Example 2.9b. To illustrate Theorem 2.9c, let

$$\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \quad \text{and} \quad \mathbf{B} = \begin{pmatrix} 3 & -2 \\ 1 & 2 \end{pmatrix}.$$

Then

$$\mathbf{AB} = \begin{pmatrix} 5 & 2 \\ 13 & 2 \end{pmatrix}, \qquad |\mathbf{AB}| = -16,$$

$$|\mathbf{A}| = -2, \qquad |\mathbf{B}| = 8, \qquad |\mathbf{A}|\,|\mathbf{B}| = -16.$$

2.10 ORTHOGONAL VECTORS AND MATRICES

Two $n \times 1$ vectors **b** and **b** are said to be *orthogonal* if
직교한다.

$$\mathbf{a}'\mathbf{b} = a_1b_1 + a_2b_2 + \cdots + a_nb_n = 0. \tag{2.80}$$

Note that the term *orthogonal* applies to *two* vectors, not to a single vector.

Geometrically, two orthogonal vectors are perpendicular to each other. This is illustrated in Figure 2.3 for the vectors $\mathbf{x}_1 = (4, 2)'$ and $\mathbf{x}_2 = (-1, 2)'$. Note that $\mathbf{x}_1'\mathbf{x}_2 = (4)(-1) + (2)(2) = 0$.

수직
To show that two orthogonal vectors are perpendicular, let θ be the angle between vectors **a** and **b** in Figure 2.4. The vector from the terminal point of **a** to the terminal point of **b** can be represented as $\mathbf{c} = \mathbf{b} - \mathbf{a}$. The law of cosines for the relationship of
b-a : a의 끝점에서 b의 끝점까지의 벡터

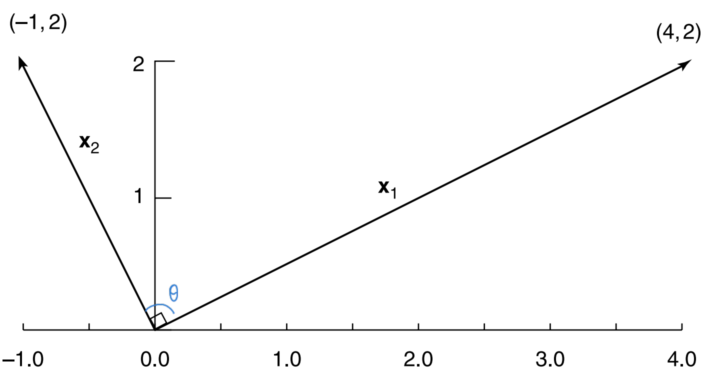


Figure 2.3 Two orthogonal (perpendicular) vectors.

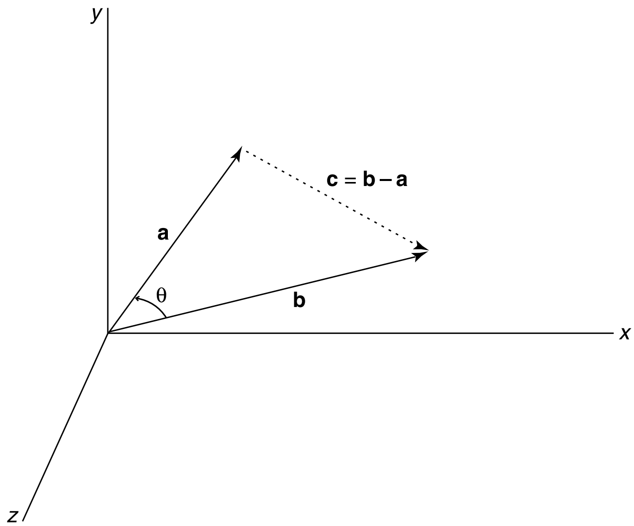
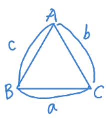


Figure 2.4 Vectors **a** and **b** in 3-space.

3차원공간

θ to the sides of the triangle can be stated in vector form as

코사인 법칙



$\cos A = \frac{b^2+c^2-a^2}{2bc}$

$$\begin{aligned}\cos\theta &= \frac{\mathbf{a}'\mathbf{a} + \mathbf{b}'\mathbf{b} - (\mathbf{b}-\mathbf{a})'(\mathbf{b}-\mathbf{a})}{2\sqrt{(\mathbf{a}'\mathbf{a})(\mathbf{b}'\mathbf{b})}} \\ &= \frac{\mathbf{a}'\mathbf{a} + \mathbf{b}'\mathbf{b} - (\mathbf{b}'\mathbf{b} + \mathbf{a}'\mathbf{a} - 2\mathbf{a}'\mathbf{b})}{2\sqrt{(\mathbf{a}'\mathbf{a})(\mathbf{b}'\mathbf{b})}} \\ &= \frac{\mathbf{a}'\mathbf{b}}{\sqrt{(\mathbf{a}'\mathbf{a})(\mathbf{b}'\mathbf{b})}}. \end{aligned} \tag{2.81}$$

When $\theta = 90°$, $\mathbf{a}'\mathbf{b} = 0$ since $\cos(90°) = 0$. Thus **a** and **b** are *perpendicular* when $\mathbf{a}'\mathbf{b} = 0$.

