Chapter 8 Eigenvalues and Eigenvectors

The eigenvalues and eigenvectors are essential in illuminating certain characteristics of a matrix in the analysis of linear dynamic systems, control theory and numerical analysis and computation. They can show if a square matrix is full rank, or if a dynamic system will converge to an equilibrium. In control theory, they determined if a state can be attained within a finite number of steps.

Definition 8.1 Let **A** be a square matrix in $\mathbb{R}^{n\times n}$. An ***eigenvector*** of **A** is a nonzero vector **e** where,

$$\mathbf{Ae} = \lambda\mathbf{e}. \tag{8.1}$$

That is, the multiplication of **A** by **e** results in the same vector **e** multiplied by a scalar λ, called ***eigenvalue***.

8.1 Characteristic Polynomial and Characteristic Roots

From equation (8.1), we can write,

$$\mathbf{Ae} - \lambda\mathbf{e} = \mathbf{0}$$
$$(\mathbf{A} - \lambda\mathbf{I})\mathbf{e} = \mathbf{0},$$

where **I** is an $n \times n$ identity matrix.

The eigenvector **e** is just a nonzero solution to the homogeneous system of linear equations with the coefficient matrix $\mathbf{A} - \lambda\mathbf{I}$, which is the matrix **A** with its diagonal elements subtracted by some scalar λ. By Corollary 4.4, for the nonzero solution **e** to exist, the matrix $\mathbf{A} - \lambda\mathbf{I}$ has to be singular. Thus, the eigenvalue λ has to be some constant such that,

$$|\mathbf{A} - \lambda\mathbf{I}| = 0. \tag{8.2}$$

For an $n \times n$ matrix **A**, the determinant $|\mathbf{A} - \lambda\mathbf{I}|$ is a polynomial function of order n of the unknown λ. A root of the polynomial function is the value of λ which makes the polynomial function vanish to zero. There will be n possible roots, not necessarily distinct or real. This polynomial function is called ***characteristic polynomial*** function, and its roots ***characteristic roots***.

Example Let **A** be a 2×2 matrix, where

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}.$$

The characteristic polynomial is

$$\begin{aligned}|\mathbf{A}-\lambda\mathbf{I}| &= \begin{vmatrix} a_{11}-\lambda & a_{12} \\ a_{21} & a_{22}-\lambda \end{vmatrix} \\ &= (a_{11}-\lambda)(a_{22}-\lambda)-a_{21}a_{12} \\ &= \lambda^2-(a_{11}+a_{22})\lambda+(a_{11}a_{22}-a_{21}a_{12}).\end{aligned}$$

There are two possible roots of the polynomial, which are the values of λ that satisfy the equation

$$\lambda^2-(a_{11}+a_{22})\lambda+(a_{11}a_{22}-a_{21}a_{12})=0.$$

Example The characteristic polynomial of the matrix $\begin{bmatrix} 1 & 2 \\ 3 & 2 \end{bmatrix}$ is given by

$$\begin{aligned}\lambda^2-3\lambda-4&=0 \\ (\lambda-4)(\lambda+1)&=0 \\ \lambda&=-1,4.\end{aligned}$$

The two eigenvalues of the matrix are $\lambda_1=-1$ and $\lambda_2=4$.

Problem Find a matrix whose characteristic roots are not distinct. Can you find a matrix whose characteristic roots are complex numbers?

Once the eigenvalues are determined, the eigenvector associated with each of the eigenvalues can be found by subtracting that eigenvalue from the diagonal elements of **A** and solve the resulting homogeneous system of linear equations.

Example Continuing with the previous example, for $\lambda_1=-1$ the homogeneous system of linear equations are given by,

$$\begin{bmatrix} 1-\lambda_1 & 2 \\ 3 & 2-\lambda_1 \end{bmatrix}\begin{bmatrix} e_1 \\ e_2 \end{bmatrix}=\begin{bmatrix} 2 & 2 \\ 3 & 3 \end{bmatrix}\begin{bmatrix} e_1 \\ e_2 \end{bmatrix}=\begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

One of the nonzero solutions is $e_1=1$ and $e_2=-1$. The eigenvector associated with this eigenvalue $\lambda_1=-1$ is thus $\mathbf{e}_1=\begin{bmatrix} 1 \\ -1 \end{bmatrix}$.

For $\lambda_2=4$, the homogeneous system of linear equations are,

$$\begin{bmatrix} 1-\lambda_2 & 2 \\ 3 & 2-\lambda_2 \end{bmatrix}\begin{bmatrix} e_1 \\ e_2 \end{bmatrix} = \begin{bmatrix} -3 & 2 \\ 3 & -2 \end{bmatrix}\begin{bmatrix} e_1 \\ e_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

and similarly, the eigenvector is $\mathbf{e}_2 = \begin{bmatrix} 1 \\ 1.5 \end{bmatrix}$.

Observe that, the columns of the matrix $\mathbf{A} - \lambda\mathbf{I}$ are necessarily linearly dependent, because the eigenvalues λ are obtained from specifying that $|\mathbf{A} - \lambda\mathbf{I}| = 0$. Thus the columns of $\mathbf{A} - \lambda\mathbf{I}$ are linearly dependent and there are nontrivial solutions to the system $(\mathbf{A} - \lambda\mathbf{I})\mathbf{e} = \mathbf{0}$.

Note also that if **e** is an eigenvector, then so is any multiple of the vector **e**. That is, if $\mathbf{Ae} = \lambda\mathbf{e}$, so is **d**, if $\mathbf{d} = c\mathbf{e}$, for any nonzero scalar c. Thus, associated with each eigenvalue, there are infinite number of eigenvectors each being a multiple of others.

> **Problem** Let $\lambda_1, \lambda_2, \ldots, \lambda_n$ be eigenvalues and $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ be eigenvectors of a matrix **A**. If $\mathbf{B} = 0.5\mathbf{A}$, what are the eigenvalues and eigenvectors of **B**?

Theorem 8.1 A matrix $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ is singular if, and only if, $\lambda = 0$ is one of its eigenvalues.

Proof If $\lambda = 0$ is a root to the polynomial $|\mathbf{A} - \lambda\mathbf{I}| = 0$, then

$$|\mathbf{A} - \lambda\mathbf{I}| = |\mathbf{A} - 0\mathbf{I}| = |\mathbf{A}| = 0.\square$$

We can thus state a summarizing corollary as follows.

Corollary 8.1 Let **A** be a matrix in $\mathbb{R}^{n\times n}$. The following statements are equivalent.

1. All the eigenvalues of **A** are nonzero.
2. $Row(\mathbf{A}) = Col(\mathbf{A}) = \mathbb{R}^n$
3. $\dim(Row(\mathbf{A})) = \dim(Col(\mathbf{A})) = n$
4. The columns of **A** are a basis of $\mathbb{R}^n$.
5. The rows of **A** are a basis of $\mathbb{R}^n$.
6. The columns of **A** are linearly independent.
7. The rows of **A** are linearly independent.
8. $\mathbf{A}^{-1}$ exists.
9. $|\mathbf{A}| \neq 0$, i.e., **A** is nonsingular.
10. $rank\ \mathbf{A} = rank\ \mathbf{A}^{\mathrm{T}} = n$.
11. $\mathbf{Ax} = \mathbf{b}$ has a unique solution for any given vector $\mathbf{b} \in \mathbb{R}^n$.

Proof Exercise.❑

Problem State a summarizing corollary similar to Corollary 8.1 when **A** has at least one zero eigenvalue.

Problem Leon [1994], page 290, #2,4, 6-9,16, 25.

