Differential Equations, Part IV: Systems of First Order Linear Equations

Jay Havaldar

We now generalize to first-order systems of linear equations, in other words systems of the form:

$$x_1' = a_{11}x_1 + ... + a_{1n}x_n + b_1$$
$$x_2' = a_{21}x_1 + ... + a_{2n}x_n + b_2$$
$$\vdots$$
$$x_n' = a_{n1}x_1 + ... + a_{nn}x_n + b_n$$

Where $a_{ij}(t)$ and $b_i(t)$ are functions. Of course, we can write this in matrix form as:

$$\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}' = \begin{bmatrix} a_{11} & ... & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & ... & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}$$

Or more succinctly, writing x as the vector with x_i as its entries, we can write:

$$\mathbf{x'} = A\mathbf{x} + \mathbf{b}$$

By analogy, we say that the equation is **homogeneous** if $\mathbf{b} = 0$. A solution is written as a vector with the ith entry representing a solution for the ith equation.

As an analogy with the dimension 1 case, we have the **superposition principle**, which states that for solutions x_1, x_2 to a homogeneous linear system, $c_1x_1 + c_2x_2$ yields a solution for any constants c_1, c_2.

From here on out, we will focus on homogeneous systems, and later on discuss non-homogeneous systems using the methods outlined earlier.

As before, we start with the guarantee of existence and uniqueness of solutions.

0.1 Existence and Uniqueness

Suppose we have a linear system defined on an interval:

$$\mathbf{x'} = A\mathbf{x} + \mathbf{b}$$

With initial conditions given by the vector equation $\mathbf{x}(t_0) = x_0$. Then, if the entries of A and the entries of $\mathbf{b}$ all continuous functions in the interval, there is a unique solution vector $\mathbf{x}(t)$ which exists throughout the interval.

0.2 Linear Independence

Perhaps more concretely, we can see that a homogeneous system of equations should have n linearly independent solutions (and in fact, this implies our earlier result that an order n linear equation ought to have n linearly independent solutions).

This makes sense because, if we evaluate both sides of the matrix equation at a particular point, we have nothing more than an $n \times n$ matrix system of equations.

Definition: Let $x_1, ..., x_n$ be a set of solutions to a homogeneous linear system. Then the matrix $X(t)$ with these vectors as its columns is called the **fundamental matrix**:

$$X = \begin{bmatrix} x_1 & x_2 & \ldots & x_n \end{bmatrix}$$

The determinant of X is called the **Wronskian** of $x_1, \ldots, x_n$.

We now have essentially the same criteria for independence: **a set of solutions is linearly independent at a point if and only if their Wronskian is nonzero there.**

If the Wronskian is nonzero, we call $x_1, \ldots, x_n$ a set of **fundamental solutions**, as before.

Furthermore, we want to say something like the earlier theorem which guarantees that every solution can be written as a combination of enough linearly independent solutions.

Theorem

> If the vector functions $x_1, \ldots, x_n$ are linearly independent solutions of an order n linear system in an interval, then each solution for the system can be written uniquely as a combination (there is only one choice for each c_i):

$$x = c_1 x_1 + \ldots c_n x_n$$

We accordingly will call the form $c_1 x_1 + \ldots c_n x_n$ the **general solution**, and given a vector of initial conditions $x(0) = x_0$, we can indeed set $t = 0$ and solve for the coefficients in the general linear algebra way.

Theorem

> Given a set of solutions to an order n linear system $x_1, \ldots, x_n$, then in an interval $\alpha < t < \beta$, the Wronskian is either zero everywhere or never vanishes.

This theorem makes our lives a lot easier. We know how to check linear independence at a point by taking the Wronskian at that point. Now we can further say that all we have to do is check a single point, and we have linear independence in an entire interval.

The proof, while interesting, will be ommitted since it doesn't add much to our dsicussion of differential equations.

The last thing we should do is, as promised, guarantee that a set of n fundamental solutions exists, i.e. that there are n linearly independent solutions to our equation.

