Chapter 1

ANALYSIS OF LINEAR SYSTEMS IN STATE SPACE FORM

This course focuses on the state space approach to the analysis and design of control systems. The idea of state of a system dates back to classical physics. Roughly speaking, the state of a system is that quantity which, together with knowledge of future inputs to the system, determine the future behaviour of the system. For example, in mechanics, knowledge of the current position and velocity (equivalently, momentum) of a particle, and the future forces acting it, determines the future trajectory of the particle. Thus, the position and the velocity together qualify as the state of the particle.

In the case of a continuous-time linear system, its state space representation corresponds to a system of first order differential equations describing its behaviour. The fact that this description fits the idea of state will become evident after we discuss the solution of linear systems of differential equation.

To illustrate how we can obtain state equations from simple differential equation models, consider the second order linear differential equation

$$\ddot{y}(t) = u(t) \tag{1.1}$$

This is basically Newton's law $F = ma$ for a particle moving on a line, where we have used u to denote $\frac{F}{m}$. It also corresponds to a transfer function from u to y to be $\frac{1}{s^2}$. Let

$$\begin{aligned} x_1(t) &= y(t) \\ x_2(t) &= \dot{y}(t) \end{aligned}$$

Set $x^T(t) = [x_1(t) \;\; x_2(t)]$. Then we can write

$$\frac{dx(t)}{dt} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} x(t) + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u(t) \tag{1.2}$$

which is a system of first order linear differential equations.

The above construction can be readily extended to higher order linear differential equations of the form

$$y^{(n)}(t) + a_1 y^{(n-1)}(t) + \cdots + a_{n-1}\dot{y}(t) + a_n y(t) = u(t)$$

Define $x(t) = \begin{bmatrix} y & \frac{dy}{dt} & \cdots & \frac{d^{n-1}y}{dt^{n-1}} \end{bmatrix}^T$. It is straightforward to check that $x(t)$ satisfies the differential equation

$$\frac{dx(t)}{dt} = \begin{bmatrix} 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \cdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & & & 1 \\ -a_n & -a_{n-1} & & \cdots & -a_2 & -a_1 \end{bmatrix} x(t) + \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} u(t)$$

On the other hand, it is not quite obvious how we should define the state x for a differential equation of the form

$$\ddot{y}(t) + 3\dot{y}(t) + 2y(t) = \dot{u}(t) + 4u(t) \tag{1.3}$$

The general problem of connecting state space representations with transfer function representations will be discussed later in this chapter. We shall only consider differential equations with constant coefficients in this course, although many results can be extended to differential equations whose coefficients vary with time.

Once we have obtained a state equation, we need to solve it to determine the state. For example, in the case of Newton's law (1.1), solving the corresponding state equation (1.2) gives the position and velocity of the particle. We now turn our attention, in the next few sections, to the solution of state equations.

1.1 The Transition Matrix for Linear Differential Equations

We begin by discussing the solution of the linear constant coefficient homogeneous differential equation

$$\begin{aligned} \dot{x}(t) &= Ax(t) \\ x(t) &\in \mathbb{R}^n \\ x(0) &= x_0 \end{aligned} \tag{1.4}$$

Very often, for simplicity, we shall suppress the time dependence in the notation. We shall now develop a systematic method for solving (1.4) using a matrix-valued function called the **transition matrix**.

It is known from the theory of ordinary differential equations that a linear systems of differential equations of the form

$$\begin{aligned} \dot{x}(t) &= Ax(t) \quad t \geq 0 \\ x(0) &= x_0 \end{aligned} \tag{1.5}$$

has a unique solution passing through x_0 at $t = 0$. To obtain the explicit solution for x, recall that the scalar linear differential equation

$$\begin{aligned} \dot{y} &= ay \\ y(0) &= y_0 \end{aligned}$$

has the solution

$$y(t) = e^{at} y_0$$

where the scalar exponential function e^{at} has the infinite series expansion

$$e^{at} = \sum_{k=0}^{\infty} \frac{a^k t^k}{k!} \tag{1.6}$$

and satisfies

$$\frac{d}{dt} e^{at} = a e^{at}$$

By analogy, let us define the **matrix exponential**

$$e^A = \sum_{n=0}^{\infty} \frac{A^n}{n!}$$

where A^0 is defined to be I, the identity matrix (analogous to $a^0 = 1$). We can then define a function e^{At}, which we call the **transition matrix** of (1.4), by

$$e^{At} = \sum_{k=0}^{\infty} \frac{A^k t^k}{k!} \tag{1.7}$$

The infinite sum can be shown to converge always, and its derivative can be evaluated by differentiating the infinite sum term by term. The analogy with the solution of the scalar equation suggests that $e^{At}x_0$ is the natural candidate for the the solution of (1.4).

1.2 Properties of the Transition Matrix and Solution of Linear ODEs

We now prove several useful properties of the transition matrix

(Property 1)
$$e^{At}|_{t=0} = I \tag{1.8}$$

Follows immediately from the definition of the transition matrix (1.7).

(Property 2)
$$\frac{d}{dt}\, e^{At} = Ae^{At} \tag{1.9}$$

To prove this, differentiate the infinite series in (1.7) term by term to get

$$\begin{aligned} \frac{d}{dt} e^{At} &= \frac{d}{dt} \sum_{k=0}^{\infty} \frac{A^k t^k}{k!} \\ &= \sum_{k=1}^{\infty} \frac{A^k t^{k-1}}{(k-1)!} \\ &= A \sum_{k=0}^{\infty} \frac{A^k t^k}{k!} \\ &= Ae^{At} \end{aligned}$$

Properties (1) and (2) together can be considered the defining equations of the transition matrix. In other words, e^{At} is the unique solution of the linear matrix differential equation

$$\frac{dF}{dt} = AF, \text{ with } F(0) = I \tag{1.10}$$

To see this, note that (1.10) is a linear differential equation. Hence there exists a unique solution to the initial value problem. Properties (1) and (2) show that e^{At} is a solution to (1.10). By uniqueness, e^{At} is the only solution. This shows (1.10) can be considered the defining equation for e^{At}.

(Property 3) If A and B commute, i.e. $AB = BA$, then

$$e^{(A+B)t} = e^{At}e^{Bt}$$

Proof: Consider $e^{At}e^{Bt}$

$$\frac{d}{dt}\,(e^{At}e^{Bt}) = Ae^{At}e^{Bt} + e^{At}Be^{Bt} \tag{1.11}$$

If A and B commutes, $e^{At}B = Be^{At}$ so that the R.H.S. of (1.11) becomes $(A+B)e^{At}e^{Bt}$. Furthermore, at $t = 0$, $e^{At}e^{Bt} = I$. Hence $e^{At}e^{Bt}$ satisfies the same differential equation as well as initial condition as $e^{(A+B)t}$. By uniqueness, they must be equal.
Setting $t = 1$ gives

$$e^{A+B} = e^A e^B \quad \text{whenever} \quad AB = BA$$

In particular, since At commutes with As for all t, s, we have

$$e^{A(t+s)} = e^{At}e^{As} \qquad \forall\ t, s$$

This is often referred to as the semigroup property.

