f_A(\lambda) = \lambda^2 - 4\lambda + 5 = 0 \Rightarrow \lambda_{1,2} = 2 \pm i$

$$f_A(A) = A^2 - 4A + 5I = 0 \Rightarrow A^2 = 4A - 5I$$

$A^3 = A^2A = (4A-5I)A = 4A^2-5A = 4(4A-5I)-5A = 11A-20I$ etc.

2. If A is orthogonal with $\det A = 1$ and $\lambda_1 = -\frac{1}{2} - \frac{\sqrt{3}}{2}i$, then $\lambda_2 = -\frac{1}{2} + \frac{\sqrt{3}}{2}i$ and $\lambda_3 = 1$.

$f_A(\lambda) = (-\lambda)^3 + (-1+1)(-\lambda)^2 + \det A \Rightarrow -A^3 + I = 0 \Rightarrow A^3 = I$

$$A^{214} = (A^3)^{71}A = I^{71}A = A$$

Cayley-Hamilton Theorem: Order Reduction

3. Order Reduction

Represent $g(t) = h(t)f_A(t) + r(t) \Rightarrow g(\lambda) = r(\lambda)$

$$g(A) = h(A)f_A(A) + r(A) \Rightarrow g(A) = r(A)$$

Let $g(A) = A^5 + 2A^4 - A^3 + A^2 - 2A + I$ with $A = \begin{pmatrix} -3 & 1 \\ 0 & -2 \end{pmatrix}$

$$f_A(\lambda) = \lambda^2 + 5\lambda + 6$$

$$\Rightarrow g(\lambda) = (\lambda^3 - 3\lambda^2 + 8\lambda - 21)(\lambda^2 + 5\lambda + 6) + \underbrace{65\lambda + 127}_{r(\lambda)}$$

$$g(A) = 65A + 127I$$

Cayley-Hamilton Theorem: Analytic Functions of a Matrix

4. Let a complex-valued function $g(t)$ be analytic in some region of the complex plane $\Rightarrow g(t) = \sum_{k=0}^{\infty} a_k t^k$

$$g(t) = h(t)f_A(t) + r(t) \Rightarrow g(\lambda_i) = r(\lambda_i) = \sum_{k=0}^{n-1} b_k \lambda_i^k$$

$$g(A) = \sum_{k=0}^{n-1} b_k A^k$$

Cayley-Hamilton Theorem: Analytic Functions of a Matrix

Let $A = \begin{pmatrix} -3 & 1 \\ 0 & -2 \end{pmatrix}$, $f_A(\lambda) = \lambda^2 + 5\lambda + 6 \Rightarrow r(t) = b_0 + b_1 t$

Let $g(t) = \sin t \Rightarrow \begin{array}{l} \sin(\lambda_1) = b_0 + b_1\lambda_1 \\ \sin(\lambda_2) = b_0 + b_1\lambda_2 \end{array}$

$$\Rightarrow \begin{array}{rc} b_0 = & 3\sin(-2) - 2\sin(-3) \\ b_1 = & \sin(-2) - \sin(-3) \end{array}$$

$$\sin A = g(A) = r(A) = b_0 I + b_1 A$$

$$\sin A = (3\sin(-2) - 2\sin(-3))I + (\sin(-2) - \sin(-3))A$$

$$\sin \begin{pmatrix} -3 & 1 \\ 0 & -2 \end{pmatrix} = \begin{pmatrix} \sin(-3) & \sin(-2) - \sin(-3) \\ 0 & \sin(-2) \end{pmatrix}$$

Cayley-Hamilton Theorem: Analytic Functions of a Matrix

4. Matrix Exponential

$$g(t) = e^{tx} \Rightarrow e^{Ax} = \sum_{k=0}^{n-1} b_k A^k, \quad e^{\lambda_i x} = \sum_{k=0}^{n-1} b_k \lambda_i^k$$

