INTRODUCTION TO LINEAR ALGEBRA

Hiroshi SUZUKI*
Division of Natural Sciences
International Christian University

October 21, 2010

11 Eigenvalues and Eigenvectors

Definition 11.1 Let A be an $n \times n$ matrix.

1. A nonzero vector $\boldsymbol{x} \in \boldsymbol{R}^n$ is called an *eigenvector* of A if $A\boldsymbol{x}$ is a scalar multiple of $\boldsymbol{x}$, that is if

$$A\boldsymbol{x} = \lambda \boldsymbol{x}$$

for some scalar λ. The scalar λ is called an *eigenvalue* of A, and $\boldsymbol{x}$ is said to be an eigenvector of A corresponding to it.

2. $p(x) = \det(xI - A) = x^n + c_1 x^{n-1} + c_2 x^{n-2} + \cdots + c_{n-1}x + c_n$ is called the *characteristic polynomial* of A, and $\det(xI - A) = 0$ is called the *characteristic equation* of A.

Example 11.1

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

$$\boldsymbol{u} = \begin{bmatrix} 1 \\ 6 \\ 8 \end{bmatrix}, \quad \boldsymbol{v} = \begin{bmatrix} 1 \\ -3 \\ 2 \end{bmatrix}, \quad \boldsymbol{w} = \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix}$$

Then

$$A\boldsymbol{u} = \begin{bmatrix} 0 & 1 & 0 \\ 6 & 1 & 3 \\ 0 & 4 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 6 \\ 8 \end{bmatrix} = \begin{bmatrix} 6 \\ 36 \\ 48 \end{bmatrix} = 6 \cdot \begin{bmatrix} 1 \\ 6 \\ 8 \end{bmatrix}$$

$$\begin{aligned} & \det(xI - A) \\ = & \begin{vmatrix} x & -1 & 0 \\ -6 & x-1 & -3 \\ 0 & -4 & x-3 \end{vmatrix} \\ = & \; x((x-1)(x-3) - 12) - (-1)(-6)(x-3) \\ = & \; x^3 - 4x^2 - 15x + 18 \\ = & \; (x-6)(x+3)(x-1). \end{aligned}$$

Theorem 11.1 (Theorem 7.1.2) *If A is an $n \times n$ matrix and λ is a real number, then the following are equivalent.*

(a) *λ is an eigenvalue of A.*

(b) *The system of equations $(\lambda I - A)\boldsymbol{x} = \boldsymbol{0}$ has nontrivial solution.*

(c) *There is a nonzero vector $\boldsymbol{x} \in \boldsymbol{R}^n$ such that $A\boldsymbol{x} = \lambda\boldsymbol{x}$.*

(d) *λ is a solution of the characteristic equation $\det(xI - A) = 0$.*

Example 11.2 Find nontrivial solutions of $(6I - A)\boldsymbol{x} = \boldsymbol{0}$, $((-3)I - A)\boldsymbol{x} = \boldsymbol{0}$ and $(I - A)\boldsymbol{x} = \boldsymbol{0}$.

12 Applications of Eigenvalues and Diagonalization

Review Let A be an $n \times n$ matrix.

1. A nonzero vector $\boldsymbol{x} \in \boldsymbol{R}^n$ is called an *eigenvector* of A if $A\boldsymbol{x}$ is a scalar multiple of $\boldsymbol{x}$, that is if

$$A\boldsymbol{x} = \lambda \boldsymbol{x}$$

for some scalar λ. The scalar λ is called an *eigenvalue* of A, and $\boldsymbol{x}$ is said to be an eigenvector of A corresponding to it.

2. $p(x) = \det(xI - A) = x^n + c_1 x^{n-1} + c_2 x^{n-2} + \cdots + c_{n-1}x + c_n$ is called the *characteristic polynomial* of A, and $\det(xI - A) = 0$ is called the *characteristic equation* of A.

3. Let λ be a real number. Then the following are equivalent.

(a) λ is an eigenvalue of A.

(b) The system of equations $(\lambda I - A)\boldsymbol{v} = \boldsymbol{0}$ has nontrivial solution.

