Chapter 4

Linear first-order systems

H:LINEAR-SYSTEMS

In this chapter we study ***linear systems*** of differential equations of the form

$$\begin{aligned} \frac{dy_1}{dt} &= ay_1 + by_2 \\ \frac{dy_2}{dt} &= cy_1 + dy_2, \end{aligned} \tag{4.0.1}$$

linear:generic-details

where a, b, c, d are constants. We can also write this system in column vector notation:

$$\frac{d}{dt}\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} ay_1 + by_2 \\ cy_1 + dy_2 \end{pmatrix} \tag{4.0.2}$$

linear:generic-vector-equa

Throughout this chapter we focus only on systems of two equations having two unknowns. However, the theory that we develop can be easily extended to systems having more unknowns.

4.1 Vectors

section:vectors

In preparation for studying linear systems of the form (4.0.1), we first develop some theory of column vectors.

Suppose we have two column vectors

$$X = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \qquad \text{and} \qquad Y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}.$$

We define the ***addition of vectors*** by

$$X + Y = X = \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \end{pmatrix}.$$

Furthermore, if a is some constant, then we define the ***scaling*** of X by a by

$$aX = \begin{pmatrix} ax_1 \\ ax_2 \end{pmatrix}.$$

Using the operations of adding and scaling, we can write a given vector in a number of ways. For example,

$$\begin{pmatrix} 5 \\ -1 \end{pmatrix} = 2\begin{pmatrix} 1 \\ 1 \end{pmatrix} + 3\begin{pmatrix} 1 \\ -1 \end{pmatrix} = 5\begin{pmatrix} 1 \\ 0 \end{pmatrix} - \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Activity 4.1.1. *I claim that any vector Y can be expressed in the form*

$$Y = \alpha_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \alpha_2 \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

for constants α_1, α_2. Verify this claim by finding formulas for α_1, α_2 such that

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \alpha_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \alpha_2 \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

We now make several useful definitions:

- Suppose we have two of vectors Y_1, Y_2 and two constants α_1, α_2. Vectors of the form

  $$Y = \alpha_1 Y_1 + \alpha_2 Y_2 \tag{4.1.1}$$

  linear:generic-c

  are called ***linear combinations*** of Y_1, Y_2.

- Two vectors Y_1, Y_2 is called ***independent*** if linear combinations of the vectors are unique in the following sense: Suppose Y is given by both (4.1.1) and also by

  $$Y = \beta_1 Y_1 + \beta_2 Y_2;$$

  then $\alpha_1 = \beta_1$ and $\alpha_2 = \beta_2$. This condition is equivalent to the statement that if

  $$\alpha_1 Y_1 + \alpha_2 Y_2 = 0 \qquad \longrightarrow \qquad \alpha_1 = 0, \alpha_2 = 0.$$

  There is a nice way to geometrically interpret what it means for two vectors to be independent: Two vectors are independent if they do not point in either the same direction or in the opposite direction.

It is a convenient fact that if we have two vectors X and Y that are independent and not zero, then any other vector can be constructed as a linear combination of X and Y. In other words, if Z is any vector, then there exists constants α and β such that

$$Z = \alpha X + \beta Y.$$

Example 4.1.1. *The vectors*

$$X = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad \text{and} \qquad Y = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

are independent. We can easily see this geometrically; it is also easy to verify this algebraically.

Suppose now that

$$Z = \begin{pmatrix} a \\ b \end{pmatrix}.$$

Then it is easy to see that

$$Z = (a - b)X + bY$$

and thus Z is a linear combination of X and Y.