Let $A = \begin{pmatrix} -3 & 1 \\ 0 & -2 \end{pmatrix}$, $f_A(\lambda) = \lambda^2 + 5\lambda + 6 \Rightarrow r(t) = b_0 + b_1 t$

$$e^{-2x} = b_0 - 2b_1,\ e^{-3x} = b_0 - 3b_1$$

$$\begin{aligned} b_0 &= \quad 3e^{-2x} - 2e^{-3x} \\ b_1 &= \quad e^{-2x} - e^{-3x} \end{aligned}$$

$$e^{Ax} = (3e^{-2x} - 2e^{-3x})I + (e^{-2x} - e^{-3x})A$$

$$exp\begin{pmatrix} -3 & 1 \\ 0 & -2 \end{pmatrix} x = \begin{pmatrix} e^{-3x} & e^{-2x} - e^{-2x} \\ 0 & e^{-2x} \end{pmatrix}$$

Cayley-Hamilton Theorem: Minimal Polynomial

Definition: The smallest degree polynomial $m_A(t) \neq 0$ such that $m_A(A) = 0$ is called the minimal polynomial of A.
The minimal polynomial of

$$D = \begin{pmatrix} 1 & & & & \\ & 2 & & & \\ & & 1 & & \\ & & & 1 & \\ & & & & 2 \end{pmatrix}$$

is $m_D(t) = (t-1)(t-2)$.

Is the polynomial $t-1$ minimal for $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$?no!

Cayley-Hamilton Theorem $\Rightarrow \forall A \in M_{n\times n}(\mathbb{K}) \quad m_A | f_A$

$$\Rightarrow m_A(t) = (t-1)^2$$

Remark: Let $g(t)$ be a polynomial with coefficients from $\mathbb{K}$. g has multiple roots iff $\gcd(g, g')$ is not a constant.

Orthogonally Diagonalizable Matrices

1. Question: What are the conditions on a square matrix to guarantee that it is diagonalizable?
   Answer: The existence of an eigenbasis.
2. Question: For which matrices is there an orthonormal eigenbasis?
   Answer: For which matrices is there an orthogonal matrix S such that $S^{-1}AS = S^TAS$ is diagonal?

Definition: A matrix A is orthogonally diagonalizable if there exists an orthogonal matrix S such that

$$S^{-1}AS = S^TAS$$

is diagonal.

Example 1

$$A = \begin{pmatrix} 4 & 2 \\ 2 & 7 \end{pmatrix} \quad A \quad \text{is symmetric}$$

$$|A - \lambda I| = 0 \Rightarrow (4-\lambda)(7-\lambda) - 4 = 0 \Rightarrow \quad \underbrace{\lambda_1 = 3,\ \lambda_2 = 8}_{\text{2 different eigenvalues}}$$

$$A \quad \text{is diagonalizable}$$

$$\bar{v}_1 = \begin{pmatrix} 2 \\ -1 \end{pmatrix}, \bar{v}_2 = \begin{pmatrix} 1 \\ 2 \end{pmatrix} \Rightarrow \bar{u}_1 = \begin{pmatrix} 2/\sqrt{5} \\ -1/\sqrt{5} \end{pmatrix}, \bar{u}_2 = \begin{pmatrix} 1/\sqrt{5} \\ 2/\sqrt{5} \end{pmatrix}$$

$$\Rightarrow \bar{u}_1,\ \bar{u}_2 \text{ is the orthonormal basis}$$

$$S = \frac{1}{\sqrt{5}} \begin{pmatrix} 2 & 1 \\ -1 & 2 \end{pmatrix} \quad \text{is orthogonal and} \quad D = S^{-1}AS = \begin{pmatrix} 3 & 0 \\ 0 & 8 \end{pmatrix}$$

$$A \text{ is orthogonally diagonalizable}$$

Example 2

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \quad |A - \lambda I| = 0 \Rightarrow \lambda^2(3-\lambda) = 0 \Rightarrow \lambda_{1,2} = 0,\ \lambda_3 = 3$$

$$\lambda_{1,2} = 0\colon\ \bar{v}_1 = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \bar{v}_2 = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix},\ \lambda_3 = 3\colon\ \bar{v}_3 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

$$\Rightarrow \bar{u}_1 = \begin{pmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{pmatrix}, \bar{u}_2 = \begin{pmatrix} -\frac{1}{\sqrt{6}} \\ -\frac{1}{\sqrt{6}} \\ 2 \end{pmatrix}, \bar{u}_3 = \begin{pmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{pmatrix}$$

$$S = \begin{pmatrix} -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & 2 & \frac{1}{\sqrt{3}} \\ 0 & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{3}} \end{pmatrix},\ D = S^{-1}AS = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 3 \end{pmatrix}$$

A is orthogonally diagonalizable

Useful Statements

Lemma 1: If A is orthogonally diagonalizable, then $A^T = A$.

