Vv214 Linear Algebra

RC6

Li Yuzhou

SJTU-UM Joint Institute
Shanghai Jiao-Tong University

July 23, 2019

Table of Contents

Table of Contents

Eigenvalues and Eigenvectors

Consider an $n \times n$ matrix A. A nonzero vector $\vec{v}$ in $\mathbb{R}^n$ is called an *eigenvector* of A if $A\vec{v}$ is a scalar multiple of $\vec{v}$, that is, if

$$A\vec{v} = \lambda\vec{v},$$

for some scalar λ. Note that this scalar λ may be zero.

The scalar λ is called the *eigenvalue* associated with the eigenvector $\vec{v}$.

Geometrically, eigenvectors are the vectors whose direction are not changed after the transformation A and the eigenvalue associated with the eigenvector is the change in its amplitude.

Exercise

Let T be the orthogonal projection onto a line L in $\mathbb{R}^2$. Describe the eigenvectors of T geometrically and find all eigenvalues of T.

Let T from $\mathbb{R}^2$ to $\mathbb{R}^2$ be the rotation in the plane through an angle of 90° in the counterclockwise direction. Find all eigenvalues and eigenvectors of T.

What are the possible real eigenvalues of an orthogonal[5] matrix A?

Discrete Dynamical Systems: Definition

$$\vec{x}(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_n(t) \end{bmatrix}.$$

Suppose that the state of the system at time $t+1$ is determined by the state at time t and that the transformation of the system from time t to time $t+1$ is linear, represented by an $n \times n$ matrix A:

$$\vec{x}(t+1) = A\vec{x}(t).$$

Then

$$\vec{x}(t) = A^t\vec{x}_0.$$

Our goal is to find close formulas for $x_n(t)$ (opposed to recursive formula).

Discrete Dynamical Systems: Method

Suppose we can find a basis

$$\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n \quad \text{of } \mathbb{R}^n$$

consisting of eigenvectors of A, with

$$A\vec{v}_1 = \lambda_1\vec{v}_1,\ A\vec{v}_2 = \lambda_2\vec{v}_2,\ \ldots,\ A\vec{v}_n = \lambda_n\vec{v}_n.$$

Find the coordinates $c_1, c_2, \ldots, c_n$ of vector $\vec{x}_0$ with respect to basis $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n$:

$$\vec{x}_0 = c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots + c_n\vec{v}_n.$$

Then

$$\vec{x}(t) = c_1\lambda_1^t\vec{v}_1 + c_2\lambda_2^t\vec{v}_2 + \cdots + c_n\lambda_n^t\vec{v}_n.$$

We can write this equation in matrix form as

$$\vec{x}(t) = \begin{bmatrix} | & | & & | \\ \vec{v}_1 & \vec{v}_2 & & \vec{v}_n \\ | & | & & | \end{bmatrix} \begin{bmatrix} \lambda_1^t & 0 & & 0 \\ 0 & \lambda_2^t & & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & & \lambda_n^t \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix}$$

$$= S \begin{bmatrix} \lambda_1 & 0 & & 0 \\ 0 & \lambda_2 & & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & & \lambda_n \end{bmatrix}^t S^{-1}\vec{x}_0.$$

Discrete Dynamical Systems: Method

We then only need to know the method to find the eigenvectors and eigenvalues of a $n \times n$ matrix A, and to judge when there exist a basis of R^n consisting of the eigenvectors of A.

Finding eigenvalues

Consider an $n \times n$ matrix A and a scalar λ. Then λ is an eigenvalue[6] of A if (and only if)

$$\det(A - \lambda I_n) = 0.$$

This is called the *characteristic equation* (or the *secular equation*) of matrix A. ∎

Exercise

Find the characteristic equation for an arbitrary 2×2 matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$.

Solution

$$\det(A - \lambda I_2) = \det \begin{bmatrix} a - \lambda & b \\ c & d - \lambda \end{bmatrix}$$

$$= (a - \lambda)(d - \lambda) - bc = \lambda^2 - (a + d)\lambda + (ad - bc) = 0$$

This is a quadratic equation. The constant term of $\det(A - \lambda I_2)$ is $ad - bc = \det A$, the value of $\det(A - \lambda I_2)$ at $\lambda = 0$. The coefficient of λ is $-(a + d)$, the opposite of the sum of the diagonal entries a and d of A. Since this sum is important in many other contexts as well, we introduce a name for it. ∎

Characteristic Polynomial

If A is an $n \times n$ matrix, then $\det(A - \lambda I_n)$ is a polynomial of degree n, of the form

$$\begin{aligned}(-\lambda)^n &+ (\operatorname{tr} A)(-\lambda)^{n-1} + \cdots + \det A \\ &= (-1)^n \lambda^n + (-1)^{n-1}(\operatorname{tr} A)\lambda^{n-1} + \cdots + \det A.\end{aligned}$$

This is called the *characteristic polynomial* of A, denoted by $f_A(\lambda)$.

$$f_A(\lambda) = \det(A - \lambda I_n) = (\lambda_1 - \lambda)(\lambda_2 - \lambda) \cdots (\lambda_n - \lambda).$$

Note that $f_A(0) = detA = \lambda_1 \lambda_2 ... \lambda_n$.

