Chapter 2
Mathematical Foundations

This chapter contains a collection of useful mathematical concepts and tools, which are useful, directly or indirectly, for the subsequent development to be covered in the main portion of the book. While much of the material is standard and can be found in classical textbooks, we also present a number of useful items that are not commonly found elsewhere. Essentially, this chapter serves as a brief overview and as a convenient reference when necessary.

2.1 Introduction

Hybrid systems are certainly pervasive today. Recently, we have witnessed a resurgence in examining quantization effects and a heightened interest in analog computation. There has also been recent progress in analyzing switched, hierarchical, and discretely controlled continuous-variable systems. It is time to focus on developing formal modeling, analysis, and control methodologies for *hybrid systems*. Therefore, hybrid systems research [357–359] is devoted to modeling, design, and validation of interacting systems of continuous process and computer programs. Therefore, the identifying characteristic of hybrid systems is that they incorporate both continuous components, usually called plants, which are governed by ordinary or functional differential equations, and also digital components such as digital computers, sensors, and actuators controlled by programs. Moreover, the growing demands for control systems that are capable of controlling complex nonlinear continuous plants with discrete intelligent controllers can be addressed by the method of hybrid systems.

Throughout this book, by a switched system we mean a class of hybrid dynamical systems consisting of a family of continuous-time subsystems and a rule that orchestrates the switching between them. An integral part of this book surveys recent developments in three basic problems regarding stability and design of switched systems. These problems are:

- stability for arbitrary switching sequences,
- stability for certain useful classes of switching sequences, and
- construction of stabilizing switching sequences.

M.S. Mahmoud, *Switched Time-Delay Systems*,
DOI 10.1007/978-1-4419-6394-9_2,

We also provide motivation for studying these problems within the framework of time-delay systems. In practice, many systems encountered exhibit switching between several subsystems (are inherently multimodal) that is dependent on various environmental factors. Another source of motivation for studying switched systems comes from the rapidly developing area of switching control. Control techniques based on switching between different controllers have been applied extensively in recent years, particularly in the adaptive context, where they have been shown to achieve stability and improve transient response. The importance of such control methods also stems in part from the existence of systems that cannot be asymptotically stabilized by a single continuous feedback control law. Additionally, the fact that some of intelligent control methods are based on the idea of switching between different controllers. The existence of systems that cannot be asymptotically stabilized by a single static continuous feedback controller [47] also motivates the study. A survey of basic problems in stability and design of switched systems is given in [193].

In this book, we treat switched systems as a class of hybrid systems consisting of a family of subsystems and a switching law that specifies which subsystem will be activated along the system trajectory at each instant of time. Switched systems deserve investigation for theoretical development as well as for practical applications. To switch between different system structures is an essential feature of many control systems, for example, in power systems and power electronics [47]. There have been many studies for switched systems without uncertainties, primarily on stability analysis and design [358]. But for robust stability analysis of uncertain switched systems, there has been comparatively little work. A notable exception is the study of quadratic stability and stabilization by state-based feedback for both continuous-time and discrete-time switched linear systems composed of polytopic uncertain systems in [357]. For performance analysis of switched systems, authors of [357] investigated the disturbance attenuation properties of time-controlled switched systems consisting of several linear time invariant subsystems by using an average dwell-time approach incorporated with a piecewise Lyapunov function. Reference [133] computed the $\mathcal{L}_2$-induced norm of a switched linear system when the interval between consecutive switching is large. However, uncertainty is not considered in these two papers although it is ubiquitous in the system model due to the complexity of the system itself, exogenous disturbance, measurement errors, and so on. During the past decade, there have also been many papers concerning robust (or quadratic) stability, stabilization, and robust $\mathcal{H}_\infty$ control of uncertain systems without switchings [331, 441].

2.2 Basic Mathematical Concepts

Let $x_j,\ \ y_j,\ \ j = 1, 2, \ldots, n \in \Re(or\ \mathbf{C})$. Then the n-dimensional vectors $x,\ \ y$ are defined by $x = [x_1\ x_2\ \ldots\ x_n]^t,\ \ y = [y_1\ y_2\ \ldots\ y_n]^t\ \in \Re^n$, respectively.

A nonempty set $\mathcal{X}$ of elements $x, \; y, \ldots$ is called the *real* (*or complex*) *vector space* (*or real* (*complex*) *linear space*) by defining two algebraic operations, *vector additions and scalar multiplication, in* $x = [x_1, \; x_2, \; \ldots, \; x_n]^t$ [46]

2.2.1 Euclidean Space

The n-dimensional Euclidean space, denoted in the sequel by $\Re^n$ is the linear vector space $\Re^n$ equipped by the inner product

$$\langle x, y\rangle = x^t \; y = \sum_{j=1}^{n} x_j y_j$$

Let $\mathcal{X}$ be a linear space over the *field* **F** (typically **F** is the field of real numbers $\Re$ or complex numbers **C**). Then a function

$$||.|| : \mathcal{X} \rightarrow \Re$$

that maps $\mathcal{X}$ into the real numbers $\Re$ is a norm on $\mathcal{X}$ iff

