A. Review of Matrix Theory

A.1. Matrix Notation and Operations

A.1.1. A. Definitions:

1. An $m \times n$ matrix **A** is a rectangular array of elements having m rows and n columns and is denoted as

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} = [a_{ij}]_{m\times n}$$

(A.1)

When $m = n$, **A** is called a *square matrix of order n*.

2. A $1 \times n$ matrix is called an n-dimensional *row vector*:

$$[a_{11} \quad a_{12} \quad \cdots \quad a_{1n}]$$

(A.2)

An $m \times 1$ matrix is called an m-dimensional *column vector*:

$$\begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix}$$

(A.3)

3. A *zero* matrix **0** is a matrix having all its elements zero.

4. A *diagonal* matrix **D** is a square matrix in which all elements not on the main diagonal are zero:

$$\mathbf{D} = \begin{bmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_n \end{bmatrix}$$

(A.4)

Sometimes the diagonal matrix **D** in Eq. (A.4) is expressed as

$$\mathbf{D} = \mathrm{diag}(d_1 \quad d_2 \quad \cdots \quad d_n)$$

(A.5)

5. The *identity* (or *unit*) matrix **I** is a diagonal matrix with all of its diagonal elements equal to 1.

$$\mathbf{I}=\begin{bmatrix}1 & 0 & \cdots & 0\\ 0 & 1 & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & 1\end{bmatrix}$$

(A.6)

A.1.2. B. Operations:

Let **A** = $[a_{ij}]_{m\times n}$, **B** = $[b_{ij}]_{m\times n}$, and **C** = $[c_{ij}]_{m\times n}$.

a. Equality of Two Matrices:

$$\mathbf{A}=\mathbf{B}\Rightarrow a_{ij}=b_{ij}$$

(A.7)

b. Addition:

$$\mathbf{C}=\mathbf{A}+\mathbf{B}\Rightarrow c_{ij}=a_{ij}+b_{ij}$$

(A.8)

c. Multiplication by a Scalar:

$$\mathbf{B}=\alpha\,\mathbf{A}\Rightarrow b_{ij}=\alpha a_{ij}$$

(A.9)

If $\alpha = -1$, then **B** = −**A** is called the *negative* of **A**.

EXAMPLE A.1 Let

$$\mathbf{A}=\begin{bmatrix}1 & 2 & 3\\ -1 & 0 & 4\end{bmatrix}\qquad \mathbf{B}=\begin{bmatrix}2 & 0 & -1\\ 4 & 1 & -2\end{bmatrix}$$

Then

$$\mathbf{A}+\mathbf{B}=\begin{bmatrix}1+2 & 2+0 & 3-1\\ -1+4 & 0+1 & 4-2\end{bmatrix}=\begin{bmatrix}3 & 2 & 2\\ 3 & 1 & 2\end{bmatrix}$$

$$-\mathbf{B}=(-1)\,\mathbf{B}=\begin{bmatrix}-2 & 0 & 1\\ -4 & -1 & 2\end{bmatrix}$$

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

Notes:

$$\begin{aligned}
&1. \quad \mathbf{A} = \mathbf{B} \text{ and } \mathbf{B} = \mathbf{C} \Rightarrow \mathbf{A} = \mathbf{C} \\
&2. \quad \mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A} \\
&3. \quad (\mathbf{A} + \mathbf{B}) + \mathbf{C} = \mathbf{A} + (\mathbf{B} + \mathbf{C}) \\
&4. \quad \mathbf{A} + \mathbf{0} = \mathbf{0} + \mathbf{A} = \mathbf{A} \\
&5. \quad \mathbf{A} - \mathbf{A} = \mathbf{A} + (-\mathbf{A}) = \mathbf{0} \\
&6. \quad (\alpha + \beta)\mathbf{A} = \alpha\mathbf{A} + \alpha\mathbf{B} \\
&7. \quad \alpha(\mathbf{A} + \mathbf{B}) = \alpha\mathbf{A} + \alpha\mathbf{B} \\
&8. \quad \alpha(\beta\mathbf{A}) = (\alpha\beta)\mathbf{A} = \beta(\alpha\mathbf{A})
\end{aligned}$$

(A.10)

d. Multiplication:

Let **A** = $[a_{ij}]_{m\times n}$, **B** = $[b_{ij}]_{n\times p}$, and **C** = $[c_{ij}]_{m\times p}$.

$$\mathbf{C} = \mathbf{AB} \Rightarrow c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}$$

(A.11)

The matrix product **AB** is defined only when the number of columns of **A** is equal to the number of rows of **B**. In this case **A** and **B** are said to be *conformable*.

EXAMPLE A.2 Let

$$\mathbf{A} = \begin{bmatrix} 0 & -1 \\ 1 & 2 \\ 2 & -3 \end{bmatrix} \qquad \mathbf{B} = \begin{bmatrix} 1 & 2 \\ 3 & -1 \end{bmatrix}$$

Then

$$\mathbf{AB} = \begin{bmatrix} 0 & -1 \\ 1 & 2 \\ 2 & -3 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & -1 \end{bmatrix} = \begin{bmatrix} 0(1) + (-1)3 & 0(2) + (-1)(-1) \\ 1(1) + 2(3) & 1(2) + 2(-1) \\ 2(1) + (-3)3 & 2(2) + (-3)(-1) \end{bmatrix} = \begin{bmatrix} -3 & 1 \\ 7 & 0 \\ -7 & 7 \end{bmatrix}$$

but **BA** is not defined.