Algebra Multiplicity

We say that an eigenvalue λ_0 of a square matrix A has *algebraic multiplicity* k if λ_0 is a root of multiplicity k of the characteristic polynomial $f_A(\lambda)$, meaning that we can write

$$f_A(\lambda) = (\lambda_0 - \lambda)^k g(\lambda)$$

for some polynomial $g(\lambda)$ with $g(\lambda_0) \neq 0$.

An $n \times n$ matrix has *at most* n real eigenvalues, even if they are counted with their algebraic multiplicities.

If n is odd, then an $n \times n$ matrix has *at least* one real eigenvalue. ■

If n is even, an $n \times n$ matrix A need not have any real eigenvalues. Consider

$$A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix},$$

Finding eigenvectors

Consider an eigenvalue λ of an $n \times n$ matrix A. Then the kernel of the matrix $A - \lambda I_n$ is called the *eigenspace* associated with λ, denoted by E_λ:

$$E_\lambda = \ker(A - \lambda I_n) = \{\vec{v} \text{ in } \mathbb{R}^n : A\vec{v} = \lambda\vec{v}\}.$$

Note that the eigenvectors with eigenvalue λ are the *nonzero* vectors in the eigenspace E_λ.

Consider the eigenspace E_1 and E_0 of an orthogonal projection matrix.

Geometric Multiplicity

Consider an eigenvalue λ of an $n \times n$ matrix A. The dimension of eigenspace $E_\lambda = \ker(A - \lambda I_n)$ is called the *geometric multiplicity* of eigenvalue λ. Thus, the geometric multiplicity is the nullity of matrix $A - \lambda I_n$, or $n - \operatorname{rank}(A - \lambda I_n)$.

Geometric multiplicity of an eigenvalue λ is always less than or equal to its algebra multiplicity.

Eigenbasis

We then learn how to judge whether there exists a basis formed by eigenvectors.

Consider an $n \times n$ matrix A. A basis of $\mathbb{R}^n$ consisting of eigenvectors of A is called an *eigenbasis* for A.

a. Consider an $n \times n$ matrix A. If we find a basis of each eigenspace of A and concatenate all these bases, then the resulting eigenvectors $\vec{v}_1, \ldots, \vec{v}_s$ will be linearly independent. (Note that s is the sum of the geometric multiplicities of the eigenvalues of A.)

b. There exists an eigenbasis for an $n \times n$ matrix A if (and only if) the geometric multiplicities of the eigenvalues add up to n (meaning that $s = n$ in part a).

If an $n \times n$ matrix A has n distinct eigenvalues, then there exists an eigenbasis for A. We can construct an eigenbasis by finding an eigenvector for each eigenvalue.

Typical Exercise

Consider an Anatolian mountain farmer who raises goats. This particular breed of goats has a maximum life span of three years. At the end of each year t, the farmer conducts a census of his goats. He counts the number of young goats $j(t)$, born in the year t; the middle-aged goats $m(t)$, born the year before; and the old ones $a(t)$, born in the year $t-2$. The state of the herd can be described by the vector

$$\vec{x}(t) = \begin{bmatrix} j(t) \\ m(t) \\ a(t) \end{bmatrix}.$$

Suppose that for this breed and this environment the evolution of the system can be modeled by the equation

$$\vec{x}(t+1) = A\vec{x}(t), \quad \text{where} \quad A = \begin{bmatrix} 0 & 0.95 & 0.6 \\ 0.8 & 0 & 0 \\ 0 & 0.5 & 0 \end{bmatrix}.$$

For example, $m(t+1) = 0.8j(t)$, meaning that 80% of the young goats will survive to the next census. We leave it as an exercise to the reader to interpret the other 3 nonzero entries of A as reproduction and survival rates.

Suppose the initial populations are $j_0 = 750$ and $m_0 = a_0 = 200$. What will the populations be after t years, according to this model? What will happen in the long term?