1. $||x|| \geq 0, \; \forall x \in \mathcal{X}$ (nonnegativity)
2. $||x|| = 0, \; \Longleftrightarrow x = 0$ (positive definiteness)
3. $||\alpha \; x|| = |\alpha|||x||\forall x \in \mathcal{X}$ (homogeneity with respect to $|\alpha|$)
4. $||x + y|| \leq ||x|| + ||y||, \quad \forall x, y \in \mathcal{X}$ (triangle inequality)

Given a linear space $\mathcal{X}$, there are many possible norms on it. For a given norm $||.||$ *on* $\mathcal{X}$, the pair $(\mathcal{X}, \; ||.||)$ is used to indicate $\mathcal{X}$ endowed with the norm $||.||$.

2.2.2 Norms of Vectors

The class of L_p-norms is defined by

$$||x||_p = \Big(\sum_{j=1}^{n} |x_j|^p\Big)^{1/p}, \quad f \text{ or } 1 \leq p < \infty$$

$$||x||_\infty = \max 1 \leq j \leq n \; |x_j|$$

The three most commonly used norms are $||x||_1, \ ||x||_2$, and $||x||_\infty$. All p-norms are equivalent in the sense that if $||x||_{p1}$ and $||x||_{p2}$ are two different p-norms, then there exist positive constants c_1 and c_s such that

$$c_1 \ ||x||_{p1} \ \le \ ||x||_{p2} \ c_2 \ ||x||_{p1}, \qquad \forall x \in \Re^n$$

2.2.2.1 Induced Norms of Matrices

For a matrix $A \in \Re^{n \times n}$, the *induced p-norm* of A is defined by

$$||A||_p \overset{\Delta}{=} \sup_{x \neq 0} \frac{||Ax||_p}{||x||_p} = \sup_{||x||_p = 1} \ ||Ax||_p$$

Obviously, for matrices $A \in \Re^{m \times n}$ and $A \in \Re^{n \times r}$, we have *the triangle inequality*:

$$||A + B||_p \ \le \ ||A|||_p + ||B||_p$$

It is easy to show that the *induced norms* are also equivalent in the same sense as for the vector norms, and satisfying

$$||AB||_p \ \le \ ||Ax||_p \ ||B||_p, \ \ \forall A \in \Re^{n \times m}, \ \ B \in \Re^{m \times r}$$

which is known as the *submultiplicative property*. For $p = 1, 2, \ldots \infty$, we have the corresponding induced norms as follows:

$$||A||_1 = \max_j \sum_{s=1}^{n} \ |a_{sj}|, \quad \text{(column sum)}$$
$$||A||_2 = \max_j \sqrt{\lambda_j (A^t A)}$$
$$||A||_\infty = \max_s \sum_{j=1}^{n} \ |a_{sj}|, \quad \text{(row sum)}$$

2.2.3 *Convex Sets*

A set $\mathbf{S} \subset \Re^n$ is said to be *open* if every vector $x \in \mathbf{S}$, there is an ϵ-neighborhood of x

$$\mathcal{N}(x, \epsilon) = \{z \in \ \Re^n |||z - x|| \ < \ \epsilon\}$$

such that $\mathcal{N}(x, \epsilon) \subset \mathbf{S}$.

A set is *closed* iff its complement in $\Re^n$ is open; *bounded* if there is $r > 0$ such that $||x|| < r, \; \forall x \in \mathbf{S}$; and *compact* if it is closed and bounded; *convex* if for every $x, y \in \mathbf{S}$, and every real number $\alpha, \; 0 < \alpha < 1$, the point $\alpha \; x + (1 - \alpha) \; x \in \mathbf{S}$.

A set $\mathbf{K} \subset \Re^n$ is said to be *convex* if for any two vectors x and y in $\mathbf{K}$ any vector of the form $(1 - \lambda)x + \lambda y$ is also in $\mathbf{K}$, where $0 \leq \lambda \leq 1$. This simply means that given two points in a convex set, the line segment between them is also in the set. Note, in particular, that subspaces and linear varieties (a linear variety is a translation of linear subspaces) are convex. Also the empty set is considered convex. The following facts provide important properties for convex sets .

1. Let $\mathcal{C}_j, \; j = 1, \ldots, m$ be a family of m convex sets in $\Re^n$. Then the intersection $\mathcal{C}_1 \cap \mathcal{C}_2 \cap ... \cap \mathcal{C}_m$.
2. Let $\mathcal{C}$ be a convex set in $\Re^n$ and $x_o \in \Re^n$. Then the set $\{x_o + x : x \in \mathcal{C}\}$ is convex.
3. A set $\mathbf{K} \subset \Re^n$ is said to be *convex cone* with vertex x_o if $\mathbf{K}$ is convex, and $x \in \mathbf{K}$ implies that $x_o + \lambda x \in \mathbf{K}$ for any $\lambda \geq 0$.

