Appendix WB
Summary of Matrix Theory

In the text, we assume you are already somewhat familiar with matrix theory and with the solution of linear systems of equations. However, for the purposes of review we present here a brief summary of matrix theory with an emphasis on the results needed in control theory. For further study, see Strang (2006) and Gantmacher (1959).

WB.1 Matrix Definitions

An array of numbers arranged in rows and columns is referred to as a **matrix**. If $\mathbf{A}$ is a matrix with m rows and n columns, an $m \times n$ (read "m by n") matrix, it is denoted as

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}, \tag{WB.1}$$

where the entries a_{ij} are its elements. If $m = n$, then the matrix is **square**; otherwise it is **rectangular**. Sometimes a matrix is simply denoted by $\mathbf{A} = [a_{ij}]$. If $m = 1$ or $n = 1$, then the matrix reduces to a **row vector** or a **column vector**, respectively. A **submatrix** of $\mathbf{A}$ is the matrix with certain rows and columns removed.

WB.2 Elementary Operations on Matrices

If $\mathbf{A}$ and $\mathbf{B}$ are matrices of the same dimension, then their sum is defined by

$$\mathbf{C} = \mathbf{A} + \mathbf{B}, \tag{WB.2}$$

where

$$c_{ij} = a_{ij} + b_{ij}. \tag{WB.3}$$

That is, the addition is done element by element. It is easy to verify the following properties of matrices:

Commutative law for addition

Associative law for addition

$$\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}, \tag{WB.4}$$

$$(\mathbf{A} + \mathbf{B}) + \mathbf{C} = \mathbf{A} + (\mathbf{B} + \mathbf{C}). \tag{WB.5}$$

Two matrices can be multiplied if they are compatible. Let $\mathbf{A} = m \times n$ and $\mathbf{B} = n \times p$. Then the $m \times p$ matrix

$$\mathbf{C} = \mathbf{AB}, \tag{WB.6}$$

is the product of the two matrices, where

$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}. \tag{WB.7}$$

Associative law for multiplication

Matrix multiplication satisfies the associative law

$$\mathbf{A}(\mathbf{BC}) = (\mathbf{AB})\mathbf{C}, \tag{WB.8}$$

but not the commutative law; that is, in general,

$$\mathbf{AB} \neq \mathbf{BA}. \tag{WB.9}$$

WB.3 Trace

The **trace** of a square matrix is the sum of its diagonal elements:

$$\text{trace } \mathbf{A} = \sum_{i=1}^{n} a_{ii}. \tag{WB.10}$$

WB.4 Transpose

The $n \times m$ matrix obtained by interchanging the rows and columns of $\mathbf{A}$ is called the **transpose of matrix A**:

$$\mathbf{A}^T = \begin{bmatrix} a_{11} & a_{21} & \dots & a_{m1} \\ a_{12} & a_{22} & \dots & a_{m2} \\ \vdots & \vdots & & \vdots \\ a_{1n} & a_{2n} & \dots & a_{mn} \end{bmatrix}.$$

A matrix is said to be **symmetric** if

$$\mathbf{A}^T = \mathbf{A}. \tag{WB.11}$$

Transposition

It is easy to show that

$$(\mathbf{AB})^T = \mathbf{B}^T\mathbf{A}^T, \tag{WB.12}$$

$$(\mathbf{ABC})^T = \mathbf{C}^T\mathbf{B}^T\mathbf{A}^T, \tag{WB.13}$$

$$(\mathbf{A}+\mathbf{B})^T = \mathbf{A}^T + \mathbf{B}^T. \tag{WB.14}$$

WB.5 Determinant and Matrix Inverse

The **determinant** of a square matrix is defined by Laplace's expansion

$$\det \mathbf{A} = \sum_{j=1}^{n} a_{ij}\gamma_{ij} \quad \text{for any} \quad i = 1, 2, \dots, n, \tag{WB.15}$$

where γ_{ij} is called the **cofactor** and

$$\gamma_{ij} = (-1)^{i+j} \det M_{ij}, \tag{WB.16}$$

where the scalar det M_{ij} is called a **minor**. M_{ij} is the same as the matrix $\mathbf{A}$ except that its ith row and jth column have been removed. Note M_{ij} is always an $(n-1) \times (n-1)$ matrix, and the minors and cofactors are identical except possibly for a sign.