거변

If $\mathbf{a}'\mathbf{a} = 1$, the vector **a** is said to be *normalized*. A vector **b** can be normalized by dividing by its length, $\sqrt{\mathbf{b}'\mathbf{b}}$. Thus

정규화된 벡터

b도 정규화하려면 그 길이 ($\|b\| = \sqrt{b'b}$)로 나누어야 함.

벡터의 길이가 1이면,

$$\mathbf{c} = \frac{\mathbf{b}}{\sqrt{\mathbf{b}'\mathbf{b}}} \tag{2.82}$$

b의 정규화된 벡터 c

is normalized so that $\mathbf{c}'\mathbf{c} = 1$.

A set of $p \times 1$ vectors $\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_p$ that are normalized ($\mathbf{c}_i'\mathbf{c}_i = 1$ for all i) and mutually orthogonal ($\mathbf{c}_i'\mathbf{c}_j = 0$ for all $i \neq j$) is said to be an *orthonormal set* of vectors. If the $p \times p$ matrix $\mathbf{C} = (\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_p)$ has orthonormal columns, $\mathbf{C}$ is called an *orthogonal* matrix. Since the elements of $\mathbf{C}'\mathbf{C}$ are products of columns of

C [see Theorem 2.2c(i)], an orthogonal matrix **C** has the property

$$\mathbf{C}'\mathbf{C} = \mathbf{I}. \tag{2.83}$$

C의 열 벡터들이 서로
직교정규이기 때문.

It can be shown that an orthogonal matrix **C** also satisfies

$$\mathbf{C}\mathbf{C}' = \mathbf{I}. \tag{2.84}$$

C의 행 벡터들도 서로 직교정규.

Thus an orthogonal matrix **C** has orthonormal rows as well as orthonormal columns. It is also clear from (2.83) and (2.84) that $\mathbf{C}' = \mathbf{C}^{-1}$ if **C** is orthogonal.

Example 2.10. To illustrate an orthogonal matrix, we start with

$$\mathbf{A} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & -2 & 0 \\ 1 & 1 & -1 \end{pmatrix},$$

whose columns are mutually orthogonal but not orthonormal. To normalize the three columns, we divide by their respective lengths, $\sqrt{3}$, $\sqrt{6}$, and $\sqrt{2}$, to obtain the matrix 각 열 벡터의 길이가 1이 되도록 하면

$$\mathbf{C} = \begin{pmatrix} 1/\sqrt{3} & 1/\sqrt{6} & 1/\sqrt{2} \\ 1/\sqrt{3} & -2/\sqrt{6} & 0 \\ 1/\sqrt{3} & 1/\sqrt{6} & -1/\sqrt{2} \end{pmatrix},$$

whose columns are orthonormal. Note that the rows of **C** are also orthonormal, so that **C** satisfies (2.84) as well as (2.83). ∎

벡터를 직교 행렬로 곱하면 / 축이 회전하는 효과

Multiplication of a vector by an orthogonal matrix has the effect of rotating axes; that is, if a point **x** is transformed to $\mathbf{z} = \mathbf{C}\mathbf{x}$, where **C** is orthogonal, then the distance from the origin to **z** is the same as the distance to **x**:

원점 ~z = 원점 ~ x

$$\mathbf{z}'\mathbf{z} = (\mathbf{C}\mathbf{x})'(\mathbf{C}\mathbf{x}) = \mathbf{x}'\mathbf{C}'\mathbf{C}\mathbf{x} = \mathbf{x}'\mathbf{I}\mathbf{x} = \mathbf{x}'\mathbf{x}. \tag{2.85}$$

Hence, the transformation from **x** to **z** is a rotation.

Some properties of orthogonal matrices are given in the following theorem.

Theorem 2.10. If the $p \times p$ matrix **C** is orthogonal and if **A** is any $p \times p$ matrix, then

(i) $|\mathbf{C}| = +1$ or -1.

행렬식

$C'C = I$ 이므로

$|C'C| = |I| = 1$

행렬식의 곱셈 성질

$|C'||C| = |C|^2 = 1.$

$\therefore |C| = \pm 1$

$|C'||A||C| = |C|^2|A| = |A|$

(ii) $|\mathbf{C}'\mathbf{A}\mathbf{C}| = |\mathbf{A}|$.

(iii) $-1 \leq c_{ij} \leq 1$, where c_{ij} is any element of $\mathbf{C}$.

∵ 정규화된 벡터들의 집합

2.11 TRACE

The *trace* of an $n \times n$ matrix $\mathbf{A} = (a_{ij})$ is a scalar function defined as the sum of the diagonal elements of $\mathbf{A}$; that is, $\mathrm{tr}(\mathbf{A}) = \sum_{i=1}^{n} a_{ii}$. For example, suppose

$$\mathbf{A} = \begin{pmatrix} 8 & 4 & 2 \\ 2 & -3 & 6 \\ 3 & 5 & 9 \end{pmatrix}.$$

Then

$$\mathrm{tr}(\mathbf{A}) = 8 - 3 + 9 = 14.$$

Some properties of the trace are given in the following theorem.

