Lecture notes on Banach's contraction principles and Picard - Lindelöf theorem.

1. To show existense and uniquness of solutions to an ODE we need a structure on a space of functions: a vector space structure and a norm to measure distance between functions and to speak about convergence in the space of functions. To solve an equation we will introduce some normed vector space. It will be in particular a space of continuous functions (on a compact) with supremum norm $\|f\|_C = \sup_{x \in I} |f(x)|$, I compact (interval) in $\mathbb{R}$.

2. Considering convergence of approximations to solutions we need to have this space of functions complete. It means that Cauchy sequences $\{x_n\}_{n=1}^{\infty}$ such that $\|x_m - x_n\| \xrightarrow[m,n \to \infty]{} 0$ must be convergent (have a limit $\bar{x}$: $\lim_{n \to \infty} x_n = \bar{x}$). An advantage of this property is that we do not need to know the limit $\bar{x}$ to show it existens, only estimate $\|x_m - x_n\|$

A complete normed space is called Banach space after famous polish mathematician Stefan Banach (1892-1945) who established grounds for functional analysis.

We can reformulate the initial value problem for an ODE:

$$\dot{x} = f(t,x) \; ; \qquad x(t_0) = x_0$$

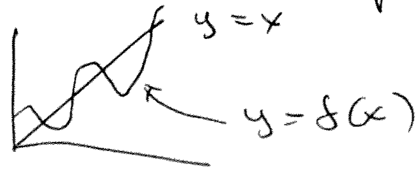
as an integral equation, integrating the ODE from t_0 to t:

$$x(t) = x_0 + \int_{t_0}^{t} f(s, x(s))\,ds.$$

Introducing notation $K(x)$ for the operator $x_0 + \int_{t_0}^{t} f(s, x(s))\,ds = K(x)$ we get a reformulation of the problem to a fixed point problem for operator $K(x)$

$$x = K(x).$$

A non linear operator can have many or no fixed points at ale. A simple example is a scalar equation in $\mathbb{R}$: $x = f(x)$



These observations lead us to considering Banach's contraction principle: the simpes theorem giving existence (and at the same time uniqueness) of a fixed point to an operator in a Banach space.

Definition of contraction.

Operator $K: C \to C$, $C \subset X$ - Banach space (or just normed space) is called contraction (on C) if $\exists\ 0 \le \theta < 1$ such that

$$\|K(x) - K(y)\| \le \theta \|x-y\| \quad \text{for all } x, y \in C$$

Geometric meaning of that is that the operator K maps a pair of points x, y to a pair of points $K(x)$, $K(y)$ that have shorter distance between them comparing with the initial distance between x and y.
Operator K does not need to have this property on a larger set than C.
It is important to specify on which set this property takes place.

Theorem 2.1 Banach's contraction principle.

Let C be a nonempty closed subset of a Banach space X and $K: C \to C$ - a contraction operator ($\|K(x) - K(y)\| \le \theta \|x-y\|$, $0 \le \theta < 1$). Then $\exists$ a unique fixed point $\bar{x} \in C$ such that

$$\|K^n(x) - \bar{x}\| \le \frac{\theta^n}{1-\theta} \|K(x) - x\|$$

for any $x \in C$. $K^n(x)$ is the operator K applied n-times to an elemen $x \in C$.

$$K^n(x) = \underbrace{K(K(\dots K(x))\dots)}_{n\text{-times}}.$$

Proof. Proof uses the following ideas: a) approximating the fixed point by iterations $x_{n+1} = K(x_n)$; $x_0 = x$
x - arbitrary $x \in C$

b) Chauchy principle: show that $\{x_n\}$ - sequence of approximations is a Cauchy sequence.

c) telescoping sums: express $x_n - x_m$ as
$$x_n - x_m = (x_n - x_{n-1}) + (x_{n-1} - x_{n-2}) \dots (x_1 - x_0)$$

d) estimates for $\|x_n - x_{n-1}\| \le \Theta^n \|x_1 - x_0\|$ using that $x_n - x_{n-1} = K(x_{n-1}) - K(x_{n-2})$ ... and K is a contraction

e) Estimate $\|x_n - x_m\| = \|(x_n - x_{n-1}) + \dots (x_1 - x_0)\|$ by a convergen partial sum of a convergent number series.

