## Chapter 1

## Introduction

## Topics :

1. Motivation and Basic Concepts
2. Mathematical Formulation of the Control Problem
3. Examples
4. Matrix Theory (review)

Copyright © Claudiu C. Remsing, 2006.
All rights reserved. $\circ$

A control system can be viewed informally as a dynamical object (e.g. ordinary differential equation) containing a parameter (control) which can be manipulated to influence the behaviour of the system so as to achieve a desired goal. In order to implement this influence, engineers build devices that incorporate various mathematical techniques. Mathematical control theory is today a well-established branch of application-oriented mathematics that deals with the basic principles underlying the analysis and design of control systems.

### 1.1 Motivation and Basic Concepts

Mathematical control theory is a rapidly growing field which provides theoretical and computational tools for dealing with a variety of problems arising in electrical and aerospace engineering, automatics, robotics, management, economics, applied chemistry, biology, ecology, medicine, etc. Selected such problems, to mention but a few, are the following : stable performance of motors and machinery, optimal guidance of rockets, optimal exploitation of natural resources, optimal investment or production strategies, regulation of physiological functions, and fight against insects, epidemics.

All these (and many other) problems require a specific approach, the aim being to compel or control a system to behave in some desired fashion.

## Systems

A system is something having parts which is perceived as a single entity.
Note : Not everything is a system (for instance, a point or the empty set). However, most things can usefully be seen as systems of some kind. A system is, so to speak, a world.

The parts making up a system may be clearly or vaguely defined. The interesting thing about a system is the way the parts are related to each other. For the systems studied in mathematics, the parts and their relations must be so clearly defined that we can single out a particular set of relations as completely characterizing the state of the system; then we identify the system with the collection of all its conceivable states. It seems to be necessary that the state space be clearly and unambiguously defined. Unfortunately this usually means that the mathematical system is drastically oversimplified in comparison with the natural system being modelled.

When attempting to study the behaviour of certain systems, it is convenient to consider the ideal case of an "isolated system" - i.e. a number of interacting elements which do not have any interaction with the rest of the
world. In reality, no system is ever completely isolated, but in many cases the interactions with the rest of the world can be neglected in a reasonable approximation. In such "isolated systems" the conditions are simpler, and therefore easier to study.

Note : Our Universe is, by definition, an isolated system.

## Dynamical and control systems

A dynamical system is one which changes in time (in some well defined way); what changes is the state of the system. For such systems, the basic problem is to predict the future behaviour. For this purpose the differential equations are exactly tailored. The differential equation itself represents the (physical or otherwise) law governing the evolution of the system; this plus the initial conditions should determine uniquely the future evolution of the system.

Note : Philosophically this leads to determinism, and is independent of any (human) observer. Modern physics changed this view, to some extent, by making the observer a much more active participant in the possible outcome of future measurements. But even within classical physics, the prediction of future evolution is not the only meaningful problem to be posed. The whole field of engineering and technology deals, to a large extent, with the "inverse" problem : given a desired future evolution, how should we construct the system?

One can introduce some way of acting upon a (dynamical) system and influence its evolution (behaviour). We think of this outside action, also called input (or control), as the result of decisions of a "controller" (possibly human), who may have some definite goal in mind or not. But this last is irrelevant and the important information we need is a rule, within the description of the system, of which inputs are possible and which are not; the possible inputs will then be called "admissible". A simple example of such systems (with inputs) is a car, whose motion depends on the input of all the actions by which we drive it. In most cases, it is possible to change the state of the system in any
prescribed fashion by properly choosing the inputs, at least within reasonable limits. In other words, one may exert influence on the system state by means of intelligent manipulation of its inputs. This then, in a general sense, constitutes a control system.

The control engineer develops the techniques and hardware necessary for the implementation of the control laws to the specific systems in question.

Note : The complexity of many systems in present-day world is such that it is often desirable for control to be carried out automatically, without direct human intervention. To take a simple example, the room thermostat in a domestic central heating system turns the boiler on and off so as to maintain room temperature at a predetermined level. Nature provides many examples of remarkable self-regulation, such as the way in which body temperature is kept constant despite large variations in external conditions.

## Some basic control-theoretic concepts

We summarize some of the main features of a control system.
The state variables $x_{1}, x_{2}, \ldots, x_{m}$ describe the condition (or state) of the system, and provide the information which (together with the knowledge of the equations describing the system) enables us to calculate the future behaviour from the knowledge of the control variables (or inputs) $u_{1}, u_{2}, \ldots, u_{\ell}$. In practice, it is often not possible to determine the values of the state variables directly; instead, only a set of controlled variables (or outputs) $y_{1}, y_{2}, \ldots, y_{n}$, which depend in some way on the state variables, is measured. In general, the aim is to make a system perform in some required way by suitably manipulating the inputs, this being done by some controlling device (or "controller").

If the controller operates according to some pre-set pattern without taking account of the output or state, the system is called open loop. If, however, there is feedback of information concerning the outputs to controller, which then appropriately modifies its course of action, the system is called closed loop.

We assume that our system models have the property that, given an initial state and any input, the resulting state and output at some specified later time are uniquely determined.

### 1.2 Mathematical Formulation of the Control Problem

Roughly speaking, a control system is a dynamical system together with a class of "admissible inputs". We whish to make this idea more precise, without striving for full generality.

## Control systems

We assume that the dynamics of the system, that is, the evolution of the state vector $x(t)=\left[\begin{array}{c}x_{1}(t) \\ \vdots \\ x_{m}(t)\end{array}\right] \in \mathbb{R}^{m \times 1}=\mathbb{R}^{m}$ under a given input (or control) vector $u(t)=\left[\begin{array}{c}u_{1}(t) \\ \vdots \\ u_{\ell}(t)\end{array}\right] \in \mathbb{R}^{\ell \times 1}=\mathbb{R}^{\ell}$ is determined by a (vector) ordinary differential equation

$$
\begin{equation*}
\dot{x}=F(t, x, u) . \tag{1.1}
\end{equation*}
$$

The input vector $u(\cdot)$ is assumed to be an "arbitrary" vector-valued mapping, but some restrictions must be imposed. First of all, its components - the input functions $u_{1}(\cdot), u_{2}(\cdot), \ldots, u_{\ell}(\cdot)$ - must be measurable (think of piecewise-continuous functions), since otherwise the differential equation (1.1) wouldn't make sense.

Note : To restrict the control to be a continuous mapping, would be too much, since in many cases the piecewise-continuous inputs (with some points of discontinuity) are the most interesting controls.

Another restriction of the input vector is the requirement that the values of $u(\cdot)$ belong to a specified set $U$. (For example, when turning a steering wheel of a car, we are restricted to a maximum turning angle to either side.) Such a restriction is then of the form

$$
\begin{equation*}
u(t) \in U \subseteq \mathbb{R}^{\ell} . \tag{1.2}
\end{equation*}
$$

An admissible input is therefore a piecewise-continuous (vector-valued) mapping $u(\cdot)$ satisfying (1.2). We denote by $\mathcal{U}$ the set of all these admissible inputs.

Furthermore, it is assumed that the vector-valued mapping $F: J \times \mathbb{R}^{m} \times$ $U \rightarrow \mathbb{R}^{m}, J \subseteq \mathbb{R}$ satisfies certain standard conditions (such as having continuous first order partial derivatives).

NOTE : This assumption guarantees local existence and uniqueness of the solution of (1.1) (subject to initial condition $x\left(t_{0}\right)=x_{0}$ ) for a given $u(\cdot) \in \mathcal{U}$.