In order to discuss differential equations, we need to discuss vectors where the entries are functions (rather than just constants). The definitions regarding adding and scaling vectors work just fine if the entries in the vectors are functions. For example,

$$e^t \begin{pmatrix} 1 \\ -3 \end{pmatrix} + \begin{pmatrix} \cos(t) \\ 5\sin(t) \end{pmatrix} = \begin{pmatrix} e^t + \cos(t) \\ -3e^t + 5\sin(t) \end{pmatrix}.$$

The notions of linear combinations and independence is a little bit trickier. In this course we say that a linear combination of time-dependent vectors $X(t)$ and $Y(t)$ is a vector of the form

$$\alpha X(t) + \beta Y(t),$$

where α and β are *constants* (and not functions of t). We say that two time-dependent vectors $X(t)$ and $Y(t)$ are ***independent*** if they are independent for each time t.

Activity 4.1.2.

1. *Show that the vectors*

$$\begin{pmatrix} \cos(t) \\ -\sin(t) \end{pmatrix} \qquad \textit{and} \qquad \begin{pmatrix} \sin(t) \\ \cos(t) \end{pmatrix}$$

are independent by showing that they are independent at all times t.

2. *Show that the vectors*

$$\begin{pmatrix} \cos(t) \\ \sin(t) \end{pmatrix} \qquad \textit{and} \qquad \begin{pmatrix} \sin(t) \\ \cos(t) \end{pmatrix}$$

are not independent by finding a time at which they are not independent.

Exercise 4.1.1. Are the vectors

$$\begin{pmatrix} 1 \\ 4 \end{pmatrix} \qquad \text{and} \qquad \begin{pmatrix} 2 \\ -2 \end{pmatrix}$$

independent? Explain your reasoning.

Exercise 4.1.2. Are the vectors

$$\begin{pmatrix} 1 \\ -2 \end{pmatrix} \qquad \text{and} \qquad \begin{pmatrix} -2 \\ 4 \end{pmatrix}$$

independent? Explain your reasoning.

Exercise 4.1.3. Explain why the vectors

$$X = \begin{pmatrix} 1 \\ 2 \end{pmatrix} \qquad \text{and} \qquad Y = \begin{pmatrix} 2 \\ -1 \end{pmatrix}$$

are independent. Then express the vectors

$$A = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \qquad \text{and} \qquad B = \begin{pmatrix} 0 \\ 5 \end{pmatrix}$$

as linear combinations of X and Y.

Exercise 4.1.4. Suppose X_* and Y_* are independent vectors. Suppose also that $X(t) = a(t)X_*$ and $Y(t) = b(t)Y_*$ for some functions $a(t)$ and $b(t)$ that are never zero at the same time. Show that $X(t)$ and $Y(t)$ are independent for all times t.

4.2 Linear functions and matrices

section:matrices

We now begin the process of applying the properties of vectors to differential equations of the form (4.0.2), which we write as

$$\frac{d}{dt}\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = F\begin{pmatrix} y_1 \\ y_2 \end{pmatrix},$$

where

$$F\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} ay_1 + by_2 \\ cy_1 + dy_2 \end{pmatrix}.$$

Our goal for this section is to understand the fundamental properties of the functions F that take this form.

We motivate our understanding of the properties of F by noting that the derivative operation respects addition and scaling by constants in the sense that if $X(t)$ and $Y(t)$ are two vectors, and α, β are constants, then

$$\frac{d}{dt}\left[\alpha X(t) + \beta Y(t)\right] = \alpha\frac{d}{dt}\left[X(t)\right] + \beta\frac{d}{dt}\left[Y(t)\right].$$

We claim that the function F respects addition and scaling in the same sense, namely that

- F respects addition, in the sense that for any two vectors $X(t)$ and $Y(t)$ we have

  $$F\Big(X(t) + Y(t)\Big) = F(X(t)) + F(Y(t)), \qquad (4.2.1)$$

  define-linear-1

  and

- F respects scaling, in the sense that for any vector $Y(t)$ and constant α

  $$F(\alpha Y(t)) = \alpha F(Y(t)). \qquad (4.2.2)$$

  define-linear-2

We verify the first property and leave the verification of the second property as a homework problem.

