Applied Linear Algebra

OTTO BRETSCHER

http://www.prenhall.com/bretscher

Chapter 7
Eigenvalues and Eigenvectors

Chia-Hui Chang
Email: chia@csie.ncu.edu.tw
National Central University, Taiwan

7.1 DYNAMICAL SYSTEMS AND EIGENVECTORS: AN INTRODUCTORY EXAMPLE

Consider a dynamical system:

$$x(t+1) = 0.86x(t) + 0.08y(t)$$

$$y(t+1) = -0.12x(t) + 1.14y(t)$$

Let

$$\vec{v}(t) = \begin{bmatrix} x(t) \\ y(t) \end{bmatrix}$$

be the **state vector** of the system at time t.

We can write the matrix equation as

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

where

$$A = \begin{bmatrix} 0.86 & 0.08 \\ -0.012 & 1.14 \end{bmatrix}$$

Suppose we know the initial state, we wish to find $\vec{v}(t)$, for any time t.

Case 1: Suppose $\vec{v}(0) = \begin{bmatrix} 100 \\ 300 \end{bmatrix}$

Case 2: Suppose $\vec{v}(0) = \begin{bmatrix} 200 \\ 100 \end{bmatrix}$

Case 3: Suppose $\vec{v}(0) = \begin{bmatrix} 1000 \\ 1000 \end{bmatrix}$

Case 1:

$$\vec{v}(1) = A\vec{v}(0) = \begin{bmatrix} 0.86 & 0.08 \\ -0.012 & 1.14 \end{bmatrix} \begin{bmatrix} 100 \\ 300 \end{bmatrix} = \begin{bmatrix} 110 \\ 330 \end{bmatrix}$$

$\vec{v}(1) = A\vec{v}(0) = 1.1\vec{v}(0)$
$\vec{v}(2) = A\vec{v}(1) = A(1.1\vec{v}(0)) = 1.1^2\vec{v}(0)$
$\vec{v}(3) = A\vec{v}(2) = A(1.1^2\vec{v}(0)) = 1.1^3\vec{v}(0)$
$\vdots$
$\vec{v}(t) = 1.1^t\vec{v}(0)$

Case 2:

$$\vec{v}(1) = A\vec{v}(0) = \begin{bmatrix} 0.86 & 0.08 \\ -0.012 & 1.14 \end{bmatrix} \begin{bmatrix} 200 \\ 100 \end{bmatrix} = \begin{bmatrix} 180 \\ 90 \end{bmatrix}$$

$\vec{v}(1) = A\vec{v}(0) = 0.9\vec{v}(0)$
$\vec{v}(t) = 0.9^t\vec{v}(0)$

Case 3:

$$\vec{v}(1) = A\vec{v}(0) = \begin{bmatrix} 0.86 & 0.08 \\ -0.012 & 1.14 \end{bmatrix} \begin{bmatrix} 1000 \\ 1000 \end{bmatrix} = \begin{bmatrix} 940 \\ 1020 \end{bmatrix}$$

The state vector $\vec{v}(1)$ is not a scalar multiple of the initial state $\vec{v}(0)$. We have to look for another approach.

Consider the two vectors

$$\vec{v}_1 = \begin{bmatrix} 100 \\ 300 \end{bmatrix} \text{ and } \vec{v}_2 = \begin{bmatrix} 200 \\ 100 \end{bmatrix}$$

Since the system is linear and

$$\vec{v}(0) = \begin{bmatrix} 1000 \\ 1000 \end{bmatrix} = 2\vec{v}_1 + 4\vec{v}_2$$

Therefore,

$$\vec{v}(t) = A^t\vec{v}(0) = A^t(2\vec{v}_1 + 4\vec{v}_2) = 2A^t\vec{v}_1 + 4A^t\vec{v}_2$$
$$= 2(1.1)^t\vec{v}_1 + 4(0.9)^t\vec{v}_2$$
$$= 2(1.1)^t \begin{bmatrix} 100 \\ 300 \end{bmatrix} + 4(0.9)^t \begin{bmatrix} 200 \\ 100 \end{bmatrix}$$

$$x(t) = 200(1.1)^t + 800(0.9)^t$$

$$y(t) = 600(1.1)^t + 400(0.9)^t$$

Since the terms involving 0.9^t approach zero as t increases, $x(t)$ and $y(t)$ eventually grow by about 10% each time, and their ratio $y(t)/x(t)$ approaches 600/200=3.