Lemma 2: Let $A^T = A$. If $\bar{v}_1$, $\bar{v}_2$ are eigenvectors of A with distinct eigenvalues λ_1, λ_2, then $\bar{v}_1 \perp \bar{v}_2$.

Lemma 3: A *symmetric* matrix $A_{n \times n}$ has *n real* eigenvalues counted with their algebraic multiplicities.

Theorem (Spectral Theorem): A matrix A is orthogonally diagonalizable iff A is symmetric ($A^T = A$).

Rietz Representation Theorem

Let $\dim X < +\infty$ *and* $\varphi : X \to \mathbb{K}$ *be a linear functional:*
$\varphi \in L(X, \mathbb{R}) = X'$.
There is a unique vector $y \in X$ *such that*

$$\varphi(x) = (x, y) \quad \forall x \in X$$

Adjoint Operators

Definition: Let $T\colon V \to W$ be linear, i.e. $T \in L(V, W)$.
The adjoint of T is the operator $T^\star\colon W \to V$ such that

$$(Tv, w) = (v, T^\star w) \quad \forall v \in V \, \forall w \in W$$

Remark: The adjoint $T^\star$ is well defined:

1. Fix $w \in W$. Define a map $V \to \mathbb{K}$ such that $v \to (Tv, w)$
2. The map $v \to (Tv, w)$ is linear $\Rightarrow$ it is a linear functional.
3. By the Riesz Representation Theorem, there exists a unique element $b \in V\colon (Tv, w) = (v, b)$
4. Denote $b = T^\star w$. Since b exists, so $T^\star w$ exists as well.
5. Therefore, there exists a map $w \to T^\star w$.

Example: Let $T\colon \mathbb{R}^3 \to \mathbb{R}^2$

$$T(x_1, x_2, x_3) = (x_2 + 3x_3, 2x_1)$$

$$(T\bar{x}, \bar{y}) = ((x_2 + 3x_3, 2x_1), (y_1, y_2)) = x_2 y_1 + 3x_3 y_1 + 2x_1 y_2$$

$$= ((x_1, x_2, x_3), (2y_2, y_1, 3y_1)) \Rightarrow T^\star \bar{y} = (2y_2, y_1, 3y_1)$$

Properties of Adjoint Operators

1. $T \in L(V, W) \Rightarrow T^\star \in L(W, V)$

$$(v, T^\star(w_1 + w_2)) = (Tv, w_1 + w_2) = (Tv, w_1) + (Tv, w_2)$$

$$= (v, T^\star w_1) + (v, T^\star w_2) = (v, T^\star w_1 + T^\star w_2)$$

$$(v, T^\star(\lambda w)) = (Tv, \lambda w) = \bar{\lambda}(Tv, w) = \bar{\lambda}(v, T^\star w) = (v, \lambda T^\star w)$$

2. $(T_1 + T_2)^\star = T_1^\star + T_2^\star \quad \forall S, T \in L(V, W)$
3. $(\lambda T)^\star = \bar{\lambda} T^\star \quad \forall \lambda \in \mathbb{K} \, \forall T \in L(V, W)$
4. $(T^\star)^\star = T \quad \forall T \in L(V, W)$

$$(w, (T^\star)^\star v) = (T^\star w, v) = \overline{(v, T^\star w)} = \overline{(Tv, w)} = (w, Tv) \, \forall v \in V$$

5. $I^\star = I$
6. $(T_1 T_2)^\star = T_2^\star T_1^\star \quad \forall T_1 \in L(W, U),\ T_2 \in L(V, W)$

Properties of Adjoint Operators

7. $Ker\ T^\star = (Im\ T)^\perp$

$$w \in Ker\ T^\star \Leftrightarrow T^\star w = 0 \Leftrightarrow (v, T^\star w) = 0 \quad \forall v \in V$$

$$\Leftrightarrow (Tv, w) = 0 \quad \forall v \in V \Leftrightarrow w \in (Im\ T)^\perp$$