Consider difference between to consequent approximations:
$$\|x_{n+1} - x_n\| = \|K(x_n) - K(x_{n-1})\| \le \theta \|x_n - x_{n-1}\| \le$$
$$= \Theta \|K(x_{n-1}) - K(x_{n-2})\| \le \Theta^2 \|x_{n-1} - x_{n-2}\| \le \dots$$
$$\dots \le \Theta^n \|x_1 - x_0\| = \Theta^n (\|K(x_0) - x_0\|)$$
Show that $\{x_n\}$ is a Cauchy sequence:

Use telescoping sums:
$$\|x_m - x_n\| = \|(x_m - x_{m-1}) + (x_{m-1} - x_{m-2}) + \dots + (x_{n+1} - x_n)\|$$
$$\le \sum_{j=m+1}^{n} \|x_j - x_{j-1}\| \quad \text{(by triangle inequality } \|a+b\| \le \|a\| + \|b\| \text{)}$$

$$\sum_{j=m+1}^{n} \|x_j - x_{j-1}\| \le \left(\sum_{j=m+1}^{n} \theta^j\right) \|x_1 - x_0\|$$

Series $\sum_{j=0}^{\infty} \theta^j$ is convergent geometric series: $\sum_{j=0}^{\infty} \theta^j = \frac{1}{1-\theta}$ if $0 \le \theta < 1$

It implies that partial sum $\sum_{j=m+1}^{n} \theta^j \underset{m,n \to \infty}{\longrightarrow} 0$ and therefore $\|x_m - x_n\| \underset{m,n \to \infty}{\longrightarrow} 0$

One can also write an explicit formula for $\sum_{j=m+1}^{n} \theta^j = \theta^m \sum_{j=0}^{n-m-1} \theta^j = \theta^m \left(\frac{1-\theta^{n-m}}{1-\theta}\right)$

and observe that it goes to zero when $n, m \to \infty$, $n > m$. for $0 \le \theta < 1$.

We have shown that $\{x_n\}$ is a Cauchy sequence, therefore $x_n \underset{n \to \infty}{\longrightarrow} \bar{x}$.

All $x_n \in C$ because $K(x) \in C$ if $x \in C$. Therefore the limit $\lim_{n \to \infty} x_n = \bar{x} \in \bar{C}$ - closure of C (it can go to the boundary of C even if $x_n \in$ interior of C.

But C is closed by requirements in the theorem $\Rightarrow$ $\bar{x} \in C$.

The limit $\bar{x}$ is a fixed point of K. We see it by going to the limit in the relation

$$x_{n+1} = K x_n \underset{n \to \infty}{\longrightarrow} \bar{x} = K \bar{x}$$

(a contraction is a continuous mapping because $\|K(x) - K(y)\| \le \theta \|x - y\|$, so if $x \to y$ $K(x) \to K(y)$)

Uniqueness of the fixed point follows by the estimate of the difference between two fixed points $\bar{x} = K(\bar{x})$ and $\hat{x} = K(\hat{x})$:

$$\|\hat{x} - \bar{x}\| = \|K(\hat{x}) - K(\bar{x})\| \le \theta \|\hat{x} - \bar{x}\|$$

$0 \le \theta < 1$!

$\Rightarrow$ can be true only if $\|\hat{x} - \bar{x}\| = 0$

$\Rightarrow \bar{x} = \hat{x}$.

("Our" $\bar{x} = \lim x_n$ belongs to C, so we can apply the estimate for K)

Uniqueness is valid only for fixed points in C. Outside C if K is defined on a larger set can be other fixed points, if K is not a contraction on some larger set.

