Eigenvalue and Eigenvector

Len Fu

11.27.2024

Abstract

This is the note made by Len Fu during his learning progress in BIT. The main content is from Linear Algebra Done Right and Linear Algebra, J.K.Wan, BIT 2024 Fall.

The follow part is a summary of the note.

1. Eigenvalues and Eigenvectors
   - The concept and properties of eigenvalues and eigenvectors, computation of eigenvalues and eigenvectors.
2. Diagonalization of Matrices
   - Criteria for diagonalizability, the process of diagonalization.
3. Diagonalization of Real Symmetric Matrices using Orthogonal Matrices
   - Properties of eigenvalues and eigenvectors of real symmetric matrices, the process of orthogonal similarity diagonalization.
4. Jordan Canonical Form of Matrices
   - The method for finding the Jordan canonical form $\boldsymbol{J}$ and the similarity transformation matrix $\boldsymbol{P}$.

Contents

1 Eigenvalue and Eigenvector of the Matrix

1.1 Basis

Definition 1.1. Set A as a $n \times n$ square, if there exists a number λ and $n - nonzero$ vector X, satisfying

$$AX = \lambda X \ or \ (\lambda I - A)X = 0$$

then we say that λ is an eigenvalue of A, and X is an eigenvector of A with eigenvalue λ.

Note.

1. Only squares have eigenvectors and eigenvalues.
2. Eigenvector must be nonvector and eigenvalue can be zero.

Definition 1.2. $(\lambda I - A)$ is the eigrnmatrix of A. $|\lambda I - A|$ is the eigenpolynomial of A. $|\lambda I - A| = 0$ is the eigenequation of the matrix A.

Then the eigenvector of A with eigenvalue λ is the combination of the solution vectors of $(\lambda I - A)X = 0$.

Since $(\lambda I - A)X = 0$ and X is nonzero vector, then $\det(\lambda I - A)$ should be zero to ensure X is nonzero vector of the solution.

Consider the solution of $(\lambda I - A)X = 0$. The characteristic polynomial of A is

$$b_n\lambda^n + b_{n-1}\lambda^{n-1} + \cdots + b_1\lambda + b_0.$$

To solve the polynomial,

$$\begin{vmatrix} \lambda - a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & \lambda - a_{22} & \cdots & -a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ -a_{n1} & -a_{n2} & \cdots & \lambda - a_{nn} \end{vmatrix}$$
$$= b_n\lambda^n + b_{n-1}\lambda^{n-1} + \cdots + b_1\lambda + b_0$$

Consider the expasion of the determinant, except for

$$(\lambda - a_{11})(\lambda - a_{22})\cdots(\lambda - a_{nn})$$

other terms' highest order of λ is $n-2$. Then the coefficents

$$\begin{cases} b_n = 1 \\ b_{n-1} = -(a_{11} + a_{22} + \cdots + a_{nn}) = tr() \end{cases}$$

And we divide the determinant into two parts and one is

$$\begin{vmatrix} -a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & \lambda - a_{22} & \cdots & -a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ -a_{n1} & -a_{n2} & \cdots & \lambda - a_{nn} \end{vmatrix}$$

the other one doesn't contribute to the b_0, thus

$$b_0 = (-1)^n |A|.$$

From the polynomial theorem,

$$f(\lambda) = (\lambda - \lambda_1)(\lambda - \lambda_2)\cdots(\lambda - \lambda_n) = 0.$$

Then

$$b_0 = (-1)^n \lambda_1 \lambda_2 \cdots \lambda_n \ b_{n-1} = -(\lambda_1 + \lambda_2 + \cdots + \lambda_n)$$

And there are some properties of the Eigenvalue and Eigenvector.

Properties 1.1.
1. $|A| = \lambda_1 \lambda_2 \cdots \lambda_n$.
2. $trA = \sum_{i=1}^{n} \lambda_i$
3. If $X_1, X_2, \cdots, X_s$ are eigenvectors of A that belong to eigenvalue λ_0, then the linear combination of $X_1, X_2, \cdots, X_s$ is also an eigenvector of A that belongs to eigenvalue λ_0. And all eigenvectors plus zero vector forms an eigenspace of A with eigenvalue λ_0, denoted as V_{λ_0} and it's a solution space of $(\lambda_0 I - A)X = 0$.

