On nice enough ODEs and dependency on initial conditions

Duarte Maia

1 Introduction

This document is a repository for me to write down the sequence of results necessary to prove, from scratch, that the flow of a vector field exists and is smooth, in some sense. The flow is obtained by solving an ODE, which means that I want to show that, under some niceness assumptions, the solutions of an ODE exist locally and vary smoothly when the initial conditions change. In this document, I will only solve the case for $\mathbb{R}^n$, as translating that to the original goal (proving vector field flows exist) is beyond the scope of this.

2 Grunbaum's lemma

Grunbaum's lemma is a tool that allows us to take differential (or, to be more precise, differential) inequalities and turn them into actual inequalities.

The integral statement is as follows: let g and h be two positive-valued continuous functions satisfying the inequality

$$g(t) \leq c + \int_{t_0}^{t} g(s)h(s)\mathrm{d}s$$

for some positive c, for $t \geq t_0$. Then, we conclude $g(t) \leq c\exp(\int_{t_0}^{t} h(s)\mathrm{d}s)$.

The proof goes as follows. Define $G(t) = c + \int_{t_0}^{t} gh$. Then, G is a strictly positive function for $t \geq t_0$, and thus $\log G(t)$ makes sense. Furthermore, it is differentiable by the fundamental theorem of calculus, with $(\log G)'(t) = G'(t)/G(t) = g(t)h(t)/G(t)$. By hypothesis, $g \leq G$ and so this is at most $h(t)$. In conclusion, $\log G(t) - \log G(t_0) \leq \int_{t_0}^{t} h$, and taking the exponential and applying $G \geq g$ again we reach the desired conclusion.

Note that this proof relies on $c > 0$ to make sense of $\log G$ and division by G. However, if $c = 0$ we can take the limit as $c \to 0$ of the strictly positive result, leading to the result: if $g(t) \leq \int_{t_0}^{t} gh$ then $g = 0$.

Another note: our proof used $t \geq t_0$, but it also works for $t < t_0$ as long as the sign of the integral is kept positive.

A lot of the integral inequalities in what follow assume implicitly $t \geq t_0$ (for example, the inequality $|\int_{t_0}^t f| \leq \int_{t_0}^t |f|$ only goes through in this case; otherwise the limits in the integral must be swapped). This is of no consequence. I hope.

3 Picard-Lindelöf

The Picard-Lindelöf theorem states that, fixed initial conditions, the solution of an ODE exists locally and is unique globally.

To be more precise: suppose $f : D \to \mathbb{R}^n$, D open subset of $\mathbb{R} \times \mathbb{R}^n$, is a *continuous* function satisfying the following property: for any compact $K \subseteq D$, $f|_K$ is (uniformly) Lipschitz in x. That is, there exists a constant L such that for all $(t, x) \in K$ and $(t, y) \in K$ we have $|f(x) - f(y)| \leq L|x - y|$. We will later give general enough conditions for this to happen (for example, f C^1 is enough).

Fix $(t_0, x_0) \in D$. We will show that the ODE given by

$$\begin{cases} x'(t) = f(t, x) \\ x(t_0) = x_0 \end{cases} \tag{1}$$

has a unique solution.

We begin by showing existence. Fix $(t_0, x_0) \in D$. Let $I \times R$ be a compact neighborhood of this point. Let M be the maximum of $|f|$ over this set and L the Lipschitz constant of f. Pick $\alpha > 0$ small enough such that

$$B_\alpha(t_0) \times \overline{B}_{M\alpha}(x_0) \subseteq I \times R \text{ and } \alpha L < 1.$$

Now put

$$X = \{\, \varphi \in C(B_\alpha(t_0), \mathbb{R}^n) \mid d(\varphi(t), x_0) \leq Md(t, t_0) \,\}.$$

We may define $T : X \to X$ given by

$$T(\varphi)(t) = x_0 + \int_{t_0}^t f(s, \varphi(s)) \mathrm{d}s.$$

The conditions imposed by $\varphi \in X$ guarantee that this is well-defined for all $t \in B_\alpha(t_0)$ and remains in X. Now, given two functions $\varphi, \psi \in X$ we conclude

$$\begin{aligned} |T\varphi - T\psi| &= |\int_{t_0}^t f(s, \varphi) - f(s, \psi) \mathrm{d}s| \\ &\leq \int_{t_0}^t L|\varphi - \psi| \\ &\leq \alpha L \|\varphi - \psi\|_\infty, \end{aligned}$$

which shows that T is contracting in the sup norm. Standard arguments (Cauchy sequences) show that this implies $T^n x_0$ converges to some function φ, and note that φ must be a fixed point of T, for

$$\|TT^n x_0 - T\varphi\| \leq \alpha L \|T^n x_0 - \varphi\| \to 0,$$

and so $T^{n+1}x_0$ converges to $T\varphi$, but since it is a subsequence of $T^n x_0$ it also converges to φ, which shows equality.