We are told that the eigenvalues of A are $\lambda_1 = 1$, $\lambda_2 = -0.6$, and $\lambda_3 = -0.4$.

Solution

$$E_1 = \ker \begin{array}{c} \begin{array}{ccc} 5 & 4 & 2 \end{array} \\ \begin{bmatrix} -1 & 0.95 & 0.6 \\ 0.8 & -1 & 0 \\ 0 & 0.5 & -1 \end{bmatrix} \end{array} = \operatorname{span} \begin{bmatrix} 5 \\ 4 \\ 2 \end{bmatrix}.$$

$$E_{-0.6} = \operatorname{span} \begin{bmatrix} 9 \\ -12 \\ 10 \end{bmatrix}, \quad E_{-0.4} = \operatorname{span} \begin{bmatrix} -2 \\ 4 \\ -5 \end{bmatrix}.$$

Ve have constructed an eigenbasis: $\vec{v}_1 = \begin{bmatrix} 5 \\ 4 \\ 2 \end{bmatrix}, \vec{v}_2 = \begin{bmatrix} 9 \\ -12 \\ 10 \end{bmatrix}, \vec{v}_3 = \begin{bmatrix} -2 \\ 4 \\ -5 \end{bmatrix}.$

Solution

Next, we need to express the initial state vector

$$\vec{x}_0 = \begin{bmatrix} j_0 \\ m_0 \\ a_0 \end{bmatrix} = \begin{bmatrix} 750 \\ 200 \\ 200 \end{bmatrix}$$

as a linear combination of the eigenvectors: $\vec{x}_0 = c_1\vec{v}_1 + c_2\vec{v}_2 + c_3\vec{v}_3$. A somewhat tedious computation reveals that $c_1 = 100$, $c_2 = 50$, $c_3 = 100$.

Now that we know the eigenvalues λ_i, the eigenvectors $\vec{v}_i$, and the coefficients c_i, we are ready to write down the solution:

$$\vec{x}(t) = A^t\vec{x}_0 = c_1A^t\vec{v}_1 + c_2A^t\vec{v}_2 + c_3A^t\vec{v}_3 = c_1\lambda_1^t\vec{v}_1 + c_2\lambda_2^t\vec{v}_2 + c_3\lambda_3^t\vec{v}_3$$

$$= 100\begin{bmatrix} 5 \\ 4 \\ 2 \end{bmatrix} + 50(-0.6)^t\begin{bmatrix} 9 \\ -12 \\ 10 \end{bmatrix} + 100(-0.4)^t\begin{bmatrix} -2 \\ 4 \\ -5 \end{bmatrix}.$$

The individual populations are

$$\begin{aligned} j(t) &= 500 + 450(-0.6)^t - 200(-0.4)^t, \\ m(t) &= 400 - 600(-0.6)^t + 400(-0.4)^t, \\ a(t) &= 200 + 500(-0.6)^t - 500(-0.4)^t. \end{aligned}$$

In the long run, the populations approach the equilibrium values

$$j = 500, \quad m = 400, \quad a = 200.$$

Similar matrices

Suppose matrix A is similar to B. Then

a. Matrices A and B have the same characteristic polynomial, that is, $f_A(\lambda) = f_B(\lambda)$.

b. $\text{rank}(A) = \text{rank}(B)$ and $\text{nullity}(A) = \text{nullity}(B)$.

c. Matrices A and B have the same eigenvalues, with the same algebraic and geometric multiplicities. (However, the eigenvectors need not be the same.)

d. Matrices A and B have the same determinant and the same trace: $\det A = \det B$ and $\operatorname{tr} A = \operatorname{tr} B$.

Table of Contents

D matrix

Consider a linear transformation $T(\vec{x}) = A\vec{x}$, where A is a square matrix. Suppose $\mathfrak{D} = (\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n)$ is an eigenbasis for T, with $A\vec{v}_i = \lambda_i \vec{v}_i$. Then the $\mathfrak{D}$-matrix D of T is

$$D = S^{-1}AS = \begin{bmatrix} \lambda_1 & 0 & & 0 \\ 0 & \lambda_2 & & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & & \lambda_n \end{bmatrix}, \quad \text{where} \quad S = \begin{bmatrix} \vec{v}_1 & \vec{v}_2 & & \vec{v}_n \end{bmatrix}.$$

Matrix D is diagonal, and its diagonal entries are the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ of T. ■

Eigenbasis and diagonlization

An $n \times n$ matrix A is called *diagonalizable* if A is similar to some diagonal matrix D, that is, if there exists an invertible $n \times n$ matrix S such that $S^{-1}AS$ is diagonal.

a. Matrix A is diagonalizable if (and only if) there exists an eigenbasis for A.

b. If an $n \times n$ matrix A has n distinct eigenvalues, then A is diagonalizable.