Picard Lindelöf theorem

We consider a reformulation of the IVP

$$\dot{x} = f(t,x), \quad f \in C(U, \mathbb{R}^n); \qquad x(t_0) = x_0 \qquad U\text{-open} \subset \mathbb{R}^{n+1}; \quad (t_0, x_0) \in U$$

to an integral equation

$$x(t) = x_0 + \int_{t_0}^{t} f(s, x(s))\,ds$$

or in abstract form

$$x = K(x) \quad \text{for} \quad K(x) = x_0 + \int_{t_0}^{t} f(s, x(s))\,ds$$

We will prove existence and uniqueness of solution as limit of iterations

$$x_{n+1} = K(x_n); \quad x_0 = x_0.$$

Convergence of these iterations is possible to show for solutions on a "short" time interval, depending on the "size" of the function $f(t,x)$.
Another limitation of the theorem that we formulate is that values of the solution $x(t)$ that we will approximate will all belong to a small ball $B_\delta(x_0)$ around the initial point $x_0 \in \mathbb{R}^n$.
Solution to IVP we construct will be local in this sense: both in space and in time. Proof of existence and uniqueness will be given in two different forms: using contraction principle for K and

using a more general estimate for iterations: not by geometric series, but by series for the exponential function. The last method will give a longer time interval for solution to exist.

Motivation for limitations in the formulation of the existence theorem.

We need to choose several structures and limitations for formulation of the theorem.

1. A Banach space X for operator K
2. A closed set $C \subset X$ such that $K: C \to C$
3. Domains of definition V for f and for functions in C, such that the property $K: C \to C$ is really valid.
4. Check estimates for iterations
$$K^n x = K(K(K\ldots K(x))) \text{ for } x \in C$$
to see if one can get convergence of
$$x_{n+1} = K(x_n) \quad \text{when } n \to \infty.$$

The right hand side $f(t,x)$ in the ODE is a continuous function, so a natural choice of Banach space for $x(t)$ is a space of continuous functions on some time interval $[t_0, t_0+T]$.

$X = C([t_0, t_0+T], \mathbb{R}^n)$ with norm

$$\|x\|_C = \sup_{t\in[t_0,t_0+T]} |x(t)|.$$

To compute $f(t, x(t))$ in the expression for $K(x)$ we need to choose a proper (smaller) domain for f. It will be the set $V = [t_0, t_0+T] \times \overline{B_\delta}(x_0)$ where $\overline{B_\delta}(x_0) = \{x \in \mathbb{R}^n : |x - x_0| \leq \delta\}$ - is a closed ball around x_0 in $\mathbb{R}^n$ and the length of the time interval for $x(t)$ and $f(t,x)$ is chosen now so that $V \subset U$ - (original domain for f).

We observe that $\sup_{(t,x)\subset V} |f(t,x)| = M < \infty$ because f is continuous and V is compact.

Next we need to choose a closed set $C \subset X$ such that $K : C \to C$ does not map points in C outside C.

We estimate $|K(x) - x_0|$ for $x(t)$ with graph in V: $\{(t, x(t)) \,|\, t \in [0,T]\} \subset V$

$$\left|K(x)(t) - x_0\right|_{t = t_0+\Delta t} \leq \int_{t_0}^{t_0+\Delta t} |f(s, x(s))|\, ds \leq \Delta t\, M \leq$$
$$\leq T M$$

If we choose $T = T_0$ in the definition of V such that $T_0 = \min\left\{T, \frac{\delta}{M}\right\}$ and introduce notation $V_0 = [t_0, t_0+T_0] \times \overline{B_\delta}(x_0)$ we get

$$\sup_{t\in[t_0, t_0+\Delta t]} |K(x)(t) - x_0| \leq \delta \text{ for } (t, x(t)) \subset V_0$$

Or $\|K(x) - x_0\|_C \leq \delta$ for $(t, x(t)) \subset V_0$

Here $\| \quad \|_C$ is a supremum norm in the set of continuous functions in $C([t_0, t_0+T_0], \mathbb{R}^n)$ and $(t, x(t)) \subset V_0$ implies that $\sup_{t\in[t_0,\, t_0+T_0]} |x(t)| \leq \delta$.