We conclude that φ is a continuous function satisfying $\varphi(t) = x_0 + \int_{t_0}^{t} f(s, \varphi(s))\mathrm{d}s$, and differentiating both sides we obtain it is a solution of the original ODE.

Now for uniqueness. Let x and y be two solutions defined in a common interval $[t_0, t_1]$ without loss of generality. Then,

$$\begin{aligned} |x(t) - y(t)| &\leq |\int_{t_0}^{t} f(s, x) - f(s, y)\mathrm{d}s| \\ &\leq \int_{t_0}^{t} L|x - y|, \end{aligned}$$

where L is the Lipschitz constant that exists on the compact set given by the union of the curves given by x and y. By Gronwall's lemma, this implies $x = y$ on the interval.

4 Maximality of solutions

Consider the ODE (1). We will show that there is a maximal solution, where we take solutions to be defined in an open interval containing t_0.

Define I_M as the union of the intervals of definition of all possible (contiguous) solutions of the ODE. By the uniqueness part of Picard-Lindelöf, all solutions agree where mutually defined, so there is an unambiguous x defined on I_M, which is a solution of the ODE, and clearly the maximal possible.

It is possible to show that if I_M is bounded, say, from above, then in some sense x is 'exploding' or leaving D. Indeed, in this case, call the supremum of I_M by the name t_f. We assert that either $f(t, x)$ is unbounded as $t \to t_f$ or, in the negative case, $(t, x(t))$ converges to a limit (t_f, x_f) as $t \to t_f$, and this limit lies outside D.

Suppose, then, $f(t, x)$ is bounded as $t \to t_f$. Then x' is bounded and so x converges as $t \to t_f$, because it is also bounded and so has a converging subsequence, but boundedness of the derivative implies that any two sublimits coincide. As such, let x_f be the limit.

If $(t_f, x_f) \in D$ then it would be possible to find $\varphi : [t_f, t_f + \varepsilon[\to \mathbb{R}^n$ that is also a solution of the ODE. It is easy to see (calculating the left-derivative) that gluing φ to the end of x gives us another (bigger) solution of the ODE, contradicting x's maximality. Therefore, $(t_f, x_f) \notin D$, concluding the proof.

5 Continuity in initial conditions

Define $x(t_0, x_0)(t)$ to be the solution of the ODE (1) with the given initial conditions. Given t_0, x_0 fixed, suppose $x(t_0, x_0)$ is defined on an interval $[a, b]$. We will show that for t_1, x_1 close enough to t_0, x_0 the solution is also defined in $[a, b]$, and x is continuous as a function from this neighbourhood to $C[a, b]$ with the sup norm.

To this effect, we begin by defining a tubular neighbourhood of x. For each t, define $\delta(t)$ as the distance from $(t, x(t))$ to D^c. Since $[a, b]$ is compact, this is minimized at some value $\delta > 0$, and so we define T as the closed tubular neighbourhood of x with radius, say, $\delta_1 < \delta$. It is easy to check that T is compact, and so f is maximized there with some value M and has a Lipschitz constant L.

Now, let (t_1, x_1) be in this tubular neighbourhood, and let us investigate how the solution starting at this point develops, starting with how long it takes to leave the tubular neighbourhood. Let φ be the original solution and x the (maximal) solution with these initial conditions.

$$\begin{aligned}|x(t) - \varphi(t)| &= |x_1 - \varphi(t_1) + \int_{t_1}^{t} f(s, x) - f(s, \varphi)\mathrm{d}s| \\ &\leq |x_1 - \varphi(t_1)| + \int_{t_1}^{t} L|x(s) - \varphi(s)|\mathrm{d}s.\end{aligned}$$

As such, Gronwall's Lemma guarantees us that

$$|x - \varphi| \leq |x_1 - \varphi(t_1)| \exp(L(b - a)).$$

Therefore, if $|x_1 - \varphi(t_1) < \exp(-L(b-a))\delta_1$ we can be sure that x is defined over the whole interval, because it doesn't leave the tubular neighbourhood and for it to stop being defined at some point eiher f would need to become unbounded (which doesn't happen because M) or x would need to converge to some point outside D (which doesn't happen because it never leaves the tubular neighbourhood, which is closed and contained in D).

As a consequence, $x(t_1, x_1)$ is well-defined in a tubular neighbourhood of (t_0, x_0), so it remains to show that it is continuous in $C[a, b]$. But the above argument serves to show that if x and y are two solutions (with initial conditions t_1, x_1 and s_1, y_1) we have

$$\|x - y\| \leq |x_1 - y(t_1)| \exp(L(b - a)) \leq C_1|x_1 - y_1| + C_2|t_1 - s_1| \tag{2}$$

which is enough to guarantee continuity in the initial conditions.