2. Show that the eigenvalues of a triangular matrix are the diagonal elements of the matrix.
4. Let **A** be a nonsingular matrix and let λ be an eigenvalue of **A**. Show that $\frac{1}{\lambda}$ is an eigenvalue of $\mathbf{A}^{-1}$.
6. An $n \times n$ matrix **A** is said to be idempotent if $\mathbf{A}^2 = \mathbf{A}$. Show that the eigenvalues of the matrix **A** must be either 0 or 1.
8. Let **A** be an $n \times n$ matrix and let $\mathbf{B} = \mathbf{A} - \alpha\mathbf{I}$ for some scalar α. How do the eigenvalues of **A** and **B** compare? Explain.
9. Show that **A** and $\mathbf{A}^{\mathrm{T}}$ have the same eigenvalues. Do they necessarily have the same eigenvectors? Explain.
16. Let **A** be an $n \times n$ matrix. Show that a vector **x** in $\mathbb{R}^n$ is an eigenvector of **A** if and only if the subspace $\boldsymbol{S}$ of $\mathbb{R}^n$ spanned by **x** and **Ax** has dimension 1.
25. Let **A** and **B** be $n \times n$ matrices. Show that:
    a. If λ is a nonzero eigenvalue of **AB**, then it is also an eigenvalue of **BA**.
    b. If $\lambda = 0$ is an eigenvalue of **AB**, then $\lambda = 0$ is also an eigenvalue of **BA**.

Problem Fraleigh & Beauregard [1995], page 301, #23, 25, 35.

23. Mark each of the following True or False and justify your answer.

______**a)** There can be only one eigenvalue associated with an eigenvector.

______**b)** There can be only one eigenvector associated with an eigenvalue.

______**c)** If **v** is an eigenvector of a matrix **A**, then **v** is an eigenvector of $\mathbf{A} + c\mathbf{I}$ for all scalars c.

______**d)** If λ is an eigenvalue of a matrix **A**, then λ is an eigenvalue of for all scalars c.

______**e)** If **v** is an eigenvector of an invertible matrix **A**, then $c\mathbf{v}$ is an eigenvector of $\mathbf{A}^{-1}$ for all nonzero scalars c.

25. Prove that if **A** is a square matrix, then $\mathbf{AA}^{\mathrm{T}}$ and $\mathbf{A}^{\mathrm{T}}\mathbf{A}$ have the same eigenvalues.

35. (Principle of biorthogonality) Let $\mathbf{A}$ be an $n \times n$ real matrix. Let $\mathbf{v}$ be an eigenvector of $\mathbf{A}$ with corresponding eigenvalue λ, and let $\mathbf{w}$ be an eigenvector of $\mathbf{A}^{\mathrm{T}}$ with corresponding eigenvalue α. Prove that if $\lambda \neq \alpha$, then $\mathbf{v}$ and $\mathbf{w}$ are perpendicular, i.e., $\mathbf{w}^{\mathrm{T}}\mathbf{v} = 0$. (Hint: Compute $\mathbf{w}^{\mathrm{T}}\mathbf{A}\mathbf{v}$ in two ways, using associativity of matrix multiplication).

Problem Show that an eigenvector of a square matrix cannot be associated to two distinct eigenvalues.

Problem Let $\boldsymbol{V}$ be a supspace of $\mathbb{R}^n$, and λ be an eigenvalue of a matrix $\mathbf{A}$ in $\mathbb{R}^{n\times n}$. Show that the ***eigenspace***, given by $\{\mathbf{x}|\mathbf{x} \in \boldsymbol{V}, \mathbf{A}\mathbf{x} = \lambda\mathbf{x}\}$, is a subspace of $\boldsymbol{V}$.

Problem Show that if λ is an eigenvalue of a matrix $\mathbf{A}$ in $\mathbb{R}^{n\times n}$ with associated eigenvector $\mathbf{e}$, then λ^k is an eigenvalue of $\mathbf{A}^k$ and has the same eigenvector $\mathbf{e}$.

8.2 Cayley-Hamilton Theorem

Theorem 8.2 (**Cayley-Hamilton Theorem**) Every square matrix $\mathbf{A}$ in $\mathbb{R}^{n\times n}$ satisfies its characteristic equation. That is, if

$$p(\lambda) = |\mathbf{A} - \lambda\mathbf{I}| = p_n\lambda^n + p_{n-1}\lambda^{n-1} + \cdots + p_1\lambda + p_0 = 0$$

then

$$p(\mathbf{A}) = p_n\mathbf{A}^n + p_{n-1}\mathbf{A}^{n-1} + \cdots + p_1\mathbf{A} + p_0\mathbf{I} = \mathbf{0}.$$

Proof For any square matrix $\mathbf{C}$, we can show the property of the adjoint that $|\mathbf{C}|\mathbf{I} = \mathbf{C} \cdot adj(\mathbf{C})$. For any scalar, not necessarily such that $p(\lambda) = 0$,

$$p(\lambda)\mathbf{I} = |\mathbf{A} - \lambda\mathbf{I}|\mathbf{I} = (\mathbf{A} - \lambda\mathbf{I}) \cdot adj\ (\mathbf{A} - \lambda\mathbf{I})$$

The matrix $adj\ (\mathbf{A} - \lambda\mathbf{I})$ is also a square matrix in $\mathbb{R}^{n\times n}$, and it can be shown inductively that each is a polynomial function of λ of order $n - 1$. That is, we can write

$$adj\ (\mathbf{A} - \lambda\mathbf{I}) = \mathbf{B}_{n-1}\lambda^{n-1} + \mathbf{B}_{n-2}\lambda^{n-2} + \cdots + \mathbf{B}_1\lambda + \mathbf{B}_0,$$

where $\mathbf{B}_i$ is a matrix of coefficients of λ^i, and

$$
\begin{aligned}
(\mathbf{A}-\lambda\mathbf{I})\ adj\ (\mathbf{A}-\lambda\mathbf{I})
&= (\mathbf{A}-\lambda\mathbf{I})(\mathbf{B}_{n-1}\lambda^{n-1}+\mathbf{B}_{n-2}\lambda^{n-2}+\cdots+\mathbf{B}_1\lambda+\mathbf{B}_0)\\
&= \mathbf{AB}_{n-1}\lambda^{n-1}+\mathbf{AB}_{n-2}\lambda^{n-2}+\cdots+\mathbf{AB}_1\lambda+\mathbf{AB}_{\underline{0}}\\
&\quad -(\mathbf{B}_{n-1}\lambda^{n}+\mathbf{B}_{n-2}\lambda^{n-1}+\cdots+\mathbf{B}_1\lambda^2+\mathbf{B}_0\lambda)\\
&= -\mathbf{B}_{n-1}\lambda^{n}+(\mathbf{AB}_{n-1}-\mathbf{B}_{n-2})\lambda^{n-1}+\cdots+(\mathbf{AB}_1-\mathbf{B}_0)\lambda+\mathbf{AB}_0\\
&= p_n\mathbf{I}\lambda^n+p_{n-1}\mathbf{I}\lambda^{n-1}+\cdots+p_1\mathbf{I}\lambda+p_0\mathbf{I}.
\end{aligned}
$$

Since the last equality holds for any value of the scalar λ, we have that all the matrices $\mathbf{B}_{n-1}, \mathbf{AB}_{n-1}-\mathbf{B}_{n-2}, \dots, \mathbf{AB}_1-\mathbf{B}_0, \mathbf{AB}_0$ must be diagonal (Why?), and

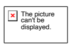
$$
\left.\begin{aligned}
p_n\mathbf{I} &= -\mathbf{B}_{n-1}\\
p_{n-1}\mathbf{I} &= \mathbf{AB}_{n-1}-\mathbf{B}_{n-2}\\
&\ \vdots\\
p_1\mathbf{I} &= \mathbf{AB}_1-\mathbf{B}_0
\end{aligned}\right\}
\left\{\begin{aligned}
p_n\mathbf{A}^n &= -\mathbf{A}^n\mathbf{B}_{n-1}\\
p_{n-1}\mathbf{A}^{n-1} &= \mathbf{A}^{n-1}\mathbf{B}_{n-1}-\boldsymbol{A}^{n-1}\boldsymbol{B}_{n-2}\\
&\ \vdots\\
p_1\mathbf{A} &= \mathbf{A}^2\mathbf{B}_1-\mathbf{AB}_0
\end{aligned}\right.
$$

$$
p_0\mathbf{I} = \mathbf{AB}_0
$$

Thus,

$$
\begin{aligned}
p(\mathbf{A}) &= p_n\mathbf{A}^n+p_{n-1}\mathbf{A}^{n-1}+\cdots+p_1\mathbf{A}+p_0\mathbf{I}\\
&= -\mathbf{A}^n\mathbf{B}_{n-1}+(\mathbf{A}^{n-1}\mathbf{B}_{n-1}-\boldsymbol{A}^{n-1}\boldsymbol{B}_{n-2})+\cdots\\
&\qquad +(\mathbf{A}^2\mathbf{B}_1-\mathbf{AB}_0)+\mathbf{AB}_0\\
&= \mathbf{0}
\end{aligned}
$$

This completes the proof.❑

Definition 8.2 Let **A** and **B** be matrices in 

. The matrices **A** and **B** are ***similar*** if $\mathbf{B} = \boldsymbol{S}^{-1}\boldsymbol{AS}$, for some nonsingular matrix **S**.