Furthermore, even if both **AB** and **BA** are defined, in general

$$\mathbf{AB} \neq \mathbf{BA}$$

(A.12)

EXAMPLE A.3 Let

$$\mathbf{A} = \begin{bmatrix} 0 & -1 \\ 1 & 2 \end{bmatrix} \qquad \mathbf{B} = \begin{bmatrix} 1 & 2 \\ 3 & -1 \end{bmatrix}$$

Then

$$\mathbf{AB} = \begin{bmatrix} 0 & -1 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 3 & -1 \end{bmatrix} = \begin{bmatrix} -3 & 1 \\ 7 & 0 \end{bmatrix}$$

$$\mathbf{BA} = \begin{bmatrix} 1 & 2 \\ 3 & -1 \end{bmatrix}\begin{bmatrix} 0 & -1 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 2 & 3 \\ -1 & -5 \end{bmatrix} \neq \mathbf{AB}$$

An example of the case where **AB** = **BA** follows.

EXAMPLE A.4 Let

$$\mathbf{A} = \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix} \qquad \mathbf{B} = \begin{bmatrix} 2 & 0 \\ 0 & 4 \end{bmatrix}$$

Then

$$\mathbf{AB} = \mathbf{BA} = \begin{bmatrix} 2 & 0 \\ 0 & 12 \end{bmatrix}$$

Notes:

$$\begin{aligned}
&1. \quad \mathbf{A0} = \mathbf{0A} = \mathbf{0} \\
&2. \quad \mathbf{AI} = \mathbf{IA} = \mathbf{A} \\
&3. \quad (\mathbf{A} + \mathbf{B})\mathbf{C} = \mathbf{AC} + \mathbf{BC} \\
&4. \quad \mathbf{A}(\mathbf{B} + \mathbf{C}) = \mathbf{AB} + \mathbf{AC} \\
&5. \quad (\mathbf{AB})\mathbf{C} = \mathbf{A}(\mathbf{BC}) = \mathbf{ABC} \\
&6. \quad \alpha(\mathbf{AB}) = (\alpha\mathbf{A})\mathbf{B} = \mathbf{A}(\alpha\mathbf{B})
\end{aligned}$$

(A.13)

It is important to note that **AB** = **0** does not necessarily imply **A** = **0** or **B** = **0**.

EXAMPLE A.5 Let

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

Then

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

A.2. Transpose and Inverse

A.2.1. A. Transpose:

Let **A** be an $n \times m$ matrix. The *transpose* of **A**, denoted by $\mathbf{A}^T$, is an $m \times n$ matrix formed by interchanging the rows and columns of **A**.

$$\mathbf{B} = \mathbf{A}^T \Rightarrow b_{ij} = a_{ji}$$

(A.14)

EXAMPLE A.6

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ -1 & 0 & 4 \end{bmatrix} \qquad \mathbf{A}^T = \begin{bmatrix} 1 & 2 & 3 \\ -1 & 0 & 4 \end{bmatrix}^T = \begin{bmatrix} 1 & -1 \\ 2 & 0 \\ 3 & 4 \end{bmatrix}$$

If $\mathbf{A}^T = \mathbf{A}$, then **A** is said to be *symmetric*, and if $\mathbf{A}^T = -\mathbf{A}$, then **A** is said to be *skew-symmetric*.

EXAMPLE A.7 Let

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & -1 \\ 3 & -1 & 5 \end{bmatrix} \qquad \mathbf{B} = \begin{bmatrix} 0 & 1 & -2 \\ -1 & 0 & 3 \\ 2 & -3 & 0 \end{bmatrix}$$

Then **A** is a symmetric matrix and **B** is a skew-symmetric matrix.

Note that if a matrix is skew-symmetric, then its diagonal elements are all zero.

Notes:

$$\begin{aligned} &1. \quad (\mathbf{A}^T)^T = \mathbf{A} \\ &2. \quad (\mathbf{A} + \mathbf{B})^T = \mathbf{A}^T + \mathbf{B}^T \\ &3. \quad (\alpha\mathbf{A})^T = \alpha\mathbf{A}^T \\ &4. \quad (\mathbf{AB})^T = \mathbf{B}^T\mathbf{A}^T \end{aligned}$$

(A.15)

A.2.2. B. Inverses:

A matrix **A** is said to be *invertible* if there exists a matrix **B** such that

$$\mathbf{BA} = \mathbf{AB} = \mathbf{I}$$

(A.16a)

The matrix **B** is called the *inverse* of **A** and is denoted by $\mathbf{A}^{-1}$. Thus,

$$\mathbf{A}^{-1}\mathbf{A} = \mathbf{A}\mathbf{A}^{-1} = \mathbf{I}$$

(A.16b)

EXAMPLE A.8

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

Thus,

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

Notes:

$$\begin{aligned}
&1.\quad (\mathbf{A}^{-1})^{-1} = \mathbf{A} \\
&2.\quad (\mathbf{A}^{-1})^{T} = (\mathbf{A}^{T})^{-1} \\
&3.\quad (\alpha\,\mathbf{A})^{-1} = \frac{1}{\alpha}\,\mathbf{A}^{-1} \\
&4.\quad (\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}
\end{aligned}$$

(A.17)

Note that if **A** is invertible, then **AB** = **0** implies that **B** = **0** since

$$\mathbf{A}^{-1}\mathbf{AB} = \mathbf{IB} = \mathbf{B} = \mathbf{A}^{-1}\mathbf{0} = \mathbf{0}$$

A.3. Linear Independence and Rank

A.3.1. A. Linear independence:

Let **A** = $[\mathbf{a}_1\mathbf{a}_2 \cdots \mathbf{a}_n]$, where $\mathbf{a}_i$ denotes the ith column vector of **A**. A set of column vectors $\mathbf{a}_i (i = 1, 2, \ldots, n)$ is said to be *linearly dependent* if there exist numbers α_i ($i = 1, 2, \ldots, n$) not all zero such that

$$\alpha_1\mathbf{a}_1 + \alpha_2\mathbf{a}_2 + \cdots + \alpha_n\mathbf{a}_n = \mathbf{0}$$

(A.18)

If Eq. (A.18) holds only for all $\alpha_i = 0$, then the set is said to be *linearly independent*.