See Figure 3, The state vector $\vec{x}(t)$ approaches the line L_1, with the slope 3.

Connect the tips of the state vector $\vec{v}(i), i = 1, 2, ..., t$, the trajectory is shown in Figure 4.

Sometimes, we are interested in the state of the system in the past at times -1, -2, ...

For different $\vec{v}(0)$, the trajectory is different. Figure 5 shows the trajectory that starts above L_1 and one that starts below L_2.

From a mathematical point of view, it is informative to sketch a phase portrait of this system in the whole $c-r$-plane (see Figure 6), even though the trajectories outside the first quadrant are meaningless in terms of population study.

Eigenvectors and Eigenvalues

How do we find the initial state vector $\vec{v}$ such that $A\vec{v}$ is a scalar multiple of $\vec{v}$, or

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

for some scalar λ?

Definition 7.1.1

Eigenvectors and eigenvalues Consider an $n \times n$ matrix A. A **nonzero** vector $\vec{v}$ in 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 $\vec{v}$.

EXAMPLE 1

Find all eigenvectors and eigenvalues of the identity matrix I_n.

Solution All nonzero vectors in R^n are eigenvectors, with eigenvalue 1.

EXAMPLE 2

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

Solution (See Figure 8.) (a). Any vector $\vec{v}$ on L is a eigenvector with eigenvalue 1. (b). Any vector $\vec{w}$ perpendicular to L is a eigenvector with eigenvalue 0.

EXAMPLE 3

Let T from R^2 to R^2 be the rotation in the plane through an angle of 90° in the counter-clockwise direction. Find all eigenvalues and eigenvectors of T. (See Figure 9)

Solution There are no eigenvectors and eigen-values here.

EXAMPLE 4

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

Solution The possible real eigenvalue is 1 or -1 since orthogonal transformation preserves length.

Dynamical Systems and Eigenvectors

Fact 7.1.3 Discrete dynamical systems
Consider the dynamical system

$$\vec{x}(t+1) = A\vec{x}(t) \text{ with } \vec{x}(0) = \vec{x}_0$$

Then

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

Suppose we can find a basis $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n$ of 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 $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n$ of R^n:

$$\vec{x}(0) = c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots + c_n\vec{v}_n.$$

$$= \begin{bmatrix} \vec{v}_1 & \vec{v}_2 & \ldots & \vec{v}_n \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ . \\ c_n \end{bmatrix}$$

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

Then $\vec{x}_0 = S \begin{bmatrix} c_1 \\ c_2 \\ . \\ c_n \end{bmatrix}$ so that $\begin{bmatrix} c_1 \\ c_2 \\ . \\ c_n \end{bmatrix} = S^{-1}\vec{x}_0$.

Consider

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

We can write this equation in matrix form as

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

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

Definition 7.1.4

Discrete trajectories and phase portraits

Consider a discrete dynamical system

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

with initial value $\vec{x}(0) = \vec{x}_0$ where A is a 2×2 matrix. In this case, the state vector $\vec{x}(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}$ can be represented geometrically in the $x_1 - x_2$-plane.

The endpoints of state vectors $\vec{x}(0) = \vec{x}_0$, $\vec{x}(1) = A\vec{x}_0$, $\vec{x}(2) = A^2\vec{x}_0$, . . . form the (discrete) *trajectory* of this system, representing its evolution in the future. Sometimes we are interested in the past states $\vec{x}(-1) = A^{-1}\vec{x}_0$, $\vec{x}(-2) = (A^2)^{-1}\vec{x}_0$, . . . as well. It is suggestive to "connect the dots" to create the illusion of a continuous trajectories. Take another look at Figure 4.

A (discrete) *phase portrait* of the system $\vec{x}(t+1) = A\vec{x}(t)$ shows discrete trajectories for various initial states, capturing all the qualitatively different scenarios (as in Figure 6).