An important class of convex cones is the one defined by the positive semidefinite ordering of matrices, that is, $A_1 \geq A_2 \geq A_3$. Let $P \in \Re^{n \times n}$ be a positive semidefinite matrix. The set of matrices $X \in \Re^{n \times n}$, such that $X \geq P$ is a convex cone in $\Re^{n \times n}$.

2.2.4 Continuous Functions

A function $f : \Re^n \longrightarrow \Re^m$ is said to be *continuous* at a point x if $f(x + \delta x) \longrightarrow f(x)$ whenever $\delta x \longrightarrow 0$. Equivalently, f is continuous at x if, given $\epsilon > 0$, there is $\delta > 0$ such that

$$||x - y|| < \epsilon \implies ||f(x) - f(y)|| < \epsilon$$

A function f is continuous on a set of $\mathbf{S}$ if it is continuous at every point of $\mathbf{S}$, and it is uniformly continuous on $\mathbf{S}$ if given $\epsilon > 0$, there is $\delta(\epsilon) > 0$ (dependent only on ϵ), such that the inequality holds for all $x, y \in \mathbf{S}$

A function $f : \Re \longrightarrow \Re$ is said to be *differentiable* at a point x if the limit

$$\dot{f}(x) = \lim_{\delta x \to 0} \frac{f(x + \delta x) - f(x)}{\delta x}$$

exists. A function $f : \Re^n \longrightarrow \Re^m$ is *continuously differentiable* at a point x (a set $\mathbf{S}$) if the partial derivatives $\partial f_j / \partial x_s$ exist and continuous at x (at every point of $\mathbf{S}$) for $1 \leq j \leq m, \; 1 \leq s \leq n$ and the *Jacobian matrix* is defined as

$$\mathbf{J} = \left[\frac{\partial f}{\partial x}\right] = \begin{bmatrix} \partial f_1/\partial x_1 & \cdots & \partial f_1/\partial x_n \\ \vdots & \ddots & \vdots \\ \partial f_m/\partial x_1 & \cdots & \partial f_m/\partial x_n \end{bmatrix} \in \Re^{m \times n}$$

2.2.5 Function Norms

Let $f(t) : \Re_+ \longrightarrow \Re$ be a continuous function or piecewise continuous function. The p-norm of f is defined by

$$||f||_p = \left(\int_0^\infty |f(t)|^p \, dt \right)^{1/p}, \quad f \text{ or } p \in [1, \infty)$$
$$||f||_\infty = \sup \; t \in [0, \infty)|f(t)|, \quad f \text{ or } p = \infty$$

By letting $p = 1, 2, \infty$, the corresponding normed spaces are called $\mathbf{L_1}$, $\mathbf{L_2}$, $\mathbf{L_\infty}$, respectively. More precisely, let $f(t)$ be a function on $[0, \infty)$ of the signal spaces, they are defined as

$$\mathbf{L_1} \stackrel{\Delta}{=} \left\{ f(t) : \Re_+ \longrightarrow \Re | ||f||_1 = \int_0^\infty |f(t)| \, dt < \infty, \text{ convolution kernel} \right\}$$
$$\mathbf{L_2} \stackrel{\Delta}{=} \left\{ f(t) : \Re_+ \longrightarrow \Re | ||f||_2 = \int_0^\infty |f(t)|^2 \, dt < \infty, \text{ finite energy} \right\}$$
$$\mathbf{L_\infty} \stackrel{\Delta}{=} \left\{ f(t) : \Re_+ \longrightarrow \Re | ||f||_\infty = \sup_{t \in [0,\infty)} |f(t)| < \infty, \text{ bounded signal} \right\}$$

From a signal point of view, the 1-*norm*, $||x||_1$ of the signal $x(t)$ is the integral of its absolute value, the square $||x||_2^2$ of the *2-norm* is often called the energy of the signal $x(t)$, and the ∞-*norm* is its absolute maximum amplitude or peak value. It must be emphasized that the definitions of the norms for vector functions are not unique.

In the case of $f(t) : \Re_+ \longrightarrow \Re^n$, $f(t) = [f_1(t) \quad f_2(t) \ldots f_n(t)]^t$ which denote a continuous function or piecewise continuous vector function, the corresponding p-norm spaces are defined as

$$\mathbf{L_p^n} \stackrel{\Delta}{=} \left\{ f(t) : \Re_+ \longrightarrow \Re^n | ||f||_p = \int_0^\infty ||f(t)||^p \, dt < \infty, \; f \text{ or } p \in [1, \infty) \right\}$$
$$\mathbf{L_\infty^n} \stackrel{\Delta}{=} \left\{ f(t) : \Re_+ \longrightarrow \Re^n | ||f||_\infty = \sup_{t \in [0,\infty)} ||f(t)|| < \infty \right\}$$