(c) There is a nonzero vector $\boldsymbol{v} \in \boldsymbol{R}^n$ such that $A\boldsymbol{v} = \lambda\boldsymbol{v}$.

(d) λ is a solution of the characteristic equation $\det(xI - A) = 0$.

Example 12.1 [Theorem 7.1.1] If A is an $n \times n$ triangular matrix, then the egenvalues of A are the entries on the main diagonal of A.

*E-mail:hsuzuki@icu.ac.jp

Definition 12.1 A square matrix A is called *diagonalizable* if there is an invertible matrix P such that $P^{-1}AP$ is a diagonal matrix; the matrix P is said to diagonalize A.

Example 12.2

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

$$\boldsymbol{u} = \begin{bmatrix} 1 \\ 6 \\ 8 \end{bmatrix}, \quad \boldsymbol{v} = \begin{bmatrix} 1 \\ -3 \\ 2 \end{bmatrix}, \quad \boldsymbol{w} = \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix}$$

$$\det(xI - A) = (x-6)(x+3)(x-1),$$

$$A\boldsymbol{u} = 6\boldsymbol{u},\ A\boldsymbol{v} = -3\boldsymbol{v},\ A\boldsymbol{w} = \boldsymbol{w}.$$

$$\begin{aligned} AT &= \begin{bmatrix} 0 & 1 & 0 \\ 6 & 1 & 3 \\ 0 & 4 & 3 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 6 & -3 & 1 \\ 8 & 2 & -2 \end{bmatrix} \\ &= \begin{bmatrix} 6 & -3 & 1 \\ 36 & 9 & 1 \\ 48 & -6 & -2 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 1 & 1 \\ 6 & -3 & 1 \\ 8 & 2 & -2 \end{bmatrix} \begin{bmatrix} 6 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & 1 \end{bmatrix} \\ &= TD \end{aligned}$$

$$T^{-1}AT = D = \begin{bmatrix} 6 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \text{ and } A = TDT^{-1}.$$

Proposition 12.1 (Theorems 7.1.3, 7.1.4) *Let A be an $n \times n$ matrix and P an invertible matrix. Suppose $A\boldsymbol{v} = \lambda\boldsymbol{v}$ for some nonzero $\boldsymbol{v}$. Then the following hold.*

(i) *$A^k\boldsymbol{v} = \lambda^k\boldsymbol{v}$ for any positive integer k and $\boldsymbol{v}$ is a eivenvector of A^k corresponding to an eigenvalue λ^k.*

(ii) *A is invertible if and only if 0 is an eigenvalue.*

(iii) $\det(xI - A) = \det(xI - P^{-1}AP)$.

Theorem 12.2 (Theorem 7.2.1) *If A is an $n \times n$ matrix, then the following are equivalent.*

(a) *A is diagonalizable.*

(b) *There are n eigenvectors $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n \in \boldsymbol{R}^n$ such that $P = [\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n]$ is invertible.*

Proof. (a)⇒(b): Suppose A is diagonalizable. Then there is an invertible matrix $P = [\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n]$ such that $D = P^{-1}AP$ is a diagonal matrix, where $\boldsymbol{v}_i$ is the ith column of P. Let $D = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$, where λ_i is the ith diagonal entry. Then

$$\begin{aligned} &[A\boldsymbol{v}_1, A\boldsymbol{v}_2, \ldots, A\boldsymbol{v}_n] \\ &= A[\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n] \\ &= AP \\ &= PD \\ &= [\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n]\mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n) \\ &= [\lambda_1\boldsymbol{v}_1, \lambda_2\boldsymbol{v}_2, \ldots, \lambda_n\boldsymbol{v}_n] \end{aligned}$$

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

(b)⇒(a): Suppose there are n eigenvectors $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n \in \boldsymbol{R}^n$ such that $P = [\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n]$ is invertible. Let $D = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$. Then

$$\begin{aligned} AP &= A[\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n] \\ &= [A\boldsymbol{v}_1, A\boldsymbol{v}_2, \ldots, A\boldsymbol{v}_n] \\ &= [\lambda_1\boldsymbol{v}_1, \lambda_2\boldsymbol{v}_2, \ldots, \lambda_n\boldsymbol{v}_n] \\ &= [\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n]\mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n) \\ &= PD \end{aligned}$$

