Concourse 18.03 – Lecture #2

Existence and uniqueness of solutions and what this tells us about integral curves

In the previous lecture, we stated a very fundamental fact about 1st order ordinary differential equations:

Existence and Uniqueness Theorem: Suppose $F(x,y)$ and the partial derivative F_y are continuous in some rectangle R containing the point $(x_0, y_0) = (a,b)$. Then for some open interval I contain a, the initial value problem $\frac{dy}{dx} = F(x,y)$, $y(a) = b$ has a unique solution defined on the interval I.

A solution to such an initial value problem is called an **integral curve**. The theorem and some parts of its proof can be interpreted in terms of the geometry of integral curves.

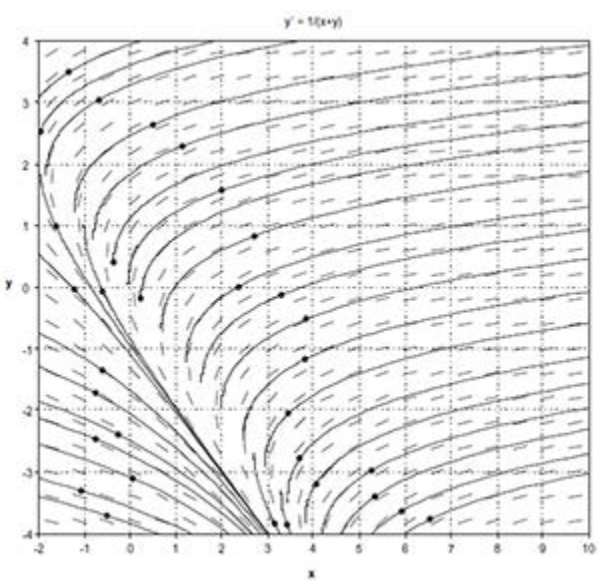
Integral Curve Theorem:

(a) Whenever $F(x,y)$ is defined, integral curves of $\frac{dy}{dx} = F(x,y)$ cannot cross at a positive angle. [This is essentially why the curves have a "parallel" nature.]

(b) If the partial derivative $\frac{\partial F}{\partial y} = F_y$ is continuous in a region, then the integral curves cannot even be tangent to each other at any point in that region.

It sometimes happens when analyzing the slope field of a 1st order ODE that certain integral curves separate other integral curves that are qualitatively fundamentally different. For example, a circular integral curve might separate those curves which spiral inward from those that spiral outward. Such an integral curve is called a **separatrix**. Discovering a separatrix often allows us to separately analyze the behavior of an ODE for initial conditions in different regions.

Example: Analyze the ODE $\frac{dy}{dx} = \frac{1}{x+y}$ using its slope field, and solve it analytically if possible.



Solution: It's easy to produce isoclines for this example. The slope will be constant wherever $x + y = \text{constant}$, and these are just lines with slope $m = -1$. The line where $x + y = 0$, i.e. the line $y = -x$ is somewhat problematic in that the slope becomes vertical along this line. [The Existence and Uniqueness Theorem will therefore break down everywhere on this line – not surprising as the slope field suggests curves "folding over" at all such points. The isocline where $x + y = -1$ is unusual in that $\frac{dy}{dx} = -1$ everywhere along this line which also has slope -1. This line is, in fact, an integral curve which you can verify by differentiating $y = -x - 1$ and substituting it into the ODE. The integral curves above and below this line are fundamentally different, so this line is a separatrix.

This differential equation does not yield simple analytic solutions of the form $y = y(x)$. However, we can turn things sideways and see if it's possible to solve for $x = x(y)$. Basic calculus permits us to rewrite the differential equation as $\frac{dx}{dy} = x + y$. This can then be written in the form $\frac{dx}{dy} - x = y$, a first order inhomogeneous linear differential equation. We will investigate two approaches to solving such a first order linear ODE.

Definition: A differential equation of the form $\frac{d^n x}{dt^n} + p_{n-1}(t)\frac{d^{n-1}x}{dt^{n-1}} + \cdots + p_1(t)\frac{dx}{dt} + p_0(t)x(t) = q(t)$, where $p_{n-1}(t), \ldots, p_1(t), p_0(t), q(t)$ are functions of the independent variable *t*, is called an ***n*th order linear ordinary differential equation**. In the case where $q(t) = 0$ for all *t*, we call the equation **homogeneous**. Otherwise we call it **inhomogeneous**.