Note that this also allows us to conclude that $x(t, t_0, x_0)$ is continuous as a three real-variable function.

In conclusion, the set where $x(t, t_0, x_0)$ is well-defined is open (as a subset of $\mathbb{R}^{2+n}$) and x is continuous in this domain.

6 Smoothness (in general)

Before investigating smoothness of the solution as a function of the initial conditions, it is perhaps useful to show that smoothness of f is enough to guarantee f Lipschitz on compacts.

Suppose, then, that f is continuous in t and C^1 in x. Then, $\partial_x f$ is continuous, and so bounded on compacts. In particular, we may consider a compact rectangular neighbourhood of an arbitrary (t_0, x_0), wherein all partial derivatives of f are bounded by, say, M. Then, for any (applicable) t, the difference $f(t,x) - f(t,y)$ can be bounded using the mean value theorem by $n^2 M|x-y|$, which shows that any point has a Lipschitz neighbourhood. It remains to show that f is Lipschitz on any compact.

The following idea is due to Manel. Fix a compact $K \subseteq D$. Let

$$A = \{\, (t,x,y) \mid (t,x) \in K, (t,y) \in K, x \neq y \,\}.$$

Consider the function $g : A \to \mathbb{R}$ defined by

$$g(t,x,y) = \frac{|f(t,x) - f(t,y)|}{|x-y|}.$$

It is clear that this is well-defined and continuous. We wish to show it is bounded from above. To do so, suppose for contradiction that (t_n, x_n, y_n) is a sequence of elements of A such that $g(t_n, x_n, y_n) \to \infty$. Since A is bounded, we may without loss of generality suppose that the sequence converges, and by continuity of g it must converge to a point (t,x,y) not in A.

It is easy to check that $(t,x) \in K$ and $(t,y) \in K$, so the only way for this point not to be in A is for $x = y$. But this contradicts the local Lipschitz condition, which guarantees that, for (t', x', y') close enough to (t,x,x), g is bounded. This concludes the proof that locally Lipschitz implies Lipschitz on compacts, and so C^1 implies Lipschitz on compacts.

7 Derivative in x_0

In what follows, we suppose to simplify that f is C^1 and does not depend on t. In other words, the ODE is reduced to

$$\begin{cases} x' = f(x) \\ x(t_0) = x_0. \end{cases}$$

We will show that the function $x(t, t_0, x_0)$ is C^1. Smoothness in t is obvious because f is continuous and smoothness in t_0 is simply a consequence of $x(t, t_0, x_0) = x(t - t_0, 0, x_0)$ together with the chain rule. Differenciability wrt x_0 is the hardest part.

The first step is to try to guess what the derivative would be. Let us refer to it as $J(t) = \partial_{x_0} x(t, t_0, x_0)$. It is easy to check that $J(t_0) = I$, and (assuming x is C^1) we could conclude that

$$\begin{aligned} \partial_t J(t) = \partial_t \partial_{x_0} x(t, t_0, x_0) = \partial_{x_0} \partial_t x(t, t_0, x_0) = \partial_{x_0} f(x(t, t_0, x_0)) = \\ = f'(t, x(t, t_0, x_0)) \partial_{x_0} x(t, t_0, x_0) = f'(t, x(t, t_0, x_0)) J(t). \end{aligned}$$

As a consequence, we guess that the derivative (wrt x_0) of x is described by the matrix J satisfying the ODE (in $\mathbb{R}^{n^2}$)

$$\begin{cases} J(t_0) = I \\ J'(t) = A(t) J(t) \end{cases}$$

where $A(t) = f'(x(t))$. Note that, by hypothesis on f, A is a continuous function of t and therefore the function $J, t \mapsto A(t)J$ is continuous in t and locally Lipschitz in J, and so J exists, at least locally.

Let us show that J is defined for all t that x is. Consider an interval $[a, b]$ in which x is defined. Then, so is A, and A has all entries bounded in this interval by some M. As a consequence,

$$|J(t)| \leq n + \int_{t_0}^{t} |A(t)J(t)| \leq n + \int_{t_0}^{t} n^3 M |J|,$$

and therefore by Grunbaum's lemma we conclude $|J| \leq n \exp((b-a)n^3 M)$, and so J is bounded. Therefore, for J to stop being extendable after some time t_f, it would need to 'go outside the domain', but for every $t \in [a, b]$ the function $J \mapsto AJ$ is defined for all J and therefore J must be extendable to all $t \in [a, b]$.