The **adjugate** of a matrix is the transpose of the matrix of its cofactors:

$$\operatorname{adj} \mathbf{A} = [\gamma_{ij}]^T. \tag{WB.17}$$

It can be shown that

$$\mathbf{A} \operatorname{adj} \mathbf{A} = (\det \mathbf{A})\mathbf{I}, \tag{WB.18}$$

Identity matrix

where $\mathbf{I}$ is called the **identity matrix**:

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

that is, $\mathbf{I}$ has ones along the diagonal and zeros elsewhere. If $\det \mathbf{A} \neq 0$, then the **inverse** of a matrix $\mathbf{A}$ is defined by

$$\mathbf{A}^{-1} = \frac{\operatorname{adj} \mathbf{A}}{\det \mathbf{A}}, \tag{WB.19}$$

and has the property

$$\mathbf{A}\mathbf{A}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}. \tag{WB.20}$$

Note a matrix has an inverse—that is, it is **nonsingular**—if its determinant is nonzero.

The inverse of the product of two matrices is the product of the inverse of the matrices in reverse order:

$$(\mathbf{A}\mathbf{B})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1} \tag{WB.21}$$

Inversion

and

$$(\mathbf{A}\mathbf{B}\mathbf{C})^{-1} = \mathbf{C}^{-1}\mathbf{B}^{-1}\mathbf{A}^{-1}. \tag{WB.22}$$

WB.6 Properties of the Determinant

When dealing with determinants of matrices, the following elementary (row or column) operations are useful:

1. If any row (or column) of $\mathbf{A}$ is multiplied by a scalar α, the resulting matrix $\bar{\mathbf{A}}$ has the determinant

$$\det \bar{\mathbf{A}} = \alpha \det \mathbf{A}. \tag{WB.23}$$

   Hence

$$\det(\alpha \mathbf{A}) = \alpha^n \det \mathbf{A}. \tag{WB.24}$$

2. If any two rows (or columns) of $\mathbf{A}$ are interchanged to obtain $\bar{\mathbf{A}}$, then

$$\det \bar{\mathbf{A}} = -\det \mathbf{A}. \tag{WB.25}$$

3. If a multiple of a row (or column) of $\mathbf{A}$ is added to another to obtain $\bar{\mathbf{A}}$, then
$$\det \bar{\mathbf{A}} = \det \mathbf{A}. \tag{WB.26}$$
4. It is also easy to show that
$$\det \mathbf{A} = \det \mathbf{A}^T \tag{WB.27}$$
and
$$\det \mathbf{AB} = \det \mathbf{A} \det \mathbf{B}. \tag{WB.28}$$
Applying Eq. (WB.28) to Eq. (WB.20), we have
$$\det \mathbf{A} \det \mathbf{A}^{-1} = 1. \tag{WB.29}$$
If $\mathbf{A}$ and $\mathbf{B}$ are square matrices, then the determinant of the block triangular matrix is the product of the determinants of the diagonal blocks:
$$\det \begin{bmatrix} \mathbf{A} & \mathbf{C} \\ 0 & \mathbf{B} \end{bmatrix} = \det \mathbf{A} \det \mathbf{B}. \tag{WB.30}$$
If $\mathbf{A}$ is nonsingular, then
$$\det \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix} = \det \mathbf{A} \det(\mathbf{D} - \mathbf{CA}^{-1}\mathbf{B}). \tag{WB.31}$$
Using this identity, we can write the transfer function of a scalar system in a compact form:
$$G(s) = \mathbf{C}(s\mathbf{I} - \mathbf{A})^{-1}\mathbf{B} + D = \frac{\det \begin{bmatrix} s\mathbf{I} - \mathbf{A} & \mathbf{B} \\ -\mathbf{C} & D \end{bmatrix}}{\det(s\mathbf{I} - \mathbf{A})}. \tag{WB.32}$$