Theorem 8.3 If matrices **A** and **B** are similar, then they have the same characteristic polynomial, and thus have the same eigenvalues.

Proof Exercise.❑

> **Problem** If matrices **A** and **B** are similar, do they have the same eigenvectors?
>
> **Problem Leon** [1994], page 292, #18, 19.

18. Let $\mathbf{x}_1, \mathbf{x}_2, \dots \mathbf{x}_r$ be eigenvectors of an $n \times n$ matrix **A** and let $\boldsymbol{S}$ be the subspace of spanned by $\mathbf{x}_1, \mathbf{x}_2, \dots \mathbf{x}_r$. Show that $\boldsymbol{S}$ is invariant under **A** (i.e., show that $\mathbf{Ax} \in \boldsymbol{S}$ whenever $\mathbf{x} \in \boldsymbol{S}$).
19. Let $\mathbf{B} = \boldsymbol{S}^{-1}\boldsymbol{AS}$ and let **x** be an eigenvector of **B** belonging to an eigenvalue λ. Show that **Sx** is an eigenvector of **A** belonging to λ.

Problem Johnson, Riess & Arnold [1998], page 286, #27. Suppose that **B** is similar to **A**. Show each of the following.

a. $\mathbf{B} + \alpha\mathbf{I}$ is similar to $\mathbf{A} + \alpha\mathbf{I}$.
b. $\mathbf{B}^{\mathrm{T}}$ is similar to $\mathbf{A}^{\mathrm{T}}$.
c. If **A** is nonsingular, then **B** is nonsingular and, moreover, $\mathbf{B}^{-1}$ is similar to $\mathbf{A}^{-1}$.

8.3 Distinct Eigenvalues and Linear Independence of Eigenvectors

The next theorem states that the eigenvectors corresponding to distinct eigenvalues are linearly independent.

Theorem 8.4 Let $\lambda_1, \lambda_2, \dots, \lambda_k$ be distinct eigenvalues of an $n \times n$ matrix **A** . The eigenvectors $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_k$ corresponding to these distinct eigenvalues are linearly independent.

Proof Suppose that the eigenvectors $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_k$ are linearly dependent. Then, there exists a minimum number h vectors out of k vectors $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_k$ that are linearly dependent. (Why?) We can assume without loss of generality that these h vectors are $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_h$, and write

$$c_1\mathbf{e}_1 + c_2\mathbf{e}_2 + \cdots + c_h\,\mathbf{e}_h = \mathbf{0},$$

for some nonzero constants $c_1, c_2, \dots, c_h$. By multiplying this equality by the matrix **A,**

$$\begin{aligned}\mathbf{A}(c_1\mathbf{e}_1 + c_2\mathbf{e}_2 + \cdots + c_h\,\mathbf{e}_h) &= c_1\mathbf{A}\mathbf{e}_1 + c_2\mathbf{A}\mathbf{e}_2 + \cdots + c_h\,\mathbf{A}\mathbf{e}_h \\ &= c_1\lambda_1\mathbf{e}_1 + c_2\lambda_2\mathbf{e}_2 + \cdots + c_h\,\lambda_h\mathbf{e}_h = \mathbf{0}.\end{aligned}$$

However, if we multiply $c_1\mathbf{e}_1 + c_2\mathbf{e}_2 + \cdots + c_h\,\mathbf{e}_h = \mathbf{0}$, by λ_1 and subtract it from the last equality we have,

$$c_2(\lambda_2 - \lambda_1)\mathbf{e}_2 + \cdots + c_h\,(\lambda_h - \lambda_1)\mathbf{e}_h = \mathbf{0}$$

Thus, the $h-1$ eigenvectors are linearly dependent since all eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_h$ are assumed to be distinct. This is a contradiction and we can conclude that $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_k$ are linearly independent.❑

Problem Redo the proof of the previous theorem by induction.

Solution For any given square matrix **A** in $\mathbb{R}^{n\times n}$, if we consider only one eigenvector, it must be linearly independent because eigenvectors are nonzero. We can now assume the induction hypothesis for the case of k that the eigenvectors $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_k$ corresponding to distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_k$ are linearly independent. Then we have to show for the case of $k+1$. We have to show that for any linear combination,

$$c_1\mathbf{e}_1 + c_2\mathbf{e}_2 + \cdots + c_{k+1}\,\mathbf{e}_{k+1} = \mathbf{0},, \tag{8.3}$$

implies that $c_1 = c_2 = \cdots = c_{k+1} = 0$. We can assume without loss of generality that $\lambda_{k+1} \neq 0$. (Why?) Multiplying the linear combination (8.3) by the eigenvalue λ_{k+1} and the matrix **A** respectively, we have

$$c_1\lambda_{k+1}\mathbf{e}_1 + c_2\lambda_{k+1}\mathbf{e}_2 + \cdots + c_{k+1}\lambda_{k+1}\,\mathbf{e}_{k+1} = \mathbf{0},, \tag{8.4}$$

and

$$\begin{aligned}&\mathbf{A}(c_1\mathbf{e}_1 + c_2\mathbf{e}_2 + \cdots + c_{k+1}\,\mathbf{e}_{k+1})\\ &= c_1\lambda_1\mathbf{e}_1 + c_2\lambda_2\mathbf{e}_2 + \cdots + c_{k+1}\lambda_{k+1}\,\mathbf{e}_{k+1}\mathbf{0}.\end{aligned} \tag{8.5}$$

Subtracting (8.4) from (8.5), we obtain a linear combination of just the first k eigenvectors $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_k$ as

$$\begin{aligned}c_1(\lambda_1{-}\lambda_{k+1})\mathbf{e}_1 + c_2(\lambda_2{-}\lambda_{k+1})\mathbf{e}_2 + \cdots\\ + c_k\,(\lambda_k{-}\lambda_{k+1})\mathbf{e}_k = \mathbf{0}\end{aligned}$$

Since all the eigenvalues are distinct and by the induction hypothesis, we can conclude that the coefficients $c_1 = c_2 = \cdots = c_k = 0$, and then the only possibility that the linear combination $c_1\mathbf{e}_1 + c_2\mathbf{e}_2 + \cdots + c_{k+1}\,\mathbf{e}_{k+1} = \mathbf{0}$, is that $c_{k+1} = 0$ as the eigenvector $\mathbf{e}_{k+1}$ is by definition nonzero.❑

Problem Would Theorem 8.4 still hold true if one of the eigenvalues is zero?

Problem If the eigenvectors $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_k$ are linearly independent, is it true that their corresponding eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_k$ must be distinct?

If the eigenvalues are distinct and nonzero, we also have additional result as stated in the following problem.

Problem Let $\lambda_1, \lambda_2, \dots, \lambda_k$ be nonzero distinct eigenvalues of an $n \times n$ matrix $\mathbf{A}$. Show that $\mathbf{Ae}_1, \mathbf{Ae}_2, \dots, \mathbf{Ae}_k$ are linearly independent, where $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_k$ are the corresponding eigenvectors of $\lambda_1, \lambda_2, \dots, \lambda_k$.

