Well-posedness for ODEs

Harald Hanche-Olsen
hanche@math.ntnu.no

This note is about well-posedness of the initial-value problem for a system of ordinary differential equations:

$$\begin{aligned} \dot{x}(t) &= f\big(x(t)\big), \\ x(0) &= a. \end{aligned} \tag{1}$$

Here $f\colon \Omega \to \mathbb{R}^n$ is a mapping defined on an open set $\Omega \subseteq \mathbb{R}^n$. The initial value a is supposed to belong to Ω, and the unknown function x is to be defined on an open interval containing $t = 0$.

By *well-posedness* of the problem we mean a positive answer to three questions: (1) Does a solution exist? (2) Is the solution unique? (3) Does the solution depend continuously on the data (a and the function f)?

Some preliminary definitions and results

Lipschitz continuity. The answer to the question of well-posedness is in general negative. It turns out that the natural requirement to obtain a well-posed problem is Lipschitz continuity of the righthand side f. The function f is called *Lipschitz continuous* if there exists a finite constant L so that

$$\big|f(x) - f(y)\big| \le L|x - y|, \qquad \text{for all } x, y \in \Omega.$$

The best such constant L is called the *Lipschitz constant* for f on Ω.

Lipschitz continuity is not uncommon. For example, assume that f is a C^1 function, by which we mean that its first order partial derivatives exist and are continuous. We write Df for the Jacobian matrix of f:

$$Df = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \cdots \\ \frac{\partial f_n}{\partial x_1} & \cdots & \frac{\partial f_n}{\partial x_n} \end{pmatrix}$$

Then, if the whole line segment $[x, y]$ with end points x and y lies within Ω, we can write

$$f(y) - f(x) = \int_0^1 \frac{d}{dt} f\big((1-t)x + ty\big)\,dt = \int_0^1 Df\big((1-t)x + ty\big)\,dt \cdot (y - x)$$

with the result that, if $\|Df(z)\| \le L$ for all z, (where the norm is the operator norm of the matrix, seen as an operator on $\mathbb{R}^n$), then $|f(x) - f(y)| \le L|x-y|$.

If f belongs to C^1 then f is *locally Lipschitz continuous*, which means that every point $x \in \Omega$ has a neighbourhood in which f is Lipschitz continuous.

Grönwall's inequality. In its simplest form, it is this simple fact:

1 Proposition. *Let u be a real, differential function on some interval. Assume that $\dot{u}(t) \le au(t)$ in this interval. Then $e^{-at}u(t)$ is a nonincreasing function of t.*

Proof: Just differentiate:

$$\frac{d}{dt}\big(e^{-at}u(t)\big) = e^{-at}\big(\dot{u}(t) - au(t)\big) \le 0,$$

and we're done. ∎

We shall not need the general form of Grönwall's inequality, but for the sake of completeness, here it is:

2 Proposition. (Grönwall's inequality) *Let u be a real, differential function on some interval. Assume that $\dot{u}(t) \le g(t)u(t)$ in this interval. Then $e^{-G(t)}u(t)$ is a nonincreasing function of t, where $\dot{G}(t) = g(t)$.*

In particular, for $t > 0$ we find the traditional form of Grönwall's inequality:

$$u(t) \le u(0)\exp\Big(\int_0^t g(\tau)\,d\tau\Big),$$

which is just a difficult way of writing $e^{-G(t)}u(t) \le e^{-G(0)}u(0)$.

The proof is just as easy as for the simplified version above.

Uniqueness

The basic idea relies on the following calculation. We assume that x and y are two solutions of (1), and note that

$$\frac{d}{dt}\big|x(t) - y(t)\big| \le \big|\dot{x}(t) - \dot{y}(t)\big| = \big|f\big(x(t)\big) - f\big(y(t)\big)\big| \le L\big|x(t) - y(t)\big|$$

if f is Lipschitz continuous.

This implies that $e^{-Lt}\left|x(t)-y(t)\right|$ is non-increasing. But for $t=0$, this quantity is zero, since $x(0)=a=y(0)$, and so it must be zero for all positive t. (The same argument holds for negative t, by time reversal: If $x(t)$ solves (1) then $\tilde{x}(t)=x(-t)$ solves a similar problem with f replaced by $-f$. So if we have uniqueness forward in time, the same must hold backward in time.)