WB.7 Inverse of Block Triangular Matrices

If $\mathbf{A}$ and $\mathbf{B}$ are square invertible matrices, then
$$\begin{bmatrix} \mathbf{A} & \mathbf{C} \\ \mathbf{0} & \mathbf{B} \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{A}^{-1} & -\mathbf{A}^{-1}\mathbf{CB}^{-1} \\ \mathbf{0} & \mathbf{B}^{-1} \end{bmatrix}. \tag{WB.33}$$

WB.8 Special Matrices

Diagonal matrix

Some matrices have special structures and are given names. We have already defined the identity matrix, which has a special form. A **diagonal matrix** has (possibly) nonzero elements along the main diagonal and zeros elsewhere:
$$\mathbf{A} = \begin{bmatrix} a_{11} & & \mathbf{0} & & \\ & a_{22} & & & \\ & & a_{33} & & \\ & & & \ddots & \\ & \mathbf{0} & & & a_{nn} \end{bmatrix}. \tag{WB.34}$$

Upper triangular matrix

A matrix is said to be **(upper) triangular** if all the elements below the main diagonal are zeros:

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & & \cdots & a_{1n} \\ 0 & a_{22} & & & \\ \vdots & 0 & & & \vdots \\ 0 & \vdots & \ddots & \ddots & \\ 0 & 0 & \cdots & 0 & a_{nn} \end{bmatrix}. \tag{WB.35}$$

The determinant of a diagonal or triangular matrix is simply the product of its diagonal elements.

A matrix is said to be in the **(upper) companion form** if it has the structure

$$\mathbf{A}_c = \begin{bmatrix} -a_1 & -a_2 & & \cdots & -a_n \\ 1 & 0 & & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & \cdots & \cdots & 1 & 0 \end{bmatrix}. \tag{WB.36}$$

Note all the information is contained in the first row. Variants of this form are the lower, left, or right companion matrices. A **Vandermonde matrix** has the following structure:

$$\mathbf{A} = \begin{bmatrix} 1 & a_1 & a_1^2 & \cdots & a_1^{n-1} \\ 1 & a_2 & a_2^2 & \cdots & a_2^{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & a_n & a_n^2 & \cdots & a_n^{n-1} \end{bmatrix}. \tag{WB.37}$$

WB.9 Rank

The **rank** of a matrix is the number of its linearly independent rows or columns. If the rank of $\mathbf{A}$ is r, then all $(r+1) \times (r+1)$ submatrices of $\mathbf{A}$ are singular, and there is at least one $r \times r$ submatrix that is nonsingular. It is also true that

$$\text{row rank of } \mathbf{A} = \text{column rank of } \mathbf{A}. \tag{WB.38}$$

WB.10 Characteristic Polynomial

The **characteristic polynomial** of a matrix $\mathbf{A}$ is defined by

$$\begin{aligned} a(s) &\triangleq \det(s\mathbf{I} - \mathbf{A}) \\ &= s^n + a_1 s^{n-1} + \cdots + a_{n-1}s + a_n, \end{aligned} \tag{WB.39}$$

where the roots of the polynomial are referred to as **eigenvalues** of $\mathbf{A}$. We can write

$$a(s) = (s - \lambda_1)(s - \lambda_2) \cdots (s - \lambda_n), \tag{WB.40}$$

where $\{\lambda_i\}$ are the eigenvalues of **A**. The characteristic polynomial of a companion matrix [for example, Eq. (WB.36)] is

$$\begin{aligned} a(s) &= \det(s\mathbf{I} - \mathbf{A}_c) \\ &= s^n + a_1 s^{n-1} + \cdots + a_{n-1}s + a_n. \end{aligned} \tag{WB.41}$$