By Theorem 8.4, we can decompose a square matrix into a product of matrices if the eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_n$ are distinct. This immensely facilitates matrix computations and has many theoretical implications. The decomposition when $\lambda_1, \lambda_2, \dots, \lambda_n$ are not distinct is also possible but is more complicated and will not be discussed here. See **Strang** [1976], *Linear Algebra and Its Applications*, for the decomposition when eigenvalues are repeated, and **Noble & Daniel** [1977], *Applied Linear Algebra*, for many other kinds of decompositions.

8.4 Matrix Diagonalization and Decomposition

Theorem 8.5 Let $\mathbf{A}$ be a matrix in $\mathbb{R}^{n\times n}$ with distinct eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_n$ and eigenvectors $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$. Then,

a) $\mathbf{AP} = \mathbf{P\Lambda}$, where $\mathbf{P} = [\mathbf{e}_1 \quad \mathbf{e}_2 \quad \cdots \quad \mathbf{e}_n]$ and $\mathbf{\Lambda}$ is a diagonal matrix whose diagonal elements are $\lambda_1, \lambda_2, \dots, \lambda_n$.
b) $\mathbf{P}^{-1}\mathbf{AP} = \mathbf{\Lambda}$
c) $\mathbf{A} = \mathbf{P\Lambda P}^{-1}$.

Proof (a)

$$\begin{aligned}\mathbf{AP} &= \mathbf{A}[\mathbf{e}_1 \quad \mathbf{e}_2 \quad \cdots \quad \mathbf{e}_n] \\ &= [\mathbf{Ae}_1 \quad \mathbf{Ae}_2 \quad \cdots \quad \mathbf{Ae}_n] \\ &= [\lambda_1\mathbf{e}_1 \quad \lambda_2\mathbf{e}_2 \quad \cdots \quad \lambda_n\mathbf{e}_n] \\ &= \mathbf{P\Lambda}.\end{aligned}$$

b. Since $\lambda_1, \lambda_2, \dots, \lambda_n$ are distinct, the vectors $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ are linearly independent and $\mathbf{P}^{-1}$ exists. Premultiplying $\mathbf{AP} = \mathbf{P\Lambda}$ by $\mathbf{P}^{-1}$, we have $\mathbf{P}^{-1}\mathbf{AP} = \mathbf{\Lambda}$.

(c) Postmultiplying $\mathbf{AP} = \mathbf{P\Lambda}$ by $\mathbf{P}^{-1}$, we have $\mathbf{A} = \mathbf{P\Lambda P}^{-1}$.❑

Note that the matrices **A** and **Λ** are similar.

Example The matrix $\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 3 & 2 \end{bmatrix}$ has $\lambda_1 = -1$ and $\lambda_2 = 4$ as eigenvalues with associated eigenvectors $\mathbf{e}_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ and $\mathbf{e}_2 = \begin{bmatrix} 1 \\ 1.5 \end{bmatrix}$, respectively. Then,

$$\mathbf{P} = [\mathbf{e}_1 \quad \mathbf{e}_2] = \begin{bmatrix} 1 & 1 \\ -1 & 1.5 \end{bmatrix}$$

and $\mathbf{P}^{-1} = \begin{bmatrix} 3/5 & -2/5 \\ 2/5 & 2/5 \end{bmatrix}$. According to Theorem 8.5, we have as expected

$$\mathbf{\Lambda} = \begin{bmatrix} -1 & 0 \\ 0 & 4 \end{bmatrix} = \begin{bmatrix} 3/5 & -2/5 \\ 2/5 & 2/5 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 2 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -1 & 1.5 \end{bmatrix}$$
$$= \mathbf{P\Lambda P}^{-1}.$$

Corollary 8.2 With the same assumptions as in Theorem 8.5 above, $|\mathbf{\Lambda}| = |\mathbf{A}|$.

Proof By the property of the determinant of product of matrices stated in Theorem 4.4,

$$|\mathbf{A}| = |\mathbf{P\Lambda P}^{-1}| = |\mathbf{P}||\mathbf{\Lambda}||\mathbf{P}^{-1}| = |\mathbf{\Lambda}|$$

❑

Problem Define the trace of matrix **A** as $tr\mathbf{A} = \sum_{i=1}^{n} a_{ii}$. Show that $tr\mathbf{AB} = tr\mathbf{BA}$, and use it to show that $tr\mathbf{A} = tr\mathbf{\Lambda} = \sum_{i=1}^{n} \lambda_i$.

Problem Show that $\mathbf{A}^k = \mathbf{P\Lambda}^k\mathbf{P}^{-1}$.

Problem Let $\lambda_1, \lambda_2, \ldots, \lambda_n$ be eigenvalues and $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ be eigenvectors of a matrix $\mathbf{A} \in \mathbb{R}^{n\times n}$. What are the eigenvalues and eigenvectors of $\mathbf{A}^2 + 3\mathbf{A} - 7\mathbf{I}$?

Problem If a matrix **A** can be diagonalized by a matrix **P**, how can $\mathbf{A}^{-1}$ be diagonalized? How about $\mathbf{A}^{\mathrm{T}}$?

Problem **Leon** [1994], page 318, #6, 10, 17, 18.

6. Let $\mathbf{A}$ be a diagonalizable matrix whose eigenvalues are all either 1 or -1. Show that $\mathbf{A}^{-1} = \mathbf{A}$.
10. Let $\mathbf{A}$ be an $n \times n$ matrix with positive real eigenvalues $\lambda_1 > \lambda_2 > \cdots > \lambda_n > 0$. Let $\mathbf{x}_i$ be an eigenvector belonging to λ_i for each i and let $\mathbf{x} = \alpha_1\mathbf{x}_1 + \alpha_2\mathbf{x}_2 + \cdots + \alpha_n\mathbf{x}_n$.
    a. Show that $\mathbf{A}^m\mathbf{x} = \sum_{i=1}^{n} \alpha_i\lambda_i^m\mathbf{x}_i$.
    b. If $\lambda_1 = 1$, show that $\lim_{m\to\infty} \mathbf{A}^m\mathbf{x} = \alpha_1\mathbf{x}_1$.
17. Let $\mathbf{A}$ be a diagonalizable $n \times n$ matrix. Prove that if $\mathbf{B}$ is any matrix that is similar to $\mathbf{A}$, then $\mathbf{B}$ is diagonalizable.
18. Show that if $\mathbf{A}$ and $\mathbf{B}$ are two $n \times n$ matrices that both have the same diagonalizing matrix $\mathbf{X}$, then .

Problem Fraleigh & Beauregard [1995], page 315, #13, 16, 22.

13. Mark each of the following True or False.

______**a)** Every matrix is diagonalizable.

______**b)** If an $n \times n$ matrix has n distinct real eigenvalues, it is diagonalizable.

______**d)** An $n \times n$ matrix is diagonalizable if and only if it has n distinct eigenvalues.

______**f)** Every invertible matrix is diagonalizable.

______**g)** Every triangular matrix is diagonalizable.

______**i)** If an $n \times n$ matrix $\mathbf{A}$ is diagonalizable, there is a unique diagonal matrix $\mathbf{D}$ that is similar to $\mathbf{A}$.

______**j)** If $\mathbf{A}$ and $\mathbf{B}$ are similar square matrices, then $|\mathbf{A}| = |\mathbf{B}|$.

16. Let $\mathbf{P}$, $\mathbf{Q}$, and $\mathbf{R}$ be $n \times n$ matrices. Recall that $\mathbf{P}$ is similar to $\mathbf{Q}$ if there exists and invertible $n \times n$ matrix $\mathbf{C}$ such that $\mathbf{C}^{-1}\mathbf{PC} = \mathbf{Q}$. This exercise shows that similarity is an equivalence relation.
    a. (*Reflexive*) Show that $\mathbf{P}$ is similar to itself.
    b. (*Symmetric*) Show that, if $\mathbf{P}$ is similar to $\mathbf{Q}$, then $\mathbf{Q}$ is similar to $\mathbf{P}$.
    c. (*Transitive*) Show that, if $\mathbf{P}$ is similar to $\mathbf{Q}$ and $\mathbf{Q}$ is similar to $\mathbf{R}$, then $\mathbf{P}$ is similar to $\mathbf{R}$.