2.3 Calculus and Algebra of Matrices

In this section, we solicit some basic facts and useful relations from linear algebra and calculus of matrices. The materials are stated along with some hints whenever needed but without proofs unless we see the benefit of providing a proof. Reference is made to matrix M or matrix function $M(t)$ in the form

$$M = \begin{bmatrix} M_{11} & \cdots & M_{1n} \\ \vdots & \ddots & \cdots \\ M_{m1} & \cdots & M_{mn} \end{bmatrix}, \quad \text{or} \quad M(t) = \begin{bmatrix} M_{11}(t) & \cdots & M_{1n}(t) \\ \vdots & \ddots & \cdots \\ M_{m1}(t) & \cdots & M_{mn}(t) \end{bmatrix}$$

2.3.1 Fundamental Subspaces

A nonempty subset $\mathcal{G} \subset \Re^n$ is called a *linear subspace* of $\Re^n$ if $x+y$ and αx are in $\mathcal{G}$ whenever x and y are in $\mathcal{G}$ for any scalar α. A set of elements $X = \{x_1, \ x_2, \ \ldots, \ x_n\}$ is said to be a *spanning set* for a linear subspace $\mathcal{G}$ of $\Re^n$ if every element $g \in \mathcal{G}$ can be written as a linear combination of the $\{x_j\}$. That is, we have

$$\mathcal{G} = \{g \in \Re : \ g = \alpha_1 x_1 + \alpha_2 x_2 + \ \ldots \ \alpha_n x_n$$

for some scalars $\alpha_1, \ \alpha_2, \ldots, \ \alpha_n$.

A spanning set X is said to be a *basis* for $\mathcal{G}$ if no element x_j of the spanning set X of $\mathcal{G}$ can written as a linear combination of the remaining elements $x_1, \ x_2, \ldots, \ x_{j-1}, \ x_{j+1}, \ldots, \ x_n$, that is, $x_j, \ 1 \le i \le n$ form a linearly independent set. It is frequent to use $x_j = [0 \ 0 \ \ldots \ 0 \ 1 \ 0 \ \ldots \ 0]^t$ the kth unit vector.

The geometric ideas of linear vector spaces had led to the concepts of *spanning a space* and a *basis for a space*. The idea now is to introduce four important subspaces which are useful. The entire linear vector space of a specific problem can be decomposed into the sum of these subspaces.

The *column space* of a matrix $A \in Re^{n \times m}$ is the space spanned by the columns of A, also called the *range space* of A, denoted by $\mathcal{R}[A]$. Similarly, the *row space* of A is the space spanned by the rows of A. Since the column rank of a matrix is the dimension of the space spanned by the columns and the row rank is the dimension of the space spanned by the rows, it is clear that the spaces $\mathcal{R}[A]$ and $\mathcal{R}[A^t]$ have the same dimension $r = \text{rank}(A)$.

The *right null space* of $A \in Re^{n \times m}$ is the space spanned by all vectors x that satisfy $A \ x = 0$, and is denoted by $\mathcal{N}[A]$. The right null space of A is also called the *kernel* of A. The *left null space* of A is the space spanned by all vectors y that satisfy $y^t \ A = 0$. This space is denoted by $\mathcal{N}[A^t]$, since it is also characterized by all vectors y such that $A^t \ y = 0$.

The dimensions of the four spaces $\mathcal{R}[A], \ \mathcal{R}[A^t], \ \mathcal{N}[A]$, and $\mathcal{N}[A^t]$ are to be determined in the sequel. Since $A \in \Re^{n \times m}$, we have the following

$$\begin{aligned} r \ &\stackrel{\Delta}{=} \ \text{rank}(A) = \text{dimension of column space } \mathcal{R}[A] \\ \dim \mathcal{N}[A] \ &\stackrel{\Delta}{=} \ \text{dimension of right null space } \mathcal{N}[A] \\ n \ &\stackrel{\Delta}{=} \ \text{total number of columns of } A \end{aligned}$$

Hence the dimension of the null space $\dim \mathcal{N}[A] = n - r$. Using the fact that $\text{rank}(A) = \text{rank}(A^t)$, we have

$$\begin{aligned} r &\triangleq \text{rank}(A^t) = \text{dimension of row space } \mathcal{R}[A^t] \\ \dim \mathcal{N}[A^t] &\triangleq \text{dimension of left null space } \mathcal{N}[A^t] \\ m &\triangleq \text{total number of rows of } A \end{aligned}$$

Hence the dimension of the null space $\dim \mathcal{N}[A^t] = m - r$. These facts are summarized below.

Note from these facts that the entire n-dimensional space can be decomposed into the sum of the two subspaces $\mathcal{R}[A^t]$ and $\mathcal{N}[A]$. Alternatively, the entire m-dimensional space can be decomposed into the sum of the two subspaces $\mathcal{R}[A]$ and $\mathcal{N}[A^t]$.