(Property 4)
$$(e^{At})^{-1} = e^{-At}$$

Proof: We have

$$e^{-At}e^{At} = e^{A(t-t)} = I$$

Since e^{At} is a square matrix, this shows that (Property 4) is true. Thus, e^{At} is invertible for all t.

Using properties (1) and (2), we immediately verify that

$$x(t) = e^{At}x_0 \tag{1.12}$$

satisfies the differential equation

$$\dot{x}(t) = Ax(t) \tag{1.13}$$

and the initial condition $x(0) = x_0$. Hence it is the unique solution.

More generally,

$$x(t) = e^{A(t-t_0)}x_0 \quad t \geq t_0 \tag{1.14}$$

is the unique solution to (1.13) with initial condition $x(t_0) = x_0$.

1.3 Computing e^{At}: Linear Algebraic Methods

The above properties of e^{At} do not provide a direct way for its computation. In this section, we develop linear algebraic methods for computing e^{At}.

Let us first consider the special case where A is a diagonal matrix given by

$$A = \begin{bmatrix} \lambda_1 & 0 & & 0 \\ 0 & \lambda_2 & & \\ & & \ddots & 0 \\ 0 & & 0 & \lambda_n \end{bmatrix}$$

The λ_i's are not necessarily distinct and they may be complex, provided we interpret (1.13) as a differential equation in C^n. Since A is diagonal,

$$A^k = \begin{bmatrix} \lambda_1^k & 0 & & 0 \\ 0 & \lambda_2^k & & \\ & & \ddots & 0 \\ 0 & & 0 & \lambda_n^k \end{bmatrix}$$

If we directly evaluate the sum of the infinite series in the definition of e^{At} in (1.7), we find for example that the $(1,1)$ entry of e^{At} is given by

$$\sum_{k=0}^{\infty} \frac{\lambda_1^k t^k}{k!} = e^{\lambda_1 t}$$

Hence, in the diagonal case,

$$e^{At} = \begin{bmatrix} e^{\lambda_1 t} & 0 & & 0 \\ 0 & e^{\lambda_2 t} & & \\ & & \ddots & 0 \\ 0 & & 0 & e^{\lambda_n t} \end{bmatrix} \tag{1.15}$$

We can also interpret this result by examining the structure of the state equation. In this case (1.13) is completely decoupled into n differential equations

$$\dot{x}_i(t) = \lambda_i x_i(t) \qquad i = 1, ..., n \tag{1.16}$$

so that

$$x_i(t) = e^{\lambda_i t} x_{0i} \tag{1.17}$$

where x_{0i} is the ith component of x_0. It follows that

$$x(t) = \begin{bmatrix} e^{\lambda_1 t} & 0 & & 0 \\ 0 & e^{\lambda_2 t} & & \\ & & \ddots & 0 \\ 0 & & 0 & e^{\lambda_n t} \end{bmatrix} x_0 \tag{1.18}$$

Since x_0 is arbitrary, e^{At} must be given by (1.15).

We now extend the above results for a general A. We examine 2 cases.

Case I: A Diagonalizable

By definition, A diagonalizable means there exists a nonsingular matrix P (in general complex) such that $P^{-1}AP$ is a diagonal matrix, say

$$P^{-1}AP = \Lambda = \begin{bmatrix} \lambda_1 & 0 & & 0 \\ 0 & \lambda_2 & & \\ & & \ddots & 0 \\ 0 & & 0 & \lambda_n \end{bmatrix} \tag{1.19}$$

where the $\lambda_i' s$ are the eigenvalues of the A matrix. Then

$$e^{At} = e^{P\Lambda P^{-1}t} \tag{1.20}$$

Now notice that for any matrix B and nonsingular matrix P,

$$(PBP^{-1})^2 = PBP^{-1}PBP^{-1} = PB^2P^{-1}$$

so that in general

$$(PBP^{-1})^k = PB^kP^{-1} \tag{1.21}$$

We then have the following

$$e^{PBP^{-1}t} = Pe^{Bt}P^{-1} \qquad \text{for any } n \times n \text{ matrix } B \tag{1.22}$$

Proof:

$$\begin{aligned} e^{PBP^{-1}t} &= \sum_{k=0}^{\infty} \frac{(PBP^{-1})^k t^k}{k!} \\ &= \sum_{k=0}^{\infty} P \frac{B^k t^k}{k!} P^{-1} \\ &= P e^{Bt} P^{-1} \end{aligned}$$

Combining (1.15), (1.20) and (1.22), we find that whenever A is diagonalizable,

$$e^{At} = P \begin{bmatrix} e^{\lambda_1 t} & 0 & & 0 \\ 0 & e^{\lambda_2 t} & & \\ & & \ddots & 0 \\ 0 & & 0 & e^{\lambda_n t} \end{bmatrix} P^{-1} \tag{1.23}$$

From linear algebra, we know that a matrix A is diagonalizable if and only if it has n linearly independent eigenvectors. Each of the following two conditions is sufficient, though not necessary, to guarantee diagonalizability:

(i) A has distinct eigenvalues

(ii) A is a symmetric matrix

In these two cases, A has a complete set of n independent eigenvectors $v_1, v_2, \cdots, v_n$, with corresponding eigenvalues $\lambda_1, \lambda_2, \cdots, \lambda_n$. Write

$$P = \begin{bmatrix} v_1 & v_2 & \cdots & v_n \end{bmatrix} \tag{1.24}$$

Then

$$P^{-1}AP = \Lambda = \begin{bmatrix} \lambda_1 & 0 & & 0 \\ 0 & \lambda_2 & & \\ & & \ddots & 0 \\ 0 & & 0 & \lambda_n \end{bmatrix} \tag{1.25}$$

The equations (1.24), (1.25), and (1.23) together provide a linear algebraic method for omputing e^{At}. From (1.23), we also see that the dynamic behaviour of e^{At} is completely characterized by the behaviour of $e^{\lambda_i t}, i = 1, 2, \cdots, n$.

