Spectral Theory and Jordan Canonical Form

Yan Zeng

Version 1.0, last revised on 2014-05-04.

Abstract

Summary of spectral theory (including Jordan canonical form) as presented in Lax[2].

Contents

1 Eigenvalues, Eigenvectors **2**

2 Characteristic Polynomial **2**

3 Spectral Mapping Theorem **3**

4 Cayley-Hamilton Theorem **3**

5 Spectral Theorem **4**

6 Minimal Polynomial **4**

7 Index and Multiplicities **5**

8 When Are Two Matrices Similar (and Jordan Canonical Form) **6**

The content of this note is based on Lax[2] Chapter 6 and Appendix 15. To make use of the Fundamental Theorem of Algebra so that characteristic polynomials have a full set of roots, the spectral theory of linear maps is formulated in linear spaces over the field of complex numbers.

1 Eigenvalues, Eigenvectors

Proposition 1. *Every $n \times n$ matrix over the field of complex numbers has an eigenvector in $\mathbb{C}^n$.*

Proof. Choose any nonzero vector w, the following set of $n+1$ vectors must be linearly dependent:

$$w, Aw, A^2w, \cdots, A^nw.$$

This implies the existence of a polynomial $p(t) = \sum_{j=0}^{n} c_j t^j$ over the complex numbers, such that $p(A)w = 0$. By the Fundamental Theorem of Algebra, $p(t)$ can be written as a product of linear factors:

$$p(t) = c\prod_{j=1}^{n}(x - a_j),\ a_j, c \in \mathbb{C}, c \neq 0,$$

and consequently, $p(A)w = 0$ can be rewritten as

$$c\prod_{j=1}^{n}(A - a_jI)w = 0.$$

This implies at least one of the matrices $(A - a_jI)$ is not invertible; such a matrix $(A - a_{j_0}I)$ has a nontrivial nullspace. So, there is an eigenvector pertaining to the eigenvalue a_{j_0}. □

Proposition 2. *Eigenvectors of a matrix A corresponding to distinct eigenvalues are linearly independent.*

Proof. Consider a nontrivial linear relation among the eigenvectors that involves the least number m of eigenvectors:

$$\sum_{j=1}^{m} b_jh_j = 0, b_j \neq 0, j = 1, \cdots, m;$$

here h_j is the eigenvector pertaining to eigenvalue a_j ($a_j \neq a_k$ for $j \neq k$). Applying A to the equation, we get

$$\sum_{j=1}^{m} b_ja_jh_j = 0,$$

and hence

$$a_m\sum_{j=1}^{m} b_jh_j - \sum_{j=1}^{m} b_ja_jh_j = \sum_{j=1}^{m} b_j(a_m - a_j)h_j = 0.$$

Clearly the coefficient of h_m is zero and none of the others is zero, so we have a nontrivial linear relation among the h_j involving only $(m-1)$ of the vectors, contrary to the m being the smallest number of vectors satisfying such a relation. □

2 Characteristic Polynomial

Proposition 1 shows that every matrix A has at least one eigenvalue, but it does not show how many or how to calculate them. Recall Lax[2] Chapter 5, Corollary 3 states that "an $n \times n$ matrix A is invertible iff $\det A \neq 0$". This implies the following

Proposition 3. *For a to be an eigenvalue of A it is necessary and sufficient that*

$$\det(aI - A) = 0.$$

The polynomial $p_A(\lambda) = \det(\lambda I - A)$ is called the **characteristic polynomial** of the matrix A.

3 Spectral Mapping Theorem

Theorem 1 (**Spectral Theory, part 1: spectral mapping theorem**)**.** *(a) Let q be any polynomial, A a square matrix, a an eigenvalue of A. Then $q(a)$ is an eigenvalue of $q(A)$.*

(b) Every eigenvalue of $q(A)$ is of the form $q(a)$, where a is an eigenvalue of A.