Theorem 2.11

(i) If $\mathbf{A}$ and $\mathbf{B}$ are $n \times n$, then

$$\mathrm{tr}(\mathbf{A} \pm \mathbf{B}) = \mathrm{tr}(\mathbf{A}) \pm \mathrm{tr}(\mathbf{B}). \tag{2.86}$$

(ii) If $\mathbf{A}$ is $n \times p$ and $\mathbf{B}$ is $p \times n$, then

proof

$$\mathrm{tr}(\mathbf{AB}) = \mathrm{tr}(\mathbf{BA}). \tag{2.87}$$

Note that in (2.87) n can be less than, equal to, or greater than p.

(iii) If $\mathbf{A}$ is $n \times p$, then

$$\mathrm{tr}(\mathbf{A}'\mathbf{A}) = \sum_{i=1}^{p} \mathbf{a}_i'\mathbf{a}_i, \tag{2.88}$$

$$\begin{pmatrix} a_{1.} \\ a_{2.} \end{pmatrix} (a_{.1}, a_{.2}) = \begin{pmatrix} a_{1.}a_{.1} & a_{1.}a_{.2} \\ a_{2.}a_{.1} & a_{2.}a_{.2} \end{pmatrix}$$

where $\mathbf{a}_i$ is the ith *column* of $\mathbf{A}$.

(iv) If $\mathbf{A}$ is $n \times p$, then

$$\mathrm{tr}(\mathbf{AA}') = \sum_{i=1}^{n} \mathbf{a}_i'\mathbf{a}_i, \tag{2.89}$$

where $\mathbf{a}_i'$ is the ith *row* of $\mathbf{A}$.

(v) If $\mathbf{A} = (a_{ij})$ is an $n \times p$ matrix with representative element a_{ij}, then

$$\mathrm{tr}(\mathbf{A}'\mathbf{A}) = \mathrm{tr}(\mathbf{A}\mathbf{A}') = \sum_{i=1}^{n}\sum_{j=1}^{p} a_{ij}^2. \tag{2.90}$$

(vi) If $\mathbf{A}$ is any $n \times n$ matrix and $\mathbf{P}$ is any $n \times n$ nonsingular matrix, then

$$\mathrm{tr}(\mathbf{P}^{-1}\mathbf{A}\mathbf{P}) = \mathrm{tr}(\mathbf{A}). \tag{2.91}$$

(vii) If $\mathbf{A}$ is any $n \times n$ matrix and $\mathbf{C}$ is any $n \times n$ orthogonal matrix, then

$$\mathrm{tr}(\mathbf{C}'\mathbf{A}\mathbf{C}) = \mathrm{tr}(\mathbf{A}). \tag{2.92}$$

(viii) If $\mathbf{A}$ is $n \times p$ of rank r and $\mathbf{A}^-$ is a generalized inverse of $\mathbf{A}$, then

$$\mathrm{tr}(\mathbf{A}^-\mathbf{A}) = \mathrm{tr}(\mathbf{A}\mathbf{A}^-) = r. \tag{2.93}$$

PROOF. We prove parts (ii), (iii), and (vi).

(ii) By (2.13), the ith diagonal element of $\mathbf{E} = \mathbf{AB}$ is $e_{ii} = \sum_k a_{ik}b_{ki}$. Then

$$\mathrm{tr}(\mathbf{AB}) = \mathrm{tr}(\mathbf{E}) = \sum_i e_{ii} = \sum_i\sum_k a_{ik}b_{ki}.$$

Similarly, the ith diagonal element of $\mathbf{F} = \mathbf{BA}$ is $f_{ii} = \sum_k b_{ik}a_{ki}$, and

$$\mathrm{tr}(\mathbf{BA}) = \mathrm{tr}(\mathbf{F}) = \sum_i f_{ii} = \sum_i\sum_k b_{ik}a_{ki}$$
$$= \sum_k\sum_i a_{ki}b_{ik} = \mathrm{tr}(\mathbf{E}) = \mathrm{tr}(\mathbf{AB}).$$

(iii) By Theorem 2.2c(i), $\mathbf{A}'\mathbf{A}$ is obtained as products of columns of $\mathbf{A}$. If $\mathbf{a}_i$ is the ith column of $\mathbf{A}$, then the ith diagonal element of $\mathbf{A}'\mathbf{A}$ is $\mathbf{a}_i'\mathbf{a}_i$.

(vi) By (2.87) we obtain

tr(AB) = tr(BA)

$$\mathrm{tr}(\mathbf{P}^{-1}\mathbf{A}\mathbf{P}) = \mathrm{tr}(\mathbf{A}\mathbf{P}\mathbf{P}^{-1}) = \mathrm{tr}(\mathbf{A}).$$

$\square$

Example 2.11. We illustrate parts (ii) and (viii) of Theorem 2.11.