A control system is a 4 -tuple

$$
\Sigma=(M, U, \mathcal{U}, F) .
$$

In this case, the set $M=\mathbb{R}^{m}$ is the state space, the set $U \subseteq \mathbb{R}^{\ell}$ is the control set, $\mathcal{U}$ is the input class, and the mapping $F$ is the dynamics of $\Sigma$. We say that the control system $\Sigma$ is defined (or described) by the state equation (1.1) and write (in classical notation) :

$$
\Sigma: \quad \dot{x}=F(t, x, u), \quad x \in M, u \in U \subseteq \mathbb{R}^{\ell} .
$$

NOTE : (1) In fact, such a system is a continuous-time, time-varying, finite dimensional, differentiable (nonlinear) control system.
(2) The state space $M$ carries certain (geometric) "structure". It is natural to assume that $M$ is a differentiable manifold (think of an open subset of some Euclidean space). The dynamics $F$ is then best viewed as a family of (nonautonomous) vector fields on (the manifold) $M$, parametrized by controls.

The control system $\Sigma$ is linear if $U=\mathbb{R}^{\ell}$ and the dynamics $F: \mathbb{R} \times$ $\mathbb{R}^{m} \times \mathbb{R}^{\ell} \rightarrow \mathbb{R}^{m}$ has the form

$$
F(t, x, u)=A(t) x+B(t) u
$$

where $A(t) \in \mathbb{R}^{m \times m}$ and $B(t) \in \mathbb{R}^{m \times \ell}$ are matrices each of whose entries is a (continuous) function $\mathbb{R} \rightarrow \mathbb{R}$; that is, the dynamics $F$ is linear in $(x, u)$ for each fixed $t \in \mathbb{R}$.

One distinguishes controls of two types : open and closed loop. An open loop control can be basically an "arbitrary" function $u:\left[t_{0}, \infty\right) \rightarrow U$ for which the initial value problem (IVP)

$$
\dot{x}=F(t, x, u), \quad x\left(t_{0}\right)=x_{0}
$$

has a well defined solution.
A closed loop control can be identified with a mapping $k: M \rightarrow U$ (which may depend on $t \geq t_{0}$ ) such that the initial value problem (IVP)

$$
\dot{x}=F(t, x, k(x(\cdot))), \quad x\left(t_{0}\right)=x_{0}
$$

has a well defined solution. The mapping $k(\cdot)$ is called feedback.
One of the main aims of control theory is to find a strategy (input) such that the corresponding output has desired properties. Depending on the properties involved one gets more specific questions. Concepts like controllability, observability, stabilizability, realization, as well as optimality are fundamental in control theory.

## Controllability

One say that a state $x_{f} \in \mathbb{R}^{n}$ is reachable from $x_{0}$ in time $T$ if there exists an open loop control $u(\cdot)$ such that

$$
x(0)=x_{0} \quad \text { and } \quad x(T)=x_{f} .
$$

If an arbitrary state $x_{f}$ is reachable from an arbitrary state $x_{0}$ in time $T$, then the control system $\Sigma$ is called (completely) controllable. In several situations one requires a weaker property of transfering an arbitrary state into a given one, in particular the origin. A formulation of effective characterizations of controllable systems is an important task of control theory.

## Observability

In many situations of practical interest one observes not the state $x(\cdot)$ but its function $t \mapsto h(t, x(t)), t \geq t_{0}$. It is therefore often necessary to investigate the pair of equations (i.e. the state equation and the observation equation)

$$
\left\{\begin{array}{l}
\dot{x}=F(t, x, u) \\
y=h(t, x) .
\end{array}\right.
$$

This is a control system with outputs; that is, a 6 -tuple

$$
\Sigma=\left(M, U, \mathcal{U}, F, \mathbb{R}^{n}, h\right)
$$

In this case, $(M, U, \mathcal{U}, F)$ is the underlying control system and $h$ is the measurement mapping. We use the same symbol for the control system with outputs and its underlying control system. The mapping $h=\left(h_{1}, h_{2}, \ldots h_{n}\right)$ : $\mathbb{R} \times M \rightarrow \mathbb{R}^{n}$ represents the vector of $n$ measurements (observations).

The control system with outputs $\Sigma$ is linear if its underlying system is linear and the measurement mapping $h: \mathbb{R} \times \mathbb{R}^{m} \rightarrow \mathbb{R}^{n}$ is linear for each $t \in \mathbb{R}$.

This new system is said to be (completely) observable if, knowing a control $u(\cdot)$ and an observation $y(\cdot)$, on a given interval $\left[t_{0}, T\right]$, one can determine uniquely the initial condition $x_{0}$.

## Stabilizability

Another important issue is that of stabilizability. Assume that for some $\bar{x} \in \mathbb{R}^{n}$ and $\bar{u} \in U, F(\bar{x}, \bar{u})=0$. A function $k: M \rightarrow U$ such that $k(\bar{x})=\bar{u}$ is called a stabilizing feedback if $\bar{x}$ is a stable equilibrium for the system

$$
\dot{x}=F(t, x, k(x(\cdot)))
$$

In the theory of (ordinary) differential equations there exist several methods to determine whether a given equilibrium state is a stable one.

## Realization

For a given initial condition $x_{0} \in \mathbb{R}^{n}$, the control system with outputs

$$
\left\{\begin{array}{l}
\dot{x}=F(t, x, u), \quad x\left(t_{0}\right)=x_{0} \\
y=h(t, x)
\end{array}\right.
$$

defines a mapping which transforms open loops controls $u(\cdot)$ onto outputs

$$
y(t)=h(t, x(t)), \quad t \in\left[t_{0}, T\right] .
$$

Denote this transformation by $\mathcal{R}$. What are its properties? What conditions should a transformation $\mathcal{R}$ satisfy to be given by such a control system ? How, among all the possible "realizations" $\Sigma$ of a transformation $\mathcal{R}$, do we find the simplest one ?

## Optimality

Besides the above problems of structural character, in control theory one also asks optimality questions. In the so-called time-optimal problem one is looking for control which not only transfers a state $x_{0}$ onto $x_{f}$ but does it in the minimal time $T$. More generally, one is looking for a control $u(\cdot)$ which minimizes a functional of the form

$$
\mathcal{J}:=\phi(x(T), T)+\int_{t_{0}}^{T} L(t, x, u) d t
$$

### 1.3 Examples

We shall mention several specific models of control systems.
Example 1. (Moving car) Suppose a car is to be driven along a straight road, and let its distance from an initial point 0 be $s(t)$ at time $t$. For simplicity, assume that the car is controlled only by the throttle, producing an accelerating force of $u_{1}(t)$ per unit mass, and by brake wich produces a retarding force of $u_{2}(t)$ per unit mass. Suppose that the only factors of interest are the car's position $x_{1}(t):=s(t)$ and velocity $x_{2}(t):=\dot{s}(t)$. Ignoring other forces such as road friction, wind resistance, etc. the equations which describe the state of the car at time $t$ are

$$
\left\{\begin{array}{l}
\dot{x}_{1}=x_{2} \\
\\
\dot{x}_{2}=u_{1}-u_{2}
\end{array}\right.
$$

or, in matrix form,

$$
\dot{x}=A x+B u(t)
$$

where

$$
x=\left[\begin{array}{l}
x_{1} \\
x_{2}
\end{array}\right], \quad u=\left[\begin{array}{l}
u_{1} \\
u_{2}
\end{array}\right], \quad A=\left[\begin{array}{ll}
0 & 1 \\
0 & 0
\end{array}\right], \quad B=\left[\begin{array}{rr}
0 & 0 \\
1 & -1
\end{array}\right] .
$$

It may be required to start from rest at 0 and reach some fixed point in the least possible time, or perhaps with minimum consumption of fuel. The mathematical problems are firstly to determine whether such objectives are achievable with the selected control variables, and if so, to find appropriate expressions for $u_{1}(\cdot)$ and $u_{2}(\cdot)$ as functions of time and/or $x_{1}(\cdot)$ and $x_{2}(\cdot)$.

NOTE : The complexity of the model could be increased so as to take into account factors such as engine speed and temperature, vehicle interior temperature, and so on.