Theorem

> In particular, suppose we have a linear system $x' = Ax$ and solve this system with the initial condition that $X(t_0) = I$ (i.e, $x_i(t_0)$ is the ith column of the identity matrix). Then, $x_1, \ldots, x_n$ form a fundamental set of solutions.

To prove this theorem, first we know that for each row I_j of the identity matrix, we have the existence and uniqueness of a solution to the equation $x_j' = Ax_j$ with $x(t_0) = I_j$. And it isn't hard to see that if we pick $t = t_0$, our fundamental matrix is defined to be the identity matrix, which has nonzero determinant. From our earlier theorem, this tells us that $x_1, \ldots, x_n$ are linearly independent throughout the interval, and thus form a fundamental set of solutions.

Our last theorem guarantees a nice result for complex solutions.

Theorem

> Consider a linear system $x' = Ax$, where each entry of A is continuous and real-valued in an interval. Then, if we have a complex solution $x = u(t) + iv(t)$, both its real parts and its complex parts are solutions to the original equation.

This theorem becomes clear by taking the derivative of x, and then noticing both sides of the equation have to agree in both the real and imaginary parts.

Summary The general theory for linear systems tells us:

- If A, b are continuous and initial conditions are given, then a solution $x' = Ax + b$ exists and is unique in an interval.

- For a linear system, there exists at least one set of n fundamental solutions $x_1, \ldots, x_n$, for which the Wronskian is nonzero at a point. If the Wronskian is nonzero at a point, it is nonzero throughout an interval.

- Every solution can be written as a linear combination of fundamental solutions, with the coefficients determined by the initial conditions.

0.3 Connection to Higher Order Linear Equations

We can transform a higher order linear equation into a matrix system as above, and thus everything we proved in the last system about the theory of higher order linear equations is just a special case. As an example, look at the second order homogeneous equation:

$$y'' + a_1 y' + a_0 y = 0$$

We can pick a vector $y = \begin{bmatrix} y \\ y' \end{bmatrix}$. Solving for y'', we obtain $y'' = -(by + ay')$, so we get a system of equations:

$$\begin{bmatrix} y \\ y' \end{bmatrix}' = \begin{bmatrix} 0 & 1 \\ -a_0 & -a_1 \end{bmatrix} \begin{bmatrix} y \\ y' \end{bmatrix}$$

Note that the determinant of the coefficient matrix is simply a_0. We could do the same thing for a third order equation to obtain:

$$\begin{bmatrix} y \\ y' \\ y'' \end{bmatrix}' = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -a_0 & -a_1 & -a_2 \end{bmatrix} \begin{bmatrix} y \\ y' \\ y'' \end{bmatrix}$$

And here the determinant is $-a_0$.

It is not hard to see that in general, we can turn an order n linear differential equation into an order n linear system. The determinant of the coefficient matrix will always be $\pm a_0$, In particular, if $a_0 = 0$ at a point, we are indeed left with an order $n - 1$ equation, since y does not appear in the equation at all.

We now move on to solving systems of equations with constant coefficients.

0.4 Homogeneous Linear Systems With Constant Coefficients

As before, we start with the simplest case, in which all the coefficients are constant and the right hand side is zero. We come up with a trial solution.

Suppose that v is an eigenvector of the matrix A with eigenvalue λ. Then any multiple of v is also an eigenvector of x with eigenvalue λ.

In particular, the equation $v' = Av$ becomes $v' = \lambda v$ for any eigenvector. So we pick $x = e^{\lambda t}v$, thus solving our equation since $x' = \lambda e^{\lambda t}v = \lambda x$.

This means that if A has n real, distinct eigenvalues, then we're done, since eigenvectors for distinct eigenvalues are linearly independent -- so we have found a fundamental set of solutions. But as you might expect, there are other cases than that.

0.4.1 Complex Eigenvalues

First, we know that complex eigenvalues occur in conjugate pairs. We can tell this because when finding eigenvalues, we solve a polynomial with real coefficients, and such polynomials have conjugate pair roots. It is not hard to see that eigenvectors, as well, occur in conjugate pairs.