8. The matrix of the adjoint $T^\star$ w.r.t orthonormal bases $e_1, \ldots, e_m \in V$; $f_1, \ldots f_n \in W$ is the conjugate transpose of the matrix of T.

$$A_T = (Te_1\ Te_2 \ldots Te_m) \quad Te_k \in W,\ f_1, \ldots, f_n \text{ is orthonormal}$$

$$\Rightarrow Te_k = (Te_k, f_1)f_1 + \ldots + (Te_k, f_n)f_n \Rightarrow (A_T)_{jk} = (Te_k, f_j)$$

$$A_{T^\star} = (T^\star f_1\ T^\star f_2 \ldots T^\star f_m) \quad T^\star f_k \in V,\ e_1, \ldots, e_m \text{ is orthonormal}$$

$$\Rightarrow T^\star f_k = (T^\star f_k, e_1)e_1 + \ldots + (T^\star f_k, e_m)e_m$$

$$\Rightarrow (A_{T^\star})_{jk} = (T^\star f_k, e_j) = \overline{(e_j, T^\star f_k)} = \overline{(Te_j, f_k)}$$

Self-Adjoint Operators

Definition: The operator $T \in L(V, V)$ is called self-adjoint if $T^\star = T$:

$$(Tv, w) = (v, Tw) \quad \forall v, w \in V$$

Hermitian=self-adjoint

Remarks:

1. Let $T\colon \mathbb{R}^2 \Rightarrow \mathbb{R}^2$ be defined by the matrix $\begin{pmatrix} 1 & a \\ 2 & 3 \end{pmatrix} \Rightarrow T$ is self-adjoint iff $a = 2$, that is, its matrix is symmetric.
2. Every eigenvalue of a self-adjoint operator is real.

   $$\lambda||v||^2 = (\lambda v, v) = (Tv, v) = (v, Tv) = (v, \lambda v) = \bar{\lambda}(v, v) = \bar{\lambda}||v||^2$$

3. T is self-adjoint iff $(Tv, v) \in \mathbb{R} \quad \forall v \in V$

   $$(Tv.v) - \overline{(Tv, v)} = (Tv, v) - (v, Tv) = (Tv, v) - (T^\star v, v)$$

   $$= ((T - T^\star)v, v)$$

Normal Operators

Definition: An operator on an inner product space is called normal if it commutes with its adjoint, i.e $TT^\star = T^\star T$.

Remarks:

1. Every self-adjoint operator is normal.
2. T is normal iff $||Tv|| = ||T^\star v|| \, \forall v$

$$T \text{ is normal} \Leftrightarrow T^\star T - TT^\star = 0 \Leftrightarrow ((T^\star T - TT^\star)v, v) = 0$$
$$\Leftrightarrow (T^\star Tv, v) = (TT^\star v, v) \Leftrightarrow ||Tv||^2 = ||T^\star v||^2$$

3. Let $Tv = \lambda v$. T is normal $\Leftrightarrow$ $T - \lambda I$ is also normal.
4. Let $Tv = \lambda v$. Then $T^\star v = \bar{\lambda} v$

$$0 = ||(T - \lambda I)v|| = ||(T - \lambda I)^\star v|| = ||(T^\star - \bar{\lambda} I)v|| \Rightarrow T^\star v = \bar{\lambda} v$$

5. Suppose $T \in L(V, V)$ is normal. Then eigenvectors of T corresponding to distinct eigenvalues are orthogonal.

$$Tu = \alpha u, \; Tv = \beta v \Rightarrow T^\star v = \bar{\beta} v$$

$$(\alpha - \beta)(u, v) = \alpha(u, v) - \beta(u, v) = (Tu, v) - (u, T^\star v) = 0$$

Complex Spectral Theorem

Let $\mathbb{K} = \mathbb{C}$ and $T \in L(V, V)$. Then the following are equivalent:
(a) T is normal.
(b) V has an orthonormal basis consisting of eigenvectors of T.
(c) T has a diagonal matrix with respect to some orthonormal basis of V.