Example 1:

$$A = \begin{bmatrix} 1 & 2 \\ -1 & 4 \end{bmatrix}$$

The characteristic equation is given by

$$\det(sI - A) = s^2 - 5s + 6 = 0$$

giving eigenvalues $2, 3$. For the eigenvalue 2, its eigenvector v satisfies

$$(2I - A)v = \begin{bmatrix} 1 & -2 \\ 1 & -2 \end{bmatrix} v = 0$$

We can choose $v = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$. For the eigenvalue 3, its eigenvector w satisfies

$$(3I - A)w = \begin{bmatrix} 2 & -2 \\ 1 & -1 \end{bmatrix} w = 0$$

We can choose $w = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. The diagonalizing matrix P is given by

$$P = \begin{bmatrix} v & w \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$$

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

Finally

$$\begin{aligned} e^{At} &= Pe^{\Lambda t}P^{-1} \\ &= \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} e^{2t} & 0 \\ 0 & e^{3t} \end{bmatrix} \begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix} \\ &= \begin{bmatrix} 2e^{2t} - e^{3t} & -2e^{2t} + 2e^{3t} \\ e^{2t} - e^{3t} & -e^{2t} + 2e^{3t} \end{bmatrix} \end{aligned}$$

Example 2:

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

Owing to the form of A, we see right away that the eigenvalues are 1, 2 and -1 so that

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

The eigenvector corresponding to the eigenvalue 2 is given by $\begin{bmatrix} 0 & 1 & 0 \end{bmatrix}^T$, while that of -1 is $\begin{bmatrix} 0 & 0 & 1 \end{bmatrix}^T$. To find the eigenvector for the eigenvalue 1, we have

$$\begin{bmatrix} 1 & 0 & 0 \\ 1 & 2 & 0 \\ 1 & 0 & -1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}$$

so that $\begin{bmatrix} 2 \\ -2 \\ 1 \end{bmatrix}$ is an eigenvector. The diagonalizing matrix P is then

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

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

$$e^{At} = \begin{bmatrix} 2 & 0 & 0 \\ -2 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} e^t & 0 & 0 \\ 0 & e^{2t} & 0 \\ 0 & 0 & e^{-t} \end{bmatrix} \begin{bmatrix} \frac{1}{2} & 0 & 0 \\ 1 & 1 & 0 \\ -\frac{1}{2} & 0 & 1 \end{bmatrix} = \begin{bmatrix} e^t & 0 & 0 \\ -e^t + e^{2t} & e^{2t} & 0 \\ \frac{1}{2}(e^t - e^{-t}) & 0 & e^{-t} \end{bmatrix}$$

Model Decomposition:

We can give a dynamical interpretation to the above results. Suppose A has distinct eigenvalues $\lambda_1, ..., \lambda_n$ so that it has a set of linearly independent eigenvectors $v_1, ..., v_n$. If $x_0 = v_j$, then

$$\begin{aligned} x(t) &= e^{At} x_0 = e^{At} v_j \\ &= \sum_{k=0}^{\infty} \frac{A^k t^k}{k!} v_j = \sum_{k=0}^{\infty} \frac{\lambda_j^k t^k}{k!} v_j \\ &= e^{\lambda_j t} v_j \end{aligned}$$

This means that if we start along en eigenvector of A, the solution x will stay in the direction of the eigenvector, with length being stretched or shrunk by $e^{\lambda_j t}$. In general, because the $v_i's$ are linearly independent, an arbitrary initial condition x_0 can be expressed as

$$x_0 = \sum_{j=1}^{n} \xi_j v_j = P\xi \tag{1.26}$$

where

$$P = \begin{bmatrix} v_1 & v_2 & \cdots & v_n \end{bmatrix}$$

and $\xi = \begin{bmatrix} \xi_1 \\ \vdots \\ \xi_n \end{bmatrix}$ Using this representation, we have

$$\begin{aligned} x(t) &= e^{At} x_0 = e^{At} \sum_{j=1}^{n} \xi_j v_j \\ &= \sum_{j=1}^{n} \xi_j e^{At} v_j = \sum_{j=1}^{n} \xi_j e^{\lambda_j t} v_j \end{aligned} \tag{1.27}$$

so that $x(t)$ is expressible as a (time-varying) linear combination of the eigenvectors of A.

We can connect this representation of $x(t)$ with the representation of e^{At} in (1.23). Note that the P in (1.26) is the diagonalizing matrix for A, i.e.

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

and

$$\begin{aligned} e^{At} x_0 &= P e^{\Lambda t} P^{-1} x_0 = P e^{\Lambda t} \xi \\ &= P \begin{bmatrix} \xi_1 e^{\lambda_1 t} \\ \xi_2 e^{\lambda_2 t} \\ \vdots \\ \xi_n e^{\lambda_n t} \end{bmatrix} \\ &= \sum_j \xi_j e^{\lambda_j t} v_j \end{aligned}$$

the same result as in (1.27). The representation (1.27) of the solution $x(t)$ in terms of the eigenvectors of A is often called the modal decomposition of $x(t)$.

Case II: A not Diagonalizable

If A is not diagonalizable, then the above procedure cannot be carried out. If A does not have distinct eigenvalues, it is in general not diagonalizable since it may not have n linearly independent eigenvectors. For example, the matrix

$$A_{nl} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

which arises in the state space representation of Newton's Law, has 0 as a repeated eigenvalue and is not diagonalizable. On the other hand, note that this A_{nl} satisfies $A_{nl}^2 = 0$, so that we can just sum the infinite series

$$e^{A_{nl}t} = I + A_{nl}t = \begin{bmatrix} 1 & t \\ 0 & 1 \end{bmatrix}$$

The general situation is quite complicated, and we shall focus on the 3×3 case for illustration. Let us consider the following 3×3 matrix A which has the eigenvalue λ with multiplicity 3:

$$A = \begin{bmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{bmatrix} \tag{1.28}$$

Write $A = \lambda I + N$ where

$$N = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \tag{1.29}$$

Direct calculation shows that

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

and

$$N^3 = 0$$

A matrix B satisfying $B^k = 0$ for some $k \geq 1$ is called **nilpotent**. The smallest integer k such that $B^k = 0$ is called the index of nilpotence. Thus the matrix N in (1.29) is nilpotent with index 3. But then the transition matrix e^{Nt} is easily evaluated to be

$$e^{Nt} = \sum_{j=0}^{2} \frac{N^j t^j}{j!} = \begin{bmatrix} 1 & t & \frac{t^2}{2!} \\ 0 & 1 & t \\ 0 & 0 & 1 \end{bmatrix} \tag{1.30}$$

Once we understand the 3×3 case, it is not difficult to check that an $l \times l$ matrix of the form

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

i.e., $1's$ along the superdiagonal, and 0 everywhere else, is nilpotent and satisfies $N^l = 0$

To compute e^{At} when A is of the form (1.28), recall Property 4 of matrix exponentials:

If A and B commute, i.e. $AB = BA$, then $e^{(A+B)t} = e^{At}e^{Bt}$.

If A is of the form (1.28), then since λI commutes with N, we can write, using Property 3 of transition matrices,

$$\begin{aligned} e^{At} &= e^{(\lambda I+N)t} = e^{(\lambda I)t}e^{Nt} = (e^{\lambda t}I)(e^{Nt}) \\ &= e^{\lambda t}\begin{bmatrix} 1 & t & \frac{t^2}{2!} \\ 0 & 1 & t \\ 0 & 0 & 1 \end{bmatrix} \end{aligned} \tag{1.31}$$

using (1.30).