Suppose we have such a complex pair of eigenvalues, $\lambda \pm i\mu$. Then using the same methods outlined above, we have a solution:

$$x = (a + bi)e^{(\lambda + i\mu)t}$$

Where a, b are vectors which are the real and imaginary parts of an eigenvector. But we can use Euler's identity $e^{it} = \cos t + i \sin t$, and rearrange to get:

$$x = e^{\lambda t}(a \cos \mu t - b \sin \mu t) + ie^{\lambda t}(a \sin \mu t + b \cos \mu t)$$

We know from an earlier theorem that both the real and the imaginary parts of this expression are solutions to the differential equation. So instead of writing two complex solutions, we can write two real solutions instead:

$$x_1 = e^{\lambda t}(a \cos \mu t - b \sin \mu t)$$
$$x_2 = e^{\lambda t}(a \sin \mu t + b \cos \mu t)$$

0.4.2 Repeated Eigenvalues

Say we have an eigenvalue λ which is repeated in the characteristic equation. If there are n linearly independent eigenvectors associated with λ, we are done; the problem is when there are not. Earlier in the 1 dimensional case, when we had an issue of repeated roots in a characteristic equation, we picked $te^{\lambda t}$ as a trial solution. We did this by performing reduction of order to get a linearly independent solution of the form $c_1 e^{rt} + c_2 t e^{rt}$, and setting $c_1 = 0$. We will apply a similar trick here.

Suppose $x_0 = e^{\lambda t} v$ is a solution to the equation $x' = Ax$, where v is an eigenvector for eigenvalue λ. Then we try to construct another solution of the form:

$$x_1 = te^{\lambda t}v + e^{\lambda t}w$$

Setting $x_1' = Ax$, and noting $A\left(e^{\lambda t}v\right) = \lambda \left(e^{\lambda t}v\right)$ we have:

$$x_1' = e^{\lambda t}v + \lambda te^{\lambda t}v + \lambda e^{\lambda t}w$$
$$Ax_1 = \lambda te^{\lambda t}v + Ae^{\lambda t}w$$

Cancelling the equal terms and setting the two equal, we get:

$$v = (A - \lambda I)\,w$$

Because we know that $(A - \lambda I)v = 0$, we can equivalently write:

$$(A - \lambda I)^2\,w = 0$$

This is the essential idea behind constructing solutions to matrix systems of differential equations with repeated eigenvalues. Our solution looks like:

$$x_1 = te^{\lambda t}v + e^{\lambda t}w$$

Where $(A - \lambda I)w = v$.

Definition: A vector w is called a **generalized eigenvector** of rank m of A if we have:

$$(A - \lambda I)^m\,w = 0$$
$$(A - \lambda I)^{m-1}\,w \neq 0$$

Note that if w is a generalized eigenvector of rank m, then $(A-\lambda I)w$ is a generalized eigenvector of rank $m - 1$, and so on. By repeating this process, we can get a chain of $m - 1$ generalized eigenvectors called a **Jordan chain**. Jordan chains are linearly independent. Note that the last vector in this chain is a generalized eigenvector of rank 1, which is just an eigenvector.

Theorem

> Every $n \times n$ matrix has a set of linearly independent generalized eigenvectors.

This theorem allows us to construct n linearly independent solutions. In fact, if you're familiar with linear algebra, this is equivalent to stating the existence of the Jordan canonical form.

Method For Solving   We aim to solve $x' = Ax$, in general. For each eigenvalue λ, we look at the equation $(A - \lambda I)v = 0$.

- Compute the solutions to $(A - \lambda I)^m v = 0$ for each m until $(A - \lambda I)^m = 0$.

- Pick a vector v_m such that $(A - \lambda I)^m v = 0$ but $(A - \lambda I)^{m-1} v \neq 0$. This is called a generalized eigenvector of rank m.

- Compute the associated Jordan chain, with v_m as chosen above, $v_{m-1} = (A - \lambda I)v_m$, $v_{m-2} = (A - \lambda I)^2 v_m$ etc.