Properties 1.2. If λ is an eigenvalue of A with eigenvector X, then we have

1. $k\lambda$ is the eigenvalue of kA.
2. λ^m is the eigenvalue of $A^m (m \in N^*)$.
3. $f(\lambda)$ is the eigenvalue of $f(A)$ if f is a polynomial transformation.
4. When A is invertible, λ^{-1} is the eigenvalue of A^{-1}

And X is the eigenvector of matrices above with corresponding eigenvalue.

Properties 1.3. The matrix A and A^T have the same spectrum.

1.2 Algebraic Multiplicity and Geometric Multiplicity

Definition 1.3 (Geometric Multiplicity). As for the eigenvalue λ_i of A, its all eigenvectors are the nonzero-solutions of the equation $(\lambda_i I - A)X = 0$. Thus the number of independent eigenvectors of A with eigenvalues λ_i is no more than $n - rank(\lambda_i - A)$. The number is the dimension of the eigenspace V_{λ_i} and the number of the solution vectors of the fundamental system of solutions of A. We call this number as the geometric multiplicity of λ_i, equals to

$$q_i = n - rank(\lambda_i I - A), \ i = 1, 2, \cdots, s.$$

I

Definition 1.4 (Algebraic Multiplicity). From the theorem of the polynomial, n-square A on $\mathbb{C}$ can be divided

$$f_A(\lambda) = (\lambda - \lambda_1)^{p_1}(\lambda - \lambda_2)^{p_2}\cdots(\lambda - \lambda_s)^{p_s}$$

where $\lambda_1, \lambda_2, \cdots, \lambda_s$ are all distinct eigenvalues of A. We call p_i the algebraic multiplicity of λ_i.

Theorem 1.1. The geometric multiplicity of any eigenvalue λ_i of A is not larger than the algebraic multiplicity of λ_i.

Proof. Set

$$X_{i1}, X_{i2}, \cdots, X_{iq}$$

as a basis of the eigenspace V_{λ_i}, we extent it into a basis of $\mathbb{C}^{\ltimes}$:

$$X_{i1}, X_{i2}, \cdots, X_{iq_i}, Y_1, Y_2, \cdots, Y_{n-q_i},$$

then we have

$$\begin{aligned} & A[X_{i1}, X_{i2}, \cdots, X_{iq_i}, Y_1, Y_2, \cdots, Y_{n-q_i}] \\ = & [\lambda_i X_{i1}, \lambda_2 X_{i2}, \cdots, X_{iq_i}, AY_1, AY_2, \cdots, AY_{n-q_i}] \\ = & [X_{i1}, X_{i2}, \cdots, X_{iq_i}, Y_1, Y_2, \cdots, Y_{n-q_i}] \begin{bmatrix} \lambda_i I_{q_i} & * \\ 0 & A_1 \end{bmatrix} \end{aligned}$$

where A_1 is $n - q_i$ square. Now set

$$P = [X_{i1}, X_{i2}, \cdots, X_{iq_i}, Y_1, Y_2, \cdots, Y_{n-q_i}],$$

obviously P id invertible. Then

$$AP = P \begin{bmatrix} \lambda_i I_{q_i} & * \\ 0 & A_1 \end{bmatrix}$$

$$and$$

$$A \sim \begin{bmatrix} \lambda_i I_{q_i} & * \\ 0 & A_1 \end{bmatrix}$$

Thus

$$f_A(\lambda) = \det\left(\lambda I - \begin{bmatrix} \lambda_i I_{q_i} & * \\ 0 & A_1 \end{bmatrix}\right) = \det\left(\begin{bmatrix} (\lambda - \lambda_i) I_{q_i} & -* \\ 0 & \lambda I_{n-q_i} - A_1 \end{bmatrix}\right) = (\lambda - \lambda_i)^{q_i} \det(\lambda I_{n-q_i} - A_1) = (\lambda - \lambda_i)^{q_i} f_{A_i}(\lambda)$$

Thus $p_i \geq q_i, \ i = 1, 2, ..., s.$ □

2 Similarity of the Matrix

2.1 Basis

Definition 2.1. Set $A, B \in C^{n \times n}$. If there exists an n-order invertible matrix P such that

$$P^{-1} A P = B$$

, we say that A and B are similar, denoted as $A \sim B$, and P is called the similarity transformation from A to B.

Properties 2.1 (Reflectivity). $A \sim A$.

Properties 2.2 (Symmetry). If $A \sim B$, then $B \sim A$.

