MATH 5340 Fall '23

Lecture 13

Part II. Ordinary differential equations

Convergence for the implicit and explicit Euler methods and proof of the contraction mapping Theorem

Last time:

① Notion of solution: A solution to $\dot{x} = f(x,t)$ will be understood (even if $x(t)$ is only continuous) to be a solution of the integral equation:

$$x(t) = x(0) + \int_0^t f(x(s),s)\,ds$$
$$=: T(x)(t)$$

i.e. solutions are fixed points of the mapping T.

(2) The implicit/explicit Euler schemes: a sequence $\{x_n\}$ is generated for each given $x_0 \in \mathbb{R}^d$ and time step $h>0$

Explicit/Forward Euler

$$x_{n+1} = x_n + h f(x_n, t_n)$$

Implicit/Backward Euler

$$x_{n+1} = x_n + h f(x_{n+1}, t_{n+1})$$

Here, $t_n = nh$, $n = 0, 1, \ldots$ etc

To show how the sequences in (2) lead to solutions to (1), we will introduce a corresponding operator $T^{(h)}$ for each $h>0$.

$$T^{(h)} : C([0,\infty)) \longrightarrow C([0,\infty))$$

Given $x(t) \in C([0,\infty))$ the function $(T^{(h)}x)(t)$ is defined as follows:

$(k = 0, 1, 2, \ldots)$

For each k, denote $I_k = [t_k, t_{k+1}]$

(for Forward Euler)

(1.) $(T^{(h)}_{FE} x)(0) = x_0$ (given initial data)

$t \in I_k$

(2) $(T^{(h)}_{FE} x)(t) = (T^{(h)}_{FE} x)(t_k) + (t - t_k) f(x(t_k), t_k)$

Exercise: Check that (1) and (2) define the function $T^{(h)} x$ uniquely and that this function is continuous and piecewise linear

For Backward Euler

(1.) $(T^{(h)}_{BE} x)(0) = x_0$ (given initial data)

$t \in I_k$

(2) $(T^{(h)}_{BE} x)(t) = (T^{(h)}_{BE} x)(t_k) + (t - t_k) \underbrace{f(x(t_{k+1}), t_{k+1})}_{\text{(This is different from before)}}$

Observations

① Wether $x(t)$ is differentiable or not, the image $T^{(h)}x(t)$ will fail to be differentiable at the points t_n. In fact, $T^{(h)}x$ will always be a piecewise linear function.

② If $x(t)$ is piecewise linear with respect to the points $\{x_n\}$, then

$$T^{(h)}_{FE}(x) = x \quad (T^{(h)}_{BE})$$

if and only if the sequence

$$x_n := x(t_n)$$

is generated by the Forward Euler scheme (respectively, the Backward Euler scheme)

③ If L is the Lipschitz constant of $f(x,t)$ w.r.t. to x, then

The map $T^{(h)}$ will be a contraction mapping from $C([0,T])$ to $C([0,T])$ provided T and h are small enough.

<u>Proof</u> Take two continuous function $x_1(t)$ and $x_2(t)$, let's compare

$$T^{(h)}_{FE} x_1(t) - T^{(h)}_{FE} x_2(t)$$

First, suppose $t \in I_0$, then

$$T^{(h)}_{FE}(x_1)(t) = x_0 + t f(x_1(0), 0)$$
$$T^{(h)}_{FE}(x_2)(t) = x_0 + t f(x_2(0), 0)$$

$$\Rightarrow T^{(h)}_{FE}(x_1)(t) - T^{(h)}_{FE}(x_2)(t) = t\left(f(x_1(0), 0) - f(x_2(0), 0)\right)$$

$$\left| T^{(h)}_{FE}(x_1)(t) - T^{(h)}_{FE}(x_2)(t) \right| \leq |t| \left| f(x_1(0), 0) - f(x_2(0), 0) \right|$$
$$\leq h L \left| x_1(0) - x_2(0) \right|$$
$$\forall \; t \in I_0.$$

What about the other intervals? Let's bound the difference over the interval I_{n+1} in terms of the difference on the previous interval plus $x_1 - x_2$ over I_{n+1}.

$$[t_{n+1}, t_{n+2}]$$
$$\shortparallel$$

Recall (from the definition), for $t \in I_{n+1}$

$$T_{FE}^{(h)}(x_1)(t) = T_{FE}^{(h)}(x_1)(t_{n+1}) + (t - t_{n+1}) f(x_1(t_{n+1}), t_{n+1})$$
$$T_{FE}^{(h)}(x_2)(t) = T_{FE}^{(h)}(x_2)(t_{n+1}) + (t - t_{n+1}) f(x_2(t_{n+1}), t_{n+1})$$