(ii) Let

$$\mathbf{A} = \begin{pmatrix} 1 & 3 \\ 2 & -1 \\ 4 & 6 \end{pmatrix} \quad \text{and} \quad \mathbf{B} = \begin{pmatrix} 3 & -2 & 1 \\ 2 & 4 & 5 \end{pmatrix}.$$

Then

$$\mathbf{AB} = \begin{pmatrix} 9 & 10 & 16 \\ 4 & -8 & -3 \\ 24 & 16 & 34 \end{pmatrix}, \quad \mathbf{BA} = \begin{pmatrix} 3 & 17 \\ 30 & 32 \end{pmatrix},$$

$$\text{tr}(\mathbf{AB}) = 9 - 8 + 34 = 35, \quad \text{tr}(\mathbf{BA}) = 3 + 32 = 35.$$

(viii) Using $\mathbf{A}$ in (2.59) and $\mathbf{A}_1^-$ in (2.60), we obtain

$$\mathbf{A}^-\mathbf{A} = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & \frac{1}{2} \\ 0 & 0 & 0 \end{pmatrix}, \quad \mathbf{AA}^- = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 0 \end{pmatrix},$$

$$\text{tr}(\mathbf{A}^-\mathbf{A}) = 1 + 1 + 0 = 2 = \text{rank}(\mathbf{A}),$$
$$\text{tr}(\mathbf{AA}^-) = 1 + 1 + 0 = 2 = \text{rank}(\mathbf{A}).$$

2.12 EIGENVALUES AND EIGENVECTORS

2.12.1 Definition

For every square matrix $\mathbf{A}$, a scalar λ and a nonzero vector $\mathbf{x}$ can be found such that

$$\mathbf{Ax} = \lambda\mathbf{x}, \tag{2.94}$$

고유값 eigenvalue

고유 벡터 eigenvector

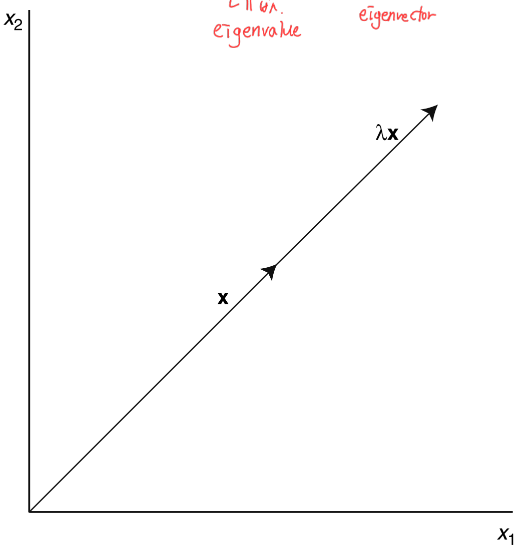


Figure 2.5 An eigenvector $\mathbf{x}$ is transformed to $\lambda\mathbf{x}$.

where λ is an *eigenvalue* of $\mathbf{A}$ and $\mathbf{x}$ is an *eigenvector*. (These terms are sometimes referred to as characteristic root and characteristic vector, respectively.) Note that in (2.94), the vector $\mathbf{x}$ is transformed by $\mathbf{A}$ onto a multiple of itself, so that the point $\mathbf{Ax}$ is on the line passing through $\mathbf{x}$ and the origin. This is illustrated in Figure 2.5.

To find λ and $\mathbf{x}$ for a matrix $\mathbf{A}$, we write (2.94) as

$$(\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = \mathbf{0}. \tag{2.95}$$

By (2.37), $(\mathbf{A} - \lambda\mathbf{I})\mathbf{x}$ is a linear combination of the columns of $\mathbf{A} - \lambda\mathbf{I}$, and by (2.40) and (2.95), these columns are linearly dependent. Thus the square matrix $(\mathbf{A} - \lambda\mathbf{I})$ is singular, and by Theorem 2.9a(iii), we can solve for λ using

If A is singular,
$|A|=0$

$$|\mathbf{A} - \lambda\mathbf{I}| = 0, \tag{2.96}$$

특성 방정식.

which is known as the *characteristic equation*.

If $\mathbf{A}$ is $n \times n$, the characteristic equation (2.96) will have n roots; that is, $\mathbf{A}$ will have n eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$. The λ's will not necessarily all be distinct, or all nonzero, or even all real. (However, the eigenvalues of a symmetric matrix are real; see Theorem 2.12c.) After finding $\lambda_1, \lambda_2, \ldots, \lambda_n$ using (2.96), the accompanying eigenvectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ can be found using (2.95). 실수,,

If an eigenvalue is 0, the corresponding eigenvector is not $\mathbf{0}$. To see this, note that if $\lambda = 0$, then $(\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = \mathbf{0}$ becomes $\mathbf{Ax} = \mathbf{0}$, which has solutions for $\mathbf{x}$ because $\mathbf{A}$ is singular, and the columns are therefore linearly dependent. [The matrix $\mathbf{A}$ is singular because it has a zero eigenvalue; see (63) and (2.107).]