- Use the Jordan chain above to construct the following linearly independent solutions to the differential equation:

$$x_1 = e^{\lambda t} v_1$$
$$x_2 = t e^{\lambda t} v_1 + e^{\lambda t} v_2$$
$$x_3 = \frac{t^2}{2!} e^{\lambda t} v_1 + t e^{\lambda t} v_2 + e^{\lambda t} v_3$$
$$\vdots$$
$$x_m = e^{\lambda t} \left(\sum_{i=1}^{m} \frac{t^{k-1}}{(k-1)!} v_i \right)$$

- If we need more solutions, pick a linearly independent generalized eigenvector of rank k and repeat this process.

With some knowledge of linear algebra, you can say that in a sense, the Jordan chains are uniquely determined. That is to say, we can say the λ has a certain number of Jordan chains with certain sizes, and these sizes are fixed. However, that's outside of the scope of these notes for now.

0.5 Non-Homogeneous Linear Systems

We now move to a discussion of linear systems of the form:

$$x' = Ax + b$$

Where once again, the entries of A, b are continous to guarantee a solution. By the same exact argument we made before, we can write the solution to any such equation in the form:

$$x = c_1 x_1 + \cdots + c_n x_n + v$$

Where x_i form a set of fundamental solutions to the associated homogeneous equation (something like a complementary solution), and v is a solution to the non-homogeneous equation (a particular solution). Furthermore, we can use matrix multiplication to rewrite:

$$x = Xc + v$$

Where $c = \begin{bmatrix} c_1 \\ \dots \\ c_n \end{bmatrix}^T$ is a vector of coefficients. We start with a simple case

0.5.1 Diagonalization

Let's say A is a diagonalizable matrix. From linear algebra, this means A can be written as $A = SDS^{-1}$, where D is a diagonal matrix. If this is the case, then we can perform a change of variables by picking $x = Sy$ or equivalently, $y = S^{-1}x$. Then we have:

$$x' = Ax + b$$
$$Sy' = SDS^{-1}Sy + b$$
$$Sy' = SDy + b$$
$$y' = Dy + b$$

Since D is a diagonal matrix, this reduces to a set of first order equations:

$$y_1' = \lambda_1 y_1 + b_1$$
$$y_2' = \lambda_1 y_2 + b_2$$
$$\vdots$$
$$y_n' = \lambda_n y_n + b_n$$

We have effectively uncoupled the equations; instead of solving them simultaneously, we can now solve them separately, and to get x we apply the transformation $x = Sy$. If $b = 0$, we have the manifest solution $y_i = c_i e^{\lambda t}$. If $b \neq 0$, we still have a set of first order linear equations, which we know we can solve from the first chapter. In the next chapter we'll see another way of solving this type of differential equation.

0.5.2 Variation of Parameters

Recall that given a set of solutions to the associated homogeneous equation, we used the variation of parameters method to find a solution for the non-homogeneous equation by writing $y = u_1 y_1 + \cdots + u_n y_n$ for some functions u_i. We can do the same thing for matrix systems.

Assume we have already found a fundamental matrix X for the associated equation $x' = Ax$. Then we search for a solution for the non-homogeneous equation of the form $x = Xu$, where u is any vector of functions. Differentiating, we get:

$$x' = Ax + b$$
$$X'u + Xu' = AXu + b$$

You can verify for yourself that the product rule does work for matrices. Since X is a fundamental matrix, we have $X' = AX$, so we can write:

$$Xu' = b$$

Since the fundamental matrix has a nonzero determinant, it has an inverse, and indeed we can solve for u:

$$u = \int X^{-1}b\,dt + c$$

Where c is any constant vector. In the first part of this chapter, we discovered that higher order linear systems are special cases of matrix systems. Using that same transformation, we get exactly the setup for variation of parameters as before. The only difference is that now, the constraints that felt like odd choices before now make more sense. As long as the Wronskian is nonzero, the fundamental matrix is invertible and the method works.