EXAMPLE A.9 Let

$$\mathbf{a}_1 = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} \qquad \mathbf{a}_2 = \begin{bmatrix} 2 \\ 1 \\ -1 \end{bmatrix} \qquad \mathbf{a}_3 = \begin{bmatrix} 4 \\ 5 \\ -3 \end{bmatrix}$$

Since $2\mathbf{a}_1 + (-3)\mathbf{a}_2 + \mathbf{a}_3 = \mathbf{0}$, $\mathbf{a}_1$, $\mathbf{a}_2$, and $\mathbf{a}_3$ are linearly dependent. Let

$$\mathbf{d}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \qquad \mathbf{d}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \qquad \mathbf{d}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

Then

$$\alpha_1 \mathbf{d}_1 + \alpha_2 \mathbf{d}_2 + \alpha_3 \mathbf{d}_3 = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

implies that $\alpha_1 = \alpha_2 = \alpha_3 = 0$. Thus, $\mathbf{d}_1$, $\mathbf{d}_2$, and $\mathbf{d}_3$ are linearly independent.

A.3.2. B. Rank of a Matrix:

The number of linearly independent column vectors in a matrix **A** is called the *column rank* of **A**, and the number of linearly independent row vectors in a matrix **A** is called the *row rank* of **A**. It can be shown that

$$\text{Rank of } \mathbf{A} = \text{column rank of A} = \text{row rank of A}$$

(A.19)

Note:

If the rank of an $N \times N$ matrix **A** is N, then **A** is invertible and $\mathbf{A}^{-1}$ exists.

A.4. Determinants

A.4.1. A. Definitions:

Let $\mathbf{A} = [a_{ij}]$ be a square matrix of order N. We associate with **A** a certain number called its *determinant*, denoted by det **A** or |**A**|. Let $\mathbf{M}_{ij}$ be the square matrix of order $(N - 1)$ obtained from **A** by deleting the ith row and jth column. The number A_{ij} defined by

$$A_{ij} = (-1)^{i+j} |\mathbf{M}_{ij}|$$

(A.20)

is called the *cofactor* of a_{ij}. Then det **A** is obtained by

$$\det \mathrm{A} = |\mathrm{A}| = \sum_{k=1}^{N} a_{ik} A_{ik} \qquad i = 1, 2, \ldots, N$$

(A.21a)

or

$$\det \mathbf{A} = |\mathbf{A}| = \sum_{k=1}^{N} a_{kj} A_{kj} \qquad j = 1, 2, \ldots, N$$

(A.21b)

Equation (A.21a) is known as the *Laplace expansion* of |**A**| along the *i*th row, and Eq. (A.21b) the Laplace expansion of |**A**| along the *j*th column.

EXAMPLE A.10 For a 1 × 1 matrix,

$$\mathbf{A} = [a_{11}] \rightarrow |\mathbf{A}| = a_{11}$$

(A.22)

For a 2 × 2 matrix,

$$\mathrm{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \rightarrow |\mathrm{A}| = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}$$

(A.23)

For a 3 × 3 matrix,

$$\mathrm{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

Using Eqs. (A.21a) and (A.23), we obtain

$$|\mathrm{A}| = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12}\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13}\begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}$$

$$= a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31}$$

(A.24)

A.4.2. B. Determinant Rank of a Matrix:

The *determinant rank* of a matrix **A** is defined as the order of the largest square submatrix **M** of **A** such that det **M** ≠ 0. It can be shown that the rank of **A** is equal to the determinant rank of **A**.

EXAMPLE A.11 Let

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

Note that |**A**| = 0. One of the largest submatrices whose determinant is not equal to zero is

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

Hence the rank of the matrix **A** is 2. (See Example A.9.)

A.4.3. C. Inverse of a Matrix:

Using determinants, the inverse of an $N \times N$ matrix **A** can be computed as

$$\mathbf{A}^{-1} = \frac{1}{\det \mathbf{A}} \operatorname{adj} \mathbf{A}$$

(A.25)

and

$$\operatorname{adj} \mathbf{A} = [A_{ij}]^T = \begin{bmatrix} A_{11} & A_{21} & \cdots & A_{N1} \\ A_{12} & A_{22} & \cdots & A_{N2} \\ \vdots & \vdots & \ddots & \vdots \\ A_{1N} & A_{2N} & \cdots & A_{NN} \end{bmatrix}$$

(A.26)

where A_{ij} is the cofactor of a_{ij} defined in Eq. (A.20) and "adj" stands for the *adjugate* (or *adjoint*). Formula (A.25) is used mainly for $N = 2$ and $N = 3$.

EXAMPLE A.12 Let

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & -3 \\ 1 & 2 & 0 \\ 3 & -1 & -2 \end{bmatrix}$$

Then

$$|\mathbf{A}| = 1\begin{vmatrix} 2 & 0 \\ -1 & -2 \end{vmatrix} - 3\begin{vmatrix} 1 & 2 \\ 3 & -1 \end{vmatrix} = -4 - 3(-7) = 17$$

$$\operatorname{adj} \mathbf{A} = \begin{bmatrix} \begin{vmatrix} 2 & 0 \\ -1 & -2 \end{vmatrix} & -\begin{vmatrix} 0 & -3 \\ -1 & -2 \end{vmatrix} & \begin{vmatrix} 0 & -3 \\ 2 & 0 \end{vmatrix} \\ -\begin{vmatrix} 1 & 0 \\ 3 & -2 \end{vmatrix} & \begin{vmatrix} 1 & -3 \\ 3 & -2 \end{vmatrix} & -\begin{vmatrix} 1 & -3 \\ 1 & 0 \end{vmatrix} \\ \begin{vmatrix} 1 & 2 \\ 3 & -1 \end{vmatrix} & -\begin{vmatrix} 1 & 0 \\ 3 & -1 \end{vmatrix} & \begin{vmatrix} 1 & 0 \\ 1 & 2 \end{vmatrix} \end{bmatrix} = \begin{bmatrix} -4 & 3 & 6 \\ 2 & 7 & -3 \\ -7 & 1 & 2 \end{bmatrix}$$

Thus,

$$\mathbf{A}^{-1} = \frac{1}{17}\begin{bmatrix} -4 & 3 & 6 \\ 2 & 7 & -3 \\ -7 & 1 & 2 \end{bmatrix}$$

For a 2 × 2 matrix,

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}^{-1} = \frac{1}{a_{11}a_{22} - a_{12}a_{21}}\begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix}$$

(A.27)

From Eq. (A.25) we see that if det **A** = 0, then $\mathbf{A}^{-1}$ does not exist. The matrix **A** is called *singular* if det **A** = 0, and *nonsingular* if det **A** ≠ 0. Thus, if a matrix is nonsingular, then it is invertible and $\mathbf{A}^{-1}$ exists.