This idea, simple as it is, is somewhat ruined by a couple ugly facts: First, $\left|x(t)-y(t)\right|$ may be non-differentiable at any point where $x(t)=y(t)$, and second, a requirement of global Lipschitz continuity is too much. However, we can adapt the idea to prove

3 Theorem. *Assume that f is locally Lipschitz continuous. Then* (1) *has at most one solution on any given interval containing* 0*.*

Proof: Assume that x and y are two solutions. Assume also that $x(t_1)\neq y(t_1)$ for some $t_1>0$ in the given interval. (We can deal with $t_1<0$ by time reversal.)

Now there is some t_0, with $0\le t_0<t$, with $x(t_0)=y(t_0)$ but $x(t)\neq y(t)$ for $t_0<t\le t_1$. There is some neighbourhood U of $x(t_0)$ on which f is Lipschitz continuous. For $t\ge t_0$ and $t-t_0$ small enough, $x(t)$ and $y(t)$ both belong to U, and so $e^{-Lt}\left|x(t)-y(t)\right|$ is non-increasing for these t. Since this quantity is continuous and zero at $t=t_0$, and strictly positive for $t>t_0$, that is nonsense. This contradiction completes the proof. ∎

Existence

For an existence proof, we rely on *Banach's fixed point theorem*: If X is a complete metric space and $\Phi\colon X\to X$ is a contraction, then Φ has a fixed point in X. This fixed point is found by iteration: Let $x_0\in X$ be arbitrary, and let $x_{n+1}=\Phi(x_n)$. The sequence (x_n) will converge to the fixed point.

To use this on (1), note that (1) is equivalent with

$$x(t)=a+\int_0^t f\bigl(x(\tau)\bigr)\,d\tau$$

which says that x is a fixed point of the mapping Φ given by

$$\Phi(x)(t)=a+\int_0^t f\bigl(x(\tau)\bigr)\,d\tau.$$

To be specific, we shall work in the metric space X consisting of all functions $x\colon[-\delta,\delta]\to B$, where B is the closed ball $B=\bigl\{x\colon |x-a|\le r\bigr\}$, and r is some positive number. We shall assume that f is Lipschitz continuous on B. Let L be the corresponding Lipschitz constant, and let M be the maximum value of $|f|$ on B.

We need to ensure that Φ really maps X into itself. To this end, estimate

$$\big|\Phi(x)(t)-a\big| = \Big|\int_0^t f\big(x(\tau)\big)\,d\tau\Big| \le \Big|\int_0^t \big|f\big(x(\tau)\big)\big|\,d\tau\Big| \le M\delta,$$

so we need to make sure that $M\delta \le r$.

Second, to make sure that Φ is a contraction, estimate

$$\begin{aligned}\big|\Phi(x)(t)-\Phi(y)(t)\big| &= \Big|\int_0^t \big(f\big(x(\tau)\big)-f\big((y(\tau)\big)\,d\tau\Big| \le \Big|\int_0^t \big|f\big(x(\tau)\big)-f\big(y(\tau)\big)\big|\,d\tau\Big|\\ &\le L\delta\|x-y\|,\end{aligned}$$

and so we need to make sure that $L\delta < 1$.

Then Φ is a contraction on X, and so we have proved:

4 Theorem. *If f is locally Lipschitz then* (1) *has a solution on some open interval containing* 0*.*

In fact, it is not hard to show that there exists a *maximal interval of existence*, that is an open interval I on which (1) has a solution, and so I contains any other open interval with a solution on it. One simply takes I to be the union of all open intervals J containing 0 so that (1) has a solution on J. For any $t \in I$, pick some J on which there exists a solution y, and define $x(t) = y(t)$. If K is another such interval, and z is a solution on K, then $J \cap K$ is yet another interval, so the uniqueness theorem shows that $y = z$ on $J \cap K$. Therefore our definition of $x(t)$ does not depend on the particular choice of J.

5 Theorem. *Let the maximal interval of existence be (a, b), where $-\infty \le a < 0 < b \le \infty$. If $b < \infty$, there is a sequence (t_k) in this interval with $t_k \to b$, so that either $|x(t_k)| \to \infty$, or $\operatorname{dist}\big(x(t),\partial\Omega\big) \to 0$.*

Similarly, if $a > -\infty$, there is a sequence with these properties converging to a.