We are now in a position to sketch the general structure of the matrix exponential. More details are given in the Appendix to this chapter. The idea is to transform a general A using a nonsingular matrix P into

$$P^{-1}AP = J = \begin{bmatrix} J_1 & & & 0 \\ & J_2 & & \\ & & \ddots & \\ 0 & & & J_k \end{bmatrix}$$

where each $J_i = \lambda_i I + N_i$, N_i nilpotent. J is called the Jordan form of A. With the block diagonal structure, we can verify

$$J^m = \begin{bmatrix} J_1^m & & & 0 \\ & J_2^m & & \\ & & \ddots & \\ 0 & & & J_k^m \end{bmatrix}$$

By applying the definition of the transition matrix, we obtain

$$e^{Jt} = \begin{bmatrix} e^{J_1 t} & & & 0 \\ & e^{J_2 t} & & \\ & & \ddots & \\ 0 & & & e^{J_k t} \end{bmatrix} \tag{1.32}$$

Each $e^{J_i t}$ can be evaluated by the method described above. Finally,

$$e^{At} = Pe^{Jt}P^{-1} \tag{1.33}$$

The above procedure in principle enables us to evaluate e^{At}, the only problem being to find the matrix P which transforms A either to diagonal or Jordan form. In general, this can be a very tedious task (See Appendix to Chapter 1 for an explanation on how it can be done). The above formula is most useful as a means for studying the qualitative dependence of e^{At} on t, and will be particularly important in our discussion of stability.

We remark that the above results hold regardless of whether the λ_i's are real or complex. In the latter case, we shall take the underlying vector space to be complex and the results then go through without modification.

1.4 Computing e^{At}: Laplace Transform Method

Another method of evaluating e^{At} analytically is to use Laplace transforms.

If we let $G(t) = e^{At}$, then the Laplace transform of $G(t)$, denoted by $\hat{G}(s)$, satisfies

$$s\hat{G}(s) = A\hat{G}(s) + I$$

or

$$\hat{G}(s) = (sI - A)^{-1} \tag{1.34}$$

If we denote the inverse Laplace transform operation by $\mathcal{L}^{-1}$

$$e^{At} = \mathcal{L}^{-1}[(sI - A)^{-1}] \tag{1.35}$$

The Laplace transform methods thus involves 2 steps:

(a) Find $(sI - A)^{-1}$. The entries of $(sI - A)^{-1}$ will be strictly proper rational functions.

(b) Determine the inverse Laplace transform for each entry of $(sI - A)^{-1}$. This is usually accomplished by doing partial fractions expansion and using the fact that $\mathcal{L}^{-1}[\frac{1}{s-\lambda}] = e^{\lambda t}$ (or looking up a mathematical table!).

Determining $(sI - A)^{-1}$ using the cofactor expansion method from linear algebra can be quite tedious. Here we give an alternative method. Let $\det(sI - A) = s^n + p_1 s^{n-1} + ... + p_n = p(s)$. Write

$$(sI - A)^{-1} = \frac{B(s)}{p(s)} = \frac{s^{n-1}B_1 + s^{n-2}B_2 + ... + B_n}{p(s)} \tag{1.36}$$

Then the B_i matrices can be determined recursively as follows:

$$\begin{aligned} B_1 &= I \\ B_{k+1} &= AB_k + p_k I \quad 1 \leq k \leq n-1 \end{aligned} \tag{1.37}$$

To illustrate the Laplace transform method for evaluating e^{At}, again consider

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

$$\begin{aligned} p(s) &= \det \begin{bmatrix} s-1 & 0 & 0 \\ -1 & s-2 & 0 \\ -1 & 0 & s+1 \end{bmatrix} = (s-1)(s-2)(s+1) \\ &= s^3 - 2s^2 - s + 2 \end{aligned}$$

The matrix polynomial $B(s)$ can then be determined using the recursive procedure (1.37).

$$B_2 = A - 2I = \begin{bmatrix} -1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & -3 \end{bmatrix}$$

$$B_3 = AB_2 - I = \begin{bmatrix} -2 & 0 & 0 \\ 1 & -1 & 0 \\ -2 & 0 & 2 \end{bmatrix}$$

giving

$$B(s) = \begin{bmatrix} s^2-s-2 & 0 & 0 \\ s+1 & s^2-1 & 0 \\ s-2 & 0 & s^2-3s+2 \end{bmatrix}$$

Putting into (1.36) and doing partial fractions expansion, we obtain

$$\begin{aligned}
(sI-A)^{-1} &= \frac{1}{(s-1)(-2)}\begin{bmatrix} -2 & 0 & 0 \\ 2 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix} + \frac{1}{(s-2)(3)}\begin{bmatrix} 0 & 0 & 0 \\ 3 & 3 & 0 \\ 0 & 0 & 0 \end{bmatrix} \\
&\quad + \frac{1}{(s+1)(6)}\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ -3 & 0 & 6 \end{bmatrix} \\
e^{At} &= \begin{bmatrix} e^t & 0 & 0 \\ -e^t & 0 & 0 \\ \frac{1}{2}e^t & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ e^{2t} & e^{2t} & 0 \\ 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ -\frac{1}{2}e^{-t} & 0 & e^{-t} \end{bmatrix} \\
&= \begin{bmatrix} e^t & 0 & 0 \\ -e^t+e^{2t} & e^{2t} & 0 \\ \frac{1}{2}(e^t-e^{-t}) & 0 & e^{-t} \end{bmatrix}
\end{aligned}$$

the same result as before.

As a second example, let $A = \begin{bmatrix} \sigma & \omega \\ -\omega & \sigma \end{bmatrix}$. Then

$$A = \begin{bmatrix} \sigma & 0 \\ 0 & \sigma \end{bmatrix} + \begin{bmatrix} 0 & \omega \\ -\omega & 0 \end{bmatrix} \quad .$$

But

$$e^{\begin{bmatrix} 0 & \omega \\ -\omega & 0 \end{bmatrix} t} = \mathcal{L}^{-1}\left\{ \begin{bmatrix} s & -\omega \\ \omega & s \end{bmatrix}^{-1} \right\}$$

where $\mathcal{L}^{-1}$ is the inverse Laplace transform operator

$$\begin{aligned}
&= \mathcal{L}^{-1}\begin{bmatrix} \frac{s}{s^2+\omega^2} & \frac{\omega}{s^2+\omega^2} \\ -\frac{\omega}{s^2+\omega^2} & \frac{s}{s^2+\omega^2} \end{bmatrix} = \begin{bmatrix} \cos\omega t & \sin\omega t \\ -\sin\omega t & \cos\omega t \end{bmatrix} \\
e^{At} &= e^{(\sigma I)t}\begin{bmatrix} \cos\omega t & \sin\omega t \\ -\sin\omega t & \cos\omega t \end{bmatrix} = \begin{bmatrix} e^{\sigma t}\cos\omega t & e^{\sigma t}\sin\omega t \\ -e^{\sigma t}\sin\omega t & e^{\sigma t}\cos\omega t \end{bmatrix}
\end{aligned}$$

In this and the previous sections, we have described analytically procedures for computing e^{At}. Of course, when the dimension n is large, these procedures would be virtually impossible to carry out. In general, we must resort to numerical techniques. Numerically stable and efficient methods of evaluating the matrix exponential can be found in the research literature.