If we know there exist an eigenbasis for A, then just follow the procedure of finding eigenvalues and vectors and follow the equation $D = S^{-1}AS$.

Exercise

Consider the linear transformation $T\big(f(x)\big) = f(2x-1)$ from P_2 to P_2. Is transformation T diagonalizable? If so, find an eigenbasis $\mathfrak{D}$ and the $\mathfrak{D}$-matrix D of T.

Solution

We will use a commutative diagram to find the matrix A of T with respect to the standard basis $\mathfrak{A} = (1, x, x^2)$.

$$\begin{tikzcd}
a + bx + cx^2 \arrow[r, "T"] \arrow[d, "L_{\mathfrak{A}}"] & \begin{array}{l} T(a + bx + cx^2) \\ = a + b(2x - 1) + c(2x - 1)^2 \\ = a - b + c + (2b - 4c)x + 4cx^2 \end{array} \arrow[d, "L_{\mathfrak{A}}"] \\
\begin{bmatrix} a \\ b \\ c \end{bmatrix} \arrow[r, swap, "A = \begin{bmatrix} 1 & -1 & 1 \\ 0 & 2 & -4 \\ 0 & 0 & 4 \end{bmatrix}"] & \begin{bmatrix} a - b + c \\ 2b - 4c \\ 4c \end{bmatrix}
\end{tikzcd}$$

The upper triangular matrix A has the three distinct eigenvalues, 1, 2, and 4, so that A is diagonalizable, by Theorem 7.4.3b. A straightforward computation produces the eigenbasis

$$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$$

for A. Transforming these vectors back into P_2, we find the eigenbasis $\mathfrak{D}$ for T consisting of

$$1, \quad x - 1, \quad x^2 - 2x + 1 = (x - 1)^2.$$

To check our work, we can verify that these are indeed eigenfunctions of T:

Solution

To check our work, we can verify that these are indeed eigenfunctions of T:

$$\begin{aligned} T(1) &= 1, \\ T(x-1) &= (2x-1)-1 = 2x-2 = 2(x-1), \\ T\left((x-1)^2\right) &= \left((2x-1)-1\right)^2 = (2x-2)^2 = 4(x-1)^2. \quad \checkmark \end{aligned}$$

The $\mathfrak{D}$-matrix of T is

$$D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 4 \end{bmatrix}.$$

Consider Figure 2, where

$$S = \begin{bmatrix} 1 & -1 & 1 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{bmatrix}$$

is the change of basis matrix from $\mathfrak{D}$ to $\mathfrak{A}$. ∎

Table of Contents

Spectral Theorem

A matrix A is *orthogonally diagonalizable* (i.e., there exists an orthogonal S such that $S^{-1}AS = S^T AS$ is diagonal) if and only if A is *symmetric* (i.e., $A^T = A$). ■

This is because only for symmetric matrix, the eigenvector of different eigenvalues are orthogonal to each other. For an eigenvalue with geometry multiplicity greater than 2, we can find a set of orthonormal basis via Gram-Schmidt process.

Orthogonal Diagonalization

Orthogonal diagonalization of a symmetric matrix A

a. Find the eigenvalues of A, and find a basis of each eigenspace.

b. Using the Gram–Schmidt process, find an *orthonormal* basis of each eigenspace.

c. Form an orthonormal eigenbasis $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n$ for A by concatenating the orthonormal bases you found in part (b), and let

$$S = \begin{bmatrix} | & | & & | \\ \vec{v}_1 & \vec{v}_2 & & \vec{v}_n \\ | & | & & | \end{bmatrix}.$$

S is orthogonal (by Theorem 8.1.2), and $S^{-1}AS$ will be diagonal. ∎

Table of Contents

Quadratic forms

A function $q(x_1, x_2, \ldots, x_n)$ from $\mathbb{R}^n$ to $\mathbb{R}$ is called a *quadratic form* if it is a linear combination of functions of the form $x_i x_j$ (where i and j may be equal). A quadratic form can be written as

$$q(\vec{x}) = \vec{x} \cdot A\vec{x} = \vec{x}^T A\vec{x},$$

for a unique symmetric $n \times n$ matrix A, called the matrix of q.

The uniqueness of matrix A will be shown in Exercise 52.

The set Q_n of quadratic forms $q(x_1, x_2, \ldots, x_n)$ is a *subspace* of the linear space of all functions from $\mathbb{R}^n$ to $\mathbb{R}$. In Exercise 42 you will be asked to think about the dimension of this space.