An important property is that $\mathcal{N}[A]$ and $\mathcal{R}[A^t]$ are *orthogonal subspaces*, that is, $\mathcal{R}[A^t]^\perp = \mathcal{N}[A]$. This has the meaning that every vector in $\mathcal{N}[A]$ is orthogonal to every vector in $\mathcal{R}[A^t]$. In the same manner, $\mathcal{R}[A]$ and $\mathcal{N}[A^t]$ are *orthogonal subspaces*, that is, $\mathcal{R}[A]^\perp = \mathcal{N}[A^t]$. The construction of the fundamental subspaces is appropriately attained by the singular value decomposition.

$$\begin{aligned} \mathcal{R}[A^t] &\triangleq \text{row space of } A: \text{dimension } r \\ \mathcal{N}[A] &\triangleq \text{right null space of } A: \text{dimension } n - r \\ \mathcal{R}[A] &\triangleq \text{column space of } A: \text{dimension } r \\ \mathcal{N}[A^t] &\triangleq \text{left null space of } A: \text{dimension } n - r \end{aligned}$$

2.3.2 Calculus of Vector–Matrix Functions of a Scalar

The differentiation and integration of time functions involving vectors and matrices arise in solving state equations, optimal control, and so on. This section summarizes the basic definitions of differentiation and integration on vectors and matrices . A number of formulas for the derivative of vector–matrix products are also included.

The derivative of a matrix function $M(t)$ of a scalar is the matrix of the derivatives of each element in the matrix

$$\frac{\mathrm{d}M(t)}{\mathrm{d}t} = \begin{bmatrix} \frac{\mathrm{d}M_{11}(t)}{\mathrm{d}t} & \cdots & \frac{\mathrm{d}M_{1n}(t)}{\mathrm{d}t} \\ \vdots & \ddots & \cdots \\ \frac{\mathrm{d}M_{m1}(t)}{\mathrm{d}t} & \cdots & \frac{\mathrm{d}M_{mn}(t)}{\mathrm{d}t} \end{bmatrix}$$

The integral of a matrix function $M(t)$ of a scalar is the matrix of the integral of each element in the matrix

$$\int_a^b M(t)\mathrm{d}t = \begin{bmatrix} \int_a^b M_{11}(t)\mathrm{d}t & \cdots & \int_a^b M_{1n}(t)\mathrm{d}t \\ \vdots & \ddots & \cdots \\ \int_a^b M_{m1}(t)\mathrm{d}t & \cdots & \int_a^b M_{mn}(t)\mathrm{d}t \end{bmatrix}$$

The Laplace transform of a matrix function $M(t)$ of a scalar is the matrix of the Laplace transform of each element in the matrix

$$\int_a^b M(t)e^{-st}\mathrm{d}t = \begin{bmatrix} \int_a^b M_{11}(t)e^{-st}\mathrm{d}t & \cdots & \int_a^b M_{1n}(t)e^{-st}\mathrm{d}t \\ \vdots & \ddots & \cdots \\ \int_a^b M_{m1}(t)e^{-st}\mathrm{d}t & \cdots & \int_a^b M_{mn}(t)e^{-st}\mathrm{d}t \end{bmatrix}$$

The scalar derivative of the product of two matrix time functions is

$$\frac{\mathrm{d}(A(t)B(t))}{\mathrm{d}t} = \frac{A(t)}{\mathrm{d}t}B(t) + A(t)\frac{B(t)}{\mathrm{d}t}$$

This result is analogous to the derivative of a product of two scalar functions of a scalar, except caution must be used in reserving the order of the product. An important special case follows:

The scalar derivative of the inverse of a matrix time function is

$$\frac{\mathrm{d}A^{-1}(t)}{\mathrm{d}t} = -A^{-1}\frac{A(t)}{\mathrm{d}t}A(t)$$

2.3.3 Derivatives of Vector–Matrix Products

The derivative of a real scalar-valued function $f(x)$ of a real vector $x = [x_1, \ldots, x_n]^t \in Re^n$ is defined by

$$\frac{\partial f(x)}{\partial x} = \begin{bmatrix} \frac{\partial f(x)}{\partial x_1} \\ \frac{\partial f(x)}{\partial x_2} \\ \vdots \\ \frac{\partial f(x)}{\partial x_n} \end{bmatrix}$$

where the partial derivative is defined by

$$\frac{\partial f(x)}{\partial x_j} \stackrel{\Delta}{=} \lim_{\Delta x_j \to 0} \frac{f(x + \Delta x) - f(x)}{\Delta x_j}, \quad \Delta x = [0 \ldots \Delta x_j \ldots 0]^t$$

An important application arises in the Taylor's series expansion of $f(x)$ about x_o in terms of $\delta x \stackrel{\Delta}{=} x - x_o$. The first three terms are

$$f(x) = f(x_o) + \left(\frac{\partial f(x)}{\partial x}\right)^t \delta x + \frac{1}{2}\delta x^t \left[\frac{\partial^2 f(x)}{\partial x^2}\right] \delta x$$

where

$$\frac{\partial f(x)}{\partial x} = \begin{bmatrix} \frac{\partial f(x)}{\partial x_1} \\ \vdots \\ \frac{\partial f(x)}{\partial x_n} \end{bmatrix}$$

$$\frac{\partial^2 f(x)}{\partial x^2} = \frac{\partial}{\partial x}\left(\frac{\partial f(x)}{\partial x}\right)^t = \begin{bmatrix} \frac{\partial^2 f(x)}{\partial x_1^2} & \cdots & \frac{\partial^2 f(x)}{\partial x_1 \partial x_n} \\ \vdots & \ddots & \cdots \\ \frac{\partial^2 f(x)}{\partial x_n \partial x_1} & \cdots & \frac{\partial^2 f(x)}{\partial x_n^2} \end{bmatrix}$$