We are specifically concerned with 1st order ODEs of the form $\boxed{\frac{dy}{dx} + p(x)y = q(x)}$ (or $\boxed{\frac{dx}{dt} + p(t)x = q(t)}$).

Integrating factors

Definition: An **integrating factor** for a given ODE is a function $v(x)$ such that when both sides of the ODE are multiplied by $v(x)$ the resulting differential equation consists of known derivatives on both sides of the equation. The ODE can then be solved by integrating both sides and then solving for the dependent variable in terms of the independent variable.

It's always possible to formally solve $\frac{dy}{dx} + p(x)y = q(x)$ via an integrating factor. We seek $v(x)$ such that we can integrate both sides of the equation $v(x)\left[\frac{dy}{dx} + p(x)y\right] = v(x)q(x)$. The left-hand-side is $v\frac{dy}{dx} + pvy$, and if we note that $\frac{d}{dx}(vy) = v\frac{dy}{dx} + v'(x)y$, we can then look for $v(x)$ such that $\cancel{v\frac{dy}{dx}} + pvy = \cancel{v\frac{dy}{dx}} + v'(x)y$ or simply $pvy = v'y \quad \Rightarrow \quad pv = v'$. This can then be rewritten as $\frac{v'(x)}{v(x)} = p(x) \quad \Rightarrow \quad \int \frac{v'(x)}{v(x)}dx = \int p(x)dx$. This gives $\ln|v(x)| = \int p(x)dx + C \quad \Rightarrow \quad \boxed{v(x) = e^{\int p(x)dx}}$ as an integrating factor. This approach, of course, works best if you can find an antiderivative of the function $p(x)$.

We then have the new ODE $\frac{d}{dx}\left(v(x)y(x)\right) = v(x)q(x)$, so integration gives $v(x)y(x) = \int v(x)q(x)dx + C$. We can then solve for $y(x) = \frac{1}{v(x)}\left[\int v(x)q(x)dx + C\right]$. If we insert the integrating factor $v(x) = e^{\int pdx}$, we can write this solution as $\boxed{y(x) = e^{-\int pdx}\left[\int q(x)e^{\int pdx}dx + C\right]}$. It may not be pretty, but it works if you can actually do the integrals.

If we switch variables in our example (just to be ever so conventional), our equation becomes $\frac{dy}{dx} = y + x$ or $\frac{dy}{dx} - y = x$. In this case, $p(x) = -1$, $\int p(x)dx = -x$ is the simplest antiderivative, and the integrating factor is then $v(x) = e^{-x}$. We then have $e^{-x}\left(\frac{dy}{dx} - y\right) = \frac{d}{dx}\left(e^{-x}y\right) = xe^{-x}$. We can then integrate using integration by parts to get $e^{-x}y = \int xe^{-x}dx = -xe^{-x} - e^{-x} + C$. If we then multiply both sides by e^x, we get $\boxed{y(x) = -x - 1 + Ce^x}$. If, for example, we had the initial condition that $y(0) = 3$, we would then have $y(0) = -1 + C = 3$, so $C = 4$ and we would get the unique solution $y(x) = -x - 1 + 4e^x$ for this initial value problem.

Considering the relatively simple expression for this solution, you might think that there could be a simpler approach. There is, but it requires us to start our way down an important path that will lead to some of the most important methods and perspectives in this entire course. This is the Linearity path.

Linearity

In the context of functions of one variable, linearity is an often abused word. In fact, a function of the form $f(x) = mx + b$ is NOT a linear function. It is more appropriately called a 1st order *affine* function. Linearity is a property most simply characterized by the fact that <u>linear functions preserve scaling and adding</u>. The linear

functions of one variable consist only of those of the form $f(x) = mx$. Note that $f(ax) = m(ax) = a(mx) = af(x)$, i.e. it preserves scaling, and $f(x+y) = m(x+y) = mx + my = f(x) + f(y)$, i.e. it preserves addition.

Definition: Formally we say that a function is **linear** if for all inputs x_1, x_2 and constants c_1, c_2 we must have $f(c_1x_1 + c_2x_2) = c_1 f(x_1) + c_2 f(x_2)$.