If we multiply both sides of (2.95) by a scalar k, we obtain

$$k(\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = k\mathbf{0} = \mathbf{0},$$

which can be rewritten as

$$(\mathbf{A} - \lambda\mathbf{I})k\mathbf{x} = \mathbf{0} \qquad \text{[by (2.12)]}.$$

Thus if $\mathbf{x}$ is an eigenvector of $\mathbf{A}$, $k\mathbf{x}$ is also an eigenvector. Eigenvectors are therefore unique only up to multiplication by a scalar. (There are many solution vectors $\mathbf{x}$ because $\mathbf{A} - \lambda\mathbf{I}$ is singular; see Section 2.8) Hence, the length of $\mathbf{x}$ is arbitrary, but its direction from the origin is unique; that is, the relative values of (ratios of) the elements of $\mathbf{x} = (x_1, x_2, \ldots, x_n)'$ are unique. Typically, an eigenvector $\mathbf{x}$ is scaled to normalized form as in (2.82), $\mathbf{x}'\mathbf{x} = 1$.

일반적으로 고유벡터는 정규화함.

Example 2.12.1. To illustrate eigenvalues and eigenvectors, consider the matrix

$$\mathbf{A} = \begin{pmatrix} 1 & 2 \\ -1 & 4 \end{pmatrix}.$$

By (2.96), the characteristic equation is

$$|\mathbf{A} - \lambda\mathbf{I}| = \begin{vmatrix} 1-\lambda & 2 \\ -1 & 4-\lambda \end{vmatrix} = (1-\lambda)(4-\lambda) + 2 = 0,$$

which becomes

$$\lambda^2 - 5\lambda + 6 = (\lambda - 3)(\lambda - 2) = 0,$$

with roots $\lambda_1 = 3$ and $\lambda_2 = 2$.

To find the eigenvector $\mathbf{x}_1$ corresponding to $\lambda_1 = 3$, we use (2.95)

$$(\mathbf{A} - \lambda_1\mathbf{I})\mathbf{x}_1 = \mathbf{0},$$
$$\begin{pmatrix} 1-3 & 2 \\ -1 & 4-3 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

which can be written as

$$\begin{aligned} -2x_1 + 2x_2 &= 0 \\ -x_1 + x_2 &= 0. \end{aligned}$$

The second equation is a multiple of the first, and either equation yields $x_1 = x_2$. The solution vector can be written with $x_1 = x_2 = c$ as an arbitrary constant:

$$\mathbf{x}_1 = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_1 \end{pmatrix} = x_1\begin{pmatrix} 1 \\ 1 \end{pmatrix} = c\begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

If c is set equal to $1/\sqrt{2}$ to normalize the eigenvector, we obtain

$$\mathbf{x}_1 = \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}.$$

Similarly, corresponding to $\lambda_2 = 2$, we obtain

$$\mathbf{x}_2 = \begin{pmatrix} 2/\sqrt{5} \\ 1/\sqrt{5} \end{pmatrix}.$$

2.12.2 Functions of a Matrix

If λ is an eigenvalue of $\mathbf{A}$ with corresponding eigenvector $\mathbf{x}$, then for certain functions $g(\mathbf{A})$, an eigenvalue is given by $g(\lambda)$ and $\mathbf{x}$ is the corresponding eigenvector of $g(\mathbf{A})$ as well as of $\mathbf{A}$. We illustrate some of these cases:

$g(\lambda)$ $g(A)$

1. If λ is an eigenvalue of $\mathbf{A}$, then $c\lambda$ is an eigenvalue of $c\mathbf{A}$, where c is an arbitrary constant such that $c \neq 0$. This is easily demonstrated by multiplying the defining relationship $\mathbf{Ax} = \lambda\mathbf{x}$ by c:

$$c\mathbf{Ax} = c\lambda\mathbf{x}. \tag{2.97}$$

   Note that $\mathbf{x}$ is an eigenvector of $\mathbf{A}$ corresponding to λ, and $\mathbf{x}$ is also an eigenvector of $c\mathbf{A}$ corresponding to $c\lambda$.

2. If λ is an eigenvalue of the $\mathbf{A}$ and $\mathbf{x}$ is the corresponding eigenvector of $\mathbf{A}$, then $c\lambda + k$ is an eigenvalue of the matrix $c\mathbf{A} + k\mathbf{I}$ and $\mathbf{x}$ is an eigenvector of $c\mathbf{A} + k\mathbf{I}$, where c and k are scalars. To show this, we add $k\mathbf{x}$ to (2.97):

$$\begin{aligned} c\mathbf{Ax} + k\mathbf{x} &= c\lambda\mathbf{x} + k\mathbf{x}, \\ (c\mathbf{A} + k\mathbf{I})\mathbf{x} &= (c\lambda + k)\mathbf{x}. \end{aligned} \tag{2.98}$$

   Thus $c\lambda + k$ is an eigenvalue of $c\mathbf{A} + k\mathbf{I}$ and $\mathbf{x}$ is the corresponding eigenvector of $c\mathbf{A} + k\mathbf{I}$. Note that (2.98) does not extend to $\mathbf{A} + \mathbf{B}$ for arbitrary $n \times n$ matrices $\mathbf{A}$ and $\mathbf{B}$; that is, $\mathbf{A} + \mathbf{B}$ does not have $\lambda_A + \lambda_B$ for an eigenvalue, where λ_A is an eigenvalue of $\mathbf{A}$ and λ_B is an eigenvalue of $\mathbf{B}$.