The derivative of a real scalar-valued function $f(A)$ with respect to a matrix

$$A = \begin{bmatrix} A_{11} & \cdots & A_{1n} \\ \vdots & \ddots & \cdots \\ A_{n1} & \cdots & A_{nn} \end{bmatrix} \in Re^{n \times n}$$

is given by

$$\frac{\partial f(A)}{\partial A} = \begin{bmatrix} \frac{\partial f(A)}{\partial A_{11}} & \cdots & \frac{\partial f(A)}{\partial A_{1n}} \\ \vdots & \ddots & \cdots \\ \frac{\partial f(A)}{\partial A_{n1}} & \cdots & \frac{\partial f(A)}{\partial A_{nn}} \end{bmatrix}$$

A vector function of a vector is given by

$$v(u) = \begin{bmatrix} v_1(u) \\ \vdots \\ \vdots \\ v_n(u) \end{bmatrix}$$

where $v_j(u)$ is a function of the vector u. The derivative of a vector function of a vector (the *Jacobian*) is defined as follows:

$$\frac{\partial v(u)}{\partial u} = \begin{bmatrix} \frac{\partial v_1(u)}{\partial u_1} & \cdots & \frac{\partial v_1(u)}{\partial u_m} \\ \vdots & \ddots & \cdots \\ \frac{\partial v_n(u)}{\partial u_1} & \cdots & \frac{\partial v_n(u)}{\partial u_m} \end{bmatrix}$$

Note that the *Jacobian* is sometimes defined as the transpose of the foregoing matrix. A special case is given by

$$\frac{\partial (S\,u)}{\partial u} = S, \quad \frac{\partial (u^t R u)}{\partial u} = 2\,u^t R$$

for arbitrary matrix S and symmetric matrix R.

The following section includes useful relations and results from linear algebra.

2.3.4 The Dini Theorem

2.3.5 Positive Definite and Positive Semidefinite Matrices

A matrix P is positive definite if P is real, symmetric, and $x^t P x > 0, \quad \forall x \neq 0$. Equivalently, all the eigenvalues of P have positive real parts. A matrix S is positive semidefinite if S is real, symmetric, and $x^t P x \geq 0, \quad \forall x \neq 0$.

Since the definiteness of the scalar $x^t P x$ is a property only of the matrix P, we need a test for determining definiteness of a constant matrix P. Define a *principal submatrix* of a square matrix P as any square submatrix sharing some diagonal elements of P. Thus the constant, real, symmetric matrix $P \in \Re^{n \times n}$ is positive definite ($P > 0$) if either of these equivalent conditions holds:

- All eigenvalues of P are positive
- The determinant of P is positive
- All successive principal submatrices of P (minors of successively increasing size) have positive determinants

2.3.6 Trace Properties

The trace of a square matrix P, trace (P), equals the sum of its diagonal elements or equivalently the sum of its eigenvalues. A basic property of the trace is invariant under cyclic perturbations, that is,

$$\text{trace}(AB) = \text{trace}(BA)$$

where AB is square. Successive applications of the above results yield

$$\text{trace}(ABC) = \text{trace}(BCA) = \text{trace}(CAB)$$

where ABC is square. In general,

$$\text{trace}(AB) = \text{trace}(B^t A^t)$$

Another result is that

$$\text{trace}(A^t BA) = \sum_{k=1}^{p} a_k^t B a_k$$

where $A \in \Re^{n\times p}$, $B \in \Re^{n\times n}$, and $\{a_k\}$ are the columns of A. The following identities on trace derivatives are noted:

$$\frac{\partial(\text{trace}(AB))}{\partial A} = \frac{\partial(\text{trace}(A^t B^t))}{\partial A} = \frac{\partial(\text{trace}(B^t A^t))}{\partial A} = \frac{\partial(\text{trace}(BA))}{\partial A} = B^t$$

$$\frac{\partial(\text{trace}(AB))}{\partial B} = \frac{\partial(\text{trace}(A^t B^t))}{\partial B} = \frac{\partial(\text{trace}(B^t A^t))}{\partial B} = \frac{\partial(\text{trace}(BA))}{\partial B} = A^t$$

$$\frac{\partial(\text{trace}(BAC))}{\partial A} = \frac{\partial(\text{trace}(B^t C^t A^t))}{\partial A} = \frac{\partial(\text{trace}(C^t A^t B^t))}{\partial A} = \frac{\partial(\text{trace}(ACB))}{\partial A} = \frac{\partial(\text{trace}(CBA))}{\partial A} = \frac{\partial(\text{trace}(A^t B^t C^t))}{\partial A} = B^t\ C^t$$

$$\frac{\partial(\text{trace}(A^t BA))}{\partial A} = \frac{\partial(\text{trace}(BAA^t))}{\partial A} = \frac{\partial(\text{trace}(AA^t B))}{\partial A} = (B + B^t)A$$