22. Let $\mathbf{A}$ and $\mathbf{C}$ be $n \times n$ matrices, and let $\mathbf{C}$ be invertible. Prove that, if $\mathbf{v}$ is an eigenvector of $\mathbf{A}$ with corresponding eigenvalue λ, then $\mathbf{C}^{-1}\mathbf{v}$

is an eigenvector of $\mathbf{C}^{-1}\mathbf{AC}$ with the corresponding eigenvalue λ.

8.5 Applications of Eigenvalues and Eigenvectors in Dynamic Systems

The following examples show the applications eigenvalues and eigenvectors in dynamic systems when eigenvalues are distinct, and so the above diagonalization is possible.

8.5.1 Application 1: *Homogeneous System of Linear Difference Equations*.

Let the state of a system at time k be represented by a vector $\mathbf{x}(k)$. The state of the system at time $\mathbf{x}(k+1)$ is given by,

$$\mathbf{x}(k+1) = \mathbf{A}\mathbf{x}(k)$$

where $\mathbf{A}$ is a $n \times n$ matrix of coefficients. This is a homogeneous system of linear difference equations where the state at time $k+1$ is determined by the state at time k.

If the state at time 0 is $\mathbf{x}(0)$, the state at time k is just $\mathbf{x}(k) = \mathbf{A}^k\mathbf{x}(0)$. This is the ***solution*** of the difference equation $\mathbf{x}(k+1) = \mathbf{A}\mathbf{x}(k)$ as can now be written only as a function of k and the starting state $\mathbf{x}(0)$, and not a function of state vector at any other time.

> **Problem** If the system happens to start at $\mathbf{x}(0)$ being one of the eigenvectors, what will the state vector $\mathbf{x}(k)$ be at any other time k after that? What if the eigenvalue associated to the eigenvector that is the starting $\mathbf{x}(0)$ equals to one?

If the matrix $\mathbf{A}$ can be diagonalized by $\mathbf{A} = \mathbf{P}\mathbf{\Lambda}\mathbf{P}^{-1}$, where columns of $\mathbf{P}$ are eigenvectors $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ associated with distinct eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_n$, the computation of the matrix $\mathbf{A}^k$ can be simplified to, $\mathbf{A}^k = \mathbf{P}\mathbf{\Lambda}^k\mathbf{P}^{-1}$. When k is large, this simplification of the matrix $\mathbf{A}$ is much easier to compute as the matrix $\mathbf{\Lambda}^k$ is still diagonal with each diagonal element being an eigenvalue to the power k.

The dynamic system can be simplified by the defining another state variable $\mathbf{z}(k)$, where

$$\mathbf{z}(k) = \mathbf{P}^{-1}\mathbf{x}(k),$$

and then

$$\begin{aligned}\mathbf{x}(k) &= \mathbf{A}^k\mathbf{x}(0)\\ \mathbf{P}^{-1}\mathbf{x}(k) &= \mathbf{P}^{-1}\mathbf{A}^k\mathbf{x}(0)\\ &= \mathbf{P}^{-1}(\mathbf{P}\mathbf{\Lambda}^k\mathbf{P}^{-1})\mathbf{x}(0)\\ &= \mathbf{\Lambda}^k\mathbf{P}^{-1}\mathbf{x}(0)\\ \mathbf{z}(k) &= \mathbf{\Lambda}^k\mathbf{z}(0).\end{aligned}$$

Thus we can compute $\mathbf{z}(k)$ easily as a collection of n ***independent*** first-order homogeneous difference equations $\mathbf{z}(k+1) = \mathbf{\Lambda}\mathbf{z}(k)$, and once we have $\mathbf{z}(k)$, we can convert it back to $\mathbf{x}(k)$ by the linear relation, $\mathbf{x}(k) = \mathbf{P}\mathbf{z}(k)$. We can then write , $\mathbf{x}(k)$ as a linear combination of eigenvectors as follows:

$$\begin{aligned}\mathbf{x}(k) = \mathbf{P}\mathbf{z}(k) &= [\mathbf{e}_1 \quad \mathbf{e}_2 \quad \cdots \quad \mathbf{e}_n]\mathbf{\Lambda}^k\mathbf{z}(0)\\ &= [\mathbf{e}_1 \quad \mathbf{e}_2 \quad \cdots \quad \mathbf{e}_n]\begin{bmatrix}\lambda_1^k z_1(0)\\ \lambda_2^k z_2(0)\\ \vdots\\ \lambda_n^k z_n(0)\end{bmatrix}\\ &= \lambda_1^k z_1(0)\mathbf{e}_1 + \lambda_2^k z_2(0)\mathbf{e}_2 + \cdots + \lambda_n^k z_n(0)\mathbf{e}_n.\end{aligned}$$

Definition 8.3 The eigenvalue λ is ***dominant*** if $|\lambda| = \max_{1\le i\le n}\{|\lambda_i|\}$. The eigenvector corresponding to the dominant eigenvalue is called the ***dominant eigenvector***.

We can now make the following conclusions about the behavior of the system $\mathbf{x}(k+1) = \mathbf{A}\mathbf{x}(k)$ when k approaches infinity. This is the question whether the system possesses an equilibrium point.

Definition 8.4 An ***equilibrium point*** $\bar{\mathbf{x}}$ of a difference equation is a point that once $\mathbf{x}(k) = \bar{\mathbf{x}}$, then for all $i > k$, $\mathbf{x}(i) = \bar{\mathbf{x}}$.

Example Every homogeneous system of linear difference equations has $\bar{\mathbf{x}} = \mathbf{0}$ as an equilibrium point.

Theorem 8.6 For the homogeneous difference equation $\mathbf{x}(k+1) = \mathbf{A}\mathbf{x}(k)$, the dominant eigenvalue and its eigenvector of the matrix **A** determine the behavior of the state vector as time k approaches infinity as follows:

a. The system $\mathbf{x}(k+1) = \mathbf{A}\mathbf{x}(k)$ diverges if, and only if, the absolute value of the dominant eigenvalue is greater than one. That is, there is no equilibrium point.

b. The system $\mathbf{x}(k+1) = \mathbf{A}\mathbf{x}(k)$ ***converges*** to zero vector if, and only if, the absolute value of the dominant eigenvalue is less than one. That is, the equilibrium point is $\bar{\mathbf{x}} = \mathbf{0}$.
c. The system $\mathbf{x}(k+1) = \mathbf{A}\mathbf{x}(k)$ ***converges*** to a nonzero vector for any $\bar{\mathbf{x}}$ if, and only if, the value of the dominant eigenvalue is equal to one. This nonzero vector $\bar{\mathbf{x}}$ is just the dominant eigenvector.
d. If the dominant eigenvalue is positive, the state vector $\mathbf{x}(k)$ will converge assymptotically to a multiple of the dominant eigenvector.
e. If the dominant eigenvalue is negative, $\mathbf{x}(k)$ will oscillate between the positive and negative of a multiple of the dominant eigenvector.