3. If λ is an eigenvalue of $\mathbf{A}$, then λ^2 is an eigenvalue of $\mathbf{A}^2$. This can be demonstrated by multiplying the defining relationship $\mathbf{Ax} = \lambda\mathbf{x}$ by $\mathbf{A}$:

$$\begin{aligned} \mathbf{A}(\mathbf{Ax}) &= \mathbf{A}(\lambda\mathbf{x}), \\ \mathbf{A}^2\mathbf{x} &= \lambda\mathbf{Ax} = \lambda(\lambda\mathbf{x}) = \lambda^2\mathbf{x}. \end{aligned} \tag{2.99}$$

   Thus λ^2 is an eigenvalue of $\mathbf{A}^2$, and $\mathbf{x}$ is the corresponding eigenvector of $\mathbf{A}^2$. This can be extended to any power of $\mathbf{A}$:

$$\mathbf{A}^k\mathbf{x} = \lambda^k\mathbf{x}; \tag{2.100}$$

   that is, λ^k is an eigenvalue of $\mathbf{A}^k$, and $\mathbf{x}$ is the corresponding eigenvector.

4. If λ is an eigenvalue of the nonsingular matrix $\mathbf{A}$, then $1/\lambda$ is an eigenvalue of $\mathbf{A}^{-1}$. To demonstrate this, we multiply $\mathbf{Ax} = \lambda\mathbf{x}$ by $\mathbf{A}^{-1}$ to obtain

$$\mathbf{A}^{-1}\mathbf{Ax} = \mathbf{A}^{-1}\lambda\mathbf{x},$$

$$\mathbf{x} = \lambda\mathbf{A}^{-1}\mathbf{x},$$

$$\mathbf{A}^{-1}\mathbf{x} = \frac{1}{\lambda}\mathbf{x}. \tag{2.101}$$

Thus $1/\lambda$ is an eigenvalue of $\mathbf{A}^{-1}$, and $\mathbf{x}$ is an eigenvector of both $\mathbf{A}$ and $\mathbf{A}^{-1}$.

5. The results in (2.97) and (2.100) can be used to obtain eigenvalues and eigenvectors of a polynomial in $\mathbf{A}$. For example, if λ is an eigenvalue of $\mathbf{A}$, then

$$\begin{aligned}(\mathbf{A}^3 + 4\mathbf{A}^2 - 3\mathbf{A} + 5\mathbf{I})\mathbf{x} &= \mathbf{A}^3\mathbf{x} + 4\mathbf{A}^2\mathbf{x} - 3\mathbf{Ax} + 5\mathbf{x} \\ &= \lambda^3\mathbf{x} + 4\lambda^2\mathbf{x} - 3\lambda\mathbf{x} + 5\mathbf{x} \\ &= (\lambda^3 + 4\lambda^2 - 3\lambda + 5)\mathbf{x}.\end{aligned}$$

Thus $\lambda^3 + 4\lambda^2 - 3\lambda + 5$ is an eigenvalue of $\mathbf{A}^3 + 4\mathbf{A}^2 - 3\mathbf{A} + 5\mathbf{I}$, and $\mathbf{x}$ is the corresponding eigenvector.

무한 급수

For certain matrices, property 5 can be extended to an infinite series. For example, if λ is an eigenvalue of $\mathbf{A}$, then, by (2.98), $1 - \lambda$ is an eigenvalue of $\mathbf{I} - \mathbf{A}$. If $\mathbf{I} - \mathbf{A}$ is nonsingular, then, by (2.101), $1/(1 - \lambda)$ is an eigenvalue of $(\mathbf{I} - \mathbf{A})^{-1}$. If $-1 < \lambda < 1$, then $1/(1 - \lambda)$ can be represented by the series

$$\frac{1}{1 - \lambda} = 1 + \lambda + \lambda^2 + \lambda^3 + \cdots.$$

Correspondingly, if all eigenvalues of $\mathbf{A}$ satisfy $-1 < \lambda < 1$, then

$$(\mathbf{I} - \mathbf{A})^{-1} = \mathbf{I} + \mathbf{A} + \mathbf{A}^2 + \mathbf{A}^3 + \cdots. \tag{2.102}$$

2.12.3 Products

It was noted in a comment following (2.98) that the eigenvalues of $\mathbf{A} + \mathbf{B}$ are not of the form $\lambda_A + \lambda_B$, where λ_A is an eigenvalue of $\mathbf{A}$ and λ_B is an eigenvalue of $\mathbf{B}$. Similarly, the eigenvalues of $\mathbf{AB}$ are not products of the form $\lambda_A\lambda_B$. However, the eigenvalues of $\mathbf{AB}$ are the same as those of $\mathbf{BA}$.

Theorem 2.12a. If $\mathbf{A}$ and $\mathbf{B}$ are $n \times n$ or if $\mathbf{A}$ is $n \times p$ and $\mathbf{B}$ is $p \times n$, then the (nonzero) eigenvalues of $\mathbf{AB}$ are the same as those of $\mathbf{BA}$. If $\mathbf{x}$ is an eigenvector of $\mathbf{AB}$, then $\mathbf{Bx}$ is an eigenvector of $\mathbf{BA}$. ∎

$AB x = \lambda x$

$B(ABx) = \lambda Bx$

$(BA)(Bx) = \lambda(Bx)$

Two additional results involving eigenvalues of products are given in the following theorem.