A.5. Eigenvalues and Eigenvectors

A.5.1. A. Definitions:

Let **A** be an $N \times N$ matrix. If

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

(A.28)

for some scalar λ and nonzero column vector **x**, then λ is called an *eigenvalue* (or *characteristic value*) of **A** and **x** is called an *eigenvector* associated with λ.

A.5.2. B. Characteristic Equation:

Equation (A.28) can be rewritten as

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

(A.29)

where **I** is the identity matrix of *N*th order. Equation (A.29) will have a nonzero eigenvector **x** only if λ**I** − **A** is singular, that is,

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

(A.30)

which is called the *characteristic equation* of **A**. The polynomial *c*(λ) defined by

$$c(\lambda) = |\lambda\mathbf{I} - \mathbf{A}| = \lambda^N + c_{N-1}\lambda^{N-1} + \cdots + c_1\lambda + c_0$$

(A.31)

is called the *characteristic polynomial* of **A**. Now if $\lambda_1, \lambda_2, \ldots, \lambda_i$ are distinct eigenvalues of **A**, then we have

$$c(\lambda) = (\lambda - \lambda_1)^{m_1}(\lambda - \lambda_2)^{m_2} \cdots (\lambda - \lambda_i)^{m_i}$$

(A.32)

where $m_1 + m_2 + \ldots + m_i = N$ and m_i is called the *algebraic multiplicity* of λ_i.

Theorem A.1:

Let λ_k ($k = 1, 2, \ldots, i$) be the distinct eigenvalues of **A** and let $\mathbf{x}_k$ be the eigenvectors associated with the eigenvalues λ_k. Then the set of eigenvectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_i$ are linearly independent.

Proof The proof is by contradiction. Suppose that $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_i$ are linearly dependent.

Then there exists $\alpha_1, \alpha_2, \ldots, \alpha_i$ not all zero such that

$$\alpha_1\mathbf{x}_1 + \alpha_2\mathbf{x}_2 + \cdots + \alpha_i\mathbf{x}_i = \sum_{K=1}^{i} \alpha_k\mathbf{x}_k = \mathbf{0}$$

(A.33)

Assuming $\alpha_1 \neq 0$, then by Eq. (A.33) we have

$$(\lambda_2\mathbf{I} - \mathbf{A})(\lambda_3\mathbf{I} - \mathbf{A})\cdots(\lambda_i\mathbf{I} - \mathbf{A})\left[\sum_{K=1}^{i} \alpha_k\mathbf{x}_k\right] = \mathbf{0}$$

(A.34)

Now by Eq. (A.28)

$$(\lambda_j\mathbf{I} - \mathbf{A})\mathbf{x}_k = (\lambda_j - \lambda_k)\mathbf{x}_k \qquad j \neq k$$

and

$$(\lambda_k\mathbf{I} - \mathbf{A})\mathbf{x}_k = \mathbf{0}$$

Then Eq. (A.34) can be written as

$$\alpha_1(\lambda_2 - \lambda_1)(\lambda_3 - \lambda_1) \cdots (\lambda_i - \lambda_1)\mathbf{x}_1 = \mathbf{0}$$

(A.35)

Since λ_k ($k = 1, 2, \ldots, i$) are distinct, Eq. (A.35) implies that $\alpha_1 = 0$, which is a contradiction. Thus, the set of eigenvectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_i$ are linearly independent.

A.6. Diagonalization and Similarity Transformation

A.6.1. A. Diagonalization:

Suppose that all eigenvalues of an $N \times N$ matrix **A** are distinct. Let $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N$ be eigenvectors associated with the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_N$. Let

$$\mathbf{P} = [\mathbf{x}_1 \quad \mathbf{x}_2 \quad \cdots \quad \mathbf{x}_N]$$

(A.36)

Then

$$\begin{aligned}\mathbf{AP} &= \mathbf{A}\begin{bmatrix}\mathbf{x}_1 & \mathbf{x}_2 & \cdots & \mathbf{x}_N\end{bmatrix}\\ &= \begin{bmatrix}\mathbf{Ax}_1 & \mathbf{Ax}_2 & \cdots & \mathbf{Ax}_N\end{bmatrix}\\ &= \begin{bmatrix}\lambda_1\mathbf{x}_1 & \lambda_2\mathbf{x}_2 & \cdots & \lambda_N\mathbf{x}_N\end{bmatrix}\\ &= \begin{bmatrix}\mathbf{x}_1 & \mathbf{x}_2 & \cdots & \mathbf{x}_N\end{bmatrix}\begin{bmatrix}\lambda_1 & 0 & \cdots & 0\\ 0 & \lambda_2 & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & \lambda_N\end{bmatrix} = \mathbf{P\Lambda}\end{aligned}$$

(A.37)

where

$$\mathbf{\Lambda} = \begin{bmatrix}\lambda_1 & 0 & \cdots & 0\\ 0 & \lambda_2 & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & \lambda_N\end{bmatrix}$$

(A.38)

By Theorem A.1, **P** has *N* linearly independent column vectors. Thus, **P** is nonsingular and $\mathbf{P}^{-1}$ exists, and hence

$$\mathbf{P}^{-1}\mathbf{AP} = \mathbf{\Lambda} = \begin{bmatrix}\lambda_1 & 0 & \cdots & 0\\ 0 & \lambda_2 & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & \lambda_N\end{bmatrix}$$

(A.39)

We call **P** the *diagonalization* matrix or *eigenvector* matrix, and **Λ** the *eigenvalue* matrix.