1.5 Differential Equations with Inputs and Outputs

The solution of the homogeneous equation (1.4) can be easily generalized to differential equations with inputs. Consider the equation

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bu(t) \\ x(0) &= x_0 \end{aligned} \tag{1.38}$$

where u is a piecewise continuous $\mathbb{R}^m$-valued function. Define a function $z(t) \;=\; e^{-At}x(t)$. Then $z(t)$ satisfies

$$\begin{aligned} \dot{z}(t) &= -e^{-At}Ax(t) + e^{-At}Ax(t) + e^{-At}Bu(t) \\ &= e^{-At}Bu(t) \end{aligned}$$

Since the above equation does not depend on $z(t)$ on the right hand side, it can be directly integrated to give

$$z(t) = z(0) + \int_0^t e^{-As}Bu(s)ds = x(0) + \int_0^t e^{-As}Bu(s)ds$$

Hence

$$\begin{aligned} x(t) &= e^{At}z(t) \\ &= e^{At}x_0 + \int_0^t e^{A(t-s)}Bu(s)ds \end{aligned} \tag{1.39}$$

More generally, we have

$$x(t) = e^{A(t-t_0)}x(t_0) + \int_{t_0}^t e^{A(t-s)}Bu(s)ds \tag{1.40}$$

This is the variation of parameters formula for solving (1.38).

We now consider linear systems with inputs and outputs.

$$\dot{x}(t) = Ax(t) + Bu(t) \tag{1.41}$$
$$x(0) = x_0$$
$$y(t) = Cx(t) + Du(t) \tag{1.42}$$

where the output y is a piecewise continuous $\mathbb{R}^p$-valued function. Using (1.39) in (1.42) give the solutions immediately

$$y(t) = Ce^{At}x_0 + \int_0^t Ce^{A(t-s)}Bu(s)ds + Du(t) \tag{1.43}$$

In the case $x_0 = 0$ and $D = 0$, we find

$$y(t) = \int_0^t Ce^{A(t-s)}Bu(s)ds$$

which is of the form $y(t) \;=\; \int_0^t h(t-s)u(s)ds$, a convolution integral. The function $h(t) = Ce^{At}B$ is called the impulse response of the system. If we allow generalized functions, we can incorporate a nonzero D term as

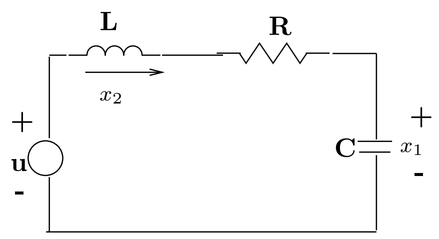
$$h(t) = Ce^{At}B + D\delta(t) \tag{1.44}$$

where $\delta(t)$ is the Dirac δ-function. Noting that the Laplace transform of e^{At} is $(sI - A)^{-1}$, the transfer function of the linear system from u to y, which is the Laplace transform of the impulse response, is given by

$$H(s) = C(sI - A)^{-1}B + D \tag{1.45}$$

The entries of $H(s)$ are proper rational functions, i.e., ratios of polynomials with degree of numerator $\leq$ degree of denominator. If $D = 0$ (no direct transmission from u to y), the entries are strictly proper. This is an important property of transfer functions arising from a state space representation of the form given by (1.41), (1.42).

As an example, consider the standard circuit below.



The choice of x_1 to denote the voltage across the capacitor and x_2 the current through the inductor gives a state description for the circuit. The equations are

$$C\dot{x}_1 = x_2 \tag{1.46}$$

$$L\dot{x}_2 + Rx_2 + x_1 = u \tag{1.47}$$

Converting into standard form, we have

$$\dot{x} = \begin{bmatrix} 0 & \frac{1}{C} \\ -\frac{1}{L} & -\frac{R}{L} \end{bmatrix} + \begin{bmatrix} 0 \\ \frac{1}{L} \end{bmatrix} u \tag{1.48}$$

Take $C = 0.5$, $L = 1$, $R = 3$. The circuit state equation satisfies

$$\dot{x} = \begin{bmatrix} 0 & 2 \\ -1 & -3 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u = Ax + Bu \tag{1.49}$$

We have

$$\begin{aligned} (sI - A)^{-1} &= \begin{bmatrix} s & -2 \\ 1 & s+3 \end{bmatrix}^{-1} \\ &= \frac{\begin{bmatrix} s+3 & 2 \\ -1 & s \end{bmatrix}}{s^2 + 3s + 2} \\ &= \begin{bmatrix} \frac{2}{s+1} - \frac{1}{s+2} & \frac{2}{s+1} - \frac{2}{s+2} \\ -\frac{1}{s+1} + \frac{1}{s+2} & -\frac{1}{s+1} + \frac{2}{s+2} \end{bmatrix} \end{aligned}$$

Upon inversion, we get

$$e^{At} = \begin{bmatrix} 2e^{-t} - e^{-2t} & 2e^{-t} - 2e^{-2t} \\ -e^{-t} + e^{-2t} & -e^{-t} + 2e^{-2t} \end{bmatrix} \tag{1.50}$$

If we are interested in the voltage across the capacitor as the output, we then have

$$y = x_1 = \begin{bmatrix} 1 & 0 \end{bmatrix} x$$

The impluse response from the input voltage source to the output voltage is given by

$$h(t) = Ce^{At}B = 2e^{-t} - 2e^{-2t}$$

If we have a constant input $u = 1$, the output voltage using (1.39), assuming initial rest ($x_0 = 0$), is given by

$$\begin{aligned} y(t) &= \int_0^t [2e^{-(t-s)} - 2e^{-2(t-s)}]ds \\ &= 2(1 - e^{-t}) - (1 - e^{-2t}) \end{aligned}$$

Armed with (1.39) and (1.43), we can see why (1.38) deserves the name of state equation. The solutions for $x(t)$ and $y(t)$ from (1.39) and (1.43) show that indeed knowledge of the current state (x_0) and future inputs $u(t), t \geq 0$ determines the future states $x(t), t \geq 0$ and outputs $y(t), t \geq 0$.

1.6 Stability

Let us consider the unforced state equation first:

$$\dot{x} = Ax, \quad x(0) = x_0 \; .$$

The vector $x(t) \rightarrow 0$ as $t \rightarrow \infty$ for every $x_0 \Leftrightarrow$ all the eigenvalues of A lie in the open left half-plane.

Idea of proof: From (1.33),

$$x(t) = e^{At}x_0 = Pe^{Jt}P^{-1}x_0$$

The entries in the matrix e^{Jt} are of the form $e^{\lambda_i t}p(t)$ where $p(t)$ is a polynomial in t, and λ_i is an eigenvalue of A. These entries converge to 0 if and only if $\mathrm{Re}\lambda_i < 0$. Thus

$$x(t) \rightarrow 0 \quad \forall x_0$$

$$\Leftrightarrow \text{all eigenvalues of } A \text{ lie in } \{\lambda : \mathrm{Re}\lambda < 0\}$$

Now let us look at the full system model:

$$\begin{aligned} \dot{x} &= Ax + Bu \\ y &= Cx + Du \; . \end{aligned}$$

The transfer matrix from u to y is

$$H(s) = C(sI - A)^{-1}B + D \; .$$

We can write this as

$$H(s) = \frac{1}{\det(sI - A)} C \cdot \mathrm{adj}(sI - A) \cdot B + D \; .$$

Notice that the elements of the matrix $\mathrm{adj}(sI - A)$ are all polynomials in s; consequently, they have no poles. Notice also that $\det(sI - A)$ is the characteristic polynomial of A. We can therefore conclude from the preceding equation that

$$\{\text{eigenvalues of } A\} \supset \{\text{poles of } H(s)\} \; .$$

Hence, if all the eigenvalues of A are in the open left half-plane, then $H(s)$ is a stable transfer matrix. The converse is not necessarily true.