Example 2. (Electrically heated oven) Let us consider a simple model of an ellectrically heated oven, which consists of a jacket with a coil directly heating the jacket and of an interior part.


Let $T_{0}$ denote the outside temperature. We make a simplifying assumption, that at an arbitrary moment $t \geq 0$, temperature in the jacket and in the interior part are uniformly distributed and equal to $T_{1}(t), T_{2}(t)$. We assume also that the flow of heat through a surface is proportional to the area of the surface and to the difference of temperature between the separated media. Let $u(t)$ be the intensity of the heat input produced by the coil at moment $t \geq 0$. Let moreover $a_{1}, a_{2}$ denote the area of exterior and interior surfaces of the jacket, respectively, $c_{1}, c_{2}$ denote heat capacities of the jacket and the interior of the oven, respectively, and $r_{1}, r_{2}$ denote radiation coefficients of the exterior and interior surfaces of the jacket, respectively. An increase of heat in the jacket is equal to the amount of heat produced by the coil reduced by the amount of heat which entered the interior and exterior of the oven. Therefore, for the interval $[t, t+\Delta t]$, we have the following balance :
$c_{1}\left(T_{1}(t+\Delta t)-T_{1}(t)\right) \approx u(t) \Delta t-\left(T_{1}(t)-T_{2}(t)\right) a_{1} r_{1} \Delta t-\left(T_{1}(t)-T_{0}\right) a_{2} r_{2} \Delta t$.

Similarly, an increase of heat in the interior of the oven is equal to the amount of heat radiated by the jacket :

$$
c_{2}\left(T_{2}(t+\Delta t)-T_{2}(t)\right)=\left(T_{1}(t)-T_{2}(t)\right) a_{1} r_{2} \Delta t .
$$

Dividing the obtained identities by $\Delta t$ and taking the limit, as $\Delta \rightarrow 0$, we
obtain :

$$
\left\{\begin{array}{l}
c_{1} \dot{T}_{1}=u-\left(T_{1}-T_{2}\right) a_{1} r_{1}-\left(T_{1}-T_{0}\right) a_{2} r_{2} \quad \text { (for the jacket) } \\
c_{2} \dot{T}_{2}=\left(T_{1}-T_{2}\right) a_{1} r_{1} \quad \text { (for the oven interior) } .
\end{array}\right.
$$

Les us notice that, according to the physical interpretation, $u(t) \geq 0$ for $t \geq 0$. Let the state variables be the excesses of temperature over the exterior, that is

$$
x_{1}:=T_{1}-T_{0} \quad \text { and } \quad x_{2}:=T_{2}-T_{0} .
$$

Then we can write the equations above in matrix form, namely

$$
\dot{x}=A x+B u(t)
$$

where
$x=\left[\begin{array}{l}x_{1} \\ x_{2}\end{array}\right], \quad u=[u], \quad A=\left[\begin{array}{cc}-\frac{a_{1} r_{1}+a_{2} r_{2}}{c_{1}} & \frac{a_{1} r_{1}}{c_{1}} \\ \frac{a_{1} r_{1}}{c_{2}} & -\frac{a_{1} r_{1}}{c_{2}}\end{array}\right], \quad B=\left[\begin{array}{c}\frac{1}{c_{1}} \\ 0\end{array}\right]$.
It is natural to limit the considerations to the case when $x_{1}(0) \geq 0$ and $x_{2}(0) \geq 0$. It is physically obvious that if $u(t) \geq 0$ for $t \geq 0$, then also $x_{1}(t) \geq 0, x_{2}(t) \geq 0$ for $t \geq 0$.

Two interesting aspects to be discussed are firstly whether it is possible to maintain the temperature of the oven interior at any desired level merely by altering $u$, and secondly, to determine whether the value of $T_{2}$ can be determined even if it is not possible to measure it directly.

Note: If the desired objective is attainable, then there may well be many different suitable control schemes, and considerations of economy, practicability of application, and so on will then determine how control is actually applied.

Example 3. (Controlled environment) Consider a controlled environment consisting of rabbits and foxes, the number of each at time $t$ being $x_{1}(t)$ and
$x_{2}(t)$, respectively. Suppose that without the presence of foxes the number of rabbits would grow exponentially, but that the rate of growth of rabbit population is reduced by an amount proportional to the number of foxes. Furthermore, suppose that, without rabbits to eat, the fox population would decrease exponentially, but the rate of growth in the number of foxes is increased by an amount proportional to the number of rabbits present. Under these assumptions, the system of equations can be written

$$
\left\{\begin{array}{c}
\dot{x}_{1}=a_{1} x_{1}-a_{2} x_{2} \\
\dot{x}_{2}=a_{3} x_{1}-a_{4} x_{2}
\end{array}\right.
$$

where $a_{1}, a_{2}, a_{3}$ and $a_{4}$ are positive constants.

Example 4. (Satellite problem) We shall consider a point mass in an inverse square law force field. The motion of a unit mass is governed by a pair of second order equations in the radius $r$ and the angle $\theta$ (polar coordinates). If we assume that the unit mass (say a satellite) has the capability of thrusting in the radial direction with the thrust $u_{1}(\cdot)$ and thrusting in the tangential direction with thrust $u_{2}(\cdot)$, then we have

$$
\left\{\begin{array}{l}
\ddot{r}=r \dot{\theta}^{2}-\frac{k}{r^{2}}+u_{1}(t) \\
\ddot{\theta}=-\frac{2 \dot{\theta} \dot{r}}{r}+\frac{1}{r} u_{2}(t) .
\end{array}\right.
$$

If $u_{1}(t)=u_{2}(t)=0$, these equations admit the solution

$$
r(t)=\sigma \quad(\sigma \text { constant }) \quad \text { and } \quad \theta(t)=\omega t \quad(\omega \text { constant }) ; \quad \sigma^{3} \omega^{2}=k .
$$

That is, circular orbits are possible. If we let $x_{1}, x_{2}, x_{3}$, and $x_{4}$ be given by

$$
x_{1}:=r-\sigma, \quad x_{2}:=\dot{r}, \quad x_{3}:=\sigma(\theta-\omega t), \quad x_{4}:=\sigma(\dot{\theta}-\omega)
$$

and normalize $\sigma$ to 1 , then it is easy to see that the linearized equations of motion about the given solution are

$$
\left[\begin{array}{c}
\dot{x}_{1} \\
\dot{x}_{2} \\
\dot{x}_{3} \\
\dot{x}_{4}
\end{array}\right]=\left[\begin{array}{cccc}
0 & 1 & 0 & 0 \\
3 \omega^{2} & 0 & 0 & 2 \omega \\
0 & 0 & 0 & 1 \\
0 & -2 \omega & 0 & 0
\end{array}\right]\left[\begin{array}{l}
x_{1} \\
x_{2} \\
x_{3} \\
x_{4}
\end{array}\right]+\left[\begin{array}{ll}
0 & 0 \\
1 & 0 \\
0 & 0 \\
0 & 1
\end{array}\right]\left[\begin{array}{l}
u_{1}(t) \\
u_{2}(t)
\end{array}\right] .
$$

Example 5. (Market economy) Suppose that the sales $S(t)$ of a product are affected by the amount of advertising $A(t)$ in such a way that the rate of change of sales decreases by an amount proportional to the advertising applied to the share of the market not already purchasing the product. If the total extent of the market is $M$, the state equation is therefore

$$
\dot{S}=-a S+b A(t)\left(1-\frac{S}{M}\right)
$$

subject to

$$
S(0)=S_{0}
$$

where $a$ and $b$ are positive constants. In practice, the amount of advertising will be limited (that is, $0 \leq A(t) \leq K$, where $K$ is a constant), and the aim would be to find the advertising schedule (that is, the function $A(t)$ which maximizes the sales over some period of time).