Proof Part **(a)** and **(b)** follow directly from

$$\mathbf{x}(k) = \lambda_1^k z_1(0)\mathbf{e}_1 + \lambda_2^k z_2(0)\mathbf{e}_2 + \cdots + \lambda_n^k z_n(0)\mathbf{e}_n.$$

Assume that λ_1 is dominant and divide both sides of this equation by λ_1^k, we have

$$\frac{\mathbf{x}(k)}{\lambda_1^k} = \left(\frac{\lambda_1}{\lambda_1}\right)^k z_1(0)\mathbf{e}_1 + \left(\frac{\lambda_2}{\lambda_1}\right)^k z_2(0)\mathbf{e}_2 + \cdots + \left(\frac{\lambda_n}{\lambda_1}\right)^k z_n(0)\mathbf{e}_n$$

$$= z_1(0)\mathbf{e}_1 + \left(\frac{\lambda_2}{\lambda_1}\right)^k z_2(0)\mathbf{e}_2 + \cdots + \left(\frac{\lambda_n}{\lambda_1}\right)^k z_n(0)\mathbf{e}_n,$$

and part **(c), (d)** and **(e)** follow.❑

Problem Show that if the initial state $\mathbf{x}(0) = \mathbf{x}_0$ is an eigenvector corresponding with an eigenvalue λ, the state of the system will always be $\mathbf{x}_0$ multiplied by a constant. What is this constant? What if eigenvalue of this eigenvector is 1?

Problem Fraleigh & Beauregard [1995], page 318, Fibonacci Sequence, page 325, #1-5.

1. Let the sequence $a_0, a_1, a_2, \ldots$ be given by $a_0 = 0$, $a_1 = 1$, and $a_k = \frac{a_{k-1}+a_{k-2}}{2}$, for $k \geq 2$.
   a. Find the matrix $\mathbf{A}$ that can be used to generate this sequence as $\mathbf{x}(k+1) = \mathbf{A}\mathbf{x}(k)$ to generate the Fibonacci sequence.
   b. Classify this generation process as stable, neutrally stable, or unstable.

c. Compute the expression $\mathbf{A}^k\mathbf{x}_0$ as a linear combination of eigenvectors of **A** for $\mathbf{x}_0 = \begin{bmatrix} a_0 \\ a_1 \end{bmatrix}$ of this process. Check computations with the first few terms of the sequence.
d. Use the answer to part **(c)** to estimate a_k for large k.

2. Repeat Exercise 1 if $a_k = a_{k-1} - \frac{3}{16}a_{k-2}$ for $k \geq 2$.
3. Repeat Exercise 1 but change the initial data to $a_0 = 1, a_1 = 0$.
4. Repeat Exercise 1 if $a_k = \frac{1}{2}a_{k-1} + \frac{3}{16}a_{k-2}$ for $k \geq 2$..
5. Repeat Exercise 1 if $a_k = a_{k-1} + \frac{3}{4}a_{k-2}$ for $k \geq 2$.

Problem Leon [1994], page 319, #19.

19. The city of Mawtookit maintains a constant population of 300,000 people from year to year. A political science study estimated that there were 150,000 independents, 90,000 Democrats, and 60,000 Republicans in the town. It was also estimated that each year 20 percent of the independents become Democrats and 10 percent become Republicans. Similarly 20 percent of the democrats become independents and 10 percent become Republicans, while 10 percent of the Republicans defect to the Democrats and 10 percent become independents each year. Let

$$\mathbf{x}_0 = \begin{bmatrix} 150{,}000 \\ 90{,}000 \\ 60{,}000 \end{bmatrix}$$

and let $\mathbf{x}(1)$ be a vector representing the number of people in each group after 1 year.

a. Find a matrix **A** such that $\mathbf{A}\mathbf{x}_0 = \mathbf{x}(1)$.
b. Show that $\lambda_1 = 1$, $\lambda_2 = 0.5$, and $\lambda_3 = 0.7$ are the eigenvalues of **A** and factor **A** into a product $\mathbf{P\Lambda P}^{-1} = \mathbf{A}$, where $\mathbf{\Lambda}$ is diagonal.
c. Which group will dominate in the long run? Justify your answer by computing $\lim_{k\to\infty} \mathbf{A}^k\mathbf{x}_0$.

8.5.2 Application 2: *Nonhomogeneous System of Linear Difference Equations*

The state of the system at time $k + 1$ is now given by

$$\mathbf{x}(k+1) = \mathbf{x}(k) + \mathbf{b}$$

where $\mathbf{A}$ is a square matrix and $\mathbf{b}$ is a nonzero vector. Then the solution of the difference equation is given by

$$\mathbf{x}(k) = \mathbf{A}^k\mathbf{x}(0) + (\mathbf{A}^{k-1} + \mathbf{A}^{k-2} + \cdots + \mathbf{A} + \mathbf{I})\mathbf{b}$$

That is, the solution of the nonhomogeneous system also contains the solution of the homogeneous one.

Theorem 8.7 If the system $\mathbf{x}(k+1) = \mathbf{x}(k) + \mathbf{b}$ has an equilibrium point $\bar{\mathbf{x}}$ for a given $\mathbf{b} \in \mathbb{R}^n$, then $\mathbf{b}$ is in the column space of the matrix $\mathbf{I} - \mathbf{A}$.

Proof By the definition of the equilibrium point, $\bar{\mathbf{x}} = \mathbf{A}\bar{\mathbf{x}} + \mathbf{b}$ and thus $(\mathbf{I} - \mathbf{A})\bar{\mathbf{x}} = \mathbf{b}$. ❑

That is, if the vector $\mathbf{b}$ is not in the column space of $\mathbf{I} - \mathbf{A}$, the system $\mathbf{x}(k+1) = \mathbf{x}(k) + \mathbf{b}$ has no equilibrium point.

Corollary 8.3 The difference equation $\mathbf{x}(k+1) = \mathbf{x}(k) + \mathbf{b}$ has an equilibrium point for any $\mathbf{b} \in \mathbb{R}^n$ if, and only if, no eigenvalue of the matrix $\mathbf{A}$ is equal to 1.

Proof Left as an exercise.❑

Even when the equilibrium point exists, there is no guarantee that starting from any $\mathbf{x}(0)$ we will eventually attain the state $\bar{\mathbf{x}}$. This is the stability question of the equilibrium.

Definition 8.5 The equilibrium point $\bar{\mathbf{x}}$ is ***stable*** if $\mathbf{x}(k)$ approaches $\bar{\mathbf{x}}$ as time k approaches infinity, for any given starting point $\mathbf{x}(0)$.

Theorem 8.8 Let $\bar{\mathbf{x}}$ be the equilibrium point of $\mathbf{x}(k+1) = \mathbf{x}(k) + \mathbf{b}$. The equilibrium point $\bar{\mathbf{x}}$ is ***stable*** if, and only if, the absolute value of the dominant eigenvalue is less than one.

Proof Let $\mathbf{y}(k) = \mathbf{x}(k) - \bar{\mathbf{x}}$, the difference of the state vector to its equilibrium. Then,

$$\begin{aligned}\mathbf{y}(k+1) &= \mathbf{x}(k+1) - \bar{\mathbf{x}} \\ &= \mathbf{A}\mathbf{x}(k) + \mathbf{b} - (\mathbf{A}\bar{\mathbf{x}} + \mathbf{b}) \\ &= \mathbf{A}(\mathbf{x}(k) - \bar{\mathbf{x}}) \\ &= \mathbf{A}\mathbf{y}(k).\end{aligned}$$

That is, $\mathbf{y}(k)$ is the state vector of a homogeneous difference equation with the same coefficient matrix $\mathbf{A}$. Thus, $\mathbf{x}(k) \to \bar{\mathbf{x}}$ if, and only if, $\mathbf{y}(k) \to \mathbf{0}$. From Theorem 8.6 (b), the homogeneous difference equation $\mathbf{y}(k+1) = \mathbf{A}\mathbf{y}(k)$ vanishes to $\mathbf{0}$ as $k \to \infty$ if, and only if, the absolute value of the dominant eigenvalue is less than one.❑

To illustrate more clearly that $\mathbf{x}(k) \to \bar{\mathbf{x}}$ for any starting point $\mathbf{x}(0)$ when the absolute value of eigenvalue is less than one, recall that since

$$\mathbf{x}(k) = \mathbf{A}^k\mathbf{x}(0) + (\mathbf{A}^{k-1} + \mathbf{A}^{k-2} + \cdots + \mathbf{A} + \mathbf{I})\mathbf{b},$$

as $k \to \infty$, $\mathbf{A}^k$ vanishes and $\mathbf{x}(k) \to (\ldots + \mathbf{A}^k + \mathbf{A} + \mathbf{I})\mathbf{b}$. It can be shown that

$$(\ldots + \mathbf{A}^2 + \mathbf{A} + \mathbf{I})\mathbf{b} = (\mathbf{I} - \mathbf{A})^{-1}\mathbf{b}\text{'}$$

and is left as an exercise.