See Figure 11, we sketch phase portraits for the case when A has two eigenvalues $\lambda_1 > \lambda_2 > 0$. (Leave out the special case when one of the eigenvalues is 1.) Let $L_1 = span(\vec{v}_1)$ and $L_2 = span(\vec{v}_2)$. Since

$$\vec{x}(t) = c_1\lambda_1^t\vec{v}_1 + c_2\lambda_2^t\vec{v}_2$$

we can sketching the trajectories for the following cases:

(a) $\lambda_1 > \lambda_2 > 1$

(b) $\lambda_1 > 1 > \lambda_2$

(c) $1 > \lambda_1 > \lambda_2$

Summary 7.1.4

Consider an $n \times n$ matrix

$$\begin{bmatrix} | & | & & | \\ \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_n \\ | & | & & | \end{bmatrix}$$

Then the following statements are equivalent:

i. A is invertible.

ii. The linear system $A\vec{x} = \vec{b}$ has a unique solution $\vec{x}$, for all $\vec{b}$ for all $\vec{b}$ in R^n.

iii. $rref(A) = I_n$.

iv. $rank(A) = n$.

v. $im(A) = R^n$.

vi. $ker(A) = \{\vec{0}\}$.

vii. The $\vec{v}_i$ are a basis of R^n.

viii. The $\vec{v}_i$ span R^n.

ix. The $\vec{v}_i$ are linearly independent.

x. $det(A) \neq 0$.

xi. 0 fails to be an eigenvalue of A.

7.2 FINDING THE EIGENVALUES OF A MATRIX

Consider an $n \times n$ matrix A and a scalar λ. By definition λ is an eigenvalue of A if there is a nonzero vector $\vec{v}$ in R^n such that

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

$$\lambda\vec{v} - A\vec{v} = \vec{0}$$

$$(\lambda I_n - A)\vec{v} = \vec{0}$$

An an eigenvector, $\vec{v}$ needs to be a nonzero vector. By definition of the kernel, that

$$ker(\lambda I_n - A) \neq \{\vec{0}\}.$$

(That is, there are other vectors in the kernel besides the zero vector.)
Therefore, the matrix $\lambda I_n - A$ is **not invertible**, and det$(\lambda I_n - A)$=0.

Fact 7.2.1 Consider an $n \times n$ matrix A and a scalar λ. Then λ is an eigenvalue of A if (and only if) $\det(\lambda I_n - A) = 0$

λ is an eigenvalue of A.

$$\Updownarrow$$

There is a nonzero vector $\vec{v}$ such that $A\vec{v} = \lambda\vec{v}$ or $(\lambda I_n - A)\vec{v} = \vec{0}$.

$$\Updownarrow$$

$ker(\lambda I_n - A) \neq \{\vec{0}\}$.

$$\Updownarrow$$

$\lambda I_n - A$ is not invertible.

$$\Updownarrow$$

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

EXAMPLE 1 Find the eigenvalues of the matrix

$$A = \begin{bmatrix} 1 & 2 \\ 4 & 3 \end{bmatrix}.$$

Solution

By Fact 7.2.1, we have to solve the equation $\det(\lambda I_2 - A)$=0:

$$det(\lambda I_2 - A) = det\left(\begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} - \begin{bmatrix} 1 & 2 \\ 4 & 3 \end{bmatrix}\right)$$

$$= det\begin{bmatrix} \lambda - 1 & -2 \\ -4 & \lambda - 3 \end{bmatrix}$$

$$= (\lambda - 1)(\lambda - 3) - 8$$

$$= \lambda^2 - 4\lambda - 5$$

$$= (\lambda - 5)(\lambda + 1) = 0$$

The matrix A have two eigenvalues 5 and -1.

EXAMPLE 2 Find the eigenvalues of

$$A = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 0 & 2 & 3 & 4 & 5 \\ 0 & 0 & 3 & 4 & 5 \\ 0 & 0 & 0 & 4 & 5 \\ 0 & 0 & 0 & 0 & 5 \end{bmatrix}.$$

Solution

Again, we have to solve the equation det($\lambda I_5 - A$)=0:

$$det(\lambda I_5 - A) = \begin{bmatrix} \lambda-1 & -2 & -3 & -4 & -5 \\ 0 & \lambda-2 & -3 & -4 & -5 \\ 0 & 0 & \lambda-3 & -4 & -5 \\ 0 & 0 & 0 & \lambda-4 & -5 \\ 0 & 0 & 0 & 0 & \lambda-5 \end{bmatrix}$$

$$= (\lambda-1)(\lambda-2)(\lambda-3)(\lambda-4)(\lambda-5) = 0$$

There are five eigenvalues 1, 2, 3, 4, and 5 for matrix A.