WB.11 Cayley–Hamilton Theorem

The Cayley–Hamilton theorem states that every square matrix **A** satisfies its characteristic polynomial. This means if **A** is an $n \times n$ matrix with characteristic equation $a(s)$, then

$$a(\mathbf{A}) \triangleq \mathbf{A}^n + a_1\mathbf{A}^{n-1} + \cdots + a_{n-1}\mathbf{A} + a_n\mathbf{I} = 0. \tag{WB.42}$$

WB.12 Eigenvalues and Eigenvectors

Any scalar λ and nonzero vector **v** that satisfy

$$\mathbf{A}\mathbf{v} = \lambda\mathbf{v}, \tag{WB.43}$$

are referred to as the eigenvalue and the associated **(right) eigenvector** of the matrix **A** [because **v** appears to the right of **A** in Eq. (WB.43)]. By rearranging terms in Eq. (WB.43), we get

$$(\lambda\mathbf{I} - \mathbf{A})\mathbf{v} = 0. \tag{WB.44}$$

Because **v** is nonzero,

$$\det(\lambda\mathbf{I} - \mathbf{A}) = 0, \tag{WB.45}$$

so λ is an eigenvalue of the matrix **A** as defined in Eq. (WB.43). The normalization of the eigenvectors is arbitrary; that is, if **v** is an eigenvector, so is $\alpha\mathbf{v}$. The eigenvectors are usually normalized to have unit length; that is, $\|\mathbf{v}\|^2 = \mathbf{v}^T\mathbf{v} = 1$.

If $\mathbf{w}^T$ is a nonzero row vector such that

$$\mathbf{w}^T\mathbf{A} = \lambda\mathbf{w}^T, \tag{WB.46}$$

then **w** is called a **left eigenvector** of **A** [because $\mathbf{w}^T$ appears to the left of **A** in Eq. (WB.46)]. Note we can write

$$\mathbf{A}^T\mathbf{w} = \lambda\mathbf{w}, \tag{WB.47}$$

so **w** is simply a right eigenvector of $\mathbf{A}^T$.

WB.13 Similarity Transformations

Consider the arbitrary nonsingular matrix **T** such that

$$\bar{\mathbf{A}} = \mathbf{T}^{-1}\mathbf{A}\mathbf{T}. \tag{WB.48}$$

The matrix operation shown in Eq. (WB.48) is referred to as a **similarity transformation**. If **A** has a full set of eigenvectors, then we can choose **T** to be the set of eigenvectors, and $\bar{\mathbf{A}}$ will be diagonal.

Consider the set of equations in state-variable form:

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}u. \tag{WB.49}$$

If we let

$$\mathbf{T}\xi = \mathbf{x}, \tag{WB.50}$$

then Eq. (WB.49) becomes

$$\mathbf{T}\dot{\xi} = \mathbf{A}\mathbf{T}\xi + \mathbf{B}u, \tag{WB.51}$$

and premultiplying both sides by $\mathbf{T}^{-1}$, we get

$$\begin{aligned}\dot{\xi} &= \mathbf{T}^{-1}\mathbf{A}\mathbf{T}\xi + \mathbf{T}^{-1}\mathbf{B}u \\ &= \bar{\mathbf{A}}\xi + \bar{\mathbf{B}}u,\end{aligned} \tag{WB.52}$$

where

$$\begin{aligned}\bar{\mathbf{A}} &= \mathbf{T}^{-1}\mathbf{A}\mathbf{T}, \\ \bar{\mathbf{B}} &= \mathbf{T}^{-1}\mathbf{B}.\end{aligned} \tag{WB.53}$$