> **Problem** If $\mathbf{y}(k) = \mathbf{x}(k) - \bar{\mathbf{x}}$, show that $\mathbf{A}^k\mathbf{y}(0) + \bar{\mathbf{x}}$ is a solution of the difference equation $\mathbf{x}(k+1) = \mathbf{x}(k) + \mathbf{b}$. That is, the solution of the nonhomogeneous system can be written as a sum of the solution of the homogeneous system $\mathbf{y}(k+1) = \mathbf{A}\mathbf{y}(k)$ and the equilibrium point $\bar{\mathbf{x}}$.

8.5.3 Application 3 Dynamics of Left and Right Eigenvectors

Definition 8.6 Let **A** be a matrix in $\mathbb{R}^{n\times n}$ with eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ and eigenvectors $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$. These eigenvectors are called ***right eigenvectors*** of **A** because $\mathbf{A}\mathbf{e}_j = \lambda_j\mathbf{e}_j$ and $\mathbf{e}_j$ is on the right of matrix **A**. A row vector $\mathbf{f}_i$ is called ***left eigenvectors*** associated with eigenvalue λ_i of **A** if $\mathbf{f}_i\mathbf{A} = \lambda_i\,\mathbf{f}_i$.

Theorem 8.9 Let $\mathbf{e}_j$ be the right eigenvector associated with eigenvalue λ_j and $\mathbf{f}_i$ be the left eigenvector associated with eigenvalue λ_i of a given matrix **A**. If $\lambda_i \neq \lambda_j$, then $\mathbf{f}_i$ and $\mathbf{e}_j$ are orthogonal.

Proof Since $\mathbf{e}_j$ and $\mathbf{f}_i$ are right and left eigenvectors of $\mathbf{A}$, $\mathbf{A}\mathbf{e}_j = \lambda_j \mathbf{e}_j$ and $\mathbf{f}_i \mathbf{A} = \lambda_i \mathbf{f}_i$. Premultiplying the first equality by $\mathbf{f}_i$ and postmultiplying the second equality by $\mathbf{e}_j$, we have $\mathbf{f}_i \mathbf{A} \mathbf{e}_j = \mathbf{f}_i \lambda_j \mathbf{e}_j = \mathbf{f}_i \lambda_i \mathbf{e}_j$. Because $\lambda_i \neq \lambda_j$, $(\lambda_i - \lambda_j)\mathbf{f}_i \mathbf{e}_j = 0$ implies that $\mathbf{f}_i \mathbf{e}_j = 0$ as required.❑

Let $\mathbf{A}$ be a matrix in $\mathbb{R}^{n\times n}$ with distinct eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_n$ and eigenvectors $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$. Then, $\mathbf{P} = [\mathbf{e}_1 \quad \mathbf{e}_2 \quad \cdots \quad \mathbf{e}_n]$ is nonsingular and exists and $\mathbf{P}\mathbf{A} = \mathbf{\Lambda}\mathbf{P}^{-1}$. That is, the rows of $\mathbf{P}^{-1} = \begin{bmatrix} \mathbf{f}_1 \\ \mathbf{f}_2 \\ \vdots \\ \mathbf{f}_n \end{bmatrix}$ are the left eigenvectors of $\mathbf{A}$.

From the diagonalization of the difference equation in Section 8.5.1 , we have

$$
\begin{aligned}
\mathbf{x}(k) &= \mathbf{A}^k \mathbf{x}(0) \\
&= \mathbf{P}\mathbf{\Lambda}^k \mathbf{P}^{-1} \mathbf{x}(0) \\
&= \mathbf{P}\mathbf{\Lambda}^k \mathbf{z}(0) \\
&= \mathbf{\Lambda}^k \mathbf{P}^{-1} \mathbf{x}(0) \\
&= \lambda_1^k z_1(0)\mathbf{e}_1 + \lambda_2^k z_2(0)\mathbf{e}_2 + \cdots + \lambda_n^k z_n(0)\mathbf{e}_n.
\end{aligned}
$$

That is, the state variable $\mathbf{x}(k)$ is a linear combination of the right eigenvectors. Also, since we define $\mathbf{z}(k) = \mathbf{P}^{-1}\mathbf{x}(k)$, we have

$$
\begin{aligned}
\mathbf{z}(k) = \mathbf{P}^{-1}\mathbf{x}(k) = \begin{bmatrix} \mathbf{f}_1 \\ \mathbf{f}_2 \\ \vdots \\ \mathbf{f}_n \end{bmatrix} \mathbf{x}(k) &= \begin{bmatrix} \sum_{j=1}^{n} f_{1j} x_j(k) \\ \sum_{j=1}^{n} f_{2j} x_j(k) \\ \vdots \\ \sum_{j=1}^{n} f_{nj} x_j(k) \end{bmatrix} \\
&= \begin{bmatrix} z_1(k) \\ z_2(k) \\ \vdots \\ z_n(k) \end{bmatrix}.
\end{aligned}
$$

Therefore, each variable $z_i(k)$ that is governed by independent first-order difference equations $\mathbf{z}(k) = \mathbf{\Lambda}^k \mathbf{z}(0)$ can be viewed as a linear combination of the state variables $x_1(k), x_2(k), \dots, x_n(k)$.

This dynamics of the right and left eigenvectors can be demonstrated by the following simple homogeneous difference equation.

Example A Simple Migration Model. See **Luenberger** [1979], *Introduction to Dynamic Systems*: *Theory, Models, and Applications*, Section 5.5, page 144-147. A population can be divided into those in rural and urban areas. Let

$\alpha =$ rate of growth of population in both segments, $\alpha > 1$
$r(k) =$ rural population at time k
$u(k) =$ urban population at time k

Thus, the total population at time k is given by $r(k) + u(k)$. Assume that γ is ideal proportion of total population to be in the rural area. The difference of the actual rural population and this ideal rural population $r(k) - \gamma(r(k) + u(k)$ will determine the migration. That is

$$r(k+1) = r(k) - \beta[r(k) - \gamma(r(k) + u(k))]$$
$$u(k+1) = u(k) + \beta[r(k) - \gamma(r(k) + u(k))]$$

where β is the fraction of this difference to migrate away from rural area. We will assume that $\beta < \alpha$. We can write the difference equation as

$$\begin{aligned}\mathbf{x}(k+1) &= \mathbf{A}\mathbf{x}(k) \\ &= \begin{bmatrix} \alpha - \beta(1-\gamma) & \beta\gamma \\ \beta(1-\gamma) & \alpha - \beta\gamma \end{bmatrix} \begin{bmatrix} r(k) \\ u(k) \end{bmatrix}.\end{aligned}$$

