110.421 DYNAMICAL SYSTEMS

Week 2 Lecture 1 Notes

Lecture 2.1: The Contraction Principle

Theorem 1 (the Contraction Principle). *Let $I \subset \mathbb{R}$ be closed (think defined by inequalities of the type $\leq$ or $\geq$), possibly infinite on one or both sides, and $f : I \to I$ a λ-contraction. Then f has a unique fixed point x_0, and*

$$|f^n(x) - x_0| \leq \lambda^n |x - x_0| .$$

Some notes before we prove this theorem:

- $\forall x \in I$, $\mathcal{O}_x \longrightarrow x_0$ exponentially due to the factor λ^n, where $0 < \lambda < 1$.
- As stated, this is only a result valid on subsets of $\mathbb{R}$, for now.
- **Question:** Can a contraction have two fixed points? This will be an exercise, but really the question is "why not?"
- For the proof of the theorem, we will need a convenient fact from Analysis: That the idea of a sequence converging in real space is the same as the fact that the sequence is *Cauchy*:

Definition 2. For $N \in \mathbb{N}$, a sequence $\{x_1, x_2, \ldots\} \in \mathbb{R}^N$ is called *Cauchy* if $\forall \epsilon > 0$, $\exists A > 0$ such that $\forall m, n \geq A$, $d(x_m, x_n) < \epsilon$.

Proposition 3. *A sequence in $\mathbb{R}^N$, $N \in \mathbb{N}$ converges iff it is Cauchy.*

Proof of Contraction Principle (in $\mathbb{R}$). We start with the idea that the orbit of an arbitrary point converges. Choose $x \in I$. It should be obviously since f is a map on I that $\mathcal{O}_x \in I$.

Date: February 5, 2013.

Now, for $m, n \in \mathbb{N}$, where we can assume without any loss of generality that $m \geq n$,

$$\begin{aligned}
|f^m(x) - f^n(x)| &= \big|f^m(x) - f^{m-1}(x) + f^{m-1}(x) - f^{m-2}(x) + f^{m-2}(x) - \dots \\
&\qquad - f^{n+1}(x) + f^{n+1}(x) - f^n(x)\big| \quad \text{(Clever form of addition by zero)} \\
&= \left|\sum_{r=n}^{m-1} \left(f^{r+1}(x) - f^r(x)\right)\right| \\
&\leq \sum_{r=n}^{m-1} \left|f^{r+1}(x) - f^r(x)\right| \quad \text{(Triangle inequality)} \\
&\leq \sum_{r=n}^{m-1} \lambda^r |f(x) - x| \quad \text{(many applications of the Lipschitz condition)} \\
&= |f(x) - x| \sum_{r=n}^{m-1} \lambda^r \\
&= |f(x) - x| \frac{\lambda^n - \lambda^m}{1 - \lambda} \quad \text{(partial sum of a geometric series)} \\
&\leq \frac{\lambda^n}{1 - \lambda} |f(x) - x|.
\end{aligned}$$

The conclusion from all of this is that, as n gets large (under the assumption that $m \geq n$, this means that m gets large also), the right hand side of the last inequality gets small. Hence, $\mathcal{O}_x$ is a Cauchy sequence. Hence it must converge to something. As I is closed, it must converge to something in I. Let's call this number x_0.

Now, applying the Lipschitz condition to the nth iterate of f leads directly to the condition that

$$|f^n(x) - f^n(y)| \leq \lambda^n |x - y|,$$

so that every two orbits also converge to each other. Hence every orbit converges to this number x_0.