The characteristic polynomial of $\bar{\mathbf{A}}$ is

$$\begin{aligned}\det(s\mathbf{I} - \bar{\mathbf{A}}) &= \det(s\mathbf{I} - \mathbf{T}^{-1}\mathbf{A}\mathbf{T}) \\ &= \det(s\mathbf{T}^{-1}\mathbf{T} - \mathbf{T}^{-1}\mathbf{A}\mathbf{T}) \\ &= \det[\mathbf{T}^{-1}(s\mathbf{I} - \mathbf{A})\mathbf{T}] \\ &= \det\mathbf{T}^{-1}\det(s\mathbf{I} - \mathbf{A})\det\mathbf{T}.\end{aligned} \tag{WB.54}$$

Using Eq. (WB.29), Eq. (WB.54) becomes

$$\det(s\mathbf{I} - \bar{\mathbf{A}}) = \det(s\mathbf{I} - \mathbf{A}). \tag{WB.55}$$

From Eq. (WB.55), we can see that $\bar{\mathbf{A}}$ and $\mathbf{A}$ both have the same characteristic polynomial, giving us the important result that a similarity transformation does not change the eigenvalues of a matrix. From Eq. (WB.50), a new state made up of a linear combination of the old state has the same eigenvalues as the old set.

WB.14 Matrix Exponential

Let $\mathbf{A}$ be a square matrix. The **matrix exponential** of $\mathbf{A}$ is defined as the series

$$e^{\mathbf{A}t} = \mathbf{I} + \mathbf{A}t + \frac{1}{2!}\mathbf{A}^2t^2 + \frac{\mathbf{A}^3t^3}{3!} + \cdots. \tag{WB.56}$$

It can be shown that the series converges. If $\mathbf{A}$ is an $n \times n$ matrix, then $e^{\mathbf{A}t}$ is also an $n \times n$ matrix and can be differentiated:

$$\frac{d}{dt}e^{\mathbf{A}t} = \mathbf{A}e^{\mathbf{A}t}. \tag{WB.57}$$

Other properties of the matrix exponential are

$$e^{\mathbf{A}t_1}e^{\mathbf{A}t_2} = e^{\mathbf{A}(t_1+t_2)} \tag{WB.58}$$

and, in general,

$$e^{\mathbf{A}}e^{\mathbf{B}} \neq e^{\mathbf{B}}e^{\mathbf{A}}. \tag{WB.59}$$

(In the exceptional case where **A** and **B** commute—that is, $\mathbf{AB} = \mathbf{BA}$—then $e^{\mathbf{A}}e^{\mathbf{B}} = e^{\mathbf{B}}e^{\mathbf{A}}$.)

WB.15 Fundamental Subspaces

The **range space** of **A**, denoted by $\mathcal{R}(\mathbf{A})$ and also called the **column space** of **A**, is defined by the set of vectors

$$\mathbf{x} = \mathbf{Ay}, \tag{WB.60}$$

for some vector **y**. The **null space** of **A**, denoted by $\mathcal{N}(\mathbf{A})$, is defined by the set of vectors **x** such that

$$\mathbf{Ax} = \mathbf{0}. \tag{WB.61}$$

If $x \in \mathcal{N}(\mathbf{A})$ and $y \in \mathcal{R}(\mathbf{A}^T)$, then $\mathbf{y}^T\mathbf{x} = 0$; that is, every vector in the null space of **A** is **orthogonal** to every vector in the range space of $\mathbf{A}^T$.

WB.16 Singular-Value Decomposition

The **singular-value decomposition (SVD)** is one of the most useful tools in linear algebra and has been widely used in control theory during the last few decades. Let **A** be an $m \times n$ matrix. Then there always exist matrices **U**, **S**, and **V** such that

$$\mathbf{A} = \mathbf{USV}^T. \tag{WB.62}$$

Here **U** and **V** are **orthogonal matrices**; that is,

$$\mathbf{UU}^T = \mathbf{I}, \mathbf{VV}^T = \mathbf{I}. \tag{WB.63}$$

S is a **quasidiagonal matrix** with singular values as its diagonal elements; that is,

$$\mathbf{S} = \begin{bmatrix} \Sigma & 0 \\ 0 & 0 \end{bmatrix}, \tag{WB.64}$$

where Σ is a diagonal matrix of nonzero singular values in descending order:

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0. \tag{WB.65}$$

The unique diagonal elements of **S** are called the **singular values**. The maximum singular value is denoted by $\bar{\sigma}(A)$, and the minimum singular value is denoted by $\underline{\sigma}(\mathbf{A})$. The rank of the matrix is the same as the number of nonzero singular values. The columns of **U** and **V**,

$$\begin{aligned} \mathbf{U} &= [\ u_1 \quad u_2 \quad \ldots \quad u_m\ ], \\ \mathbf{V} &= [\ v_1 \quad v_2 \quad \ldots \quad v_n\ ], \end{aligned} \tag{WB.66}$$

are called the left and right **singular vectors**, respectively. SVD provides complete information about the fundamental subspaces associated with a matrix:

$$\begin{aligned}
\mathcal{N}(\mathbf{A}) &= \text{span}[\ \upsilon_{r+1} \quad \upsilon_{r+2} \quad \ldots \quad \upsilon_n\ ],\\
\mathcal{R}(\mathbf{A}) &= \text{span}[\ u_1 \quad u_2 \quad \ldots \quad u_r\ ],\\
\mathcal{R}(\mathbf{A}^T) &= \text{span}[\ \upsilon_1 \quad \upsilon_2 \quad \ldots \quad \upsilon_r\ ],\\
\mathcal{N}(\mathbf{A}^T) &= \text{span}[\ u_{r+1} \quad u_{r+2} \quad \ldots \quad u_m\ ].
\end{aligned} \tag{WB.67}$$

Here $\mathcal{N}$ denotes the null space and $\mathcal{R}$, the range space, respectively.

The **norm** of the matrix $\mathbf{A}$, denoted by $\|\mathbf{A}\|_2$, is given by

$$\|\mathbf{A}\|_2 = \bar{\sigma}(\mathbf{A}). \tag{WB.68}$$

If $\mathbf{A}$ is a function of ω, then the infinity norm of $\mathbf{A}$, $\|\mathbf{A}\|_\infty$, is given by

$$\|\mathbf{A}(j\omega)\|_\infty = \max_\omega \bar{\sigma}(\mathbf{A}). \tag{WB.69}$$

WB.17 Positive Definite Matrices

A matrix $\mathbf{A}$ is said to be **positive semidefinite** if

$$\mathbf{x}^T\mathbf{A}\mathbf{x} \geq 0 \quad \text{for all } \mathbf{x}. \tag{WB.70}$$

The matrix is said to be **positive definite** if equality holds in Eq. (WB.70) only for $\mathbf{x} = 0$. A symmetric matrix is positive definite if and only if all of its eigenvalues are positive. It is positive semidefinite if and only if all of its eigenvalues are nonnegative.

An alternate method for determining positive definiteness is to test the minors of the matrix. A matrix is positive definite if all the leading principal minors are positive, and positive semidefinite if they are all nonnegative.

WB.18 Matrix Identity

If $\mathbf{A}$ is an $n \times m$ matrix and $\mathbf{B}$ is an $m \times n$ matrix, then

$$\det[\mathbf{I}_n - \mathbf{AB}] = \det[\mathbf{I}_m - \mathbf{BA}],$$

where $\mathbf{I}_n$ and $\mathbf{I}_m$ are identity matrices of size n and m, respectively.

WB.19 Cramer's Rule

Consider the solution to a linear system of equations:

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}, \tag{WB.71}$$

or

$$Ax = b, \tag{WB.72}$$

where A is a nonsingular square $n \times n$ matrix, x is an $n \times 1$ vector of the unknowns, and b is also an $n \times 1$ vector. The solution can be expressed in terms of the ratio of the two determinants

$$x_i = \frac{\det(A_i)}{\det(A)}, \qquad i = 1, 2, \ldots, n \tag{WB.73}$$

where A_i is the matrix A with its ith column replaced by b.