Using these basic ideas, a list of matrix calculus results are given below:

$$\frac{\partial(\text{trace}(AX^t))}{\partial X} = A, \quad \frac{\partial(\text{trace}(AXB))}{\partial X} = A^t\ B^t$$

$$\frac{\partial(\text{trace}(AX^t B))}{\partial X} = B\ A, \quad \frac{\partial(\text{trace}(AX))}{\partial X^t} = A$$

$$\frac{\partial(\text{trace}(AX^t))}{\partial X^t} = A^t, \quad \frac{\partial(\text{trace}(AXB))}{\partial X^t} = B\ A$$

$$\frac{\partial(\text{trace}(AX^t B))}{\partial X^t} = A^t\ B^t, \quad \frac{\partial(\text{trace}(XX))}{\partial X} = 2\ X^t$$

$$\frac{\partial(\text{trace}(XX^t))}{\partial X} = 2\ X$$

$$\frac{\partial(\text{trace}(AX^n))}{\partial X} = \left(\sum_{j=0}^{n-1} X^j\ A\ X^{n-j-1}\right)^t$$

$$\frac{\partial(\text{trace}(AXBX))}{\partial X} = A^t X^t B^t + B^t X^t A^t$$
$$\frac{\partial(\text{trace}(AXBX^t))}{\partial X} = A^t X B^t + AXB$$
$$\frac{\partial(\text{trace}(X^{-1}))}{\partial X} = -\left(X^{-2}\right)^t$$
$$\frac{\partial(\text{trace}(AX^{-1}B))}{\partial X} = -\left(X^{-1}BAX^{-1}\right)^t$$
$$\frac{\partial(\text{trace}(AB))}{\partial A} = B^t + B - \text{diag}(B)$$

2.3.7 Partitioned Matrices

Given a partitioned matrix (matrix of matrices) of the form

$$M = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$$

where $A, \ B, \ C$, and D are of compatible dimensions. Then

(1) if A^{-1} exists, a Schur complement of M is defined as $D - CA^{-1}B$, and
(2) if D^{-1} exists, a Schur complement of M is defined as $A - BD^{-1}C$.

When $A, \ B, \ C$, and D are all $n \times n$ matrices, then

$$a) \quad \det \begin{bmatrix} A & B \\ C & D \end{bmatrix} = \det(A)\det(D - CA^{-1}B), \ \det(A) \neq 0$$
$$b) \quad \det \begin{bmatrix} A & B \\ C & D \end{bmatrix} = \det(D)\det(A - BD^{-1}C), \ \det(D) \neq 0$$

In the special case, we have

$$\det \begin{bmatrix} A & B \\ C & 0 \end{bmatrix} = \det(A)\det(C)$$

where A and C are square. Since the determinant is invariant under row, it follows

$$\det\begin{bmatrix} A & B \\ C & D \end{bmatrix} = \det\begin{bmatrix} A & B \\ C - CA^{-1}A & D - CA^{-1}B \end{bmatrix}$$
$$= \det\begin{bmatrix} A & B \\ 0 & D - CA^{-1}B \end{bmatrix} = \det(A)\det(D - CA^{-1}B)$$

which justifies the forgoing result.

Given matrices $A \in \Re^{m\times n}$ and $B \in \Re^{n\times m}$, then

$$\det(I_m - AB) = \det(I_n - BA)$$

In case that A is invertible, then $\det(A^{-1}) = \det(A)^{-1}$.

2.3.8 The Matrix Inversion Lemma

Suppose that $A \in \Re^{n\times n}$, $B \in \Re^{n\times p}$, $C \in \Re^{p\times p}$, and $D \in \Re^{p\times n}$. Assume that A^{-1} and C^{-1} both exist, then

$$(A + BCD)^{-1} = A^{-1} - A^{-1}B(DA^{-1}B + C^{-1})^{-1}DA^{-1}$$

In the case of partitioned matrices, we have the following result

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix}^{-1} = \begin{bmatrix} A^{-1} + A^{-1}B\Xi^{-1}CA^{-1} & -A^{-1}B\Xi^{-1} \\ -\Xi^{-1}CA^{-1} & \Xi^{-1} \end{bmatrix}$$
$$\Xi = (D - CA^{-1}B)$$

provided that A^{-1} exists. Alternatively,

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix}^{-1} = \begin{bmatrix} \Xi^{-1} & -\Xi^{-1}BD^{-1} \\ -D^{-1}C\Xi^{-1} & D^{-1} + D^{-1}C\Xi^{-1}BD^{-1} \end{bmatrix}$$
$$\Xi = (D - CA^{-1}B)$$

provided that D^{-1} exists.