Diagonalizing a quadratic form

Consider a quadratic form $q(\vec{x}) = \vec{x} \cdot A\vec{x}$, where A is a symmetric $n \times n$ matrix. Let $\mathfrak{B}$ be an orthonormal eigenbasis for A, with associated eigenvalues $\lambda_1, \ldots, \lambda_n$. Then

$$q(\vec{x}) = \lambda_1 c_1^2 + \lambda_2 c_2^2 + \cdots + \lambda_n c_n^2,$$

where the c_i are the coordinates of $\vec{x}$ with respect to $\mathfrak{B}$.[3] ∎

Definiteness

Consider a quadratic form $q(\vec{x}) = \vec{x} \cdot A\vec{x}$, where A is a symmetric $n \times n$ matrix.

We say that A is *positive definite* if $q(\vec{x})$ is positive for all nonzero $\vec{x}$ in $\mathbb{R}^n$, and we call A *positive semidefinite* if $q(\vec{x}) \geq 0$, for all $\vec{x}$ in $\mathbb{R}^n$.

Negative definite and negative semidefinite symmetric matrices are defined analogously.

Finally, we call A *indefinite* if q takes positive as well as negative values.

A symmetric matrix A is *positive definite* if (and only if) all of its eigenvalues are positive. The matrix A is *positive semidefinite* if (and only if) all of its eigenvalues are positive or zero. ■

Proving Definiteness

Consider a symmetric $n \times n$ matrix A. For $m = 1, \ldots, n$, let $A^{(m)}$ be the $m \times m$ matrix obtained by omitting all rows and columns of A past the mth. These matrices $A^{(m)}$ are called the *principal submatrices* of A.

The matrix A is positive definite if (and only if) $\det(A^{(m)}) > 0$, for all $m = 1, \ldots, n$. ■

Consider the matrix

$$A = \begin{bmatrix} 9 & -1 & 2 \\ -1 & 7 & -3 \\ 2 & -3 & 3 \end{bmatrix}$$

from Example 2:

$$\det(A^{(1)}) = \det\begin{bmatrix} 9 \end{bmatrix} = 9 > 0$$

$$\det(A^{(2)}) = \det\begin{bmatrix} 9 & -1 \\ -1 & 7 \end{bmatrix} = 62 > 0$$

$$\det(A^{(3)}) = \det(A) = 89 > 0$$

We can conclude that A is positive definite.

Alternatively, we could find the eigenvalues of A and use Theorem 8.2.4. Using technology, we find that $\lambda_1 \approx 10.7$, $\lambda_2 \approx 7.1$, and $\lambda_3 \approx 1.2$, confirming our result.

Geometric Aspect

Consider a quadratic form $q(\vec{x}) = \vec{x} \cdot A\vec{x}$, where A is a symmetric $n \times n$ matrix with n distinct eigenvalues. Then the eigenspaces of A are called the *principal axes* of q. (Note that these eigenspaces will be one dimensional.)

Consider the curve C in $\mathbb{R}^2$ defined by

$$q(x_1, x_2) = ax_1^2 + bx_1x_2 + cx_2^2 = 1.$$

Let λ_1 and λ_2 be the eigenvalues of the matrix $\begin{bmatrix} a & b/2 \\ b/2 & c \end{bmatrix}$ of q.

If both λ_1 and λ_2 are positive, then C is an *ellipse*. If one eigenvalue is positive and the other is negative, then C is a hyperbola. ■

Exercise

Sketch the curve

$$8x_1^2 - 4x_1x_2 + 5x_2^2 = 1.$$

Answer

In Example 1, we found that we can write this equation as

$$9c_1^2 + 4c_2^2 = 1,$$

where c_1, c_2 are the coordinates of $\vec{x}$ with respect to the orthonormal eigenbasis

$$\vec{v}_1 = \frac{1}{\sqrt{5}}\begin{bmatrix} 2 \\ -1 \end{bmatrix}, \quad \vec{v}_2 = \frac{1}{\sqrt{5}}\begin{bmatrix} 1 \\ 2 \end{bmatrix},$$

for $A = \begin{bmatrix} 8 & -2 \\ -2 & 5 \end{bmatrix}$. We sketch this ellipse in Figure 4.

The c_1- and the c_2-axes are called the *principal axes* of the quadratic form $q(x_1, x_2) = 8x_1^2 - 4x_1x_2 + 5x_2^2$. Note that these are the eigenspaces of the matrix

$$A = \begin{bmatrix} 8 & -2 \\ -2 & 5 \end{bmatrix}$$

of the quadratic form. ■