Fact 7.2.2 The eigenvalues of a **triangular matrix** are its diagonal entries.

The eigenvalues of an $n \times n$ matrix A as zeros of the function

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

EXAMPLE 3 Find $f_A(\lambda)$ for the 2×2 matrix

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

Solution

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

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

The constant term is $det(A)$. Why? Because the constant term is $f_A(0)$=det($0I_2-A$)=det($-A$) =det(A).

Meanwhile, the coefficient of λ is the negative of the sum of the diagonal entries of A. Since the sum is important in many other contexts, we introduce a name for it.

Definition 7.2.3 Trace
The sum of the diagonal entries of an $n \times n$ matrix A is called the *trace* of A, denoted by tr(A).

Fact 7.2.4 If A is a 2×2, then

$$f_A(\lambda) = det(\lambda I_2 - A) = \lambda^2 - tr(A)\lambda + det(A)$$

For the matrix $A = \begin{bmatrix} 1 & 2 \\ 4 & 3 \end{bmatrix}$, we have tr($A$)=4 and det($A$)=-5, so that

$$f_A(\lambda) = \lambda^2 - 4\lambda - 5.$$

What is the format of $f_A(\lambda)$ for an $n \times n$ matrix A?

$$f_A(\lambda) = \lambda^n - tr(A)\lambda^{n-1} + (\text{a polynomial of degree} \leq (n-2)).$$

The constant term of this polynomial is $f_A(0) = det(-A) = (-1)^n det(A)$.

Fact 7.2.5 Characteristic polynomial
Consider an $n \times n$ matrix A. Then $f_A(\lambda) = det(\lambda I_n - A)$ is a polynomial of degree n of the form

$$f_A(\lambda) = \lambda^n - tr(A)\lambda^{n-1} + \cdots + (-1)^n det(A)$$

$f_A(\lambda)$ is called the *characteristic polynomial* of A

From elementary algebra, a polynomial of degree n has at most n zeros. If n is odd,

$$\lim_{\lambda \to \infty} f_A(\lambda) = \infty \text{ and } \lim_{\lambda \to -\infty} f_A(\lambda) = -\infty.$$

See Figure 1.

EXAMPLE 4 Find the eigenvalues of

$$A = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 0 & 2 & 3 & 4 & 5 \\ 0 & 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 2 & 3 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

Solution

Since $f_A(\lambda) = (\lambda - 1)^3(\lambda - 2)^2$, the eigenvalues are 1 and 2. Since 1 is a root of multiplicity 3 of the characteristic polynomial, we say that the eigenvalue 1 has **algebraic multiplicity** 3. Likewise, the eigenvalue 2 has algebraic multiplicity 2.

Definition 7.2.6

Algebraic multiplicity of an eigenvalue We say that an eigenvalue λ_0 of a square matrix A has *algebraic multiplicity* k if

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

for some polynomial $g(\lambda)$ with $g(\lambda_0) \neq 0$ (i.e.,if λ_0 is a root of multiplicity k of $f_A(\lambda)$).

EXAMPLE 5 Find the eigenvalues of

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

with their algebraic multiplicities.

Solution

$$f_A(\lambda) = det \begin{bmatrix} \lambda - 2 & 1 & 1 \\ 1 & \lambda - 2 & 1 \\ 1 & 1 & \lambda - 2 \end{bmatrix}$$

$$= (\lambda - 2)^3 + 2 - 3(\lambda - 2) = (\lambda - 3)^2 \lambda$$

We found two distinct eigenvalues, 3 and 0, with algebraic multiplicities 2 and 1, respectively.

Fact 7.2.7 An $n \times n$ matrix has at most n eigenvalues, even if they are counted with their algebraic multiplicities.

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

EXAMPLE 6 Describe all possible cases for the number of real eigenvalues of a 3×3 matrix and their algebraic multiplicities. Give an example in each case and graph the characteristic polynomial.

Solution

Case 1: See Figure 3.

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}, f_A(\lambda) = (\lambda - 1)(\lambda - 2)(\lambda - 3).$$

Case 2: See Figure 4.