$\Longrightarrow$

$$T_{FE}^{(h)}(x_1)(t) - T_{FE}^{(h)}(x_2)(t)$$

$$= T_{FE}^{(h)}(x_1)(t_{n+1}) - T_{FE}^{(h)}(x_2)(t_{n+1}) + (t - t_{n+1})\big(f(x_1(t_{n+1}), t_{n+1}) - f(x_2(t_{n+1}), t_{n+1})\big)$$

Then, by the triangle inequality,

$$\left| T_{FE}^{(h)}(x_1)(t) - T_{FE}^{(h)}(x_2)(t) \right| \qquad \forall \; t \in I_{n+1} = [t_{n+1}, t_{n+2}]$$
$$\forall \; n = 0, 1, 2, \dots$$

$$\leq \left| T_{FE}^{(h)}(x_1)(t_{n+1}) - T_{FE}^{(h)}(x_2)(t_{n+1}) \right| + hL \left| x_1(t_{n+1}) - x_2(t_{n+1}) \right|$$

We can iterate this inequality since $\forall n$ t_{n+1} is the right end of the interval I_n, this iteration produces the following inequality

$$\boxed{\begin{aligned} &\left| T_{FE}^{(h)}(x_1)(t) - T_{FE}^{(h)}(x_2)(t) \right| \qquad \forall \; t \in I_{n+1} = [t_{n+1}, t_{n+2}] \\ &\leq hL \sum_{k=0}^{n+1} \left| x_1(t_k) - x_2(t_k) \right| \end{aligned}}$$

Now observe that

$$hL \sum_{k=0}^{n+1} \left| x_1(t_k) - x_2(t_k) \right| \leq (n+1) hL \max_{0 \leq t \leq (n+1)h} \left| x_1(t) - x_2(t) \right|$$

If $t \in [0,T]$, then I only need the above inequalities up to the last $n \in \mathbb{N}$ s.t.

$$(n+1)h \leq T$$

since in that case $[0,T] \subset \bigcup_{k=0}^{n+1} I_k$, since then

$$\max_{[0,T]} \left| T^{(h)}(x_1)(t) - T^{(h)}(x_2)(t) \right|$$

$$\leq hL \sum_{k=0}^{n+1} |x_1(t_k) - x_2(t_k)|$$

$$\leq hL(n+1) \max_{[0,T]} |x_1(t) - x_2(t)|$$

Thus, $h(n+1) \leq T$, so we have

$$\| T^{(h)}(x_1) - T^{(h)}(x_2) \|_{L^\infty([0,T])} \leq TL \|x_1 - x_2\|_{L^\infty([0,T])}$$

and thus it is a contraction mapping if $TL < 1$.

(4) From item (3) and the contraction mapping theorem follows that if $TL < 1$, the $T^{(h)}_{FE}$ has a unique fixed point (for $T^{(h)}_{BE}$ you need to add that $hL < 1$). Let's call this fixed point $x^{(h)}(t)$, then by (1) and (2) this $x^{(h)}(t)$ is the same as the sequence created via FE/BE.
All that remains is investigating wether $x^{(h)}(t)$ converges to $x(t)$, the analytical solution to the IVP, as $h \to 0$.

The Local Truncation Error and convergence of the scheme

The local truncation error is a series of quantities associated to each IVP and each numerical scheme. It is defined as follows:

Let $x(t)$ be the solution to the IVP

$$\begin{cases} \dot{x}(t) = f(x(t), t) \\ x(0) = x_0 \end{cases}$$

Fix $h>0$, the local truncation error for Forward Euler is defined as

$$\varepsilon_n^{(h)} := \left(x(t_{n+1}) - x(t_n) - h f(x(t_n), t_n)\right)\frac{1}{h}$$

For $n = 0, 1, 2, \ldots$

Meanwhile, the local truncation error for Backward Euler is defined as

$$\varepsilon_n^{(h)} := \left(x(t_{n+1}) - x(t_n) - h f(x(t_{n+1}), t_{n+1})\right)\frac{1}{h}$$

Observation : Clearly, if $\varepsilon_n^{(h)} = 0 \ \forall n$, Then $x(t)$ would be a fixed point of $T_{FE}^{(h)}$ or $T_{BE}^{(h)}$, since the fixed points for these two operators are things we can compute numerically, the local truncation error ~~should~~ mean, point by point, how close our numerical solution $x^{(h)}(t)$ is to the analytical solution $x(t)$.

Theorem: Let $T^{(h)}$ denote either $T^{(h)}_{FE}$ or $T^{(h)}_{BE}$, and let $\mathcal{E}^{(h)}_n$ denote the local truncation error. Then, for any T s.t. $TL<1$ we have

analytical solution.

numerical solution

$$\| x - x^{(h)} \|_{L^\infty(0,T)} \leq \frac{1}{1-TL} T \max_{0 \leq n \leq \lfloor \frac{T}{h} \rfloor} | \mathcal{E}^{(h)}_n |$$

where $x(t)$ is the solution to the IVP and $x^{(h)}$ is the fixed point of $T^{(h)}$.