Properties 2.3 (Transitivity). If $A \sim B$ and $B \sim C$, then $A \sim C$.

Properties 2.4. 1. $P^{-1}(A_1 + A_2 + \cdots + A_n)P = P^{-1} A_1 P + P^{-1} A_2 P + \cdots + P^{-1} A_n P = P^{-1} \sum_{i=1}^{n} P^{-1} A_i P.$

2. $P^{-1}(kA)P = kP^{-1}AP$.

2.2 Conditions of Similar Digonalizablity

Definition 2.2 (Digonalizable). If there exists an invertible matrix P such that

$$P^{-1}AP = D$$

where A is a square and D is a diagonal matrix. Then A is called diagonalizable.

Theorem 2.1 (NS Condition). A n-square A is similar diagonalizable if and only if A has n linear independent eigenvectors.

Proof.

$$P^{-1}AP = diag(\lambda_1, \lambda_2, \cdots, \lambda_n)$$
$$AP = Pdiag(\lambda_1, \lambda_2, \cdots, \lambda_n)$$

Now set $P = [X_1, X_2, \cdots, X_n]$, and

$$[AX_1, AX_2, \cdots, AX_n] = [\lambda_1 X_1, \lambda_2 X_2, \cdots, \lambda_n X_n]$$

then we have

$$(A - \lambda_i)X_i = 0, \; for \; i = 1, 2, \cdots, n.$$

Since P is invertible, we can find n linear irrelative vectors $X_1, X_2, \cdots, X_n$. And $X_1, X_2, \cdots, X_n$ are n linear irrelative eigenvectors of A and $\lambda_1, \lambda_2, \cdots, \lambda_n$ are eigenvalues of A.

Inversely, if A has n linear irrelative eigenvectors $X_1, X_2, \cdots, X_n$ with eigenvalues $\lambda_1, \lambda_2, \cdots, \lambda_n$, satisfying that

$$AX_i = \lambda_i X_i, \; for \; i = 1, 2, \cdots, n.$$

Set $P = [X_1, X_2, \cdots, X_n]$ and obviously P is invertible, and

$$P^{-1}AP = diag\lambda_1, \lambda_2, \cdots, \lambda_n$$

which reveals that A is similar diagonalizable. □

Note.

1. The similar transformation matrix P is not unique.
2. The order of $X_1, X_2, \cdots, X_n$ changes as the order of $\lambda_1, \lambda_2, \cdots, \lambda_n$ changes.

Theorem 2.2. The eigenvectors of A with different eigenvalues are linearly independent.

Proof. Set $\lambda_1, \lambda_2, \cdots, \lambda_m$ are non-eqaul eigenvalues of A, $X_1, X_2, \cdots, X_n$ are eigenvectors of A with eigenvalues $\lambda_1, \lambda_2, \cdots, \lambda_m$, then

$$AX_i = \lambda_i X_i \; for \; i = 1, 2, \cdots, m.$$

Now I use the mathematical induction to prove the above statement.

Base Case:m=1 Since $X_1 \neq 0$, X_1 is linear independent.

Inductive Hypothesis:m=k-1 Suppose the statement is true for $m = k - 1$.

Inductive Step:m=k Consider the condition that $m = k$,

$$k_1X_1 + k_2X_2 + \cdots + k_mX_m = 0$$

Then we left multiply A, we get

$$k_1AX_1 + k_2AX_2 + \cdots + k_mAX_m = 0.$$

Since $AX_i = \lambda_iX_i$, we get

$$k_1\lambda_1X_1 + k_2\lambda_2X_2 + \cdots + k_m\lambda_mX_m = 0.$$

We multiply λ_m on the first equation and we get

$$k_1\lambda_mX_1 + k_2\lambda_2X_2 + \cdots + k_m\lambda_mX_m = 0.$$

Then

$$k_1(\lambda_1 - \lambda_m)X_1 + k_2(\lambda_2 - \lambda_m)X_2 + \cdots + k_{m-1}(\lambda_{m-1} - \lambda_m)X_{m-1} = 0,$$

and since $\lambda_i \neq \lambda_j$ and $X_1, X_2, \cdots, X_{m-1}$ are linearly independent, then

$$k_i = 0 \ i = 1, 2, \cdots, m-1.$$

Correspondingly, $k_m = 0$.