Example 6. (Soft landing) Let us consider a spacecraft of total mass $M$ moving vertically with the gas thruster directed toward the landing surface. Let $h(\cdot)$ be the height of the spacecraft above the surface, $u(\cdot)$ the thrust of its engine produced by the expulsion of gas from the jet. The gas is a product of the combustion of the fuel. The combustion decreases the total mass of the spacecraft, and the thrust $u$ is proportionalto the speed with which the mass decreases. Assuming that there is no atmosphere above the surface and that
$g$ is gravitational acceleration, one arrives at the following equations :

$$
\left\{\begin{aligned}
M \ddot{h} & =-g M+u(t) \\
\dot{M} & =-k u(t) \quad(k>0)
\end{aligned}\right.
$$

with the initial conditions

$$
M(0)=M_{0}, \quad h(0)=h_{0}, \quad \dot{h}(0)=h_{1} .
$$

One imposes additional constraints on the control parameter of the type

$$
0 \leq u \leq \alpha \quad \text { and } \quad M \geq m,
$$

where $m$ is the mass of the spacecraft without fuel. Let us fix $T>0$. The soft landing problem consists of finding a control $u(\cdot)$ such that for the solutions $M(\cdot), h(\cdot)$ of the above equations

$$
M(t) \geq m, \quad h(t) \geq 0, \quad t \in[0, T], \quad \text { and } \quad h(T)=\dot{h}(T)=0 .
$$

Note: A natural optimization question arises when the moment $T$ is not fixed and one is minimizing the landing time.

### 1.4 Matrix Theory (review)

## Matrices and determinants

We write a matrix as follows

$$
A=\left[a_{i j}\right] \quad(i=1,2, \ldots, m ; j=1,2, \ldots, n)
$$

where $a_{i j}$ is the element (entry) in its $i^{\text {th }}$ row and $j^{\text {th }}$ column, and $A$ thus has $m$ rows and $n$ columns; we use to say that $A$ is an $m \times n$ matrix. We shall denote by $\mathbb{R}^{m \times n}$ the set (vector space) of all $m \times n$ matrices with real entries.

Note : (1) It is convenient to identify the set (vector space) $\mathbb{R}^{m}$ of all m-tuples of real numbers with the set (vector space) $\mathbb{R}^{m \times 1}$ of all column m-matrices (or column $m$-vectors).
(2) There is a natural one-to-one correspondence between $m \times n$ matrices and linear mappings from $\mathbb{R}^{n}$ to $\mathbb{R}^{m}$ : the vector spaces $\mathbb{R}^{m \times n}$ and $L\left(\mathbb{R}^{n}, \mathbb{R}^{m}\right)$ are isomorphic.

The transpose $A^{T}$ of $A$ is obtained by interchanging the rows and columns of $A=\left[a_{i j}\right]$, so $A^{T}:=\left[a_{j i}\right]$ is an $n \times m$ matrix. If $\lambda$ is a scalar (real number), and $A$ and $B$ are matrices of appropriate size, then :
(a) $\left(A^{T}\right)^{T}=A$.
(b) $\quad(A+B)^{T}=A^{T}+B^{T}$.
(c) $\quad(\lambda A)^{T}=\lambda A^{T}$.
(d) $\quad(A B)^{T}=B^{T} A^{T}$.

If $A$ is invertible, then $A^{T}$ is invertible, too and $\left(A^{T}\right)^{-1}=\left(A^{-1}\right)^{T}$.
We can write a matrix $A=\left[a_{i j}\right] \in \mathbb{R}^{m \times n}$ in the following forms :

- $A=\left[\begin{array}{llll}a_{1} & a_{2} & \ldots & a_{n}\end{array}\right]$ with $a_{j}=\left[\begin{array}{c}a_{1 j} \\ a_{2 j} \\ \vdots \\ a_{m j}\end{array}\right] \in \mathbb{R}^{m \times 1} \quad(j=1,2, \ldots, n)$
- $A=\left[\begin{array}{c}a_{1} \\ a_{2} \\ \vdots \\ a_{m}\end{array}\right]$ with $a_{i}=\left[\begin{array}{llll}a_{i 1} & a_{i 2} & \ldots & a_{i n}\end{array}\right] \in \mathbb{R}^{1 \times n} \quad(i=1,2, \ldots, m)$.

When $m=n, A$ is said to be square of order $n$. We shall write

$$
I_{n}=\left[\delta_{i j}\right]=\left[\begin{array}{ccccc}
1 & 0 & \cdots & 0 & 0 \\
0 & 1 & \cdots & 0 & 0 \\
\vdots & \vdots & \cdots & \vdots & \vdots \\
0 & 0 & \cdots & 0 & 1
\end{array}\right] \in \mathbb{R}^{n \times n}
$$

for the unit matrix (or identity matrix) of order $n$. Here, $\delta_{i j}$ stands for Kronecker's symbol; that is,

$$
\delta_{i j}:= \begin{cases}1 & \text { if } i=j \\ 0 & \text { if } i \neq j\end{cases}
$$

The unit matrix $I_{n}$ has all its elements zero except those on the main diagonal; any (square) matrix of this form is called a diagonal matrix, written

$$
\operatorname{diag}\left(a_{11}, a_{22}, \ldots, a_{n n}\right)
$$

Two $n \times n$ matrices $A$ and $B$ related by

$$
B=S^{-1} A S
$$

are called similar. This is a basic equivalence relation on matrices.
Note : Two matrices are similar if and only if they represent the same linear mapping in different bases. Any matrix property that is preserved under similarity is a property of the underlying linear mapping.

If $A=\left[a_{i j}\right] \in \mathbb{R}^{n \times n}$, then its trace is the sum of all elements on the main diagonal; that is,

$$
\operatorname{tr}(A):=\sum_{i=1}^{n} a_{i i}
$$

The trace operator has some important properties :
(a) $\operatorname{tr}(\lambda A)=\lambda \operatorname{tr}(A)$, where $\lambda$ is a scalar.
(b) $\operatorname{tr}(A+B)=\operatorname{tr}(A)+\operatorname{tr}(B)$.
(c) $\quad \operatorname{tr}\left(I_{n}\right)=n$.
(d) $\operatorname{tr}(A B)=\operatorname{tr}(B A)$.
(e) $\operatorname{tr}\left(A^{T}\right)=\operatorname{tr}(A)$.
(f) $\operatorname{tr}\left(A^{T} A\right) \geq 0$.

Exercise 1 Given $A, S \in \mathbb{R}^{n \times n}$ with $S$ invertible, show that

$$
\operatorname{tr}\left(S A S^{-1}\right)=\operatorname{tr}(A)
$$

That is, similar matrices have the same trace.

We recall briefly the main properties of the determinant function

$$
\operatorname{det}: \mathbb{R}^{n \times n} \rightarrow \mathbb{R}, \quad A \mapsto \operatorname{det}(A)
$$

These are :
(a) $\operatorname{det}(A B)=\operatorname{det}(A) \cdot \operatorname{det}(B)$.
(b) $\operatorname{det}\left(I_{n}\right)=1$.
(c) $\operatorname{det}(A) \neq 0$ if and only if $A$ is invertible.

Note : There is a unique function det : $\mathbb{R}^{n \times n} \rightarrow \mathbb{R}$ having these three properties. For $A=\left[a_{i j}\right] \in \mathbb{R}^{n \times n}$ we have

$$
\operatorname{det}(A)=\sum_{\alpha} \operatorname{sgn}(\alpha) a_{1 \alpha(1)} a_{2 \alpha(2)} \cdots a_{n \alpha(n)}
$$

where the sum is taken over the $n!$ permutations (on $n$ elements) $\alpha \in S_{n}$.

Exercise 2 Given $A, S \in \mathbb{R}^{n \times n}$ with $S$ invertible, show that

$$
\operatorname{det}\left(S A S^{-1}\right)=\operatorname{det}(A)
$$

That is, similar matrices have the same determinant.