The estimate for $\|K(x) - x_0\| \leq \delta$ shows that the operator K maps functions from $C([t_0, t_0+T_0], B_\delta(x_0))$ to the same set or the ball $\|x - x_0\|_C \leq \delta$ in $C([t_0, t_0+T_0], \mathbb{R}^n)$ into itself. This ball (a Banach space) can be chosen as the closed set $C \subset X$ as in the Banach contraction principle to find fixed point $\bar{x}$ to K: $\bar{x} = K\bar{x}$ in C.

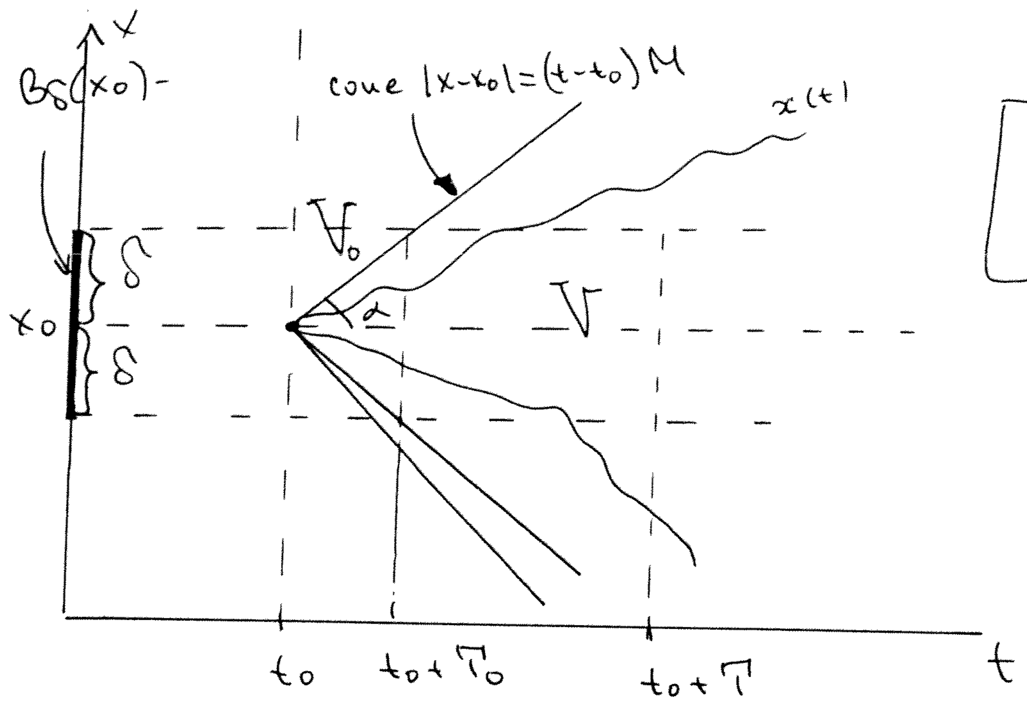
$K: C \to C$, and C is a closed set in $C([t_0, t_0+T_0], \mathbb{R}^n)$

One can interpret the limitation for the time interval: $T_0 = \min[T, \frac{\delta}{M}]$ by interpreting what it means in terms of the solution $\bar{x}(t)$ to the problem.

$$\bar{x} = x_0 + \int_{t_0}^{t} f(s, \bar{x}(s))\,ds \; ; \quad \bar{x} = K\bar{x}$$

$$|K(\bar{x})(t) - x_0| = |\bar{x}(t) - x_0| \leq \Delta t\, M \; ; \quad t = t_0 + \Delta t$$

We draw a cone $|x - x_0| = (t - t_0)M$ in "$\mathbb{R}^{n+1}$" in the set $V = [t_0, t_0+T] \times \overline{B_\delta(x_0)}$



$\tan\alpha = \dfrac{\delta}{M}$

$x(t)$ cannot leave the cone $|x-x_0|=(t-t_0)M$ because of the estimate. It defines the time interval $[t_0, t_0+T_0]$ such that the solution $x(t)$ does not leave the ball $B_\delta(x)$ around point x_0.