Notes:

1. A sufficient (but not necessary) condition that an $N \times N$ matrix **A** be diagonalizable is that **A** has *N* distinct eigenvalues.
2. If **A** does not have *N* independent eigenvectors, then **A** is not diagonalizable.
3. The diagonalization matrix **P** is not unique. Reordering the columns of **P** or multiplying them by nonzero scalars will produce a new diagonalization matrix.

A.6.2. B. Similarity Transformation:

Let **A** and **B** be two square matrices of the same order. If there exists a nonsingular matrix **Q** such that

$$\mathbf{B} = \mathbf{Q}^{-1}\mathbf{A}\mathbf{Q}$$

(A.40)

then we say that **B** is *similar* to **A** and Eq. (A.40) is called the *similarity transformation*.

Notes:

1. If **B** is similar to **A**, then **A** is similar to **B**.
2. If **A** is similar to **B** and **B** is similar to **C**, then **A** is similar to **C**.
3. If **A** and **B** are similar, then **A** and **B** have the same eigenvalues.
4. An $N \times N$ matrix **A** is similar to a diagonal matrix **D** if and only if there exist N linearly independent eigenvectors of **A**.

A.7. Functions of a Matrix

A.7.1. A. Powers of a Matrix:

We define powers of an $N \times N$ matrix **A** as

$$\mathbf{A}^n = \underbrace{\mathbf{A}\mathbf{A}\cdots\mathbf{A}}_{n}$$

$$\mathbf{A}^0 = \mathbf{I}$$

(A.41)

It can be easily verified by direct multiplication that if

$$\mathbf{D} = \begin{bmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_N \end{bmatrix}$$

(A.42)

then

$$\mathbf{D}^n = \begin{bmatrix} d_1^n & 0 & \cdots & 0 \\ 0 & d_2^n & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_N^n \end{bmatrix}$$

(A.43)

Notes:

1. If the eigenvalues of **A** are $\lambda_1, \lambda_2, \ldots, \lambda_i$, then the eigenvalues of $\mathbf{A}^n$ are $\lambda^n{}_1, \lambda^n{}_2, \ldots, \lambda^n{}_i$.
2. Each eigenvector of **A** is still an eigenvector of $\mathbf{A}^n$.
3. If **P** diagonalizes **A**, that is,

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{\Lambda} = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_N \end{bmatrix}$$

(A.44)

then it also diagonalizes $\mathbf{A}^n$, that is,

$$\mathbf{P}^{-1}\mathbf{A}^n\mathbf{P} = \mathbf{\Lambda}^n = \begin{bmatrix} \lambda_1^n & 0 & \cdots & 0 \\ 0 & \lambda_2^n & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_N^n \end{bmatrix}$$

(A.45)

since

$$\begin{aligned} (\mathbf{P}^{-1}\mathbf{A}\mathbf{P})(\mathbf{P}^{-1}\mathbf{A}\mathbf{P}) &= \mathbf{P}^{-1}\mathbf{A}^2\mathbf{P} = \mathbf{\Lambda}^2 \\ (\mathbf{P}^{-1}\mathbf{A}^2\mathbf{P})(\mathbf{P}^{-1}\mathbf{A}\mathbf{P}) &= \mathbf{P}^{-1}\mathbf{A}^3\mathbf{P} = \mathbf{\Lambda}^3 \\ &\vdots \end{aligned}$$

(A.46)

A.7.2. B. Function of a Matrix:

Consider a function of λ defined by

$$f(\lambda) = a_0 + a_1\lambda + a_2\lambda^2 + \cdots = \sum_{k=0}^{\infty} a_k\lambda^k$$

(A.47)

With any such function we can associate a function of an$N \times N$ matrix **A**:

$$f(\mathbf{A}) = a_0\mathbf{I} + a_1\mathbf{A} + a_2\mathbf{A}^2 + \cdots = \sum_{k=0}^{\infty} a_k\mathbf{A}^k$$

(A.48)

If **A** is a diagonal matrix **D** in Eq. (A.42), then using Eq. (A.43), we have

$$f(\mathbf{D}) = a_0\mathbf{I} + a_1\mathbf{D} + a_2\mathbf{D}^2 + \cdots = \sum_{k=0}^{\infty} a_k \mathbf{D}^k$$

$$= \begin{bmatrix} \sum_{k=0}^{\infty} a_k d_1^k & 0 & \cdots & 0 \\ 0 & \sum_{k=0}^{\infty} a_k d_2^k & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sum_{k=0}^{\infty} a_k d_N^k \end{bmatrix} = \begin{bmatrix} f(d_1) & 0 & \cdots & 0 \\ 0 & f(d_2) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & f(d_N) \end{bmatrix}$$

(A.49)

If **P** diagonalizes **A**, that is [Eq. (A.44)],

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

then we have

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

(A.50)

and

$$\mathbf{A}^2 = (\mathbf{P}\mathbf{\Lambda}\mathbf{P}^{-1})(\mathbf{P}\mathbf{\Lambda}\mathbf{P}^{-1}) = \mathbf{P}\mathbf{\Lambda}^2\mathbf{P}^{-1}$$
$$\mathbf{A}^3 = (\mathbf{P}\mathbf{\Lambda}^2\mathbf{P}^{-1})(\mathbf{P}\mathbf{\Lambda}\mathbf{P}^{-1}) = \mathbf{P}\mathbf{\Lambda}^3\mathbf{P}^{-1}$$
$$\vdots$$

(A.51)

Thus, we obtain

$$f(\mathbf{A}) = \mathbf{P}f(\mathbf{\Lambda})\mathbf{P}^{-1}$$

(A.52)

Replacing **D** by **Λ** in Eq. (A.49), we get

$$f(\mathbf{A}) = \mathbf{P} \begin{bmatrix} f(\lambda_1) & 0 & \cdots & 0 \\ 0 & f(\lambda_2) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & f(\lambda_N) \end{bmatrix} \mathbf{P}^{-1}$$

(A.53)

where λ_k are the eigenvalues of **A**.