Here $\partial\Omega$ is the boundary of Ω.

Proof sketch: Assume not. Then there is a constant $M < \infty$ and a $\varepsilon > 0$ so that $|x(t) \le M|$ and $\operatorname{dist}\big(x(t),\partial\Omega\big)$ whenever $0 < t < b$. That is, $x(t)$ belongs to the compact set

$$K = \big\{x \in \Omega\colon\ |x(t) \le M| \text{ and } \operatorname{dist}\big(x(t),\partial\Omega\big)\big\}.$$

By compactness, there is a sequence (t_k) with $t_k \to b$ and $x(t_k) \to z \in K$.

From the proof of the existence result above, there exists some $\delta > 0$ so that the initial value problem can be solved in $[-\delta, \delta]$ for all initial values in some neighbourhood of z. That means the same is true for an initial value $x(t_k)$ for all sufficiently large k, so the solution can be extended at least up to time $t = t_k + \delta$. Since $t_k \to b$ and the solution cannot be extended beyond $t = b$, this is absurd. ∎

Continuous dependence on data

I shall only consider the dependence on the initial value a. Assume that x solves (1), and that y solves the same system, but with initial data $y(0) = b$. If f is Lipschitz continuous with Lipschitz constant L, we have seen that $e^{-Lt}\bigl|x(t) - y(t)\bigr|$ is non-increasing, so that

$$\bigl|x(t) - y(t)\bigr| \le e^{L|t|}\bigl|x(0) - y(0)\bigr|$$

(I added a strategic absolute value in the exponent on the righthand side, so the result can also be used for $t < 0$. It's another use of time reversal.) So the solution depends continuously on the initial data. (The dependence is locally Lipschitz continuous, but that takes a bit of effort to prove, so I'll skip it.)

If f is smoother, then we can even conclude that the solution depends on the initial data in a *differentiable* way:

Write now $x(t, a)$ for the solution with initial condition a, so that (1) can be written

$$\frac{\partial x}{\partial t} = f\bigl(x(t, a)\bigr), \qquad x(0, a) = a$$

Assuming for a moment that f is differentiable with respect to a, with continuous partial derivatives, we expect to find

$$\frac{\partial}{\partial t}\frac{\partial x}{\partial a_j} = \frac{\partial}{\partial a_j}\frac{\partial x}{\partial t} = \frac{\partial}{\partial a_j} f\bigl(x(t, a)\bigr) = Df\bigl(x(t, a)\bigr)\frac{\partial x}{\partial a_j}$$

so that $\partial x/\partial a_j$ itself satisfies a differential equation. It will also satify the initial condition $\partial x/\partial a_j(0) = e_j$, where e_j is the jth unit vector.

So one can turn this argument inside out: Assuming that Df is Lipschitz continuous, the problem $\dot z_j = Df\bigl(x(t, a)\bigr)z_j$, $z_j(0) = e_j$ has a solution, and that solution can then be shown to be the partial derivative $\partial x/\partial a_j$.

Odds and ends

Non-autonomous systems. The initial value problem for a non-autonomous system

$$\begin{aligned}\dot{x}(t) &= f\bigl(x(t),t\bigr),\\ x(0) &= a.\end{aligned}$$

can be reduced to the autonomous form (1) by writing $w(t) = \bigl(x(t),t\bigr)$ and solving the autonmous system

$$\begin{aligned}\dot{w}(t) &= f\bigl(w(t)\bigr),\\ w(0) &= (a,0).\end{aligned}$$

This may not be the best way to study non-autonomous systems, but it does show that the well-posedness results extends to this case.

Continuous dependence on f. Assume that f depends on further parameters $b\in\mathbb{R}^m$:

$$\begin{aligned}\dot{x}(t) &= f\bigl(x(t),b\bigr),\\ x(0) &= a.\end{aligned}$$

A rather silly looking way to solve this is to write $w(t) = \bigl(x(t),b\bigr)$ and to solve

$$\begin{aligned}\dot{w}(t) &= \bigl(f\bigl(w(t)\bigr),0\bigr),\\ w(0) &= (a,b).\end{aligned}$$

That is, we add the components b to x and add equations saying that those components of w are constants (their derivatives are zero).

Note that the b moved from f into the initial conditions. It follows that the solution depends continuously (smoothly, if f is smooth) on b.