To get convergence of our approximations we need estimates of the operator K like $\|K(x)-K(y)\| \le \text{Const}\,\|x-y\|$ as we see from contraction principle or similar arguments. Such estimates is possible to get only if the function $f(t,x)$ has some smoothness in the second argument

Definition $f(t,x)$ is locally Lipschitz continuous in the second argument uniformly with respect to the first argument if for any compact $V_0 \subset U$

$$L = \sup_{(t,x)\neq(t,y)\in V_0} \frac{|f(t,x)-f(t,y)|}{|x-y|} < \infty, \text{ or}$$

it is the same as $|f(t,x)-f(t,y)| < L|x-y|$ with L - independent of $(t,x),(t,y)\in V_0$.

§2.2. Theorem 2.2. (Picard-Lindelöf)

Let $f \in C(U, \mathbb{R}^n)$, $U \subset \mathbb{R}^{n+1}$ - an open set, $(t_0, x_0) \in U$.
If f - locally Lipschitz continuous in the second argument uniformly with respect to t-argument, then $\exists$ a unique local solution $\bar{x}(t) \in C^1(I)$ of IVP, where I is some interval around t_0.

More specifically if $V = [t_0, t_0+T] \times \overline{B_\delta(x_0)} \subset U$ and $M = \max\limits_{(t,x)\in V} |f(t,x)|$. Then solution exists on an interval $t \in [t_0, t_0+T_0]$ with $T_0 = \min\left(T, \frac{\delta}{M}\right)$ and remains in $\overline{B_\delta(x_0)}$. Similar result is valid for $t \in [t_0 - T_0, t_0]$.

Proof. We have already observed that the operator $K(x)(t) = x_0 + \int_{t_0}^{t} f(s, x(s))\,ds$ maps ball $\|x - x_0\|_C \leq \delta$ into itself in the space $X = C([0, T_0], \mathbb{R}^n)$ - Banach space.

$$\|K(x)(t) - K(y)(t)\| = \sup_{t\in[0,T_0]} \left| \int_{t_0}^{t} f(s, x(s))\,ds - \int_{t_0}^{t} f(s, y(s))\,ds \right| \leq$$

$$\leq \sup_{t\in[t_0,T_0]} \left| \int_{t_0}^{t} |f(s,x(s)) - f(s,y(s))|\,ds \right| \leq$$

$$\leq \sup_{t\in[t_0,T_0]} L \int_{t_0}^{t} |x(s) - y(s)|\,ds \leq L \cdot \sup_{s\in[t_0,T_0]} |x(s) - y(s)| \cdot T_0$$

$$= T_0 \cdot L \cdot \|x - y\|_C .$$

Therefore if $T_0 < \frac{1}{L}$ => K is a contraction in $X = C([t_0, t_0+T_0], \mathbb{R}^n)$ and $K : C \to C$ - a ball $\|x - x_0\| \leq \delta$ in X.

Therefore if $T_0 < \frac{1}{L}$ the operator K has a unique fixed point $\bar{x}$ in C.

$$\bar{x}(t) = x_0 + \int_{t_0}^{t} f(s, \bar{x}(s))\,ds.$$

The function $f(s, \bar{x}(s))$ is continuous with respect to s - therefore $\int_{t_0}^{t} f(s, x(s))ds$ has derivative w.r.t. t and $\bar{x}(t)$ satisfies the IVP.

$$\dot{\bar{x}} = f(t, \bar{x}(t)) \; ; \quad \bar{x}(t) = x_0$$

If one uses not the contraction principle, but a more complicated estimate for $K^n x = \underbrace{K(K(K \ldots K(x)))}_{n\text{-times}}$ and estimates

$\| x_{n+1} - x_n \|$ not by geometric series, but by a series for exponential function, One can observe that convergence follows even for $T_0 = \frac{\delta}{M}$ (or T if it is smaller)

Extensions Problem 2.10. (Weissinger)

Let C be a non-empty closed set in a Banach space and K - an operator mapping C to itself $K: C \to C$.