For a square matrix Y, the matrices Y and $(I + Y)^{-1}$ commute, that is, given that the inverse exists

$$Y\ (I + Y)^{-1} = (I + Y)^{-1}\ Y$$

Two additional inversion formulas are given below:

$$\begin{aligned} Y\,(I + XY)^{-1} &= (I + YX)^{-1}\,Y \\ (I + YX)^{-1} &= I - YX\,(I + YX)^{-1} \end{aligned}$$

The following result provides conditions for the positive definiteness of a partitioned matrix in terms of its submatrices. The following three statements are equivalent:

$$\begin{aligned} &1)\ \begin{bmatrix} A_o & A_a \\ A_a^t & A_c \end{bmatrix} > 0 \\ &2)\ A_c > 0, \quad A_o - A_a A_c^{-1} A_a^t > 0 \\ &3)\ A_a > 0, \quad A_c - A_a^t A_o^{-1} A_a > 0 \end{aligned}$$

2.3.9 The Singular Value Decomposition

The singular value decomposition (SVD) is a matrix factorization that has found a number of applications to engineering problems. The SVD of a matrix $M \in Re^{n \times m}$ is

$$M = U\ S\ V^\dagger = \sum_{j=1}^{p} \sigma_j\ U_j\ V_j^\dagger$$

where $U \in Re^{\alpha \times \alpha}$ and $V \in Re^{\beta \times \beta}$ are unitary matrices ($U^\dagger\ U = U\ U^\dagger = I$ and $V^\dagger\ V = V\ V^\dagger I$); $S \in Re^{\alpha \times \beta}$ is a real, diagonal (but not necessarily square); and $pmin(\alpha, \beta)$*. The singular values* $\{\sigma_1, \sigma_2, \ldots, \sigma_\beta\}$ *of M are defined as the positive square roots of the diagonal elements of* $S^t S$, and are ordered from largest to smallest.

To proceed further, we recall a result on unitary matrices. If U is a unitary matrix ($U^\dagger\ U = I$), then the transformation U preserves length, that is,

$$\begin{aligned} ||U\ x|| &= \sqrt{(Ux)^\dagger (Ux)} = \sqrt{x^\dagger\ U^\dagger\ U\ x} \\ &= \sqrt{x^\dagger\ x} = ||x|| \end{aligned}$$

As a consequence, we have

$$\begin{aligned} ||M\ x|| &= \sqrt{x^\dagger\ M^\dagger\ M\ x} = \sqrt{x^\dagger\ V S^t U^\dagger\ U S V^\dagger\ x} \\ &= \sqrt{x^\dagger\ V S^t S V^\dagger\ x} \end{aligned}$$

To evaluate the maximum gain of matrix M, we calculate the maximum norm of the above equation to yield

$$\max_{||x||=1} ||M\ x|| = \max_{||x||=1} \sqrt{x^{\dagger}\ VS^{t}SV^{\dagger}\ x} = \max_{||\tilde{x}||=1} \sqrt{\tilde{x}^{\dagger}\ VS^{t}S\ \tilde{x}}$$

Note that maximization over $\tilde{x} = Vx$ is equivalent to maximizing over x since V is invertible and preserves the norm (equals 1 in this case). Expanding the norm yields

$$\begin{aligned} \max_{||x||=1} ||M\ x|| &= \max_{||\tilde{x}||=1} \sqrt{\tilde{x}^{\dagger}\ VS^{t}S\ \tilde{x}} \\ &= \max_{||\tilde{x}||=1} \sqrt{\sigma_1^2|\tilde{x}_1|^2 + \sigma_2^2|\tilde{x}_2|^2 + \cdots + \sigma_\beta^2|\tilde{x}_\beta|^2} \end{aligned}$$

The foregoing expression is maximized, given the constraint $||\tilde{x}|| = 1$, when $\tilde{x}$ is concentrated at the largest singular value; that is, $|\tilde{x}| = [1\ 0\ \ldots\ 0]^t$. The maximum gain is then

$$\max_{||x||=1} ||M\ x|| = \sqrt{\sigma_1^2|1|^2 + \sigma_2^2|0|^2 + \cdots + \sigma_\beta^2|0|^2} = \sigma_1 = \sigma_M$$

In words, this reads *the maximum gain of a matrix is given by the maximum singular value* σ_M. Following similar lines of development, it is easy to show that

$$\begin{aligned} \min_{||x||=1} ||M\ x|| &= \sigma_\beta = \sigma_m \\ &= \begin{cases} \sigma_p & \alpha \geq \beta \\ 0 & \alpha < \beta \end{cases} \end{aligned}$$

A property of the singular values is expressed by

$$\sigma_M(M^{-1}) = \frac{1}{\sigma_m(M)}$$

2.4 Notes and References

The topics covered in this chapter is meant to provide the reader with a general platform containing the basic mathematical information needed for further examination of switched time-delay systems. These topics are properly selected from standard books and monographs on mathematical analysis. For further details, the reader is referred to the standard texts [29, 46, 157, 160, 443] where fundamentals are provided.