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}, f_A(\lambda) = (\lambda - 1)^2(\lambda - 2).$$

Case 3: See Figure 5.

$$A = I_3, f_A(\lambda) = (\lambda - 1)^3.$$

Case 4: See Figure 6.

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}, f_A(\lambda) = (\lambda - 1)(\lambda^2 + 1).$$

It is usually impossible to find the exact eigenvalue of a matrix. To find approximations for the eigenvalues, you could graph the characteristic polynomial. The graph may give you an idea of the number of eigenvalues and their approximate values. Numerical analysts tell us that this is not a very efficient way to go; other techniques are used in practice. (See Exercise 7.5.33 for an example; another approach uses QR factorization.)

Exercises 7.2: 3, 5, 9, 11, 18, 20, 25

7.3 FINDING THE EIGENVECTORS OF A MATRIX

After we have found an eigenvalue λ of an $n \times n$ matrix A, we have to find the vectors $\vec{v}$ in R^n such that

$$A\vec{v} = \lambda\vec{v} \text{ or } (\lambda I_n - A)\vec{v} = \vec{0}$$

In other words, we have to find the *kernel* of the matrix $\lambda I_n - A$.

Definition 7.3.1 Eigenspace
Consider an eigenvalue λ of an $n \times n$ matrix A. Then the kernel of the matrix $\lambda I_n - A$ is called the $eigenspace$ associated with λ, denoted by E_λ:

$$E_\lambda = ker(\lambda I_n - A)$$

Note that E_λ consists of all solutions $\vec{v}$ of the linear system

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

EXAMPLE 1 Let $T(\vec{x}) = A\vec{v}$ be the orthogonal projection onto a plane E in R^3. Describe the eigenspaces geometrically.

Solution See Figure 1.
The nonzero vectors $\vec{v}$ in E are eigenvectors with eigenvalue 1. Therefore, the eigenspace E_1 is just the plane E.

Likewise, E_0 is simply the kernel of A ($A\vec{v} = \vec{0}$); that is, the line $E^{\perp}$ perpendicular to E.

EXAMPLE 2 Find the eigenvectors of the matrix $A = \begin{bmatrix} 1 & 2 \\ 4 & 3 \end{bmatrix}$.

Solution

See Section 7.2, Example 1, we saw the eigenvalues are 5 and -1. Then

$$E_5 = ker(5I_2 - A) = ker\begin{bmatrix} 4 & -2 \\ -4 & 2 \end{bmatrix}$$
$$= ker\begin{bmatrix} 4 & -2 \\ 0 & 0 \end{bmatrix} = span\begin{bmatrix} 2 \\ 4 \end{bmatrix} = span\begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

$$E_{-1} = ker(-I_2 - A) = ker\begin{bmatrix} -2 & -2 \\ -4 & -4 \end{bmatrix}$$
$$= span\begin{bmatrix} 2 \\ -2 \end{bmatrix} = span\begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

Both eigenspaces are lines, See Figure 2.

EXAMPLE 3 Find the eigenvectors of

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

Solution

Since

$$f_A(\lambda) = \lambda(\lambda - 1)^2$$

the eigenvalues are 1 and 0 with algebraic multiplicities 2 and 1.

$$E_1 = ker \begin{bmatrix} 0 & -1 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix}$$

To find this kernel, apply Gauss-Jordan Elimination:

$$\begin{bmatrix} 0 & -1 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix} \xrightarrow{rref} \begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

The general solution of the system

$$\left| \begin{array}{lll} x_2 & & = 0 \\ & x_3 & = 0 \end{array} \right|$$

is

$$\begin{bmatrix} x_1 \\ 0 \\ 0 \end{bmatrix} = x_1 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

Therefore,

$$E_1 = span \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

Likewise, compute the E_0:

$$E_0 = span \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$$

Both eigenspaces are lines in the x_1-x_2-plane, as shown in Figure 3.

Compare with Example 1. There, too, we have two eigenvalues 1 and 0, but one of the eigenspace, E_1, is a plane.

Definition 7.3.2 Geometric multiplicity
Consider an eigenvalue λ if a matrix A. Then the dimension of eigenvalue $E_\lambda = ker(\lambda I_n - A)$ is called the *geometric multiplicity* of eigenvalue λ. Thus, the geometric multiplicity of λ is the *nullity* of matrix $\lambda I_n - A$.