Let $\|K^n(x) - K^n(y)\| \leq \theta_n \|x-y\|$, $\theta_n \geq 0$ $x, y \in C$; $\sum_{n=0}^{\infty} \theta_n < \infty$.

$$K^n(x) = \underbrace{K(K(\dots K(x)))}_{n\text{- times}}.$$

Then K has a unique fixed point $\bar{x}$ in C and $\|K^n(x) - \bar{x}\| \leq \sum_{j=n}^{\infty} \theta_j \|K(x) - x\|$, $\forall x \in C$.

Bevis. Let $x_n = K^n(x)$

$$x_{n+1} = x_{n+1} - x_n + x_n - \dots - x_0 = \sum_{k=0}^{n+1} K^k(K(x)) - K^k(x)$$

$$\|K^k(K(x)) - K^k(x)\| \leq \theta_k \|K(x) - x\|$$

Consider series $\sum_{k=0}^{\infty} K^k(K(x)) - K^k(x) = x_{n+1}$

It converges by Weierstrass criterion because $\|K^k(K(x)) - K^k(x)\|$ is estimated by a convergent number series.

$$\sum_{k=0}^{\infty} \theta_k \cdot \|K(x) - x\| \quad \text{with } \theta_k > 0.$$

Therefore $\|K^n(x) - \bar{x}\| \leq \sum_{j=n}^{\infty} \theta_j \|K(x) - x\|$ for $\bar{x} = \lim_{n\to\infty} x_n$ and $\forall x \in C$. Uniqueness follow from the estimate $\|\bar{x} - \tilde{x}\| = \|K^n(\bar{x}) - K^n(\tilde{x})\| \leq \theta_n \|\bar{x} - \tilde{x}\|$ for n enough large, $\theta_n < 1$.

A modyfied proof of Picard-Lindelöf theorem, that is valid for the whole interval $[t_0, t_0+T_0]$, $T_0 = \frac{\delta}{M}$ without the limitation $T_0 < \frac{1}{L}$.

We repeat that $K: C \to C$ where C is a ball $\|x - x_0\|_C$ in $C([t_0, t_0+T_0], \mathbb{R}^n)$

It implies that all $x_n \in C$.

We will show by induction that

$$|x_{n+1}(t) - x_n(t)| \le \frac{M L^n (t-t_0)^{n+1}}{(n+1)!} \quad *(n+1)$$

$t_0 \le t \le t_0 + T_0$

$T_0 = \min\left(T, \frac{\delta}{M}\right)$.

Suppose that $*(n)$ is valid for $|x_n(t) - x_{n-1}(t)|$

$$x_{n+1}(t) - x_n(t) = \int_{t_0}^{t} f(s, x_n(s)) - f(s, x_{n-1}(s)) \, ds$$

$$|x_{n+1}(t) - x_n(t)| \le L \int_{t_0}^{t} |x_n(s) - x_{n-1}(s)| \, ds \le$$

$$\le L \int_{t_0}^{t} M \frac{L^{n-1} (s-t_0)^n}{n!} ds = M L^n \frac{(t-t_0)^{n+1}}{(n+1)!} \quad \blacksquare$$

$$x_n = \sum_{k=1}^{\infty} (x_k - x_{k-1}) + x_0$$

Element in series are estimated as $(*)$

Series $\sum_{k=0}^{\infty} M L^k \frac{(t-t_0)^{k+1}}{(k+1)!}$ converges to

$M \cdot \exp(L(t-t_0))$ $\Rightarrow$ therefore series and the sequense $x_n \xrightarrow[n\to\infty]{} x$

We observe that $\bar{x} = K\bar{x}$ again by going to the limit in

$$\begin{array}{ccccl} x_{n+1} & - & K(x_n) & = 0 & \quad \text{with } n \to \infty \\ \downarrow & & \downarrow & & \\ \bar{x} & - & K(\bar{x}) & = 0. & \end{array}$$

Uniqueness follows from the contraction property of K on any shorter interval around any time point $t \in [t_0, t_0 + T_0]$.

$\dot{\bar{x}} = f(t, \bar{x}(t))$ by the same argument as before.

<u>Comment</u> instead of using Weierstrass' criterion we could apply the criterion by Weissinger.