A.7.3. C. The Cayley-Hamilton Theorem:

Let the characteristic polynomial $c(\lambda)$ of an $N \times N$ matrix **A** be given by [Eq. (A.31)]

$$c(\lambda) = |\lambda \mathbf{I} - \mathbf{A}| = \lambda^N + c_{N-1}\lambda^{N-1} + \cdots + c_1\lambda + c_0$$

The *Cayley-Hamilton* theorem states that the matrix **A** satisfies its own characteristic equation; that is,

$$c(\mathbf{A}) = \mathbf{A}^N + c_{N-1}\mathbf{A}^{N-1} + \cdots + c_1\mathbf{A} + c_0\mathbf{I} = \mathbf{0}$$

(A.54)

EXAMPLE A.13 Let

$$\mathbf{A} = \begin{bmatrix} 2 & 1 \\ 0 & 3 \end{bmatrix}$$

Then, its characteristic polynomial is

$$c(\lambda) = |\lambda \mathbf{I} - \mathbf{A}| = \begin{vmatrix} \lambda - 2 & -1 \\ 0 & \lambda - 3 \end{vmatrix} = (\lambda - 2)(\lambda - 3) = \lambda^2 - 5\lambda + 6$$

and

$$\begin{aligned} c(\mathbf{A}) = \mathbf{A}^2 - 5\mathbf{A} + 6\mathbf{I} &= \begin{bmatrix} 2 & 1 \\ 0 & 3 \end{bmatrix}^2 - 5\begin{bmatrix} 2 & 1 \\ 0 & 3 \end{bmatrix} + 6\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 4 & 5 \\ 0 & 9 \end{bmatrix} - \begin{bmatrix} 10 & 5 \\ 0 & 15 \end{bmatrix} + \begin{bmatrix} 6 & 0 \\ 0 & 6 \end{bmatrix} \\ &= \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} = \mathbf{0} \end{aligned}$$

Rewriting Eq. (A.54), we have

$$\mathbf{A}^N = -c_0\mathbf{I} - c_1\mathbf{A} - \cdots - c_{N-1}\mathbf{A}^{N-1}$$

(A.55)

Multiplying through by **A** and then substituting the expression (A.55) for $\mathbf{A}^N$ on the right and rearranging, we get

$$\mathbf{A}^{N+1} = \alpha_0\mathbf{I} + \alpha_1\mathbf{A} + \cdots + \alpha_{N-1}\mathbf{A}^{N-1}$$

(A.56)

By continuing this process, we can express any positive integral power of **A** as a linear combination of **I**, **A**, ..., $\mathbf{A}^{N-1}$. Thus, $f(\mathbf{A})$ defined by Eq. (A.48) can be represented by

$$f(\mathbf{A}) = b_0\mathbf{I} + b_1\mathbf{A} + \cdots + b_{N-1}\mathbf{A}^{N-1} = \sum_{m=0}^{N-1} b_m \mathbf{A}^m$$

(A.57)

In a similar manner, if λ is an eigenvalue of **A**, then $f(\lambda)$ can also be expressed as

$$f(\lambda) = b_0 + b_1\lambda + \cdots + b_{N-1}\lambda^{N-1} = \sum_{m=0}^{N-1} b_m \lambda^m$$

(A.58)

Thus, if all eigenvalues of **A** are distinct, the coefficients b_m ($m = 0, 1, \ldots, N - 1$) can be determined by the following N equations:

$$f(\lambda_k) = b_0 + b_1\lambda_k + \cdots + b_{N-1}\lambda_k^{N-1} \qquad k = 1, 2, \ldots, N$$

(A.59)

If all eigenvalues of **A** are not distinct, then Eq. (A.59) will not yield N equations. Assume that an eigenvalue λ_i has multiplicity r and all other eigenvalues are distinct. In this case differentiating both sides of Eq. (A.58) r times with respect to λ and setting $\lambda = \lambda_i$, we obtain r equations corresponding to λ_i:

$$\left.\frac{d^{n-1}}{d\lambda^{n-1}} f(\lambda)\right|_{\lambda=\lambda_i} = \left.\frac{d^{n-1}}{d\lambda^{n-1}}\left(\sum_{m=0}^{N-1} b_m\lambda^m\right)\right|_{\lambda=\lambda_i} \qquad n = 1, 2, \ldots, r$$

(A.60)

Combining Eqs. (A.59) and (A.60), we can determine all coefficients b_m in Eq. (A.57).

A.7.4. D. Minimal Polynomial of A:

The *minimal* (or *minimum*) polynomial $m(\lambda)$ of an $N \times N$ matrix **A** is the polynomial of lowest degree having 1 as its leading coefficient such that $m(\mathbf{A}) = \mathbf{0}$. Since **A** satisfies its characteristic equation, the degree of $m(\lambda)$ is not greater than N.

EXAMPLE A.14 Let

$$\mathbf{A} = \begin{bmatrix} \alpha & 0 \\ 0 & \alpha \end{bmatrix}$$

The characteristic polynomial is

$$c(\lambda) = |\lambda\mathbf{I} - \mathbf{A}| = \begin{vmatrix} \lambda - \alpha & 0 \\ 0 & \lambda - \alpha \end{vmatrix} = (\lambda - \alpha)^2 = \lambda^2 - 2\alpha\lambda + \alpha^2$$

and the minimal polynomial is

$$m(\lambda) = \lambda - \alpha$$

since

$$m(\mathbf{A}) = \mathbf{A} - \alpha\mathbf{I} = \begin{bmatrix} \alpha & 0 \\ 0 & \alpha \end{bmatrix} - \alpha \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} = \mathbf{0}$$

Notes:

1. Every eigenvalue of **A** is a zero of $m(\lambda)$.
2. If all the eigenvalues of **A** are distinct, then $c(\lambda) = m(\lambda)$.
3. $c(\lambda)$ is divisible by $m(\lambda)$.
4. $m(\lambda)$ may be used in the same way as $c(\lambda)$ for the expression of higher powers of **A** in terms of a limited number of powers of **A**.