Example 3 shows that the geometric multiplicity of an eigenvalue may be different from the algebraic multiplicity. We have

(algebraic multiplicity of eigenvalue 1)=2,

but

(geometric multiplicity of eigenvalue 1)=1.

Fact 7.3.3
Consider an eigenvalue λ of a matrix A. Then

(geometric multiplicity of λ)$\leq$
(algebraic multiplicity of λ).

EXAMPLE 4 Consider an upper triangular matrix of the form

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

What can you say about the geometric multiplicity of the eigenvalue 4?

Solution

$$E_4 = \begin{bmatrix} 3 & \bullet & \bullet & \bullet & \bullet \\ 0 & 2 & \bullet & \bullet & \bullet \\ 0 & 0 & 0 & \bullet & \bullet \\ 0 & 0 & 0 & 0 & \bullet \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \xrightarrow{rref} \begin{bmatrix} 1 & \bullet & \bullet & \bullet & \bullet \\ 0 & 1 & \bullet & \bullet & \bullet \\ 0 & 0 & 0 & \sharp & \bullet \\ 0 & 0 & 0 & 0 & \sharp \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

The bullets on row 3 and 4 could be leading 1's. Therefore, the rank of this matrix will be between 2 and 4, and its nullity will be between 3 and 1. We can conclude that the geometric multiplicity of the eigenvalue 4 is less than the algebraic multiplicity.

Recall Fact 7.1.3, such a basis deserves a name.

Definition 7.3.4 Eigenbasis
Consider an $n\times$ n matrix A. A basis of R^n consisting of eigenvectors of A is called an $eigenbasis$ for A.

Example 1 Revisited: Projection on a plane E in R^3. Pick a basis $\vec{v}_1, \vec{v}_2$ of E and a nonzero $\vec{v}_3$ in $E^{\perp}$. The vectors $\vec{v}_1, \vec{v}_2, \vec{v}_3$ form an eigenbasis. See Figure 4.

Example 2 Revisited: $A = \begin{bmatrix} 1 & 2 \\ 4 & 3 \end{bmatrix}$.
The vectors $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$ form an eigenbasis for A, see Figure 5.

Example 3 Revisited: $A = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix}$.
There are not enough eigenvectors to form an eigenbasis. See Figure 6.

EXAMPLE 5 Consider a 3×3 matrix A with three eigenvalues, 1, 2, and 3. Let $\vec{v}_1, \vec{v}_2$, and $\vec{v}_3$ be corresponding eigenvectors. Are vectors $\vec{v}_1, \vec{v}_2$, and $\vec{v}_3$ necessarily linearly independent?

Solution See Figure 7.
Consider the plane E spanned by $\vec{v}_1$, and $\vec{v}_2$. We have to examine $\vec{v}_3$ can not be contained in this plane.

Consider a vector $\vec{x} = c_1\vec{v_1} + c_2\vec{v_2}$ in E (with $c_1 \neq 0$ and $c_2 \neq 0$). Then $A\vec{x} = c_1A\vec{v}_1 + c_2A\vec{v}_2 = c_1\vec{v_1} + 2c_2\vec{v_2}$. This vector can not be a scalar multiple of $\vec{x}$; that is, E does not contain any eigenvectors besides the multiples of $\vec{v}_1$ and $\vec{v}_2$; in particular, $\vec{v}_3$ is not contained in E.

Fact 7.3.5 Considers the eigenvectors $\vec{v}_1$, $\vec{v}_2$, ..., $\vec{v}_m$ of an $n \times n$ matrix A, with distinct eigenvalues λ_1, λ_2, ..., λ_m. Then the $\vec{v}_i$ are linearly independent.