In the case of functions $T : \mathbf{R}^n \to \mathbf{R}^m$, linearity means that the scaling of vectors and the addition of vectors is preserved via a linear transformation. All such transformations are of the form $T(\mathbf{x}) = \mathbf{Ax}$ where $\mathbf{A}$ is an $m \times n$ matrix with constant entries. Linearity then translates into the matrix algebra facts that $\mathbf{A}(k\mathbf{x}) = k(\mathbf{Ax})$ and $\mathbf{A}(\mathbf{x}+\mathbf{y}) = \mathbf{Ax} + \mathbf{Ay}$, or (combined) $\mathbf{A}(\alpha\mathbf{x} + \beta\mathbf{y}) = \alpha\mathbf{Ax} + \beta\mathbf{Ay}$ for all scalars α, β and all vectors $\mathbf{x}, \mathbf{y}$.

Our current situation involves working with functions in the same way that we looked at vectors in $\mathbf{R}^n$. Just as we can scale and add vectors, we can also scale and add functions. A transformation that acts on functions in a manner analogous to the way matrices act on vectors is known as a **linear (differential) operator**. The basic examples are differentiation and multiplication by a fixed function. We can then compose these basic operators and add them to form more complicated operators.

There are many spaces of functions in which we can seek solutions to differential equations. Perhaps the most common such space is the space of functions that are differentiable to all orders.

Multiplication by a fixed function is a linear operator

Suppose we have a fixed function $p(x)$ and we define a transformation of functions by $[T(f)](x) = p(x)f(x)$. We can easily see that for any constant *c*, $[T(cf)](x) = p(x)cf(x) = cp(x)f(x) = c[T(f)](x)$, so $T(cf) = cT(f)$, i.e. *T* preserves scaling. Similarly, if f_1 and f_2 are two functions, then

$[T(f_1 + f_2)](x) = p(x)(f_1 + f_2)(x) = p(x)(f_1(x) + f_2(x)) = p(x)f_1(x) + p(x)f_2(x) = [T(f_1)](x) + [T(f_2)](x)$.

This is really just the distributive law, but the result is that formally $T(f_1 + f_2) = T(f_1) + T(f_2)$, i.e. *T* preserves addition of functions. Together, this shows that *T* is a linear operator.

Differentiation of functions is a linear operator

Let *D* be the transformation defined by $D(f) = f'$. That is, $[D(f)](x) = f'(x)$. The old refrains you learned in first semester calculus are precisely what makes this a linear operator: (a) The derivative of a constant times a function is the constant times the derivation of the function; and (b) The derivative of a sum is the sum of the derivatives. In symbolic terms, $D(cf) = cf'$ and $D(f+g) = f' + g'$. We can put these together as a single linearity rule: $D(c_1f_1 + c_2f_2) = c_1D(f_1) + c_2D(f_2)$.

The composition of linear operators (or any linear function), where defined, is also linear

If *S* and *T* are both linear operators and if the composition $S \circ T$ is defined, then using the linearity properties of both we have that for all scalars c_1, c_2 and functions f_1, f_2,

$$\begin{aligned}(S \circ T)(c_1f_1 + c_2f_2) &= S(T(c_1f_1 + c_2f_2)) = S(c_1T(f_1) + c_2T(f_2)) \\ &= c_1S(T(f_1) + c_2S(T(f_2)) = c_1(S \circ T)(f_1) + c_2(S \circ T)(f_2)\end{aligned}$$

For example, since differentiation acts linearly, we can compose this with itself to get the 2nd derivative and this also acts linearly. The same holds for higher order derivatives.

The sum of two linear operators is also a linear operator

The sum of two operators is defined in the same way we add any functions, i.e. $(S+T)(f) = S(f) + T(f)$.