Since P is invertible, $P^{-1}AP = D$. ■

Theorem 12.3 *If an $n \times n$ matrix A has n distinct eigenvalues, then A is diagonalizable.*

Proof. Suppose $\lambda_1, \lambda_2, \ldots, \lambda_n$ be n distinct eigenvalues of A. By Theorem 11.1 (7.1.2), there exist nonzero vectors $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n$ such that $A\boldsymbol{v}_1 = \lambda_1\boldsymbol{v}_1,\ A\boldsymbol{v}_2 = \lambda_2\boldsymbol{v}_2, \ldots, A\boldsymbol{v}_n = \lambda_n\boldsymbol{v}_n$. Let $P = [\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n]$ and $D = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$. Then $AP = PD$ by the proof of Theorem 12.2. It remains to show that P is invertible.

Suppose $P\boldsymbol{x} = \boldsymbol{0}$. We show that $\boldsymbol{x} = \boldsymbol{0}$. Since $AP = PD$, $A^iP = PD^i$. Hence

$$\begin{aligned} O &= [P\boldsymbol{x}, AP\boldsymbol{x}, \ldots, A^{n-1}P\boldsymbol{x}] \\ &= [P\boldsymbol{x}, PD\boldsymbol{x}, \ldots, PD^{n-1}\boldsymbol{x}] \\ &= P[\boldsymbol{x}, D\boldsymbol{x}, \ldots, D^{n-1}\boldsymbol{x}] = PXV \\ &= [x_1\boldsymbol{v}_1, x_2\boldsymbol{v}_2, \ldots, x_n\boldsymbol{v}_n]V. \end{aligned}$$

Therefore $x_i\boldsymbol{v}_i = \boldsymbol{0}$ for all i and $\boldsymbol{x} = \boldsymbol{0}$. ■

Example 12.3 Let

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

Then

$$\det(xI - A) = \begin{vmatrix} x & -1 & 0 & -1 \\ -1 & x & -1 & 0 \\ 0 & -1 & x & -1 \\ -1 & 0 & -1 & x \end{vmatrix} = x^2(x-2)(x+2)$$

$$\boldsymbol{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \boldsymbol{v}_2 = \begin{bmatrix} 1 \\ 1 \\ -1 \\ -1 \end{bmatrix}, \boldsymbol{v}_3 = \begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \end{bmatrix}, \boldsymbol{v}_4 = \begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \end{bmatrix},$$

$$P = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \\ 1 & -1 & 1 & -1 \end{bmatrix}$$

$$\begin{aligned} AP &= \begin{bmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \\ 1 & -1 & 1 & -1 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \\ 1 & -1 & 1 & -1 \end{bmatrix} \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -2 \end{bmatrix} \\ &= PD \end{aligned}$$

Extra Results without a Proof

Theorem 12.4 (Hamilton-Cayley) *Let A be an $n \times n$ matrix and $p(x) = \det(xI - A)$ is the characteristic polynomial of A. Then $p(A) = O$.*

Proof. The proof is complicated. So we prove only when A is diagonalizable. If $A = D$ is a diagonal matrix, this is obvious. For the general case, suppose $P^{-1}AP = D$. Then $A = PDP^{-1}$ and $p(A) = Pp(D)P^{-1}$. By Proposition 12.1, the characteristic of D is equal to $p(x)$. Now clearly $p(D) = O$. ■

Theorem 12.5 (Theorem 7.3.1) *Let A be an $n \times n$ matrix. Then the following are equivalent.*

(i) *There is a matrix P such that $P^{-1} = P^T$ (orthogonal) and P^TAP is diagonal.*

(ii) *There are n eigenvectors $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n$ such that $\boldsymbol{v}_i \cdot \boldsymbol{v}_j = \delta_{i,j}$. (orthonormal)*

(iii) $A = A^T$.

Proof. (ii)⇒ (i) and (i) ⇒ (iii) are easy. ■