Example:

$$A = \begin{bmatrix} -1 & 0 \\ 0 & 2 \end{bmatrix}, \qquad B = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$
$$C = [1 \ \ 0], \qquad D = 0$$
$$H(s) = \frac{1}{s+1} \qquad .$$

So the transfer function has stable poles, but the system is obviously unstable. Any initial condition x_0 with a nonzero second component will cause $x(t)$ to grow without bound.

A more subtle example is the following

Example:

$$A = \begin{bmatrix} 0 & 1 & 1 \\ -2 & -2 & 0 \\ 2 & 1 & -1 \end{bmatrix}, \qquad B = \begin{bmatrix} -1 & 0 \\ 2 & 1 \\ -1 & -1 \end{bmatrix}$$
$$C = [2 \ \ 2 \ \ 1], \qquad D = [0 \ \ 0]$$

$$\{\text{eigenvalues of } A\} = \{0, -1, -2\}$$

$$\begin{aligned}
H(s) &= \frac{1}{\det(sI - A)} C \cdot \operatorname{adj}(sI - A) \cdot B \\
&= \frac{1}{s^2 + 3s^2 + 2s} \cdot [2 \ \ 2 \ \ 1] \cdot \begin{bmatrix} s^2 + 3s + 2 & s + 2 & s + 2 \\ -2s - 2 & s^2 + s - 2 & -2 \\ 2s + 2 & s + 2 & s^2 + 2s + 2 \end{bmatrix} \begin{bmatrix} -1 & 0 \\ 2 & 1 \\ -1 & -1 \end{bmatrix} \\
&= \frac{1}{s^2 + 3s^2 + 2s} [2s^2 + 4s + 2 \ \ 2s^2 + 5s + 2 \ \ s^2 + 4s + 2] \begin{bmatrix} -1 & 0 \\ 2 & 1 \\ -1 & -1 \end{bmatrix} \\
&= \frac{1}{s^3 + 3s^2 + 2s} [s^2 + 2s \ \ s^2 + s] \\
&= \begin{bmatrix} \frac{s+2}{s^2 + 3s + 2} & \frac{s+1}{s^2 + 3s + 2} \end{bmatrix}
\end{aligned}$$

Thus {poles of $H(s)$} = {-1,-2}. Hence the eigenvalue of A at $\lambda = 0$ does not appear as a pole of $H(s)$. Note that there is a cancellation of a factor of s in the numerator and denominator of $H(s)$. This corresponds to a pole-zero cancellation in the determination of $H(s)$, and indicates something internal in the system cannot be seen from the input-output behaviour.

1.7 Transfer Function to State Space

We have already seen, in Section 1.5, how we can determine the impulse response and the transfer function of a linear time-invariant system from its state space description. In this section, we discuss the converse, harder problem of determining a state space description of a linear time-invariant system from its transfer function. This is also called the **realization problem** in control theory. We shall only solve the single-input single-output case, as the general multivariable problem is much harder and beyond the scope of this course.

Suppose we are given the scalar-valued transfer function $H(s)$ of a single-input single-output linear time invariant system. We assume $H(s)$ is a proper rational function; otherwise we cannot find a state space representation of the form (1.41), (1.42) for $H(s)$. We also assume that there are no common factors in the numerator and denominator of $H(s)$. After long division, if necessary, we can write

$$H(s) = d + \frac{q(s)}{p(s)}$$

where degree of $q(s) <$ degree of $p(s)$. The constant d corresponds to the direct transmission term from u to y, so the realization problem reduces to finding a triple (c^T, A, b) such that

$$c^T(sI - A)^{-1}b = \frac{q(s)}{p(s)}$$

so that the transfer function $H(s)$ is given by

$$H(s) = c^T(sI - A)^{-1}b + d = \hat{H}(s) + d$$

Note that we have used the notation c^T for the $1 \times n$ matrix C, in keeping with our normal convention that lower case letters denote column vectors.

We now give 2 solutions to the realization problem. Let

$$\hat{H}(s) = \frac{q(s)}{p(s)}$$

with

$$\begin{aligned} p(s) &= s^n + p_1 s^{n-1} + \cdots + p_n \\ q(s) &= q_1 s^{n-1} + q_2 s^{n-2} + \cdots + q_n \end{aligned}$$

Then the following (c^T, A, b) triple realizes $\frac{q(s)}{p(s)}$

$$A = \begin{bmatrix} 0 & \cdots & 0 & -p_n \\ 1 & 0 & \cdots & \vdots \\ \vdots & \ddots & \vdots & \vdots \\ \cdots & \cdots & 1 & -p_1 \end{bmatrix}$$

$$b = \begin{bmatrix} q_n \\ \vdots \\ q_1 \end{bmatrix}$$

$$c^T = [0 \ \cdots \ 0 \ \ 1]$$

To see this, let

$$\xi^T(s) = [\xi_1(s) \ \ \xi_2(s) \ \ldots \ \xi_n(s)]$$

$$= c^T(sI - A)^{-1}$$

Then

$$\xi^T(s)(sI - A) = c^T \tag{1.51}$$

Writing out (1.51) in component form, we have

$$\begin{aligned} s\xi_1(s) - \xi_2(s) &= 0 \\ s\xi_2(s) - \xi_3(s) &= 0 \\ &\vdots \\ s\xi_{n-1}(s) - \xi_n(s) &= 0 \\ s\xi_n(s) + \sum_{i=1}^{n} \xi_{n-i+1}(s)p_i &= 1 \end{aligned}$$

Successively solving these equations, we find

$$\xi_k(s) = s^{k-1}\xi_1(s), \quad 2 \le k \le n$$

and

$$s^n\xi_1(s) + \sum_{i=1}^{n} p_i s^{n-i}\xi_1(s) = 1 \tag{1.52}$$

so that

$$\xi_1(s) = \frac{1}{p(s)}$$

Hence

$$\xi^T(s) = \frac{1}{p(s)}[1 \quad s \quad s^2 \cdots \quad s^{n-1}] \tag{1.53}$$

and

$$c^T(sI - A)^{-1}b = \xi^T(s)b = \frac{q(s)}{p(s)}$$

The resulting system is often said to be in observable canonical form, We shall discuss the meaning of the word "observable" in control theory later in the course.