Activity 4.2.1. *Verify that property* (4.2.1) *holds.*

A function that satisfies properties (4.2.1) and (4.2.2) is called a ***linear function***. A differential equation of the form $\frac{d}{dt}Y = F(Y)$ with F linear is called a ***linear differential equation***. Thus (4.0.2) is a linear equation.

It is important to notice that a function being a linear function is a stronger condition than the graph of that function being a line!

example:basic-linear

Example 4.2.1. *The function $f(y) = 5y$ is linear. (Activity: verify this!) However, the function $f(y) = 5y + 7$ is not linear. (Activity: verify this!)*

In Example 4.2.1, we saw that $f(y) = 5y$ is a linear function. This function takes the form

$$f(\text{unknown}) = (\text{stuff}) \text{ multiplied by } (\text{unknown}).$$

Our goal is to realize the function

$$F\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} ay_1 + by_2 \\ cy_1 + dy_2 \end{pmatrix}$$

as taking such a form. To do this we need to introduce a version of

$$(\text{stuff}) \text{ multiplied by } (\text{unknown})$$

that makes sense when the "unknown" is a vector. The mathematical gadget that does this is called ***matrix multiplication***.

Roughly speaking, we can think of a ***matrix*** is a grid containing the numbers appearing in a linear function. For the function

$$F\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} ay_1 + by_2 \\ cy_1 + dy_2 \end{pmatrix}$$

the associated matrix is

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

We define ***a matrix multiplying by a vector*** by

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} ay_1 + by_2 \\ cy_1 + dy_2 \end{pmatrix}.$$

There is a fun way to remember this that involves waving your arms in the air. . . ask Paul for details!

Example 4.2.2.

$$\begin{pmatrix} 5 & -2 \\ 7 & 0 \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 5y_1 - 2y_2 \\ 7y_1 \end{pmatrix}.$$

Activity 4.2.2. *Multiply the following matrices and vectors*

1.

$$\begin{pmatrix} 8 & 4 \\ 0 & 3 \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$$

2.

$$\begin{pmatrix} 5 & 1 \\ 1 & 5 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$$

3.

$$\begin{pmatrix} 5 & 1 \\ 1 & 5 \end{pmatrix} \begin{pmatrix} \cos(t) \\ \sin(t) \end{pmatrix}$$

It is important to be able to write linear functions in a variety of equivalent forms.

linear-functions

Example 4.2.3. *Consider the function*

$$F \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} u + 3v \\ 5u + 2v \end{pmatrix}.$$

We can write this in matrix form as

$$F \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} 1 & 3 \\ 5 & 2 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix}.$$

Activity 4.2.3. *Write the following function*

$$F \begin{pmatrix} p \\ q \end{pmatrix} = \begin{pmatrix} p + 7q \\ 2p - q \end{pmatrix}.$$

in matrix form.

Note that we can translate the linearity properties of F in to properties of matrix multiplication:

- Matrix multiplication respects addition in the sense that for any two vectors $X(t)$ and $Y(t)$ we have

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \Big(X(t) + Y(t) \Big) = \begin{pmatrix} a & b \\ c & d \end{pmatrix} X(t) + \begin{pmatrix} a & b \\ c & d \end{pmatrix} Y(t). \qquad (4.2.3)$$

linear-matrix-property-1

- Matrix multiplication respects scaling in the sense that for any vector $Y(t)$ and any function $\alpha(t)$ we have

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \Big(a(t) Y(t) \Big) = a(t) \begin{pmatrix} a & b \\ c & d \end{pmatrix} Y(t). \qquad (4.2.4)$$

linear-matrix-property-2

Exercise 4.2.1. Verify that the property (4.2.2) holds.

Exercise 4.2.2. Writing

$$X(t) = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \quad \text{and} \quad Y(t) = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix},$$

show that properties (4.2.3) and (4.2.4) hold.

Exercise 4.2.3. Re-write the following function

$$F\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x + y \\ x \end{pmatrix}$$

in a matrix form.

Exercise 4.2.4. Re-write the following function

$$F\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} y \\ -x \end{pmatrix}$$

in a matrix form.