If $\operatorname{det}(A)=0, A$ is singular, otherwise nonsingular (or invertible); in the latter case, the inverse of $A$ is

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

where $\operatorname{adj}(A):=\left[A_{i j}\right]^{T}$ is the adjoint of $A$; here,

$$
A_{j}^{i}:=(-1)^{i+j} M_{i j} \quad\left(\text { the cofactor of } a_{i j}\right)
$$

and $M_{i j}$ is the determinant of the submatrix formed by deleting the $i^{\text {th }}$ row and the $j^{\text {th }}$ column of $A$.

Note : Let $A \in \mathbb{R}^{n \times n}$. Then $\operatorname{det}(A)=0$ (the matrix $A$ is singular) if and only if $A x=0$ for some nonzero column $n$-vector $x \in \mathbb{R}^{n \times 1}$.

## Linear dependence and rank

Consider a set of column $m$-vectors

$$
a_{1}=\left[\begin{array}{c}
a_{11} \\
a_{21} \\
\vdots \\
a_{m 1}
\end{array}\right], \quad a_{2}=\left[\begin{array}{c}
a_{12} \\
a_{22} \\
\vdots \\
a_{m 2}
\end{array}\right], \quad \ldots, \quad a_{n}=\left[\begin{array}{c}
a_{1 n} \\
a_{2 n} \\
\vdots \\
a_{m n}
\end{array}\right] .
$$

If $\alpha_{1}, \alpha_{2}, \ldots, \alpha_{n}$ are scalars, then the vector

$$
\alpha_{1} a_{1}+\alpha_{2} a_{2}+\cdots+\alpha_{n} a_{n}=\left[\begin{array}{c}
\alpha_{1} a_{11}+\alpha_{2} a_{12}+\cdots+\alpha_{n} a_{1 n} \\
\alpha_{1} a_{21}+\alpha_{2} a_{22}+\cdots+\alpha_{n} a_{2 n} \\
\vdots \\
\alpha_{1} a_{m 1}+\alpha_{2} a_{m 2}+\cdots+\alpha_{n} a_{m n}
\end{array}\right]
$$

is called a linear combination of $a_{1}, a_{2}, \ldots, a_{n}$. If there exist scalars $\alpha_{1}, \alpha_{2}, \ldots, \alpha_{n}$, not all zero, such that

$$
\alpha_{1} a_{1}+\alpha_{2} a_{2}+\cdots+\alpha_{n} a_{n}=\left[\begin{array}{c}
0 \\
0 \\
\vdots \\
0
\end{array}\right]
$$

then the vectors $a_{1}, a_{2}, \ldots, a_{n}$ are said to be linearly dependent; otherwise, they are linearly independent.

Note : We can equally well consider row $n$-vectors.

Let $A=\left[a_{i j}\right] \in \mathbb{R}^{m \times n}$. The rank of $A$, denoted by $\operatorname{rank}(A)$, is defined as the maximum number of linearly independent columns (or rows) of $A$. Clearly, $\operatorname{rank}(A) \leq \min \{m, n\}$. Consider the kernel (or null-space)

$$
\operatorname{ker}(A):=\left\{x \in \mathbb{R}^{n} \mid A x=0\right\} \subseteq \mathbb{R}^{n}
$$

and the image space (or column space)

$$
\operatorname{im}(A):=\left\{A x \mid x \in \mathbb{R}^{n}\right\} \subseteq \mathbb{R}^{m} .
$$

The dimension of $\operatorname{ker}(A)$ is termed the nullity of $A$. $\operatorname{im}(A)$ is nothing more than the (vector) space spanned by the columns of $A$; that is,

$$
\operatorname{im}(A)=\operatorname{span}\left\{a_{1}, a_{2}, \ldots, a_{n}\right\}:=\left\{\alpha_{1} a_{1}+\alpha_{2} a_{2} \cdots+\alpha_{n} a_{n} \mid \alpha_{1}, \ldots, \alpha_{n} \in \mathbb{R}\right\}
$$

where $A=\left[\begin{array}{llll}a_{1} & a_{2} & \cdots & a_{n}\end{array}\right]$. Hence the dimension of $\operatorname{im}(A)$ is equal to $\operatorname{rank}(A)$.

Note : $\operatorname{im}\left(A^{T}\right)$ is also known as the row space of $A$; it is the (vector) space spanned by the rows of $A$ (i.e. the columns of $A^{T}$ ).

An important result states that (for a matrix $A \in \mathbb{R}^{m \times n}$ ) :

$$
\operatorname{rank}(A)+\operatorname{dim} \operatorname{ker}(A)=n .
$$

Rank is invariant under multiplication by a nonsingular matrix. In particular, rank is invariant under similarity. However, multiplication by rectangular or singular matrices can alter the rank, and the following formula shows exactly how much alteration occurs.

If $A \in \mathbb{R}^{m \times n}$ and $B \in \mathbb{R}^{n \times p}$, then:

$$
\operatorname{rank}(A B)=\operatorname{rank}(B)-\operatorname{dim} \operatorname{ker}(A) \cap \operatorname{im}(B) .
$$

Exercise 3 Given $A \in \mathbb{R}^{m \times n}$ and $B \in \mathbb{R}^{n \times p}$, show that :
(a) $\operatorname{rank}(A B) \leq \min \{\operatorname{rank}(A), \operatorname{rank}(B)\}$.
(b) $\operatorname{rank}(A)+\operatorname{rank}(B)-n \leq \operatorname{rank}(A B)$.

Suppose now that $a_{i j}(i=1,2, \ldots, m ; j=1,2, \ldots, n)$ are the coefficients in a set of $m$ linear algebraic equations in $n$ unknowns

$$
\sum_{j=1}^{n} a_{i j} x_{j}=b_{i}, \quad i=1,2, \ldots, m
$$

These equations can be written in matrix form as

$$
A x=b
$$

where

$$
A=\left[a_{i j}\right], \quad x=\left[\begin{array}{c}
x_{1} \\
x_{2} \\
\vdots \\
x_{n}
\end{array}\right], \quad b=\left[\begin{array}{c}
b_{1} \\
b_{2} \\
\vdots \\
b_{m}
\end{array}\right] .
$$

Such a linear system (of equations) possesses a solution if and only if

$$
\operatorname{rank}(A)=\operatorname{rank}\left[\begin{array}{cc}
A & b
\end{array}\right]
$$

where $\left[\begin{array}{ll}A & b\end{array}\right]$ is the $m \times(n+1)$ matrix obtained by appending $b$ to $A$ as an extra column.

Two particular cases should be mentioned :

- When $A \in \mathbb{R}^{n \times n}$, the linear system $A x=b$ has a unique solution if and only if $A$ is nonsingular.
- When $A \in \mathbb{R}^{m \times n}$, the homogeneous linear system $A x=0$ has a nonzero solution if and only if $\operatorname{rank}(A)<n$.

Note: Similar remarks apply to the set of equations

$$
y A=c
$$

where

$$
A=\left[a_{i j}\right], \quad y=\left[\begin{array}{llll}
y_{1} & y_{2} & \cdots & y_{m}
\end{array}\right], \quad c=\left[\begin{array}{llll}
c_{1} & c_{2} & \ldots & c_{n}
\end{array}\right]
$$

## Eigenvalues and eigenvectors

Let $A=\left[a_{i j}\right] \in \mathbb{R}^{n \times n}$. A nonzero vector $w \in \mathbb{R}^{n \times 1}$ is called an eigenvector (or characteristic vector) of $A$ if there is a scalar (real number) $\lambda$ such that

$$
A w=\lambda w .
$$

The scalar $\lambda$ is called the eigenvalue (or characteristic value) associated with the eigenvector $w$. Geometrically, $A w=\lambda w$ says that under the (linear) mapping $x \mapsto A x$ the eigenvectors experience only changes in magnitude or sign. The eigenvalue $\lambda$ is simply the amount of "stretch" or "shrink" to which the eigenvector $w$ is subjected when acted upon by $A$.