If *S* and *T* are both linear operators, then we'll have that for all scalars c_1, c_2 and functions f_1, f_2,

$$\begin{aligned}[S+T](c_1f_1 + c_2f_2) &= S(c_1f_1 + c_2f_2) + T(c_1f_1 + c_2f_2) = c_1S(f_1) + c_2S(f_2) + c_1T(f_1) + c_2T(f_2) \\ &= c_1S(f_1) + c_1T(f_1) + c_2S(f_2) + c_2T(f_2) = c_1[S(f_1) + T(f_1)] + c_2[S(f_2) + T(f_2)] = c_1[S+T](f_1) + c_2[S+T](f_2)\end{aligned}$$

If we put the fact that composition of linear operators and the addition of linear operators yields another linear operator, we see that the expression $\frac{d^n x}{dt^n} + p_{n-1}(t)\frac{d^{n-1}x}{dt^{n-1}} + \cdots + p_1(t)\frac{dx}{dt} + p_0(t)x(t)$ for functions $p_{n-1}(t), \ldots, p_1(t), p_0(t)$ represents a linear operator acting on an undetermined function $x(t)$. If we write this operator as $T(x(t)) = \frac{d^n x}{dt^n} + p_{n-1}(t)\frac{d^{n-1}x}{dt^{n-1}} + \cdots + p_1(t)\frac{dx}{dt} + p_0(t)x(t)$, we then know by linearity that $T(x_1(t) + x_2(t)) = T(x_1(t)) + T(x_2(t))$ and $T(c\,x(t)) = c\,T(x(t))$ and, more generally, $T(c_1\,x_1(t) + c_2\,x_2(t)) = c_1\,T(x_1(t)) + c_2\,T(x_2(t))$.

Now that we have paved the road to Linearity, we can apply this idea to solving linear differential equations.

Linearity method using homogeneous solutions and particular solutions

Suppose we have an inhomogeneous linear ODE of the form $T(f) = g$ where T is an *n*th order linear differential operator. We can produce ALL solutions to $T(f) = g$ as follows:

(1) First solve the homogeneous equation $T(f) = 0$ to find a general expression for all such solutions. Call this the homogeneous solution f_h. It will generally involve *n* arbitrary constants.

(2) Find a single particular solution to the inhomogeneous equation $T(f) = g$. Call this particular solution f_p.

(3) The general solution to $T(f) = g$ is then $\boxed{f = f_h + f_p}$.

Proof of the method: We know that $T(f_p) = g$, so suppose f is any other solution to $T(f) = g$. Then $T(f - f_p) = T(f) - T(f_p) = g - g = 0$. So $f - f_p$ solves the homogeneous equation and must be included among all homogeneous solution, i.e. $f - f_p = f_h$. Therefore $f = f_h + f_p$.

This fact is really the same thing that we see when solving a consistent, inhomogeneous system of linear algebraic equations. In matrix form, if the system is represented as $\mathbf{Ax} = \mathbf{b}$ where $\mathbf{A}$ is an $m \times n$ matrix, and if $\mathbf{x}_h$ represents all solutions to the homogeneous equation $\mathbf{Ax} = \mathbf{0}$ and $\mathbf{x}_p$ is a single solution to $\mathbf{Ax} = \mathbf{b}$, then all solutions to $\mathbf{Ax} = \mathbf{b}$ will be of the form $\mathbf{x} = \mathbf{x}_h + \mathbf{x}_p$. Typically, these homogeneous solutions are lines, planes or higher-dimensional analogues (subspaces) passing through the origin. This just says that the inhomogeneous solutions are parallel translates of these subspaces.

So, let's solve the problem already:

The ODE $\frac{dy}{dx} - y = x$ is first order, linear, and inhomogeneous.

(1) The homogeneous equation is just $\frac{dy}{dx} - y = 0$ or $\frac{dy}{dx} = y$. We've already solved this to get all solutions in the form $y_h = Ae^x$.

(2) We can find an inhomogeneous solution by educated guessing (formally called the method of undetermined coefficients). Try a solution of the form $y = ax + b$. Calculate $\frac{dy}{dx} = a$ and substitute into the ODE to get $\frac{dy}{dx} - y = a - (ax + b) = (a - b) - bx = x$. We can solve this by choosing $a - b = 0$ and $-b = 1$. So $\boxed{b = -1}$ and $\boxed{a = -1}$, and a particular solution is therefore $y_p = -x - 1$.

(3) By linearity, all solutions are therefore of the form $y = y_h + y_p = Ae^x - x - 1$. This agrees with our previous result.

And **our original problem** $\frac{dx}{dy} - x = y$ gives $\boxed{x = Ae^y - y - 1}$ where the constant *A* is determined by initial conditions. If you compare this with the slope field picture on the first page, you'll see that this accurately describes all of the integral curves including the separatrix which occurs where $A = 0$.

Notes by Robert Winters