As an example, we can now write down the state space representation for the equation given in (1.3). The transfer function of (1.3) is

$$H(s) = \frac{s + 4}{s^2 + 3s + 2}$$

The observable representation is therefore given by

$$\begin{aligned} \dot{x} &= \begin{bmatrix} 0 & -2 \\ 1 & -3 \end{bmatrix} x + \begin{bmatrix} 4 \\ 1 \end{bmatrix} u \\ y &= \begin{bmatrix} 0 & 1 \end{bmatrix} x \end{aligned}$$

Now note that because $H(s)$ is scalar-valued, $H(s) = H^T(s)$. We can immediately conclude that the following (c_1^T, A_1, b_1) triple also realizes $\frac{q(s)}{p(s)}$

$$A_1 = \begin{bmatrix} 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \cdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & & & 1 \\ -p_n & -p_{n-1} & & \cdots & -p_2 & -p_1 \end{bmatrix}$$

$$b_1 = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}$$

$$c_1^T = [q_n \ ... \ q_1]$$

so that $H(s)$ can also be written as $H(s) = d + c_1^T(sI - A_1)^{-1}b_1$.

The resulting system is said to be in controllable canonical form. We shall discuss the meaning of the word "controllable" in the next chapter. The controllable and observable canonical forms will be used later for control design.

1.8 Linearization

So far, we have only considered systems described by linear constant coefficient differential equations. Many systems are nonlinear, with dynamics described by nonlinear differential equations of the form

$$\dot{x} = f(x, u) \tag{1.54}$$

where f has n components, each of which is a continuously differentiable function of $x_1, \cdots x_n, u_1, \cdots, u_m$. Suppose the constant vectors (x_p, u_p) corresponds to a desired operating point, also called an equilibrium point, for the system, i.e.

$$f(x_p, u_p) = 0$$

If the initial condition $x_0 = x_p$, then by using $u(t) = u_p$ for all t, $x(t) = x_p$ for all t, i.e. the desired operating point is maintained. However, if $x_0 \neq x_p$ then to satisfy (1.54), (x, u) will be different from (x_p, u_p). However, if x_0 is close to x_p then by continuity, we expect x and u to be close to x_p and u_p as well. Let

$$\begin{aligned} x &= x_p + \delta x \\ u &= u_p + \delta u \end{aligned}$$

Taylor's series expansion shows that

$$\begin{aligned} f_i(x, u) &= f_i(x_p + \delta x, u_p + \delta u) \\ &= [\frac{\partial f_i}{\partial x_1} \ \frac{\partial f_i}{\partial x_2} \ \cdots \ \frac{\partial f_i}{\partial x_n}]_{(x_p, u_p)} \delta x + [\frac{\partial f_i}{\partial u_1} \ \frac{\partial f_i}{\partial u_2} \ \cdots \ \frac{\partial f_i}{\partial u_m}]_{(x_p, u_p)} \delta u + \text{h.o.t.} \end{aligned}$$

If we drop higher order terms (h.o.t.) involving δx_j and δu_k, we obtain a linear differential equation for δx

$$\frac{d}{dt}\delta x = A\delta x + B\delta u \tag{1.55}$$

where $[A]_{jk} = \frac{\partial f_j}{\partial x_k}(x_p, u_p)$, $[B]_{jk} = \frac{\partial f_j}{\partial u_k}(x_p, u_p)$. A is called the Jacobian matrix of f with respect to x, sometimes, simply written as $\frac{\partial f}{\partial x}(x_p, u_p)$. Similarly, $B = \frac{\partial f}{\partial u}(x_p, u_p)$.

The resulting system (1.55) is called the linearized system of the nonlinear system (1.54) about the operating point (x_p, u_p). The process is called linearization of the nonlinear system about (x_p, u_p). Keeping the original system near the operating point can often be achieved by keeping the linearized system near 0.

Example:

Consider a pendulum subjected to an applied torque. The normalized equation of motion is

$$\ddot{\theta} + \frac{g}{l}\sin\theta = u \tag{1.56}$$

where g is the gravitational constant, and l is the length of the pendulum. Let $x = \begin{bmatrix} \theta & \dot{\theta} \end{bmatrix}^T$. We have the following nonlinear differential equation

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= -\frac{g}{l}\sin x_1 + u \end{aligned} \tag{1.57}$$

We recognize

$$f(x,u) = \begin{bmatrix} x_2 \\ -\frac{g}{l}\sin x_1 + u \end{bmatrix}$$

Clearly $(x, u) = (0, 0)$ is an equilibrium point. The linearized system is given by

$$\dot{x} = \begin{bmatrix} 0 & 1 \\ -\frac{g}{l} & 0 \end{bmatrix} x + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u = A_0 x + Bu$$

The system matrix A_0 has purely imaginary eigenvalues, corresponding to harmonic motion. Note, however, that $(x_1, x_2, u) = (\pi, 0, 0)$ is also an equilibrium point, corresponding to the pendulum in the inverted vertical position. The linearized system is then given by

$$\dot{x} = \begin{bmatrix} 0 & 1 \\ \frac{g}{l} & 0 \end{bmatrix} x + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u = A_\pi x + Bu$$

The system matrix A_π has an unstable eigenvalue, and the behaviour of the linearized system about the 2 equilibrium points are quite different.

Appendix to Chapter 1: Jordan Forms

This appendix provides more details on the Jordan form of a matrix and its use in computing the transition matrix. The results are deep and complicated, and we cannot give a complete exposition here. Standard textbooks on linear algebra should be consulted for a more thorough treatment.

When a matrix A has repeated eigenvalues, it is usually not diagonalizable. However, there always exists a nonsingular matrix P such that

$$P^{-1}AP = J = \begin{bmatrix} J_1 & & & 0 \\ & J_2 & & \\ & & \ddots & \\ 0 & & & J_k \end{bmatrix} \tag{A.1}$$

where J_i is of the form

$$J_i = \begin{bmatrix} \lambda_i & 1 & 0 & \cdots & 0 \\ & \lambda_i & 1 & & 0 \\ & & & & \vdots \\ & & & & 1 \\ 0 & & & & \lambda_i \end{bmatrix}, \qquad \text{an } m_i \times m_i \text{ matrix} \tag{A.2}$$

The $\lambda_i's$ need not be all distinct, but $\sum_{i=1}^{k} m_i = n$. The special form on the right hand side of (A.1) is called the Jordan form of A, and J_i is the Jordan block associated with eigenvalue λ_i. Note that if $m_i = 1$ for all i, i.e., each Jordan block is 1×1 and $k = n$, the Jordan form specializes to the diagonal form.

There are immediately some natural questions which arise in connection with the Jordan form: for an eigenvalue λ_i with multiplicity n_i, how many Jordan blocks can there be associated with λ_i, and how do we determine the size of the various blocks? Also, how do we determine the nonsingular P which transforms A into Jordan form?

First recall that the **nullspace** of A, denoted by $\mathcal{N}(A)$, is the set of vectors x such that $Ax = 0$. $\mathcal{N}(A)$ is a subspace and has a dimension. If λ is an eigenvalue of A, the dimension of $\mathcal{N}(A - \lambda I)$ is called the geometric multiplicity of λ. The geometric multiplicity of λ corresponds to the number of independent eigenvectors associated with λ.