Note: The words eigenvalue and eigenvector are derived from the German word eigen, which means "owen by" or "peculiar to".

The set of distinct eigenvalues, denoted by $\sigma(A)$, is called the spectrum of $A$. We have

$$
\lambda \in \sigma(A) \Longleftrightarrow \lambda I_{n}-A \text { is singular } \Longleftrightarrow \operatorname{det}\left(\lambda I_{n}-A\right)=0
$$

The set (vector space) of all eigenvectors with eigenvalue $\lambda$, together with the zero vector, is called the $\lambda$-eigenspace of $A$ and is denoted by $E_{\lambda}$. That is,

$$
E_{\lambda}:=\operatorname{ker}\left(\lambda I_{n}-A\right) .
$$

The eigenvectors $w \in E_{\lambda}$ are found by solving the equation

$$
\left(\lambda I_{n}-A\right) w=0 .
$$

This matrix equation is equivalent to a system of $n$ linear algebraic equations; the solution space is exactly the $\lambda$-eigenspace $E_{\lambda}$.

Exercise 4 Let $A \in \mathbb{R}^{n \times n}$. If $\lambda_{1}, \lambda_{2}, \ldots, \lambda_{r}$ are $r(r \leq n)$ distinct eigenvalues of $A$ with corresponding eigenvectors $w_{1}, w_{2}, \ldots, w_{r}$, show that the vectors $w_{1}, w_{2}, \ldots, w_{r}$ are linearly independent.

The characteristic polynomial of $A \in \mathbb{R}^{n \times n}$ is

$$
\operatorname{char}_{A}(\lambda):=\operatorname{det}\left(\lambda I_{n}-A\right) .
$$

The algebraic equation

$$
\operatorname{char}_{A}(\lambda) \equiv \lambda^{n}+k_{1} \lambda^{n-1}+\cdots+k_{n-1} \lambda+k_{n}=0
$$

is called the characteristic equation of $A$.
Note : The degree of the characteristic polynomial (equation) is $n$ and the leading term is $\lambda^{n}$. The eigenvalues of $A$ are exactly the real roots of $\operatorname{char}_{A}(\lambda)$.

The fundamental theorem of algebra insures that the characteristic polynomial $\operatorname{char}_{A}(\lambda)$ has $n$ roots, but some roots may be complex numbers, and some roots may be repeated. A complex root of the characteristic polynomial is called a complex eigenvalue of $A$. The complex eigenvalues must occur in conjugate pairs. If $\lambda$ is a complex eigenvalue of the matrix $A$, we write $\lambda \in \sigma_{\mathbb{C}}(A)$. Henceforth, we shall refer to both sets $\sigma(A)$ and $\sigma_{\mathbb{C}}(A)$ as the spectrum of $A$. (In fact, $\sigma_{\mathbb{C}}(A)$ is the spectrum of the "complexification" of (the linear mapping) $A: \xi+i \eta \mapsto A \xi+i A \eta)$.

An important result is the following : If the matrix $A$ is symmetric (i.e. $A=A^{T}$ ), then all its eigenvalues are real.

A useful result is the Cayley-Hamilton Theorem, which states that every square matrix satisfies its own characteristic equation; that is, if $A \in$ $\mathbb{R}^{n \times n}$, then :

$$
\operatorname{char}_{A}(A) \equiv A^{n}+k_{1} A^{n-1}+\cdots+k_{n-1} A+k_{n} I_{n}=O .
$$

Let $\lambda \in \sigma_{\mathbb{C}}(A)$.

- The algebraic multiplicity $m_{\lambda}$ of $\lambda$ is the number of times it is repeated as a root of the characteristic polynomial.
- The geometric multiplicity $d_{\lambda}$ of $\lambda$ is the dimension of the $\lambda$ eigenspace $E_{\lambda}$. In other words, $d_{\lambda}$ is the maximum number of linearly independent eigenvectors associated with $\lambda$.

In general, $d_{\lambda} \leq m_{\lambda}$. The following remarkable result holds :
The matrix $A \in \mathbb{R}^{n \times n}$ is diagonalizable (that is, $A$ is similar to a diagonal matrix) if and only if $d_{\lambda}=m_{\lambda}$. If all the eigenvalues of $A$ are real and distinct, then $A$ is diagonalizable. The converse is not true.

Exercise 5 Let $A \in \mathbb{R}^{n \times n}$ with complex eigenvalues $\lambda_{1}, \lambda_{2}, \ldots, \lambda_{n}$ (listed with their algebraic multiplicities). Show that:
(a) $\lambda_{1}+\lambda_{2}+\cdots+\lambda_{n}=\operatorname{tr}(A)$.
(b) $\lambda_{1} \cdot \lambda_{2} \cdots \lambda_{n}=\operatorname{det}(A)$.

## Quadratic forms

Let $A=\left[a_{i j}\right] \in \mathbb{R}^{n \times n}$ be a symmetric matrix. A function $q: \mathbb{R}^{n}=\mathbb{R}^{n \times 1} \rightarrow$ $\mathbb{R}$ defined by

$$
\begin{aligned}
q(x) & :=x^{T} A x=\sum_{i=1}^{n} \sum_{j=1}^{n} a_{i j} x_{i} x_{j} \\
& =a_{11} x_{1}^{2}+a_{22} x_{2}^{2}+\cdots+a_{n n} x_{n}^{2}+2 a_{12} x_{1} x_{2}+2 a_{13} x_{1} x_{3}+\cdots
\end{aligned}
$$

is called a quadratic form (on $\mathbb{R}^{n}$ ). Clearly, $q(0)=0$.
A quadratic form $q$ is said to be :
(a) positive definite provided $q(x)>0$ for all nonzero $x \in \mathbb{R}^{n \times 1}$.
(b) negative definite provided $q(x)<0$ for all nonzero $x \in \mathbb{R}^{n \times 1}$.
(c) positive semi-definite provided $q(x) \geq 0$ for all $x \in \mathbb{R}^{n \times 1}$.
(d) negative semi-definite provided $q(x) \leq 0$ for all $x \in \mathbb{R}^{n \times 1}$.

Finally, we call $q$ indefinite provided $q$ takes positive as well as negative values.

Note : (1) The terms describing the quadratic form $q$ are also applied to the (symmetric) matrix $A$ associated with the form.
(2) The definitions on definiteness and semi-definiteness can be extended to scalar functions (defined on some $\mathbb{R}^{n}$ ) which are not necessarily quadratic.

One simple way of determining the sign properties of a quadratic form is the following:

The quadratic form $q: x \mapsto x^{T} A x$ (or, equivalently, the matrix $A$ ) is :

- positive definite if and only if all the eigenvalues of $A$ are positive.
- negative definite if and only if all the eigenvalues of $A$ are negative.
- positive semi-definite if and only if all the eigenvalues of $A$ are nonnegative.
- negative semi-definite if and only if all the eigenvalues of $A$ are nonpositive.

An alternative approach involves the principal minors $P_{i}$ of $A$, these being any $i^{\text {th }}$ order minors whose main diagonal is part of the main diagonal of $A$. In particular, the leading principal minors of $A$ are

$$
\Delta_{1}:=a_{11}, \quad \Delta_{2}:=\left|\begin{array}{cc}
a_{11} & a_{12} \\
a_{21} & a_{22}
\end{array}\right|, \quad \Delta_{3}:=\left|\begin{array}{ccc}
a_{11} & a_{12} & a_{13} \\
a_{21} & a_{22} & a_{23} \\
a_{31} & a_{32} & a_{33}
\end{array}\right|, \quad \text { etc. }
$$

The Sylvester conditions state that the quadratic form $q: x \mapsto x^{T} A x$ (or, equivalently, the matrix $A$ ) is :

- positive definite if and only if $\Delta_{i}>0, \quad i=1,2, \ldots, n$;
- negative definite if and only if $(-1)^{i} \Delta_{i}>0, \quad i=1,1, \ldots, n$;
- positive semi-definite if and only if $P_{i} \geq 0$ for all principal minors;
- negative semi-definite if and only if $(-1)^{i} P_{i} \geq 0$ for all principal minors.