Thus $k_1, k_2, \cdots, k_m = 0$, which reveals that $X_1, X_2, \cdots, X_m$ are linearly independent. □

Corollary 2.2.1. If n-square A has n distinct eigenvalues, then A can be similar diagonalizable.

Theorem 2.3. Set $\lambda_1, \lambda_2, \cdots, \lambda_n$ are distinct eigenvalues of A, $X_{i1}, X_{i2}, \cdots, X_{il_i}$ are linear independent eigenvectors of A belonging to eigenvalue $\lambda_i, i = 1, 2, \cdots, m$, then the vector set of $X_{11}, X_{12}, \cdots, X_{1l_1}, \cdots, X_{m1}, X_{m2}, \cdots, X_{ml_l}$ is linear independent.

Theorem 2.4 (NS Condition). Set $\lambda_1, \lambda_2, \cdots, \lambda_n$ are distinct eigenvalues of A, p_i *and* q_i are the algebraic multiplicity and geometric multiplicity of λ_i respectively. Then A can be similar diagonalizable if and only if

$$p_i = q_i \ i = 1, 2, \cdots, s.$$

Proof. As we know that

$$f_A(\lambda) = (\lambda - \lambda_i)^{q_i} f_{A_1}(\lambda)$$

If A is similar diagonalizable, then its eigenpolynomial can written as $(\lambda - \lambda_i)^{p_i}$..., then $p_i = q_i$ for all i. □

2.3 Method to Similar Diagonalize

1. Find all the eigenvalues of A, like $\lambda_1, \lambda_2, \cdots, \lambda_s$.
2. For every eigenvalue λ_i, find the rank of $(\lambda_iI - A)$, and judge that whether $q_i = n - rank(\lambda_iI - A)$ equals to its p_i. If they are equivalent, then it is diagonalizable. If they are not equivalent, then A is not diagonalizable.
3. If A is diagonalizable, then for every λ_i, solve the $(\lambda_ii - A)X = 0$, and we can have a fundamental solve

$$X_{i1}, X_{i2}, \cdots, X_{iq_i}, \ i = 1, 2, \cdots, s.$$

4. Let $P = [X_{11}, X_{12}, \cdots, X_{1q_1}, \cdots, X_{s1}, X_{s2}, \cdots, X_{sq_s}]$, then we can get $P^{-1}AP = D$, where

$$D = diag(\lambda_1, \cdots, \lambda - 1, \cdots, \lambda_s, \cdots, \lambda_s)$$

with q_i λ_i.

3 Real-Symmetric Matrix's Normalised Diagonalization

Definition 3.1 (Conjugate Matrix). Set matrix A in the field $\mathbb{C}$, and $A^\dagger$ is the Conjugate Matrix of A, where

$$a_{ij} = a_{ij}^\dagger.$$

Properties 3.1. 1. $(A^\dagger)^\dagger = A$

2. $(A^\dagger)^T = (A^T)^\dagger$
3. $(kA)^\dagger = k^\dagger A^\dagger$
4. $(A+B)^\dagger = A^\dagger + B^\dagger$
5. $(AB)^\dagger = A^\dagger B^\dagger$
6. $((AB)^T)^\dagger = (B^T)^\dagger (A^T)^\dagger$
7. If A is invertible, then $(A^{-1})^\dagger = (A^\dagger)^{-1}$
8. If A is square, then $\det A^\dagger = (\det A)^\dagger$
9. $rank(A) = rank(A^\dagger)$

Theorem 3.1. The eigenvalue of real symmetric matrix is real.

Theorem 3.2. Eigenvectors with different eigenvalues of real symmetric matrix is orthogonal.

Proof. Set λ_1, λ_2 are two distinct eigenvalues of real symmetric matrix A, their conresponding eigenvectors are X_1, X_2, then

$$AX_i = \lambda_i X_i,\ for\ i = 1, 2.$$

To prove $X_1 \perp X_2$, we need to show that $X_2^T X_1 = 0$. Then

$$\begin{aligned} X_2^T A^T &= \lambda_2 X_2^T \\ X_2^T A^T X_1 &= \lambda_2 X_2^T X_1 \\ (\lambda_2 - \lambda_1)(X_2^T X_1) &= 0 \end{aligned}$$

Since $\lambda_2 \neq \lambda_1$, then $X_2^T X_1 = 0$. □

Theorem 3.3. Set λ_0 is any eigenvalue of real symmetric matrix A, and p and q represent its algebraic and geometric multiplicity respectively, then $p = q$.