Proof. (a) is easy to prove. For (b), suppose b is an eigenvalue of $q(A)$. Then $q(A) - bI$ is not invertible. By the Fundamental Theorem of Algebra, $q(s) - b$ can be decomposed as

$$q(s) - b = c \prod (s - r_i),$$

where r_i's may repeat. By $q(A) - bI = c \prod (A - r_i I)$, at lease one $A - r_i I$ is not invertible; such an r_i is an eigenvalue of A. Note $q(r_i) = b$, we are done. □

4 Cayley-Hamilton Theorem

Theorem 2 (Cayley-Hamilton)**.** *Every matrix A satisfies its own characteristic equation:*

$$p_A(A) = 0.$$

The proof of Cayley-Hamilton Theorem relies on a clever use of the following version of Cramer's rule:

Proposition 4. *Let A be an $n \times n$ matrix and B defined as the matrix of cofactors of A; that is,*

$$B_{ij} = (-1)^{i+j} \det A_{ji},$$

where A_{ji} is the (ji)th minor of A. Then $AB = BA = \det A \cdot I_{n \times n}$.

Proof. Suppose A has the column form $A = (a_1, \cdots, a_n)$. By replacing the jth column with the ith column in A, we obtain

$$M = (a_1, \cdots, a_{j-1}, a_i, a_j, \cdots, a_n).$$

On one hand, Property (i) of a determinant gives $\det M = \delta_{ij} \det A$; on the other hand, Laplace expansion of a determinant gives

$$\det M = \sum_{k=1}^{n} (-1)^{k+j} a_{ki} \det A_{kj} = \sum_{k=1}^{n} a_{ki} B_{jk} = (B_{j1}, \cdots, B_{jn}) \begin{pmatrix} a_{1i} \\ \vdots \\ a_{ni} \end{pmatrix}.$$

Combined, we can conclude $\det A \cdot I_{n \times n} = BA$. By replacing the ith column with the jth column in A, we can get similar result for AB. □

Proof. (Cayley-Hamilton) Let $Q(s) = sI - A$ and $P(s)$ defined as the matrix of cofactors of $Q(s)$:

$$P_{ij}(s) = (-1)^{i+j} D_{ji}(s),$$

where D_{ij} is the determinant of the ijth minor of $Q(s)$. By Proposition 4,

$$P(s)Q(s) = \det Q(s) \cdot I = p_A(s) I.$$

Since $Q(A) = 0$, it follows that

$$p_A(A) = 0.$$

□

5 Spectral Theorem

Theorem 3 (**Spectral Theory, part 2: spectral theorem).** *Let A be an $n \times n$ matrix with complex entries. Every vector in $\mathbb{C}^n$ can be written as a sum of eigenvectors of A, genuine or generalized.*

Proof. The key to the proof is the following observation:

$$N_{pq} = N_p \bigoplus N_q,$$

where p and q are polynomials with complex coefficients and have no common zero, $N_p = \ker p(A)$, $N_q = \ker q(A)$, and $N_{pq} = \ker[p(A)q(A)]$. This observation is based on the existence of two polynomials a and b such that

$$ap + bq \equiv 1. \tag{1}$$

Equation (1) holds for the ring of polynomials over a field, and more generally, for a principal ideal ring or a Euclidean domain. See any textbook on abstract algebra for reference.

By induction, we can easily conclude

$$N_{p_1 \cdots p_k} = N_{p_1} \bigoplus \cdots \bigoplus N_{p_k},$$

where p_1, $\cdots$, p_k are a collection of polynomials that are pairwise without a common zero. In particular, if we denote A's distinct eigenvalues by a_1, $\cdots$, a_k, the characteristic polynomial $p_A(s)$ can be written as

$$p_A(s) = c \prod_{i=1}^{k} (s - a_i)^{m_i}.$$

and hence by Theorem 2 (Cayley-Hamilton)

$$\mathbb{C}^n = N_{p_A} = N_{m_1}(a_1) \bigoplus \cdots \bigoplus N_{m_k}(a_k),$$

where $N_{m_i}(a_i) = \ker(A - a_i I)^{m_i}$ $(i = 1, \cdots, m)$. This finishes our proof. □

Corollary 1. *Real symmetric matrices can be diagonalized.*

Proof. It suffices to show for any eigenvalue a of a real symmetric matrix A, we have $d(a) = 1$. Indeed, assume $N_1(a) \subsetneq N_2(a)$, then for any $x \in N_2(a) \setminus N_1(a)$, we have

$$(aI - A)x \neq 0, (aI - A)^2 x = 0.$$

But $(aI - A)^2 x = 0$ means

$$\langle (aI - A)x, (aI - A)x \rangle = \langle (aI - A)^2 x, x \rangle = 0,$$

which implies $(aI - A)x = 0$, contradiction. So we must have $d(a) = 1$. □

6 Minimal Polynomial

Denote by $\wp = \wp_A$ the set of all polynomials p which satisfy $p(A) = 0$. Denote by m a nonzero polynomial of smallest degree in $\wp$, then

Proposition 5. *All p in $\wp$ are multiples of m, and except for a constant factor, m is unique.*

We denote by m_A the unique m whose leading coefficient is 1, and call m_A the **minimal polynomial** of A.