   It can be shown that m(λ) can be determined by

   $$m(\lambda) = \frac{c(\lambda)}{d(\lambda)}$$

   (A.61)

   where $d(\lambda)$ is the greatest common divisor (gcd) of all elements of adj(λ**I** − **A**).

EXAMPLE A.15 Let

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

Then

$$c(\lambda) = |\lambda \mathbf{I} - A| = \begin{bmatrix} \lambda - 5 & 6 & 6 \\ 1 & \lambda - 4 & -2 \\ -3 & 6 & \lambda + 4 \end{bmatrix}$$
$$= \lambda^3 - 5\lambda^2 + 8\lambda - 4 = (\lambda - 1)(\lambda - 2)^2$$

$$\operatorname{adj}[\lambda \mathbf{I} - \mathbf{A}] = \begin{bmatrix} \begin{vmatrix} \lambda - 4 & -2 \\ 6 & \lambda + 4 \end{vmatrix} & -\begin{vmatrix} 6 & 6 \\ 6 & \lambda + 4 \end{vmatrix} & \begin{vmatrix} 6 & 6 \\ \lambda - 4 & -2 \end{vmatrix} \\ -\begin{vmatrix} 1 & -2 \\ -3 & \lambda + 4 \end{vmatrix} & \begin{vmatrix} \lambda - 5 & 6 \\ -3 & \lambda + 4 \end{vmatrix} & -\begin{vmatrix} \lambda - 5 & 6 \\ 1 & -2 \end{vmatrix} \\ \begin{vmatrix} 1 & \lambda - 4 \\ -3 & 6 \end{vmatrix} & -\begin{vmatrix} \lambda - 5 & 6 \\ -3 & 6 \end{vmatrix} & \begin{vmatrix} \lambda - 5 & 6 \\ 1 & \lambda - 4 \end{vmatrix} \end{bmatrix}$$
$$= \begin{bmatrix} (\lambda + 2)(\lambda - 2) & -6(\lambda - 2) & -6(\lambda - 2) \\ -(\lambda - 2) & (\lambda + 1)(\lambda - 2) & 2(\lambda - 2) \\ 3(\lambda - 2) & -6(\lambda - 2) & (\lambda - 2)(\lambda - 7) \end{bmatrix}$$

Thus, $d(\lambda) = \lambda - 2$ and

$$m(\lambda) = \frac{c(\lambda)}{d(\lambda)} = (\lambda - 1)(\lambda - 2) = \lambda^2 - 3\lambda + 2$$

and

$$m(\mathbf{A}) = (\mathbf{A} - \mathbf{I})(\mathbf{A} - 2\mathbf{I}) = \begin{bmatrix} 4 & -6 & -6 \\ -1 & 3 & 2 \\ 3 & -6 & -5 \end{bmatrix} \begin{bmatrix} 3 & -6 & -6 \\ -1 & 2 & 2 \\ 3 & -6 & -6 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

A.7.5. E. Spectral Decomposition:

It can be shown that if the minimal polynomial $m(\lambda)$ of an $N \times N$ matrix **A** has the form

$$m(\lambda) = (\lambda - \lambda_1)(\lambda - \lambda_2) \cdots (\lambda - \lambda_i)$$

(A.62)

then **A** can be represented by

$$\mathbf{A} = \lambda_1 \mathbf{E}_1 + \lambda_2 \mathbf{E}_2 + \cdots + \lambda_i \mathbf{E}_i$$

(A.63)

where $\mathbf{E}_j$ ($j = 1, 2, \ldots, i$) are called *constituent* matrices and have the following properties:

$$
\begin{aligned}
&1. \quad \mathbf{I} = \mathbf{E}_1 + \mathbf{E}_2 + \cdots + \mathbf{E}_i \\
&2. \quad \mathbf{E}_m \mathbf{E}_k = \mathbf{0},\ m \neq k \\
&3. \quad \mathbf{E}_k^2 = \mathbf{E}_k \\
&4. \quad \mathbf{A}\mathbf{E}_k = \mathbf{E}_k \mathbf{A} = \lambda_k \mathbf{E}_k
\end{aligned}
$$

(A.64)

Any matrix **B** for which $\mathbf{B}^2 = \mathbf{B}$ is called *idempotent*. Thus, the constituent matrices $\mathbf{E}_j$ are idempotent matrices. The set of eigenvalues of **A** is called the *spectrum* of **A**, and Eq. (A.63) is called the *spectral decomposition* of **A**. Using the properties of Eq. (A.64), we have

$$
\begin{aligned}
\mathbf{A}^2 &= \lambda_1^2 \mathbf{E}_1 + \lambda_2^2 \mathbf{E}_2 + \cdots + \lambda_i^2 \mathbf{E}_i \\
&\vdots \\
\mathbf{A}^n &= \lambda_1^n \mathbf{E}_1 + \lambda_2^n \mathbf{E}_2 + \cdots + \lambda_i^n \mathbf{E}_i
\end{aligned}
$$

(A.65)

and

$$
f(\mathbf{A}) = f(\lambda_1)\mathbf{E}_1 + f(\lambda_2)\,\mathbf{E}_2 + \cdots + f(\lambda_i)\mathbf{E}_i
$$

(A.66)