If $q$ satisfies none of the above conditions, then it is indefinite.
Let $A \in \mathbb{R}^{n \times n}$ be a symmetric matrix such that $\operatorname{rank}(A)=r$. Then $A$ is positive semi-definite ( $A^{T}=A \geq 0$ ) if and only if $A=B^{T} B$ for some matrix $B$ with $\operatorname{rank}(B)=r$.

## The matrix exponential

In order to define the exponential of a matrix, we need to discuss the convergence of (infinite) sequences and series involving matrices.

Because $\mathbb{R}^{m \times n}$ is a vector space of dimension $m n$, magnitudes of matrices $A \in \mathbb{R}^{m \times n}$ can be "measured" by employing any norm on $\mathbb{R}^{m n}$. One of the simplest matrix norms is the following :

$$
\|A\|:=\sqrt{\sum_{i=1}^{m} \sum_{j=1}^{n} a_{i j}^{2}}=\sqrt{\operatorname{tr}\left(A^{T} A\right)} \text { for } A=\left[a_{i j}\right] \in \mathbb{R}^{m \times n} \text {. }
$$

For a column matrix (vector) $x \in \mathbb{R}^{m \times 1}$ this gives the Euclidean norm

$$
\|x\|_{e}:=\sqrt{x_{1}^{2}+x_{2}^{2}+\cdots+x_{m}^{2}} .
$$

Note : (1) Other matrix norms can also be defined. For example, any norm $\|\cdot\|_{*}$ that is defined on $\mathbb{R}^{m}$ and $\mathbb{R}^{n}$ induces a matrix norm on $\mathbb{R}^{m \times n}$ by setting

$$
\|A\|_{*}:=\max _{\|x\|_{*}=1}\|A x\|_{*} \text { for } A \in \mathbb{R}^{m \times n} \text { and } x \in \mathbb{R}^{n \times 1} \text {. }
$$

(2) The matrix norm $\|\cdot\|$ and the Euclidean norm $\|\cdot\|_{e}$ are compatible:

$$
\|A x\|_{e} \leq\|A\|\|x\|_{e} .
$$

The matrix norm has all the usual properties of a norm; that is (for $A, B \in$ $\mathbb{R}^{m \times n}$ and $\left.\lambda \in \mathbb{R}\right)$ :
(a) $\quad\|A\| \geq 0$, and $\|A\|=0 \Longleftrightarrow A=0$.
(b) $\quad\|\lambda A\|=|\lambda|\|A\|$.
(c) $\|A+B\| \leq\|A\|+\|B\|$.

Also, the following two relations hold (provided that the product $A B$ is defined) :
(d) $\quad\|A B\| \leq\|A\|\|B\|$.
(e) $\quad\left\|A^{k}\right\| \leq\|A\|^{k}, \quad k=0,1,2, \ldots$

A sequence $\left(A_{k}\right)_{k \in \mathbb{N}}$ of matrices in $\mathbb{R}^{m \times n}$ is said to converge to the limit $A \in \mathbb{R}^{m \times n}$, denoted by

$$
\lim _{k \rightarrow \infty} A_{k}=A,
$$

if the sequence $\left(\left\|A_{k}-A\right\|\right)_{k \in \mathbb{N}}$ of (positive) real numbers converges to 0 ; that is, for every $\varepsilon>0$ there exists an $\nu \in \mathbb{N}$ such that for $k \geq \nu,\left\|A-A_{k}\right\|<\varepsilon$.

A necessary and sufficient condition for convergence is that each entry of $A_{k}$ tends (converges) to the corresponding entry of $A$ as $k \rightarrow \infty$.

The (infinite) matrix series $\sum_{k \geq 0} A_{k}$ converges provided the sequence of partial sums $\left(S_{k}\right)_{k \in \mathbb{N}}$, where $S_{k}:=A_{1}+A_{2}+\cdots+A_{k}$, converges to a limit $S$ (as $k \rightarrow \infty$ ); if a limit exists, then it is unique and we shall write $S=$ $\sum_{k=0}^{\infty} A_{k}$. The series is absolutely convergent if the scalar series $\sum_{k \geq 0}\left\|A_{k}\right\|$ is convergent; an absolutely convergent matrix series is convergent.

Consider now a matrix $A \in \mathbb{R}^{n \times n}$.
Exercise 6 Show that the matrix power series $\sum_{k \geq 0} \frac{t^{k}}{k!} A^{k}$ is convergent (in fact, absolutely convergent) for every $t \in \mathbb{R}$.

We define the matrix exponential of $A$ by

$$
\exp (t A):=I_{n}+t A+\frac{t^{2}}{2!} A^{2}+\cdots=\sum_{k=0}^{\infty} \frac{t^{k}}{k!} A^{k}
$$

The matrix exponential has a number of important properties.
(a) $\frac{d}{d t}(\exp (t A))=A \exp (t A)=\exp (t A) A$.
(b) $\quad \exp ((t+s) A)=\exp (t A) \cdot \exp (s A)$.
(c) $\quad \exp (A)=\lim _{k \rightarrow \infty}\left(I_{n}+\frac{A}{k}\right)^{k}$.
(d) $\quad \operatorname{det}(\exp (A))=e^{\operatorname{tr}(A)}$.

From (b) it follows that

$$
\exp (t A) \cdot \exp (-t A)=I_{n}
$$

and thus

$$
\exp (t A)^{-1}=\exp (-t A) .
$$

### 1.5 Exercises

Exercise 7 Suppose $A, S \in \mathbb{R}^{n \times n}$ and $S$ is invertible. Show that

$$
\left(S^{-1} A S\right)^{2}=S^{-1} A^{2} S .
$$

Generalize to $\left(S^{-1} A S\right)^{n}$.
Exercise 8 Set

$$
u=\left[\begin{array}{c}
u_{1} \\
u_{2} \\
\vdots \\
u_{n}
\end{array}\right], \quad n \geq 2
$$

Write $A=u u^{T} \in \mathbb{R}^{n \times n}$ and show that $A$ is singular.
Exercise 9 Let $A \in \mathbb{R}^{n \times n}$ and $0 \neq b \in \mathbb{R}^{n \times 1}$ such that

$$
A^{r} b \neq 0, \quad A^{r+1} b=0
$$

for some positive integer $r<n$. By considering the equation

$$
c_{0} b+c_{1} A b+c_{2} A^{2} b+\cdots+c_{r} A^{r} b=0
$$

where the coefficients $c_{i} \in \mathbb{R}$, deduce that the (column) vectors

$$
b, A b, A^{2} b, \cdots, A^{r} b
$$

are linearly independent.

Exercise 10 Verify that

$$
\operatorname{rank}\left(A^{T} A\right)=\operatorname{rank}(A)=\operatorname{rank}\left(A A^{T}\right)
$$

for the matrix

$$
A=\left[\begin{array}{cccc}
1 & 3 & 1 & -4 \\
-1 & -3 & 1 & 0 \\
2 & 6 & 2 & -8
\end{array}\right] \in \mathbb{R}^{3 \times 4}
$$

Exercise 11 Find the characteristic polynomial, the eigenvalues and the corresponding eigenvectors for each given matrix.
(a) $\left[\begin{array}{ll}1 & 4 \\ 2 & 3\end{array}\right]$.
(b) $\left[\begin{array}{rr}2 & 0 \\ 0 & -1\end{array}\right]$.
(c) $\left[\begin{array}{rr}2 & 2 \\ -1 & -1\end{array}\right]$.
(d) $\left[\begin{array}{rr}2 & -2 \\ 2 & 2\end{array}\right]$.
(e) $\left[\begin{array}{rr}0 & -1 \\ a b & a+b\end{array}\right]$.
(f) $\left[\begin{array}{lll}2 & 2 & 0 \\ 1 & 2 & 1 \\ 1 & 2 & 1\end{array}\right]$.
(g) $\left[\begin{array}{lll}1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1\end{array}\right]$.
(h) $\left[\begin{array}{lll}1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1\end{array}\right]$.
(i) $\left[\begin{array}{rrr}-1 & 3 & 0 \\ 3 & 7 & 0 \\ 0 & 0 & 6\end{array}\right]$.
(j) $\left[\begin{array}{rr}\cos \theta & \sin \theta \\ -\sin \theta & \cos \theta\end{array}\right]$.