Proof. Use Euclidean division. □

For given $a \in \mathbb{C}$, we denote by $N_m = N_m(a)$ the nullspace of $(A - aI)^m$. The subspaces N_m consist of generalized eigenvectors; they are indexed increasingly:

$$N_1 \subset N_2 \subset \cdots.$$

Since these are subspaces of a finite-dimensional space, they must be equal from a certain index on.[1] We denote by $d = d(a)$ the smallest such index, that is

$$N_d = N_{d+1} = \cdots$$

but

$$N_{d-1} \neq N_d;$$

$d(a)$ is called the **index** of the eigenvalue a.

Theorem 4 (Minimal Polynomial). *Let A ben an $n \times n$ matrix: denote its distinct eigenvalues by $a_1, \cdots, a_k$, and denote the index of a_j by d_j. We claim that the minimal polynomial m_A is*

$$m_A(s) = \prod_{i=1}^{k} (s - a_i)^{d_i}.$$

Proof. A number is an eigenvalue of A if and only if it's a root of the characteristic polynomial p_A. So $p_A(s)$ can be written as $p_A(s) = \prod_1^k (s - a_i)^{m_i}$ with each m_i a positive integer $(i = 1, \cdots, k)$. We have shown in the text that p_A is a multiple of m_A, so we can assume $m_A(s) = \prod_{i=1}^k (s - a_i)^{r_i}$ with each r_i satisfying $0 \leq r_i \leq m_i$ $(i = 1, \cdots, k)$. We argue $r_i = d_i$ for any $1 \leq i \leq k$.

Indeed, we have

$$\mathbb{C}^n = N_{p_A} = \bigoplus_{j=1}^{k} N_{m_j}(a_j) = \bigoplus_{j=1}^{k} N_{d_j}(a_j).$$

where the last equality comes from the observation $N_{m_j}(a_j) \subseteq N_{m_j + d_j}(a_j) = N_{d_j}(a_j)$ by the definition of d_j. This shows the polynomial $\prod_{j=1}^k (s - a_j)^{d_j} \in \wp$. By the definition of minimal polynomial, $r_j \leq d_j$ for $j = 1, \cdots, n$.

Assume for some j, $r_j < d_j$, we can then find $x \in N_{d_j}(a_j) \setminus N_{r_j}(a_j)$ with $x \neq 0$. Define $q(s) = \prod_{i=1, i \neq j}^k (s - a_i)^{r_i}$, then by Corollary 10 x can be uniquely decomposed into $x' + x''$ with $x' \in N_q$ and $x'' \in N_{r_j}(a_j)$. We have $0 = (A - a_j I)^{d_j} x = (A - a_j I)^{d_j} x' + 0$. So $x' \in N_q \cap N_{d_j}(a_j) = \{0\}$. This implies $x = x'' \in N_{r_j}(a_j)$. Contradiction. Therefore, $r_i \geq d_i$ for any $1 \leq i \leq k$.

Combined, we conclude $m_A(s) = \prod_{i=1}^k (s - a_i)^{d_i}$. □

Remark 1. *Along the way, we have shown that the index d of an eigenvalue is no greater than the multiplicity of the eigenvalue as a root of the characteristic polynomial. This is also a consequence of the theorem itself, since $m_A(s)$ divides $p_A(s)$.*

7 Index and Multiplicities

We summarize several relationships among index, algebraic multiplicity, geometric multiplicity, and the dimension of the space of generalized eigenvectors pertaining to a given eigenvalue. The first result is an algebraic proof of Lax[2, page 132], Lemma 10 of Chapter 9.

Proposition 6 (Geometric and algebraic multiplicities). *Let A be an $n \times n$ matrix over a field $\mathbb{F}$ and α an eigenvalue of A. If $m(\alpha)$ is the multiplicity of α as a root of the characteristic polynomial p_A of A, then* $\dim N_1(\alpha) \leq m(\alpha)$.