The constituent matrices $\mathbf{E}_j$ can be evaluated as follows. The partial-fraction expansion of

$$
\begin{aligned}
\frac{1}{m(\lambda)} &= \frac{1}{(\lambda - \lambda_1)(\lambda - \lambda_2)\cdots(\lambda - \lambda_i)} \\
&= \frac{k_1}{\lambda - \lambda_1} + \frac{k_2}{\lambda - \lambda_2} + \cdots + \frac{k_i}{\lambda - \lambda_i}
\end{aligned}
$$

leads to

$$
k_j = \frac{1}{\prod_{\substack{m=1 \\ m \neq j}}^{i} (\lambda_j - \lambda_m)}
$$

Then

$$
\frac{1}{m(\lambda)} = \frac{k_1 g_1(\lambda) + k_2 g_2(\lambda) + \cdots + k_i g_i(\lambda)}{(\lambda - \lambda_1)(\lambda - \lambda_2)\cdots(\lambda - \lambda_i)}
$$

where

$$
g_j(\lambda) = \prod_{\substack{m=1 \\ m \neq j}}^{i} (\lambda - \lambda_m)
$$

Let $e_j(\lambda) = k_j g_j(\lambda)$. Then the constituent matrices $\mathbf{E}_j$ can be evaluated as

$$\mathbf{E}_j = e_j(\mathbf{A}) = \frac{\prod_{\substack{m=1\\ m\neq j}}^{i} (\mathbf{A} - \lambda_m \mathbf{I})}{\prod_{\substack{m=1\\ m\neq j}}^{i} (\lambda_j - \lambda_m)}$$

(A.67)

EXAMPLE A.16 Consider the matrix **A** in Example A.15:

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

From Example A.15, we have

$$m(\lambda) = (\lambda - 1)(\lambda - 2)$$

Then

$$\frac{1}{m(\lambda)} = \frac{1}{(\lambda - 1)(\lambda - 2)} = \frac{-1}{\lambda - 1} + \frac{1}{\lambda - 2}$$

and

$$e_1(\lambda) = -(\lambda - 2) \qquad e_2(\lambda) = \lambda - 1$$

Then

$$\mathbf{E}_1 = e_1(\mathbf{A}) = -(\mathbf{A} - 2\mathbf{I}) = \begin{bmatrix} -3 & 6 & 6 \\ 1 & -2 & -2 \\ -3 & 6 & 6 \end{bmatrix}$$

$$\mathbf{E}_2 = e_2(\mathbf{A}) = \mathrm{A} - \mathbf{I} = \begin{bmatrix} 4 & -6 & -6 \\ -1 & 3 & 2 \\ 3 & -6 & -5 \end{bmatrix}$$

$$\mathrm{A} = \lambda_1 \mathbf{E}_1 + \lambda_2 \mathbf{E}_2 = \mathbf{E}_1 + 2\mathbf{E}_2$$

$$= \begin{bmatrix} -3 & 6 & 6 \\ 1 & -2 & -2 \\ -3 & 6 & 6 \end{bmatrix} + 2 \begin{bmatrix} 4 & -6 & -6 \\ -1 & 3 & 2 \\ 3 & -6 & -5 \end{bmatrix} = \begin{bmatrix} 5 & -6 & -6 \\ -1 & 4 & 2 \\ 3 & -6 & -4 \end{bmatrix}$$

A.8. Differentiation and Integration of Matrices

A.8.1. A. Definitions:

The derivative of an $m \times n$ matrix **A**(t) is defined to be the $m \times n$ matrix, each element of which is the derivative of the corresponding element of **A**; that is,

$$\frac{d}{dt}\mathbf{A}(t) = \left[\frac{d}{dt}a_{ij}(t)\right]_{m \times n} = \begin{bmatrix} \frac{d}{dt}a_{11}(t) & \frac{d}{dt}a_{12}(t) & \cdots & \frac{d}{dt}a_{1n}(t) \\ \frac{d}{dt}a_{21}(t) & \frac{d}{dt}a_{22}(t) & \cdots & \frac{d}{dt}a_{2n}(t) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{d}{dt}a_{m1}(t) & \frac{d}{dt}a_{m2}(t) & \cdots & \frac{d}{dt}a_{mn}(t) \end{bmatrix}$$

(A.68)

Similarly, the integral of an $m \times n$ matrix **A**(t) is defined to be

$$\int \mathbf{A}(t)\,dt = \left[\int a_{ij}(t)\,dt\right]_{m \times n} = \begin{bmatrix} \int a_{11}(t)\,dt & \int a_{12}(t)\,dt & \cdots & \int a_{1n}(t)\,dt \\ \int a_{21}(t)\,dt & \int a_{22}(t)\,dt & \cdots & \int a_{2n}(t)\,dt \\ \vdots & \vdots & \ddots & \vdots \\ \int a_{m1}(t)\,dt & \int a_{m2}(t)\,dt & \cdots & \int a_{mn}(t)\,dt \end{bmatrix}$$

(A.69)

EXAMPLE A.17 Let

$$\mathbf{A} = \begin{bmatrix} t & t^2 \\ 1 & t^3 \end{bmatrix}$$

Then

$$\frac{d}{dt}\mathbf{A} = \begin{bmatrix} \frac{d}{dt}t & \frac{d}{dt}t^2 \\ \frac{d}{dt}1 & \frac{d}{dt}t^3 \end{bmatrix} = \begin{bmatrix} 1 & 2t \\ 0 & 3t^2 \end{bmatrix}$$

and

$$\int_0^1 \mathrm{A}\ dt = \begin{bmatrix} \int_0^1 t\ dt & \int_0^1 t^2\ dt \\ \int_0^1 1\ dt & \int_0^1 t^3\ dt \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & \frac{1}{3} \\ 1 & \frac{1}{4} \end{bmatrix}$$

A.8.2. B. Differentiation of the Product of Two Matrices:

If the matrices **A**(t) and **B**(t) can be differentiated with respect to t, then

$$\frac{d}{dt}[\mathbf{A}(t)\mathbf{B}(t)] = \frac{d\mathbf{A}(t)}{dt}\mathbf{B}(t) + \mathbf{A}(t)\frac{d\mathbf{B}(t)}{dt}$$

(A.70)