Systems of First Order Linear Equations

1 Introduction to First Order Equations

1.1 Relevance

This chapter is about first order linear equations. But why would we devote an entire chapter to it? Well, they can be very useful, as you can transform any differential equation to a set of first order equations. For example, the n^{th} order equation

$$y^{(n)} = F(t, y, y', \ldots, y^{(n-1)}) \tag{1.1}$$

can be transformed into a system of linear equations by setting

$$x_1 = y, \qquad x_2 = y', \qquad x_3 = y'', \qquad \ldots, \qquad x_n = y^{(n-1)}. \tag{1.2}$$

This would give a system of first order differential equations, consisting of the equations

$$\begin{aligned} x_1' &= x_2 \\ x_2' &= x_3 \\ &\vdots \\ x_{n-1}' &= x_n \\ x_n' &= G(t, x_1, x_2, \ldots, x_n), \end{aligned} \tag{1.3}$$

where G is a function that depends on the original function F.

1.2 Definitions

The general form of a **system of first order differential equations** is

$$\begin{aligned} x_1' &= F_1(t, x_1, x_2, \ldots, x_n) \\ x_2' &= F_1(t, x_1, x_2, \ldots, x_n) \\ &\vdots \\ x_n' &= F_1(t, x_1, x_2, \ldots, x_n). \end{aligned} \tag{1.4}$$

The system is said to have a **solution** on the interval $I : \alpha < t < \beta$ if there exists a set of n functions

$$x_1 = \phi_1(t), \quad x_2 = \phi_2(t), \quad \ldots, \quad x_n = \phi_n(t), \tag{1.5}$$

that satisfies the corresponding system of equations on that interval.

If each of the functions $F_1, F_2, \ldots, F_n$ is a linear function of $x_1, x_2, \ldots, x_n$, then the system of equations is said to be **linear**. Otherwise it is **nonlinear**. The most general form of a system of n first order linear differential equations is therefore

$$\begin{aligned} x_1' &= p_{11}(t)x_1 + \ldots + p_{1n}(t)x_n + g_1(t) \\ x_1' &= p_{21}(t)x_1 + \ldots + p_{2n}(t)x_n + g_2(t) \\ &\vdots \\ x_1' &= p_{n1}(t)x_1 + \ldots + p_{nn}(t)x_n + g_n(t). \end{aligned} \tag{1.6}$$

If also each of the functions $g_1(t), \ldots, g_n(t)$ is zero for all t in the interval I, then the system is said to be **homogeneous**. Otherwise it is **nonhomogeneous**.

We can rewrite the general form to a much simpler form, involving matrices. This would give

$$\mathbf{x}' = P(t)\mathbf{x} + \mathbf{g}(t), \tag{1.7}$$

where $P(t)$ is the matrix formed by all the functions $p_{ij}(t)$.

1.3 Homogeneous systems

In homogeneous systems $\mathbf{g}(t) = 0$. Such systems can therefore be written like

$$\mathbf{x}' = P(t)\mathbf{x}. \tag{1.8}$$

Let's suppose we're dealing with a system of the n^{th} order. Also suppose we have n solutions $\mathbf{x}_1, \ldots, \mathbf{x}_n$ to this system. Now any linear combination $c_1\mathbf{x}_1 + \ldots + c_n\mathbf{x}_n$ of these vectors is also a solution. In fact, we can put the n solutions we had in a matrix $X(t)$, being

$$X(t) = \Big[\mathbf{x}_1(t) \ \ldots \ \mathbf{x}_n(t)\Big]. \tag{1.9}$$

Now every vector $\phi(t)$ satisfying

$$\phi(t) = c_1\mathbf{x}_1 + \ldots + c_n\mathbf{x}_n = X(t)\mathbf{c} \tag{1.10}$$

is a solution to our system of first order differential equations. If the linear combinations of the set $\mathbf{x}_1, \ldots, \mathbf{x}_n$ contain all solutions to the system, then this set is called a **general solution set**.