Proof

We argue by induction on m. Assume the claim holds for $m-1$. Consider a relation

$$c_1\vec{v}_1 + \cdots + c_{m-1}\vec{v}_{m-1} + c_m\vec{v}_m = \vec{0}$$

• apply the transformation A to both sides:

$$c_1\lambda_1\vec{v}_1 + \cdots + c_{m-1}\lambda_{m-1}\vec{v}_{m-1} + c_m\lambda_m\vec{v}_m = \vec{0}$$

• multiply both sides by λ_m:

$$c_1\lambda_m\vec{v}_1 + \cdots + c_{m-1}\lambda_m\vec{v}_{m-1} + c_m\lambda_m\vec{v}_m = \vec{0}$$

Subtract the above two equations:

$$c_1(\lambda_1 - \lambda_m)\vec{v}_1 + \cdots + c_{m-1}(\lambda_{m-1} - \lambda_m)\vec{v}_{m-1} = \vec{0}$$

Since $\vec{v}_1$, $\vec{v}_2$, ..., $\vec{v}_{m-1}$ are linearly independent by induction, $c_i(\lambda_i - \lambda_m) = 0$, for $i = 1, ..., m-1$. The eigenvalues are assumed to be distinct; therefore $\lambda_i - \lambda_m \neq 0$, and $c_i = 0$. The first equation tells us that $c_m\vec{v}_m = \vec{0}$, so that $c_m = 0$ as well.

Fact 7.3.6 If an $n \times n$ matrix A has n distinct eigenvalues, then there is an eigenbasis for A. We can construct an eigenbasis by choosing an eigenvector for each eigenvalue.

EXAMPLE 6 Is there an eigenbasis for the following matrix?

$$A = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 0 & 2 & 3 & 4 & 5 & 6 \\ 0 & 0 & 3 & 4 & 5 & 6 \\ 0 & 0 & 0 & 4 & 5 & 6 \\ 0 & 0 & 0 & 0 & 5 & 6 \\ 0 & 0 & 0 & 0 & 0 & 6 \end{bmatrix}$$

Fact 7.3.7 Consider an $n \times n$ matrix A. If the geometric multiplicities of the eigenvalues of A add up to n, then there is an eigenbasis for A: We can construct an eigenbasis by choosing a basis of each eigenspace and combining these vectors.

Proof

Suppose the eigenvalues are λ_1, λ_2, ..., λ_m, with $\dim(E_{\lambda_i})=d_i$. We first choose a basis $\vec{v}_1$, $\vec{v}_2$, ..., $\vec{v}_{d_1}$ of E_{λ_1}, and then a basis $\vec{v}_{d_1+1}$, ..., $\vec{v}_{d_1+d_2}$ of E_{λ_2}, and so on.

Consider a relation

$$\underbrace{c_1\vec{v}_1+\cdots+c_{d_1}\vec{v}_{d_1}}_{\vec{w}_1 \text{ in } E_{\lambda_1}}+\underbrace{\cdots+c_{d_1+d_2}\vec{v}_{d_1+d_2}}_{\vec{w}_2 \text{ in } E_{\lambda_2}}+\cdots+\underbrace{\cdots+c_n\vec{v}_n}_{\vec{w}_m \text{ in } E_{\lambda_m}}=\vec{0}$$

Each under-braced sum $\vec{w}_i$ must be a zero vector since if they are nonzero eigenvectors, they must be linearly independent and the relation can not hold.

Because $\vec{w}_1=0$, it follows that $c_1=c_2=\cdots=c_{d_1}=0$, since $\vec{v}_1$, $\vec{v}_2$, ..., $\vec{v}_{d_1}$ are linearly independent. Likewise, all the other c_j are zero.

EXAMPLE 7 Consider an Albanian mountain farmer who raises goats. This particular breed of goats has a 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 $j(t)$ (those born in the year t), the middle-aged ones $m(t)$ (born the year before), and the old ones $a(t)$ (born in the year $t-2$). The state of the herd can be represented by the vector

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

How do we expect the population to change from year to year? Suppose that for this breed and environment the evolution of the system can be modelled by

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

where $A = \begin{bmatrix} 0 & 0.95 & 0.6 \\ 0.8 & 0 & 0 \\ 0 & 0.5 & 0 \end{bmatrix}$

We leave it as an exercise to interpret the entries of A in terms of reproduction rates 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?

Solution

Step 1: Find eigenvalues.

Step 2: Find eigenvectors.

Step 3: Express the initial vector $\vec{v}_0 = \begin{bmatrix} 750 \\ 200 \\ 200 \end{bmatrix}$ as a linear combination of eigenvectors.

Step 4: Write the closed formula for $\vec{v}(t)$.