Theorem 3.4. For any n real symmetric matrix A, there exists a n orthogonal matrix P and a diagonal matrix D such that $A = PDP^{-1}$. where

$$D = diag(\lambda_1, \cdots, \lambda_n)\ and\ P = [\beta_1, \beta_2, \cdots, \beta_n]\ where\ \beta\ isorthogonalvectors.$$

We can use Gram-Schmidt Orthogonalization to orthogonalize the vectors :

1. $\beta_1 = \alpha_1$
2. $i = 2,3,\ldots,m$ β_i
$$\beta_i = \alpha_i - \sum_{j=1}^{i-1} \frac{(\alpha_i,\beta_j)}{(\beta_j,\beta_j)}\beta_j$$
where $(\cdot,\cdot)$ is the inner product.
3. Normalize β_i and we get the normalised vectors γ_i
$$\gamma_i = \frac{\beta_i}{\|\beta_i\|}$$

4

4.1 Jordan Canonical Form

Definition 4.1 (Jordan Block). Define a m upper triangular matrix, as a m Jordan Block

$$\begin{bmatrix} a & 1 & 0\cdots & 0 & 0 & 0 \\ 0 & a & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & a & 1 \\ 0 & 0 & 0 & \cdots & 1 & a \end{bmatrix}$$

Definition 4.2 (Jordan Form). Call a quasi-diagonal matrix A as a Jordan Form if it can be written as

$$\begin{bmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_s \end{bmatrix}$$

where J_i is a Jordan Block.

Theorem 4.1. Set $A \in C^{n\times n}$, then A can be similar to a Jordan Form.

$$J = \begin{bmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_s \end{bmatrix}$$

and except for $J_1, J_2, \cdots, J_s$'s order, J is unique for A and is Jordan Canonical Form of A.

4.1.1 Elementary Divisors

Definition 4.3 (Elementary Transformation). Set A as an n square, and let $A(\lambda) = \lambda I - A$, we call these three transformations as Elementary Transformation.:

1. Swap two rows or columns of $A(\lambda)$.

2. Multiply a row or column of $A(\lambda)$ by a nonzero scalar.

3. Use a polynomial of λ to multiply a row or column of $A(\lambda)$ and add it to another row or column of $A(\lambda)$.

Definition 4.4 (Elementary Divisors). Set A as a complex n-square, use elementary transformation to transform $A(\lambda)$ to diagonal matrix, Then decompose the diagonal elements into a product of distinct first-degree polynomial powers of λ. The powers of all first-degree factors with exponents greater than zero are called the elementary divisors of A.

Corollary 4.1.1 (m-Jordan Block).

$$\begin{bmatrix} a & 1 & 0\cdots & 0 & 0 \\ 0 & a & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & a \end{bmatrix}$$

The elementary divisors of A are $(\lambda - a)^m$.

Theorem 4.2. Set A's elementary divisors are $(\lambda - a_1)^{m_1}, \cdots, (\lambda - a_s)^{m_s}$, then A's Jordan Canonical Form is

$$\begin{bmatrix} J_1 & 0 & \cdots & 0 \\ 0 & J_2 & \cdots & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & \cdots & J_s \end{bmatrix}$$

where J_i is

$$\begin{bmatrix} a_i & 1 & 0\cdots & 0 & \\ 0 & a_i & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & a_i \end{bmatrix}_{m_i \times m_i} \qquad (i = 1, 2, \cdots, s)$$

4.2 Smith Normal Form

5 Exercise

Exercise 5.1. The eigenvalues of A are $1,2,3$, find the eigenvalues of A^2-2I.

Solution 5.1.1. We know that $A \sim diag(1,2,3)$ and then $A^2 \sim diag(1,4,9)$. Then $(A^2-2I) \sim (diag(1,4,9)-2I) = diag(-1,2,7)$, then the eigenvalues of A^2-2I are $-1,2,7$.

Exercise 5.2. Sovle the maxima of the

$$f(x_1,x_2) = 2x_1^2 + 2x_1x_2 + 2x_2^2$$

under the constraint $x_1^2 + x_2^2 = 1$ using Lagrange Multipliers.

Solution 5.2.1. Using Lagrange Multipliers,

$$L(x_1,x_2) = 2x_1^2 + 2x_1x_2 + 2x_2^2 - \lambda(x_1^2 + x_2^2 - 1)$$

and