Theorem 2.12b. Let **A** be any $n \times n$ matrix.

$Ax = \lambda x$. 새로운 벡터 y 정의. $y = Px$. $\Rightarrow x = P^{-1}y$ 가 된다.
$A(P^{-1}y) = \lambda(P^{-1}y)$ 양변에 P를 곱하면
$PAP^{-1}y = \lambda y$.

(i) If **P** is any $n \times n$ nonsingular matrix, then **A** and $\mathbf{P}^{-1}\mathbf{AP}$ have the same eigenvalues.

(ii) If **C** is any $n \times n$ orthogonal matrix, then **A** and $\mathbf{C}'\mathbf{AC}$ have the same eigenvalues. $C'C = I$, $C^{-1} = C'$

$Ax = \lambda x$. 새로운 벡터 $y = Cx$ 정의. $x = C'y$ 가 된다.
$A(C'y) = \lambda(C'y)$ 양변에 C를 곱하면 $(C'AC)y = \lambda y$

2.12.4 Symmetric Matrices

Two properties of the eigenvalues and eigenvectors of a symmetric matrix are given in the following theorem.

Theorem 2.12c. Let **A** be an $n \times n$ symmetric matrix. $A' = A$

(i) The eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ of **A** are real. 실수.

(ii) The eigenvectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k$ of **A** corresponding to distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_k$ are mutually orthogonal; the eigenvectors $\mathbf{x}_{k+1}, \mathbf{x}_{k+2}, \ldots, \mathbf{x}_n$ corresponding to the nondistinct eigenvalues can be chosen to be mutually orthogonal to each other and to the other eigenvectors; that is, $\mathbf{x}_i'\mathbf{x}_j = 0$ for $i \neq j$. ■

중복되는 고유값.

If the eigenvectors of a symmetric matrix **A** are normalized and placed as columns of a matrix **C**, then by Theorem 2.12c(ii), **C** is an orthogonal matrix. This orthogonal matrix can be used to express **A** in terms of its eigenvalues and eigenvectors.

Theorem 2.12d. If **A** is an $n \times n$ symmetric matrix with eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ and normalized eigenvectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$, then **A** can be expressed as

$$\mathbf{A} = \mathbf{CDC}' \tag{2.103}$$

$$= \sum_{i=1}^{n} \lambda_i \mathbf{x}_i \mathbf{x}_i', \tag{2.104}$$

where $\mathbf{D} = \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$ and **C** is the orthogonal matrix $\mathbf{C} = (\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n)$. The result in either (2.103) or (2.104) is often called the *spectral decomposition* of **A**.
스펙트럴 분해

Proof. By Theorem 2.12c(ii), **C** is orthogonal. Then by (2.84), $\mathbf{I} = \mathbf{CC}'$, and multiplication by **A** gives

$$\mathbf{A} = \mathbf{A}\underbrace{\mathbf{CC}'}_{I}.$$

We now substitute $\mathbf{C} = (\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n)$ to obtain

$$\begin{aligned}
\mathbf{A} &= \mathbf{A}(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n)\mathbf{C}' \\
&= (\mathbf{A}\mathbf{x}_1, \mathbf{A}\mathbf{x}_2, \ldots, \mathbf{A}\mathbf{x}_n)\mathbf{C}' \qquad \text{[by (2.28)]} \\
&= (\lambda_1\mathbf{x}_1, \lambda_2\mathbf{x}_2, \ldots, \lambda_n\mathbf{x}_n)\mathbf{C}' \qquad \text{[by (2.94)]} \\
&= \mathbf{CDC}', \qquad (2.105)
\end{aligned}$$

since multiplication on the right by $\mathbf{D} = \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$ multiplies columns of $\mathbf{C}$ by elements of $\mathbf{D}$ [see (2.30)]. Now writing $\mathbf{C}'$ in the form

$$\mathbf{C}' = (\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n)' = \begin{pmatrix} \mathbf{x}_1' \\ \mathbf{x}_2' \\ \vdots \\ \mathbf{x}_n' \end{pmatrix} \qquad \text{[by (2.39)]},$$

(2.105) becomes

$$\begin{aligned}
\mathbf{A} &= \underbrace{(\lambda_1\mathbf{x}_1, \lambda_2\mathbf{x}_2, \ldots, \lambda_n\mathbf{x}_n)}_{CD} \overbrace{\begin{pmatrix} \mathbf{x}_1' \\ \mathbf{x}_2' \\ \vdots \\ \mathbf{x}_n' \end{pmatrix}}^{C'} \\
&= \lambda_1\mathbf{x}_1\mathbf{x}_1' + \lambda_2\mathbf{x}_2\mathbf{x}_2' + \cdots + \lambda_n\mathbf{x}_n\mathbf{x}_n'.
\end{aligned}$$

∎

Corollary 1. If $\mathbf{A}$ is symmetric and $\mathbf{C}$ and $\mathbf{D}$ are defined as in Theorem 2.12d, then $\mathbf{C}$ *diagonalizes* $\mathbf{A}$:

대각화

$$\mathbf{C}'\mathbf{AC} = \mathbf{D}. \qquad (2.106)$$

∎

We can express the determinant and trace of a square matrix $\mathbf{A}$ in terms of its eigenvalues. → 이를 통해 A의 고유값으로 |A|와 tr(A)를 표현 가능.