Any general solution set that is linearly independent at the interval I is said to be a **fundamental set of solutions** for this interval. For such a set, every solution ϕ can be expressed as $\phi = X\mathbf{c}$ in exactly one way.

Let's define the **Wronskian** of the n solutions (denoted by $W[\mathbf{x}_1, \ldots, \mathbf{x}_n]$) as

$$W[\mathbf{x}_1, \ldots, \mathbf{x}_n](t) = \det X(t). \tag{1.11}$$

If $W[\mathbf{x}_1, \ldots, \mathbf{x}_n] \neq 0$ on a certain interval $I : \alpha < t < \beta$, then the vectors $\mathbf{x}_1, \ldots, \mathbf{x}_n$ are linearly independent on I and thus form a fundamental set of solutions.

1.4 Fundamental matrices

If the set $\mathbf{x}_1, \ldots, \mathbf{x}_n$ is a fundamental set of solutions, then the matrix $X(t)$ with columns $\mathbf{x}_1, \ldots, \mathbf{x}_n$ is called the **fundamental matrix**. Since this is an important matrix, it is written with a different sign, being $\Psi(t)$. Any solution $\mathbf{x}$ to the system of differential equations can now be written as

$$\mathbf{x}(t) = \Psi(t)\mathbf{c}, \tag{1.12}$$

for some constant vector $\mathbf{c}$. Now let's suppose we need to solve an initial value problem. An **initial vector** $\mathbf{x}_0 = \mathbf{x}(t_0)$ is given. We have already found $\Psi(t)$ with the methods described above. We just have to find $\mathbf{c}$ such that

$$\mathbf{x}_0 = \Psi(t_0)\mathbf{c} \qquad \Rightarrow \qquad \mathbf{c} = \Psi^{-1}(t_0)\mathbf{x}_0. \tag{1.13}$$

The solution can now be found using

$$\mathbf{x} = \Psi(t)\mathbf{c} = \Psi(t)\Psi^{-1}(t_0)\mathbf{x}_0. \tag{1.14}$$

If the matrix $\Psi(t)$ satisfies the condition $\Psi(t_0) = I$, where I is the identity matrix, then it is a **special fundamental matrix**. Such a matrix is denoted by $\Phi(t)$. Using this fact, the solution reduces to

$$\mathbf{x} = \Phi(t)\mathbf{x}_0. \tag{1.15}$$

2 Homogeneous Systems with Constant Coefficients

2.1 Relevance of eigenvalues

Let's turn our attention to **homogeneous systems with constant coefficients**, meaning that the matrix P does not depend on t. Since the matrix is now constant, we use a different sign for it, being A. A system of equations can now be expressed as $\mathbf{x}' = A\mathbf{x}$. Let's suppose every solution $\mathbf{x}(t)$ to this system of equations can be written as

$$\mathbf{x}(t) = \xi e^{rt}, \tag{2.1}$$

for some constant vector ξ. Using this, we can rewrite the system of equations to

$$r\xi e^{rt} = A\xi e^{rt} \qquad \Rightarrow \qquad (A - rI)\xi = \mathbf{0}. \tag{2.2}$$

There are nontrivial solutions (meaning $\xi \neq \mathbf{0}$) to this equation if $\det(A - rI) = 0$. This is a familiar equation in Linear Algebra. It is only true if r is an eigenvalue of A. The solution set of $\mathbf{x}' = A\mathbf{x}$ therefore depends on those eigenvalues. In the following paragraphs, a closer look is given to the eigenvalues.