And finally, we can show that x_0 is actually a fixed point solution for f. To see this, again $\forall x \in I$, and $\forall n \in \mathbb{N}$,

$$\begin{aligned}
|x_0 - f(x_0)| &= \left|x_0 - f^n(x) + f^n(x) - f^{n+1}(x) + f^{n+1}(x) - f(x_0)\right| \\
&\leq |x_0 - f^n(x)| + \left|f^n(x) - f^{n+1}(x)\right| + \left|f^{n+1}(x) - f(x_0)\right| \\
&\leq |x_0 - f^n(x)| + \lambda^n |x - f(x)| + \lambda |f^n(x) - x_0| \\
&= (1 + \lambda) |x_0 - f^n(x)| + \lambda^n |x - f(x)|.
\end{aligned}$$

Again, the steps are straightforward. The first step is another clever addition of zero, and the second is the triangle inequality. The third involves using the Lipschitz condition on two of the terms, and the last is a clean up of the leftovers. However, this last inequality is the most important. It must be valid for every choice of $n \in \mathbb{N}$. Hence choosing an increasing sequence of values for n, we see that as $n \to \infty$, both $|x_0 - f^n(x)| \longrightarrow 0$ and $\lambda^n \longrightarrow 0$. Hence, it must be the case that $|x_0 - f(x_0)| = 0$ or $x_0 = f(x_0)$. Thus, x_0 is a fixed point solution for f on I and the theorem is established. $\square$

Example 4. $f(x) = \sqrt{x}$ on the interval $[1,\infty)$ is a $\frac{1}{2}$-contraction. Why? $f \in C^1$ and $0 < f'(x) = \dfrac{1}{2\sqrt{x}} \leq \dfrac{1}{2}$. By the Contraction Principle, then, there is a unique fixed point for this discrete dynamical system. Can you find it?

Example 5. $f(x) = \ln(x-1) + 5$ on the interval $I = [2, 100]$. Is this a contraction? What is $\mathrm{Lip}(f)$? Where is the fixed point?

1. Several Variables

The Contraction Principle in several variables is basically the same as that of one variable. The only difference, really, is that the absolute signs are replaced by the more general metric in $\mathbb{R}^n$:

Theorem 6 (The Contraction Principle)**.** *Let $X \subset \mathbb{R}^n$ be closed (again, one way to avoid the technicalities of what constitutes a closed set in real space is to think of all sets which are solutions to either equations or inequalities of the form $\leq$ or $\geq$. Or think of a domain in $\mathbb{R}^3$ in the vector calculus sense, and let X be this domain with its closure), and $f : X \to X$ a λ-contraction. Then f has unique fixed point $x_0 \in X$ and $\forall x \in X$,*

$$d\left(f^n(x), x_0\right) \leq \lambda^n d\left(x, x_0\right).$$

Notes:

- We say here that the "dynamics are simple". The orbit of every point in x does exactly the same thing: Converge exponentially to the fixed point solution x_0.
- This also means that every orbit converges to every other orbit also!
- **Question:** Is it possible to have a periodic point with nontrivial period in a contraction? (Fixed points have prime period 1. Fixed points are trivial periodic points. A nontrivial periodic point – a periodic point of nontrivial period – is one whose prime period is 2 or more.)

Also, in dimension-1, the derivative of f (if it existed) can help to define the Lipschitz constant (Recall Propositions 2.2.3 and 2.2.5 in the book). How about in several dimensions? In this context, recall that for a C^1 function $f : X \subset \mathbb{R}^n \to \mathbb{R}^n$, the derivative $df : \mathbb{R}^n \to \mathbb{R}^n$ is also a map on $\mathbb{R}^n$ whose value at a point x is a linear map from $T_x\mathbb{R}^n$ to $T_{f(x)}\mathbb{R}^n$. The points in the domain X where the $n \times n$-matrix df_x is of maximal rank are called regular. At a domain point, we can define a matrix norm for df_x as the maximal stretching ability of the unit vectors in the tangent space:

$$||df_x|| = \max_{||v||=1} ||df_x(v)|| \,.$$

Note that $||\cdot||$ is a non-negative real number.

With this, there are two "derivative"-versions of the Contraction Principle of note. The first is a sort of "global version" since the result holds over the entire domain:

Theorem 7 (Global version)**.** *If X is the closure of a strictly convex set in $\mathbb{R}^n$, and $f : X \to \mathbb{R}^n$ a C^1-map with $||df_x|| \leq \lambda < 1$, then f has a unique fixed point x_0, and $\forall x \in X$,*

$$d(f^n(x), x_0) \leq \lambda^n d(x, x_0).$$

Note that here "strictly convex" means that for every $x, y \in X$, the line joining x and y lies completely within the interior of X. Hence, technically, we really want that f is differentiable on the interior of X and continuous on the boundary.