Exercise 12 Let $A \in \mathbb{R}^{n \times n}$.
(a) How are the eigenvalues of $A-\mu I_{n}$ related to those of $A$ ?
(b) How are the eigenvalues of $\mu A$ related to those of $A$ ?
(c) How are the eigenvalues of $A^{n}$ related to those of $A$ ?
(d) How are the eigenvalues of $A^{-1}$ related to those of $A$ ?

Exercise 13 Consider a matrix $A \in \mathbb{R}^{n \times n}$. Show that :
(a) The characteristic polynomials of $A$ and $A^{T}$ are the same.
(b) The characteristic polynomials of $A$ and $S^{-1} A S$ are the same.
(c) If $n=2$, the characteristic polynomial of $A$ can be written as follows

$$
\operatorname{char}_{A}(\lambda)=\lambda^{2}-\operatorname{tr}(A) \lambda+\operatorname{det}(A) .
$$

Exercise 14 Let the matrix $A \in \mathbb{R}^{n \times n}$ be invertible and $w$ an eigenvector of $A$ with associated eigenvalue $\lambda$.
(a) Is $w$ an eigenvector of $A^{3}$ ? If so, what is the eigenvalue?
(b) Is $w$ an eigenvector of $A^{-1}$ ? If so, what is the eigenvalue?
(c) Is $w$ an eigenvector of $A+2 I_{n}$ ? If so, what is the eigenvalue?
(d) Is $w$ an eigenvector of $7 A$ ? If so, what is the eigenvalue?

## Exercise 15

(a) A skew-symmetric matrix $S \in \mathbb{R}^{n \times n}$ is defined by $S^{T}=-S$. If $q=x^{T} S x$ show (by considering $q^{T}$ ) that $q=0$ for all (column) vectors $x \in \mathbb{R}^{n \times 1}$.
(b) Show that any matrix $A \in \mathbb{R}^{n \times n}$ can be written as $A=A_{1}+A_{2}$, where $A_{1}$ is symmetric and $A_{2}$ is skew-symmetric. Hence (using the result of (a)) deduce that

$$
x^{T} A x=x^{T} A_{1} x \quad \text { for all (column) vectors } x \in \mathbb{R}^{n \times 1} .
$$

Exercise 16 Prove that

$$
\exp ((t+s) A)=\exp (t A) \exp (s A)
$$

[Hint: Multiply the series in powers of $A$ formally; the legitimacy of the term-by-term multiplication is assured by the fact that $\exp (t A)$ is absolutly convergent.]

## Exercise 17

(a) Show that $\exp (t A) \cdot \exp (t B)$ does not have to be either $\exp (t(A+B))$ or $\exp (t B) \cdot \exp (t A)$ by calculating all three, where

$$
A=\left[\begin{array}{ll}
0 & 1 \\
0 & 0
\end{array}\right] \quad \text { and } \quad B=\left[\begin{array}{ll}
1 & 0 \\
0 & 0
\end{array}\right]
$$

(b) Suppose that $A B=B A$. Show that

$$
\exp (t(A+B))=\exp (t A) \cdot \exp (t B)=\exp (t B) \cdot \exp (t A)
$$

[Hint: Show that if $P(t)=\exp (t(A+B)) \cdot \exp (-t A) \cdot \exp (-t B)$, then $\dot{P}(t)=0$ for all $t$. Since $P(0)=I_{n}$, we must have $P(t)=I_{n}$.]

## Exercise 18

(a) Find $\exp (t A)$ if

$$
A=\left[\begin{array}{ll}
a & 0 \\
0 & b
\end{array}\right] .
$$

Generalize to $A=\operatorname{diag}\left(a_{1}, a_{2}, \ldots, a_{n}\right)$ (the diagonal matrix with diagonal elements $\left.a_{1}, a_{2}, \ldots, a_{n}\right)$.
(b) Consider the matrix

$$
A=\left[\begin{array}{ll}
a & b \\
0 & a
\end{array}\right]
$$

Show that

$$
\exp (t A)=\left[\begin{array}{rr}
e^{a t} & b t e^{a t} \\
0 & e^{a t}
\end{array}\right]
$$

(c) A matrix $A$ is nilpotent if some power $A^{k}$ is the zero matrix. Then the matrix exponential $\exp (t A)$ can be calculated easily because the series stops with the power $A^{k-1}$. That is, we have $A^{k}=A^{k+1}=\cdots=0$, so

$$
\exp (t A)=I_{n}+t A+\frac{t^{2}}{2!} A^{2}+\cdots+\frac{t^{k-1}}{(k-1)!} A^{k-1}
$$

Find $\exp (t A)$ for
i. $A=\left[\begin{array}{lll}0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0\end{array}\right]$.
ii. $A=\left[\begin{array}{lll}0 & 0 & 0 \\ 2 & 0 & 0 \\ 3 & 4 & 0\end{array}\right]$.

Exercise 19 Find $\exp (t A)$, where $A$ is given.
(a) $A=\left[\begin{array}{ll}1 & 0 \\ 0 & 2\end{array}\right]$.
(b) $A=\left[\begin{array}{ll}0 & 1 \\ 0 & 0\end{array}\right]$.
(c) $A=\left[\begin{array}{rr}0 & 1 \\ -1 & 0\end{array}\right]$.
(d) $A=\left[\begin{array}{rr}-1 & 1 \\ 0 & -1\end{array}\right]$.
(e) $A=\left[\begin{array}{ll}1 & 1 \\ 1 & 1\end{array}\right]$.
(f) $A=\left[\begin{array}{lll}0 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0\end{array}\right]$.
(g) $A=\left[\begin{array}{rrr}2 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & 7\end{array}\right]$.

Exercise 20 TRUE or FALSE ? Motivate your answers.
(a) If $A \in \mathbb{R}^{n \times n}$ and $\lambda \in \mathbb{R}$, then

$$
\operatorname{det}(\lambda A)=\lambda \operatorname{det}(A)
$$

(b) If $A, B \in \mathbb{R}^{n \times n}$ then

$$
\operatorname{det}(A+B)=\operatorname{det}(A)+\operatorname{det}(B)
$$

(c) If $A \in \mathbb{R}^{n \times n}$ then

$$
\operatorname{det}\left(A A^{T}\right)=\operatorname{det}\left(A^{T} A\right)
$$

(d) If $A \in \mathbb{R}^{m \times n}$ then

$$
\operatorname{rank}\left(A^{T} A\right)=\operatorname{rank}(A)=\operatorname{rank}\left(A A^{T}\right)
$$

(e) A matrix $A \in \mathbb{R}^{n \times n}$ is invertible if and only if 0 is not an eigenvalue of $A$.
(f) If $t \in \mathbb{R}$ then

$$
\exp \left(t\left[\begin{array}{rr}
0 & 1 \\
-1 & 0
\end{array}\right]\right)=\left[\begin{array}{rr}
\cos t & \sin t \\
-\sin t & \cos t
\end{array}\right]
$$

(g) If $A, B \in \mathbb{R}^{n \times n}$ then

$$
\exp (A+B)=\exp (A) \cdot \exp (B)
$$

(h) If $A \in \mathbb{R}^{n \times n}$ then

$$
\operatorname{det}(\exp (A)) \neq 0
$$