2.2 Real and different eigenvalues

If A has n eigenvalues (with n being the size of the square matrix A) that are all real and different from each other, the solutions are relatively easy to find. Let's call the eigenvalues $r_1, \ldots, r_n$ and the corresponding eigenvectors $\xi_1, \ldots, \xi_n$. The corresponding solutions are

$$\mathbf{x}_1(t) = \xi_1 e^{r_1 t}, \ldots, \mathbf{x}_n(t) = \xi_n e^{r_n t}. \tag{2.3}$$

These solutions also form a fundamental solution set. This can be shown by looking at the Wronskian of the set, being

$$W[\mathbf{x}_1, \ldots, \mathbf{x}_n](t) = \left|\xi_1 e^{r_1 t} \ \ldots \ \xi_n e^{r_n t}\right| = e^{(r_1 + \ldots + r_n)t} \left|\xi_1 \ \ldots \ \xi_n\right|. \tag{2.4}$$

The second part of this equation was derived using determinant calculation rules. Since the eigenvectors are linearly independent, the determinant of the matrix on the right side is nonzero. Also the exponential in this equation is nonzero. Therefore the Wronskian is nonzero, proving that the set of solutions $\xi_1, \ldots, \xi_n$ forms a fundamental set of solutions.

2.3 Complex eigenvalues

The solutions of $\det(A - rI) = 0$ are not always real. Sometimes the eigenvalues are complex. Let's assume A contains no complex numbers. In that case complex eigenvalues always come in pairs. In fact, if $r_1 = \lambda + \mu i$ is an eigenvalue, then its complex conjugate $r_2 = \lambda - \mu i$ is also an eigenvalue. The corresponding eigenvectors $\xi_1 = \mathbf{a} + \mathbf{b}i$ and $\xi_2 = \mathbf{a} - \mathbf{b}i$ are also complex conjugates.

Using these data, we can rewrite $\mathbf{x}_1 = \xi_1 e^{rt}$ to

$$\mathbf{x}_1(t) = e^{\lambda t}(\mathbf{a}\cos\mu t - \mathbf{b}\sin\mu t) + e^{\lambda t}(\mathbf{a}\sin\mu t + \mathbf{b}\cos\mu t)i = \mathbf{u}(t) + \mathbf{v}(t)i. \tag{2.5}$$

Equivalently $\mathbf{x}_2(t) = \mathbf{u}(t) - \mathbf{v}(t)i$. Now we have two solutions to the system of equations. But these solutions are complex, and we are looking for real solutions.

However, it turns out that $\mathbf{u}$ and $\mathbf{v}$ are also linearly independent solutions of the system of equations. Therefore the solutions belonging to the two complex eigenvalues r_1 and r_2 are

$$\begin{aligned} \mathbf{u}(t) &= e^{\lambda t}(\mathbf{a}\cos\mu t - \mathbf{b}\sin\mu t), \\ \mathbf{v}(t) &= e^{\lambda t}(\mathbf{a}\sin\mu t + \mathbf{b}\cos\mu t). \end{aligned} \tag{2.6}$$

2.4 Repeated eigenvalues

It may occur that a matrix A with size $n \times n$ has a repeated eigenvalue r (the multiplicity of r is greater than one). If the amount of eigenvectors corresponding to r is equal to the multiplicity of r, then there is no problem. The system can be solved using the methods from paragraph 2.2. If, however, an eigenvalue r doesn't have enough eigenvectors, we don't get n solutions for the entire system. So more tricks are needed to find more solutions.

Let's first look at the case where 1 eigenvector is missing. Of course the eigenvalue r always has at least one corresponding eigenvector ξ. So $\mathbf{x}_1 = \xi e^{rt}$ is already a solution. A second solution will then be of the form

$$\mathbf{x}_2 = \xi t e^{rt} + \eta e^{rt}, \tag{2.7}$$

where η is a constant vector called the **generalized eigenvector**. It can be found using

$$(A - rI)\eta = \xi. \tag{2.8}$$

This equation is always inconsistent, and therefore has infinitely many solutions. One of the components of η can therefore be taken as a constant. Which component is assumed to be a constant does not effect the final solution. Once η is determined, the missing solution can be found and the system of equations can be solved.