This next version is a local version, and present a very usable tool for the analysis of what happens near (in a neighborhood of) a fixed point:

Theorem 8 (Local version)**.** *Let f be differentiable with a fixed point x_0 such that all of the eigenvalues of the derivative df_{x_0} have absolute values less than 1. Then there exists a open neighborhood (an ϵ-ball) $U(x_0)$ such that on the closure $\overline{U}$ of U, $f(\overline{U}) \subset \overline{U}$ and f is a contraction on $\overline{U}$.*

Thus, on $\overline{U}$, which will be a strictly convex set, the global version above applies. To see these two versions in action, here are two applications:

Example 9 (local version app)**.** In the Newton-Raphson method for the location of a root of a C^3-function $f : \mathbb{R} \to \mathbb{R}$, basic local information of f near a root can ensure that the method will be successful, as long as one starts "close" enough to the root, that is.

Proposition 10. *Let $f : \mathbb{R} \to \mathbb{R}$ be C^2 with a root r. If $\exists \delta, M > 0$, such that $|f'(x)| > \delta$ and $|f''(x)| < M$ on an open interval containing r, then*

$$F(x) = x - \frac{f(x)}{f'(x)}.$$

is a contraction near r.

Proof. Since $f \in C^2[\mathbb{R})]$, it follows that F is C^1, and

$$F'(r) = \frac{d}{dx}\left[x - \frac{f(x)}{f'(x)}\right]\Bigg|_{x=r} = \frac{f(r)f''(r)}{[f'(r)]^2} = 0$$

since r is a root of f. And by continuity then, one can find a small open interval containing r, where $|F'(x)| < 1$ on its closure by the local version of the Contraction Principle above (this is what "near" means in the proposition. $\square$

Example 11 (global version app)**.** The standard proof of the existence and uniqueness of solutions to the first-order, initial value problem (IVP)

$$\dot{y} = f(t, y), \quad y(t_0) = y_0$$

in a neighborhood of (t_0, y_0) in the t, y-plane involves the use of what are called Picard Iterations. The basic theorem says that if $f(t, y)$ and $\frac{df}{dy}(t, y)$ are continuous in a neighborhood of (t_0, y_0), then there exists a unique solution passing through the initial point. The proof involves a knowledge of just how solutions behave. Any solution $y = \phi(t)$ to the IVP passing through (t_0, y_0) would satisfy

$$\phi(t) = y_0 + \int_{t_0}^{t} f(s, \phi(s))\, ds. \tag{1}$$

This comes simply from integrating the vector field given by f along the solution (that is, differentiate the equation to recover the ODE). AS long as f is continuous, the integral will exist and hence the solution will exist through (t_0, y_0), at least locally.

How the Contraction Principle applies is quite clever. In short, construct a function T on the space of all functions from $\mathbb{R}$ to $\mathbb{R}$ in the following way. Given any $\psi : \mathbb{R} \to \mathbb{R}$, define

$$T(\psi) = y_0 + \int_{t_0}^{t} f(s, \psi(s))\, ds.$$

Here, $T(\psi)$ will just be another function from $\mathbb{R}$ to $\mathbb{R}$. In this way, we can use T to define a discrete dynamical system, where T is the evolution and the space of all functions from $\mathbb{R}$ to $\mathbb{R}$ is the state space. Iterating T is what Picard developed. It turns out (and this takes some work), that T is a contraction (to show this, one needs at one point that $\frac{df}{dy}(t, y)$ is continuous near (t_0, y_0)). But thus there will be only one fixed point. And the fixed point would satisfy $T(\psi) = \psi$. But the function that is the fixed point is precisely the one that satisfies Equation 1! Nice. For more details, go to the website for the course and seek out the Contraction Principle example in the schedule or on the main webpage.