Note also that J is continuous as a function of x_0, as the ODE that specifies it can also be rewritten as to make x_0 part of the initial conditions:

$$\begin{cases} J(t_0) = I \\ X(t_0) = x_0 \\ J'(t) = f'(x(t, t_0, X(t)))J(t) \\ X'(t) = 0. \end{cases}$$

We now turn to verifying that J is in fact the derivative of x as a function of x_0. To this effect, let us call $x(t) \equiv x(t, t_0, x_0)$ and $x_h(t) \equiv x(t, t_0, x_0 + h)$, for small enough h that this is well-defined. We wish to show that, for t and t_0 fixed, the following expression converges to zero as $h \to 0$:

$$\frac{1}{|h|} \left| x_h(t) - x(t) - J(t)h \right|.$$

To this effect, let us try to bound $|x_h(t) - x(t) - J(t)h|$ from above. The first step is to expand it as an integral:

$$|x_h(t) - x(t) - J(t)h| \leq \int_{t_0}^{t} |f(x_h(s)) - f(x(s)) - f'(x(s))J(s)h| \mathrm{d}s,$$

which looks tempting because the quantity $f(x_h(s)) - f(x(s))$ should be of the form $f'(x(s))\Delta(s) + o(\Delta(s))$, where $\Delta = x_h - x$, and to a first order approximation $\Delta(s)$ should be more or less $J(s)h$... Of course, this is not a rigorous argument, but it is at least an indication that we might be in the right track.

Let us begin by applying the mean value theorem to conclude $f(x_h(s)) - f(x(s)) = (f'(x(s)) + E(h, s))\Delta(s)$, where E is an error matrix whose entries are of the form $\partial_j f_i(\xi(h, s)) - \partial_j f_i(x(s))$, where $\xi(h, s)$ is a value somewhere within $|\Delta(s)|$ of $x(s)$. Since f is C^1, for all ε there exists a δ such that $|\Delta(s)| < \delta$ implies that all entries in $E(h, s)$ are at most ε. In turn, because of inequality (2) there exists C such that $\|\Delta\| \leq C|h|$ and as a consequence if $|h| < \delta/C$

then $\|\Delta\| < \delta$, and therefore all entries of E are (for all s) at most ε. As a consequence, we have, for small enough h, the inequality

$$\begin{aligned}
|x_h(t) - x(t) - J(t)h| &\leq \int_{t_0}^{t} |f(x_h(s)) - f(x(s)) - f'(x(s))J(s)h| \mathrm{d}s \\
&= \int_{t_0}^{t} |f'(x(s))\Delta(s) + E(h,s)\Delta(s) - f'(x(s))J(s)h| \mathrm{d}s \\
&\leq (b-a)\varepsilon\|\Delta\| + \int_{t_0}^{t} |f'(x(s))\Delta(s) - f'(x(s))J(s)h| \mathrm{d}s \\
&\leq (b-a)\varepsilon C|h| + \int_{t_0}^{t} |f'(x(s))(x_h(s) - x(s) - J(s)h)| \mathrm{d}s \\
&\leq (b-a)\varepsilon C|h| + \int_{t_0}^{t} n^2 M|x_h(s) - x(s) - J(s)h| \mathrm{d}s,
\end{aligned}$$

where M is the maximum value of the entries of $f'(x(t))$ as t ranges from A To B.

As a consequence, we may apply Grunbaum's lemma to get, finally, that for all ε there exists δ such that $|h| < \delta$ implies

$$\frac{1}{|h|}|x_h(t) - x(t) - J(t)h| < (b-a)\varepsilon C \exp(n^2 M(b-a)),$$

which concludes the proof that $J(t)$ is the derivative of $x_h(t)$ at $h = 0$, or, in other words, $\partial_{x_0} x(t, t_0, x_0)$. We have already discussed continuity of J and the other derivatives of x, and so we conclude that if f is C^1 so is the solution as a function of initial data.

8 Smoothness

We will now show that x absorbs all degrees of smoothness from f. In particular, if f is C^∞, so is x.

We will show by induction that if f is C^p then so is x. We have already proven the base case $p = 1$, concluding that the jacobian of x is given by the solution of the ODE

$$\begin{cases}
J(t_0) = I \\
X(t_0) = x_0 \\
J'(t) = f'(x(t - t_0, 0, X))J(t) \\
X' = 0.
\end{cases}$$

Now, suppose we have already shown the case for some p, and let us show it for $p + 1$. If $f \in C^{p+1}$ then in particular $f \in C^p$, and so x is C^p. Therefore, the function

$$(J, X, t) \mapsto f'(x(t - t_0, 0, X))J$$

is C^p, because it is the composition of two C^p functions: f' and x. Therefore, by applying the induction hypothesis to the ODE that yields J, we conclude J is C^p and thus that x is C^{p+1} everywhere it is defined. In particular, if f is C^∞ then so is x. The proof of everything is complete.