If two eigenvectors corresponding to one eigenvalue r are missing, things get slightly more complicated. The first two solutions $\mathbf{x}_1$ and $\mathbf{x}_2$ can be found using the method described above. The third solution can be found using

$$\mathbf{x}_3 = \xi \frac{1}{2} t^2 e^{rt} + \eta t e^{rt} + \varsigma e^{rt}, \tag{2.9}$$

where ς is a constant vector, which can be found using

$$(A - rI)\varsigma = \eta. \tag{2.10}$$

If more eigenvectors are missing, this method can be expanded, analog to the method shown above.

3 Nonhomogeneous Linear Systems

3.1 Basics of nonhomogeneous systems

Let's now consider the system of differential equations given by

$$\mathbf{x}' = P(t)\mathbf{x} + \mathbf{g}(t). \tag{3.1}$$

The general solution has the form of

$$\mathbf{x} = c_1\mathbf{x}_1(t) + \ldots + c_n\mathbf{x}_n(t) + \mathbf{v}(t). \tag{3.2}$$

The first part of this solution is the general solution to the homogeneous system. It is already known how to find this. The vector $\mathbf{v}(t)$ is any specific solution to the nonhomogeneous system as given by equation 3.1. How to find that is something we'll be looking at now.

3.2 Methods of finding a solution

There are several methods of solving such a system. The first method is called **Diagonalization**, which uses matrix inverses. As finding the inverse of a matrix can be a tedious process, this method is often not preferred.

A second way of finding a solution is using the Laplace transform. This especially comes in handy when $\mathbf{g}(t)$ contains a unit step function or a unit impulse function. The method itself isn't very difficult. Let's consider the system

$$\mathbf{x}' = A\mathbf{x} + \mathbf{g}(t), \tag{3.3}$$

where A is some constant matrix. Taking the Laplace transform gives

$$sX(s) - \mathbf{x}(0) = AX(s) + \mathbf{G}(s), \tag{3.4}$$

where $X(s) = L\{\mathbf{x}\}$ is the Laplace transform of $\mathbf{x}$. This equation should then be solved for $X(s)$, which should then be transformed back to the solution $\mathbf{x}$.

The other two methods we will discuss are the **method of undetermined coefficients** and the **method of variation of parameters**.

3.3 Method of undetermined coefficients

If the components of the function $\mathbf{g}(t)$ are polynomial, exponential or sinusoidal functions (or sums or products of these), then the method of undetermined coefficients can be used.

First assume a general form of the specific solution $\mathbf{v}$, with several undetermined coefficients. Then insert this solution into equation 3.1. After this, you should try to solve for the undetermined coefficients.

If the undetermined coefficients can not be solved, then you might try a different form of a specific solution. If any part of this form is already present in the general solution of the homogeneous system, it is often worth while multiplying this part by a factor t.

3.4 Method of variation of parameters

In homogeneous systems, the general solution can be found using $\mathbf{x} = \Psi(t)\mathbf{c}$, where Ψ is the fundamental matrix of the system and $\mathbf{c}$ is some constant vector. For nonhomogeneous systems this is not possible. However, when using the method of variation of parameters, we assume that the general solution can be written like

$$\mathbf{x} = \Psi(t)\mathbf{u}(t), \tag{3.5}$$

where $\mathbf{u}(t)$ is some vector function of t. It can then be shown that

$$\Psi(t)\mathbf{u}'(t) = \mathbf{g}(t), \tag{3.6}$$

or equivalently,

$$\mathbf{u}(t) = \int \Psi^{-1}(t)\mathbf{g}(t)dt + \mathbf{c}, \tag{3.7}$$

where $\mathbf{c}$ is an arbitrary constant vector. The general solution can now be written as

$$\mathbf{x} = \Psi(t)\mathbf{u} = \Psi(t)\mathbf{c} + \Psi(t)\int \Psi^{-1}(t)\mathbf{g}(t)dt. \tag{3.8}$$

Note that the part $\Psi(t)\mathbf{c}$ is the solution to the homogeneous system.

The method of variation hasn't got many constraints and is therefore the most general method of solving systems of nonhomogeneous linear differential equations.