Fact: Suppose λ is an eigenvalue of A. The number of Jordan blocks associated with λ is equal to the geometric multiplicity of λ.

The number of times λ is repeated as a root of the characteristic equation $\det(sI - A) = 0$ is called its algebraic multiplicity (commonly simplified to just multiplicity). The geometric multiplicity of λ is always less than or equal to its algebraic multiplicity. If the geometric multiplicity is strictly less than the algebraic multiplicity, λ has **generalized eigenvectors**. A generalized eigenvector is a nonzero vector v such that for some positive integer k, $(A - \lambda I)^k v = 0$, but that $(A - \lambda I)^{k-1} v \neq 0$. Such a vector v defines a chain of generalized eigenvectors $\{v_1, v_2, \cdots, v_k\}$ through the equations $v_k = v$, $v_{k-1} = (A - \lambda I)v_k$, $v_{k-2} = (A - \lambda I)v_{k-1}, \cdots$, $v_1 = (A - \lambda I)v_2$. Note that $(A - \lambda I)v_1 = (A - \lambda I)^k v = 0$ so that v_1 is an eigenvector. This set of equations is often written as

$$\begin{aligned} Av_2 &= \lambda v_2 + v_1 \\ Av_3 &= \lambda v_3 + v_2 \\ &\vdots \\ Av_k &= \lambda v_k + v_{k-1} \end{aligned}$$

It can be shown that the generalized eigenvectors associated with an eigenvalue λ are linearly independent. The nonsingular matrix P which transforms A into its Jordan form is constructed from the set of eigenvectors and generalized eigenvectors.

The complete procedure for determining all the generalized eigenvectors, though too complex to describe precisely here, can be well-illustrated by an example.

Example:

Consider the matrix A given by

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

$$\det(sI - A) = s^3(s+3)$$

so that the eigenvalues are -3, and 0, with algebraic multiplicity 3. Let us first determine the geometric multiplicity of the eigenvalue 0. The equation

$$Aw = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 2 & -2 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ w_3 \\ w_4 \end{bmatrix} = 0 \tag{A.3}$$

gives $w_3 = 0$, $w_4 = 0$, and w_1, w_2 arbitrary. Thus the geometric multiplicity of 0 is 2, i.e., there are 2 Jordan blocks associated with 0. Since 0 has algebraic multiplicity 3, this means there are 2 eigenvectors and 1 generalized eigenvector. To determine the generalized eigenvector, we solve the homogeneous equation $A^2 v = 0$, for solutions v such that $Av \neq 0$. Now

$$\begin{aligned} A^2 &= \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 2 & -2 \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 2 & -2 \end{bmatrix} \\ &= \begin{bmatrix} 0 & 0 & -1 & 1 \\ 0 & 0 & 2 & -2 \\ 0 & 0 & 3 & -3 \\ 0 & 0 & -6 & 6 \end{bmatrix} \end{aligned}$$

It is readily seen that

$$v = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}$$

solves $A^2 v = 0$ and

$$Av = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}$$

so that we have the chain of generalized eigenvectors

$$v_2 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}, \quad v_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}$$

with v_1 an eigenvector. From (A.3), a second independent eigenvector is $w = \begin{bmatrix} 1 & 0 & 0 & 0 \end{bmatrix}^T$. This completes the determination of the eigenvectors and generalized eigenvectors associated with 0. Finally the eigenvector for the eigenvalue -3 is given by the solution of

$$(A + 3I)z = \begin{bmatrix} 3 & 0 & 1 & 0 \\ 0 & 3 & 0 & 1 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 2 & 1 \end{bmatrix} z = 0$$

yielding $z = \begin{bmatrix} 1 & -2 & -3 & 6 \end{bmatrix}^T$ The nonsingular matrix bringing A to Jordan form is then given by

$$P = \begin{bmatrix} w & v_1 & v_2 & z \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 1 & -3 \\ 0 & 0 & 1 & 6 \end{bmatrix}$$

Note that the chain of generalized eigenvectors v_1, v_2 go together. One can verify that

$$P^{-1} = \begin{bmatrix} 1 & -1 & \frac{1}{3} & -\frac{1}{3} \\ 0 & 1 & -\frac{2}{9} & -\frac{2}{9} \\ 0 & 0 & \frac{2}{3} & \frac{1}{3} \\ 0 & 0 & -\frac{1}{9} & \frac{1}{9} \end{bmatrix}$$

and

$$P^{-1}AP = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -3 \end{bmatrix} \tag{A.4}$$

The ordering of the Jordan blocks is not unique. If we choose

$$M = \begin{bmatrix} v_1 & v_2 & w & z \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 & 1 \\ 1 & 0 & 0 & -2 \\ 0 & 1 & 0 & -3 \\ 0 & 1 & 0 & 6 \end{bmatrix}$$

$$M^{-1}AM = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -3 \end{bmatrix} \tag{A.5}$$

although it is common to choose the blocks ordered in decreasing size, i.e., that given in (A.5).

Once we have transformed A into its Jordan form (A.1), we can readily determine the transition matrix e^{At}. As before, we have

$$A^m = (PJP^{-1})^m = PJ^mP^{-1}$$

By the block diagonal nature of J, we readily see

$$J^m = \begin{bmatrix} J_1^m & & & 0 \\ & J_2^m & & \\ & & \ddots & \\ 0 & & & J_k^m \end{bmatrix}$$

Thus,

$$e^{At} = P \begin{bmatrix} e^{J_1 t} & & & 0 \\ & e^{J_2 t} & & \\ & & \ddots & \\ 0 & & & e^{J_k t} \end{bmatrix} P^{-1}$$

Since the $m_i \times m_i$ block $J_i = \lambda_i I + N_i$, using the methods of Section 1.3, we can immediately write down

$$e^{J_i t} = e^{\lambda_i t} \begin{bmatrix} 1 & t & \cdots & \frac{t^{m_i-1}}{(m_i-1)!} \\ & & \ddots & \\ & & \ddots & t \\ 0 & & & 1 \end{bmatrix}$$

We have now a complete linear algebraic procedure for determining e^{At} when A is not diagonalizable.

For

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

and using the M matrix for transformation to Jordan form, we have

$$\begin{aligned} e^{At} &= M \begin{bmatrix} 1 & t & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & e^{-3t} \end{bmatrix} M^{-1} \\ &= \begin{bmatrix} 1 & 0 & \frac{1}{9} + \frac{2}{3}t - \frac{1}{9}e^{-3t} & -\frac{1}{9} + \frac{1}{3}t + \frac{1}{9}e^{-3t} \\ 0 & 1 & -\frac{2}{9} + \frac{2}{3}t + \frac{2}{9}e^{-3t} & \frac{2}{9} + \frac{1}{3}t - \frac{2}{9}e^{-3t} \\ 0 & 0 & \frac{2}{3} + \frac{1}{3}e^{-3t} & \frac{1}{3} - \frac{1}{3}e^{-3t} \\ 0 & 0 & \frac{2}{3} - \frac{2}{3}e^{-3t} & \frac{1}{3} + \frac{2}{3}e^{-3t} \end{bmatrix} \end{aligned}$$