Theorem 2.12e. If $\mathbf{A}$ is any $n \times n$ matrix with eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$, then

$A = CDC'$

$|A| = |CDC'|$

$= |C|\,|D|\,|C'| = |D|$

orthogonal

$|\text{diag}(\lambda_1, \ldots, \lambda_n)|$

(i)
$$|\mathbf{A}| = \prod_{i=1}^{n} \lambda_i. \qquad (2.107)$$

(ii)
$$\text{tr}(\mathbf{A}) = \sum_{i=1}^{n} \lambda_i. \qquad (2.108)$$

$tr(A) = tr(CDC') = tr(D)$

$tr(\text{diag}(\lambda_1, \ldots, \lambda_n))$

∎

We have included Theorem 2.12e here because it is easy to prove for a symmetric matrix $\mathbf{A}$ using Theorem 2.12d (see Problem 2.72). However, the theorem is true for any square matrix (Searle 1982, p. 278).

Example 2.12.4. To illustrate Theorem 2.12e, consider the matrix $\mathbf{A}$ in Example 2.12.1

$Ax = \lambda x,$
$(A - \lambda I) = 0$

$$\mathbf{A} = \begin{pmatrix} 1 & 2 \\ -1 & 4 \end{pmatrix},$$

$\begin{vmatrix} 1-\lambda & 2 \\ -1 & 4-\lambda \end{vmatrix} = 0$,
$4 - 5\lambda + \lambda^2 + 2 = 0$
$\lambda^2 - 5\lambda + 6 = 0$
$(\lambda - 2)(\lambda - 3) = 0$

which has eigenvalues $\lambda_1 = 3$ and $\lambda_2 = 2$. The product $\lambda_1\lambda_2 = 6$ is the same as $|\mathbf{A}| = 4 - (-1)(2) = 6$. The sum $\lambda_1 + \lambda_2 = 3 + 2 = 5$ is the same as $\mathrm{tr}(\mathbf{A}) = 1 + 4 = 5$. □

2.12.5 Positive Definite and Semidefinite Matrices

The eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ of positive definite and positive semidefinite matrices (Section 2.6) are positive and nonnegative, respectively.
양수 (>0) ↳음수가 아님 (≥ 0)

Theorem 2.12f. Let $\mathbf{A}$ be $n \times n$ with eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$.

(i) If $\mathbf{A}$ is positive definite, then $\lambda_i > 0$ for $i = 1, 2, \ldots, n$.
(ii) If $\mathbf{A}$ is positive semidefinite, then $\lambda_i \geq 0$ for $i = 1, 2, \ldots, n$. The number of eigenvalues λ_i for which $\lambda_i > 0$ is the rank of $\mathbf{A}$.
양수인 고유값의 개수 = rank(A)

Proof.

(i) For any λ_i, we have $\mathbf{Ax}_i = \lambda_i\mathbf{x}_i$. Multiplying by $\mathbf{x}_i'$, we obtain

$$\mathbf{x}_i'\mathbf{Ax}_i = \lambda_i\mathbf{x}_i'\mathbf{x}_i,$$

$$\lambda_i = \frac{\mathbf{x}_i'\mathbf{Ax}_i}{\mathbf{x}_i'\mathbf{x}_i} > 0.$$

양의 정부호라 항상 양수.
거리라서 항상 양수.

In the second expression, $\mathbf{x}_i'\mathbf{Ax}_i$ is positive because $\mathbf{A}$ is positive definite, and $\mathbf{x}_i'\mathbf{x}_i$ is positive because $\mathbf{x}_i \neq \mathbf{0}$. □

$\lambda_i > 0$ 이므로 $\sqrt{\lambda_i}$ 가 항상 존재.

If a matrix $\mathbf{A}$ is positive definite, we can find a *square root matrix* $\mathbf{A}^{1/2}$ as follows. Since the eigenvalues of $\mathbf{A}$ are positive, we can substitute the square roots $\sqrt{\lambda_i}$ for λ_i in the spectral decomposition of $\mathbf{A}$ in (2.103), to obtain

$$\mathbf{A}^{1/2} = \mathbf{CD}^{1/2}\mathbf{C}', \tag{2.109}$$

where $\mathbf{D}^{1/2} = \mathrm{diag}(\sqrt{\lambda_1}, \sqrt{\lambda_2}, \ldots, \sqrt{\lambda_n})$. The matrix $\mathbf{A}^{1/2}$ is symmetric and has the property

$$\mathbf{A}^{1/2}\mathbf{A}^{1/2} = (\mathbf{A}^{1/2})^2 = \mathbf{A}. \tag{2.110}$$