[1] To see it's impossible to have the scenario where $N_d = N_{d+1}$ but $N_d \neq N_l$ for some $l > d + 1$, we note $\forall x \in N_{d+2}$,

$$(A - aI)x \in N_{d+1} = N_d.$$

So $x \in N_{d+1} = N_d$. Then we work by induction.

$m(\alpha)$ is called the **algebraic multiplicity** of α and $\dim N_1(\alpha)$ is called the **geometric multiplicity** of α. So this result says "geometric multiplicity $\dim N_1(\alpha) \le$ algebraic multiplicity $m(\alpha)$".

Proof. Let $v_1, \cdots, v_s$ be a basis of $N_1(\alpha)$ and extend it to a basis of $\mathbb{F}^n$: $v_1, \cdots, v_s, u_1, \cdots, u_r$. Define $U = (v_1, \cdots, v_s, u_1, \cdots, u_r)$. Then

$$\begin{aligned} U^{-1}AU &= U^{-1}A(v_1, \cdots, v_s, u_1, \cdots, u_r) \\ &= U^{-1}(\alpha v_1, \cdots, \alpha v_s, Au_1, \cdots, Au_r) \\ &= (\alpha U^{-1}v_1, \cdots, \alpha U^{-1}v_s, U^{-1}Au_1, \cdots, U^{-1}Au_r). \end{aligned}$$

Because $U^{-1}U = I$, we must have $U^{-1}AU = \begin{bmatrix} \alpha I_{s\times s} & B \\ 0 & C \end{bmatrix}$ and $\det(\lambda I - A) = \det(\lambda I - U^{-1}AU) = \det \begin{bmatrix} (\lambda-\alpha)I_{s\times s} & -B \\ 0 & \lambda I_{(n-s)\times(n-s)} - C \end{bmatrix} = (\lambda-\alpha)^s \det(\lambda I - C)$.[2] So $s \le m(\alpha)$. ☐

We continue to use the notation from Proposition 6, and we define $d(\alpha)$ as the index of α. Then we have

Proposition 7 (Index and algebraic multiplicity). $d(\alpha) \le m(\alpha)$.

Proof. Use the result on minimal polynomial (Theorem 4) or its remark. ☐

Using the notations from Propositions 6 and 7, we have

Proposition 8 (Algebraic multiplicity and the dimension of the space of generalized eigenvectors). $m(\alpha) = \dim N_{d(\alpha)}(\alpha)$.

Proof. See Lax[2, page 133], Theorem 11 of Chapter 9. ☐

In summary, we have

$$\boxed{\dim N_1(\alpha), d(\alpha) \le m(\alpha) = \dim N_{d(\alpha)}(\alpha).}$$

In words, it becomes

geometric multiplicity of α, index of α
$\le$ algebraic multiplicity of α as a root of the characteristic polynomial
$=$ dim. of the space of generalized eigenvectors pertaining to α.

8 When Are Two Matrices Similar (and Jordan Canonical Form)

Theorem 5 (Spectral Theory, part 3: Jordan canonical form). *(i) Suppose the pair of matrices A and B are similar,*

$$A = SBS^{-1},$$

S some invertible matrix. Then A and B have the same eigenvalues:

$$a_1 = b_1, \cdots, a_k = b_k;$$

furthermore, the null spaces $N_m(a_j) = \ker(A - a_jI)^m$ and $M_m(a_j) = \ker(B - a_jI)^m$ have for all j and m the same dimension:

$$\dim N_m(a_j) = \dim M_m(a_j). \tag{2}$$

(ii) Conversely, if A and B have the same eigenvalues, and if condition (2) about the nullspaces having the same dimension is satisfied, then A and B are similar.

Proof. Part (i) is obvious. For part (ii), see Lax[2, page 363], Appendix 15. ☐

Corollary 2 (Jordan canonical form). *Every $n \times n$ matrix over complex field is similar to a Jordan canonical form.*

[2] For the last equality, see, for example, Munkres [1, page 24], Problem 6.

References

[1] J. Munkres. *Analysis on manifolds*, Westview Press, 1997.

[2] P. Lax. *Linear algebra and its applications*, 2nd Edition, Wiley-Interscience, 2007.