We can calculate the eigenvalues, right eigenvectors and left eigenvectors as follows:

$$\lambda_1 = \alpha, \lambda_2 = \alpha - \beta,$$
$$\mathbf{e}_1 = \begin{bmatrix} \gamma \\ 1-\gamma \end{bmatrix}, \mathbf{e}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix},$$
$$\mathbf{f}_1 = [1 \quad 1], \mathbf{f}_2 = [1-\gamma \quad -\gamma],$$

where we normalize $\mathbf{f}_1\mathbf{e}_1 = 1$ and $\mathbf{f}_2\mathbf{e}_2 = 1$. We have,

$$\mathbf{A} = \begin{bmatrix} \gamma & 1 \\ 1-\gamma & -1 \end{bmatrix} \begin{bmatrix} \alpha & 0 \\ 0 & \alpha-\beta \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1-\gamma & -\gamma \end{bmatrix}$$
$$\begin{aligned}\mathbf{A}^k &= \begin{bmatrix} \gamma & 1 \\ 1-\gamma & -1 \end{bmatrix} \begin{bmatrix} \alpha^k & 0 \\ 0 & (\alpha-\beta)^k \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1-\gamma & -\gamma \end{bmatrix} \\ &= \begin{bmatrix} \gamma\alpha^k & (\alpha-\beta)^k \\ (1-\gamma)\alpha^k & -(\alpha-\beta)^k \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1-\gamma & -\gamma \end{bmatrix}\end{aligned}$$
$$= \begin{bmatrix} \gamma\alpha^k + (1-\gamma)(\alpha-\beta)^k & \gamma\alpha^k - \gamma(\alpha-\beta)^k \\ (1-\gamma)\alpha^k - (1-\gamma)(\alpha-\beta)^k & (1-\gamma)\alpha^k + \gamma(\alpha-\beta)^k \end{bmatrix}.$$

The solution of this difference solution is thus given by,

$$\mathbf{x}(k) = \mathbf{A}^k\mathbf{x}(0) = \mathbf{A}^k \begin{bmatrix} r(0) \\ u(0) \end{bmatrix}$$
$$= \begin{bmatrix} \gamma\alpha^k + (1-\gamma)(\alpha-\beta)^k & \gamma\alpha^k - \gamma(\alpha-\beta)^k \\ (1-\gamma)\alpha^k - (1-\gamma)(\alpha-\beta)^k & (1-\gamma)\alpha^k + \gamma(\alpha-\beta)^k \end{bmatrix} \begin{bmatrix} r(0) \\ u(0) \end{bmatrix}$$
$$= \begin{bmatrix} (\gamma\alpha^k + (1-\gamma)(\alpha-\beta)^k)r(0) + (\gamma\alpha^k - \gamma(\alpha-\beta)^k)u(0) \\ \left((1-\gamma)\alpha^k - (1-\gamma)(\alpha-\beta)^k\right)r(0) + \left((1-\gamma)\alpha^k + \gamma(\alpha-\beta)^k\right)u(0) \end{bmatrix}$$
$$= \begin{bmatrix} (r(0)+u(0))\gamma\alpha^k + (\alpha-\beta)^k((1-\gamma)r(0) - \gamma u(0)) \\ \left(r(0)+u(0)\right)(1-\gamma)\alpha^k + (\alpha-\beta)^k(-(1-\gamma)r(0) + \gamma u(0)) \end{bmatrix}.$$

We have

$$z_1(0) = \mathbf{f}_1\mathbf{x}(0) = [1 \quad 1] \begin{bmatrix} r(0) \\ u(0) \end{bmatrix} = r(0) + u(0)$$
$$= \text{Total population at time 0.}$$
$$z_2(0) = \mathbf{f}_2\mathbf{x}(0) = [1-\gamma \quad -\gamma] \begin{bmatrix} r(0) \\ u(0) \end{bmatrix}$$
$$= r(0) - \gamma(r(0) + u(0))$$
$$= \text{Net rural population imbalance over ideal rural population at time 0.}$$

We have the following conclusions:

1) The solution $\mathbf{x}(k)$ can then be written as

$$\mathbf{x}(k) = \lambda_1^k z_1(0)\mathbf{e}_1 + \lambda_2^k z_2(0)\mathbf{e}_2$$
$$= \alpha^k z(0) \begin{bmatrix} \gamma \\ 1-\gamma \end{bmatrix} + (\alpha-\beta)^k z_2(0) \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$
$$= \alpha^k\left(r(0)+u(0)\right) \begin{bmatrix} \gamma \\ 1-\gamma \end{bmatrix}$$
$$+ (\alpha-\beta)^k(r(0) - \gamma(r(0)+u(0))) \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

2) Since $\alpha > \beta > 0$, the system $\mathbf{x}(k)$ is dominated by the eigenvector $\mathbf{e}_1 = \begin{bmatrix} \gamma \\ 1-\gamma \end{bmatrix}$ as $\lambda_1 = \alpha > \lambda_2 = \alpha - \beta$. The system $\mathbf{x}(k)$ does not converge such that $\mathbf{x}(k) \to \mathbf{0}$ because $\lambda_1 = \alpha > 1$.

3) If we allow $\beta > \alpha$, then $\mathbf{x}(k)$ will oscillate. The system $\mathbf{x}(k)$ is still dominated by the eigenvector $\mathbf{e}_1 = \begin{bmatrix} \gamma \\ 1-\gamma \end{bmatrix}$ if $|\lambda_2| = |\alpha - \beta| < 0$.

Example Markov Chains.

> **Problem Johnson, Riess & Arnold** [1998], page 299, #23, 25.

23. Let $\mathbf{B} = [b_{ij}]$ be an $n \times n$ matrix. Matrix **B** is called a stochastic matrix if **B** contains only nonnegative entries and $b_{i1} + b_{i2} + \cdots + b_{in} = 1$, $i = 1,2,\dots,n$. Show that $\lambda = 1$ is an eigenvalue of **B**.
25. Let **B** be a stochastic matrix, and let λ be any eigenvalue of **B**. Show that $|\lambda| \le 1$. For simplicity, assume that λ is real.

8.6 Eigenvalues and Eigenvectors of Symmetric Matrix

One sufficient condition to guarantee that the eigenvalues are real is that the matrix **A** is symmetric. This will also make the inverse of the matrix **P** equal to its transpose, i.e., **P** is orthogonal.

Theorem 8.10 (Theorem 23.16, **Simon & Blume** [1994]) Let **A** be a symmetric matrix in $\mathbb{R}^{n\times n}$. Then,

a) All n eigenvalues of **A** are real.
b) Eigenvectors $\mathbf{e}_i$ and $\mathbf{e}_j$ corresponding to distinct eigenvalues λ_i and λ_j , $\lambda_i \neq \lambda_j$, are orthogonal , i.e., $\mathbf{e}_i^{\mathrm{T}}\mathbf{e}_j = 0$.
c) Even if **A** has repeated eigenvalues, there is a orthogonal matrix **P** whose columns are eigenvectors of **A** such that,
    I. The columns of **P** are mutually orthogonal to each other.
    II. $\mathbf{P}^{-1} = \mathbf{P}^{\mathrm{T}}$, and
    III. $\mathbf{P}^{\mathrm{T}}\mathbf{A}\mathbf{P} = \mathbf{\Lambda}$.

Proof See **Simon & Blume** [1994], pages 621-623 and Exercise 23.41. See also **Johnson, Riess & Arnold** [1998], Section 3.6, pages 263-272.❑

This diagonalization of a symmetric matrix is called the ***Schur decomposition***.

Definition 8.7 A matrix **P** is an ***orthogonal matrix*** if $\mathbf{P}^{-1} = \mathbf{P}^{\mathrm{T}}$, or equivalently, $\mathbf{P}^{\mathrm{T}}\mathbf{P} = \mathbf{I}$.

Problem Show that the determinant of any orthogonal matrix is either 1 or 0.
Problem Show that any k vectors that are mutually orthogonal are linearly independent.
Problem **Johnson, Riess & Arnold** [1998], page 286, #29, 30.
29. Let **u** be a vector in $\mathbb{R}^n$ such that $\mathbf{u}^{\mathrm{T}}\mathbf{u} = 1$. Let $\mathbf{Q} = \mathbf{I} - 2\mathbf{u}\mathbf{u}^{\mathrm{T}}$. Show that **Q** is an

orthogonal matrix. Also, calculate the vector $\mathbf{Qu}$. Is $\mathbf{u}$ an eigenvector for $\mathbf{Q}$?

30. Suppose that $\mathbf{A}$ and $\mathbf{B}$ are orthogonal $n \times n$ matrices. Show that $\mathbf{AB}$ is an orthogonal matrix.