Fact 7.3.8

The eigenvalues of similar matrices Suppose matrix A is similar to B. Then

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

2. rank(A) =rank(B) and nullity(A) =nullity(B)

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

4. det(A)=det(B) and tr(A)=tr(B)

Proof

a. If $B = S^{-1}AS$, then

$$f_B(\lambda) = det(\lambda I_n - B) = det(\lambda I_n - S^{-1}AS)$$

$= det(S^{-1}(\lambda I_n - A)S) = det(S^{-1})det(\lambda I_n - A)det(S)$
$= det(\lambda I_n - A) = f_A(\lambda)$

b. See Section 3.4, exercise 45 and 46.

c. If follows from part (a) that matrices A and B have the same eigenvalues, with the same algebraic multiplicities. As for for the geometric multiplicities, note that $\lambda I_n - A$ is similar to $\lambda I_n - B$ for all λ, so that nullity($\lambda I_n - A$)=nullity($\lambda I_n - B$) by part (b).

d. These equations follow from part (a) and Fact 7.2.5. Trance and determinant are coefficients of the characteristic polynomial.

7.4 Diagonalization

Fact 7.4.1 The matrix of a linear transformation with respect to an eigenbasis is diagonal

Consider a transformation $T\vec{x} = A\vec{x}$, where A is an $n \times n$ matrix. Suppose B is an eigenbasis for T consisting of vectors $\vec{v}_1, \vec{v}_2, ..., \vec{v}_n$, with $A\vec{v}_i = \lambda_i \vec{v}_i$. Then the B-matrix D of T is

$$D = S^{-1}AS = \begin{bmatrix} \lambda_1 & 0 & . & 0 \\ 0 & \lambda_2 & . & 0 \\ . & . & . & . \\ 0 & 0 & 0 & \lambda_n \end{bmatrix}$$

Here

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

$$\begin{array}{ccc} \vec{x} & \xrightarrow{A} & A\vec{x} \\ S \uparrow & & \uparrow S \\ \left[\ \vec{x}\ \right]_S & \xrightarrow[D]{} & \left[\ A\vec{x}\ \right]_S \end{array}$$

$$\vec{x} = S\left[\ \vec{x}\ \right]_S$$

$$\left[\ \vec{x}\ \right]_S = S^{-1}\vec{x}$$

Def 7.4.2 Diagonalizable matrices
An $n \times n$ matrix A is called *diagonalizable* if A is similar to a diagonal matrix D, that is, if there is an invertible $n \times n$ matrix S such that $D = S^{-1}AS$ is diagonal.

Fact 7.4.3
Matrix A is diagonalizable iff there is an eigenbasis for A. In particular, if an $n \times n$ matrix A has n distinct eigenvalues, then A is diagonalizable.

Alg 7.4.4 Diagonalization

Suppose we are asked to decide whether a given $n \times n$ matrix A is diagonalizable, if so, to find an invertible matrix S such that $S^{-1}AS$ is diagonal. We proceed as follows:

1. Find the eigenvalues of A, i.e., solve $f(\lambda) = det(\lambda I_n - A) = 0$.

2. For each eigenvalue λ, find a basis of the eigenspace $E_\lambda = ker(\lambda I_n - A)$.

3. A is diagonalizable iff the dimensions of the eigenspaces add up to n. In this case, we find an eigenbasis $\vec{v}_1, \vec{v}_2, ..., \vec{v}_n$ for A by combining the bases of the eigenspaces. Let $S = \begin{bmatrix} \vec{v}_1 & \vec{v}_2 & ... & \vec{v}_n \end{bmatrix}$, then the matrix $S^{-1}AS$ is a diagonal matrix.

Example Diagonalize the matrix

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Solution

a. The eigenvalues are 0 and 1.

b. $E_0 = ker(A) = span(\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix})$

and $E_1 = ker(I_3 - A) = span \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$

c. If we let

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

then

$$D = S^{-1}AS = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Alg 7.4.5 Powers of a diagonalizable matrix
To compute the powers A^t of a diagonalizable matrix A (where t is a positive integer), proceed as follows:

1. Use Alg 7.4.4 to diagonalize A, i.e. find S such that $S^{-1}AS = D$.

2. Since $A = SDS^{-1}$, $A^t = SD^tS^{-1}$.

3. To compute D^t, raise the diagonal entries of D to the tth power.