MA3227 Numerical Analysis II

Lecture 14: Banach's Fixed Point Theorem

Simon Etter

2019/2020

Banach's Fixed Point Theorem

Introduction

Consider the root-finding problem $f(x^\star) = 0$ where $f : \mathbb{R}^N \to \mathbb{R}^N$.
We have seen:

- For $N = 1$, we can use bracketing intervals to show that roots exist.
- For $N > 1$, bracketing intervals no longer work.

There are (at least) two common approaches for showing that roots exist for arbitrary $f(x)$.

Minimisation approach

Use the equivalence

$$f(x^\star) = 0 \quad \iff \quad x^\star = \arg\min_x \|f(x)\|_2^2$$

and show that $g(x) = \|f(x)\|_2^2$ has a unique minimiser. The latter part can be tackled using properties like convexity of $g(x)$.

Banach's Fixed Point Theorem

Introduction (continued)

Fixed-point approach

Use the equivalence

$$f(x^\star) = 0 \quad \Longleftrightarrow \quad x^\star = x^\star + f(x^\star)$$

and show that $g(x) = x + f(x)$ has a unique fixed point, i.e. a point $x^\star \in \mathbb{R}^N$ such that $x^\star = g(x^\star)$.

Results of the form "If $g(x)$ satisfies [list of properties], then $g(x)$ has a (unique) fixed point" are called fixed-point theorems.

The aim of this lecture is to introduce a particular fixed point theorem.

Remark

For both the minimisation and the fixed-point approach, there are many other and often better ways for constructing $g(x)$. For example, Newton's method corresponds to the fixed-point equation

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

Banach's Fixed Point Theorem

Def: Contraction

A function $g : D \to \mathbb{R}^n$ with $D \subset \mathbb{R}^n$ is called a contraction if there exists a $q < 1$ such that for all $x_1, x_2 \in D$ we have

$$\left\|g(x_1) - g(x_2)\right\| \leq q \left\|x_1 - x_2\right\|.$$

Banach fixed-point theorem

Assume $D \subset \mathbb{R}^N$ is closed and $g : D \to D$ is a contraction. Then, g has a unique fixed point $x = g(x) \in D$, and this fixed point is the limit of the sequence $x_{k+1} = g(x_k)$ for any initial guess x_0.

Proof (not examinable). Uniqueness: Assume there are two fixed points $x_1, x_2 \in D$. Then,

$$\|x_1 - x_2\| = \|g(x_1) - g(x_2)\| \leq q \|x_1 - x_2\|.$$

Since $q < 1$, this bound can only be satisfied if $\|x_1 - x_2\| = 0$.

Banach's Fixed Point Theorem

Proof (not examinable, continued). Existence: We have

$$\|x_{k+1} - x_k\| = \|g(x_k) - g(x_{k-1})\| \leq q\,\|x_k - x_{k-1}\|$$

and thus by induction

$$\|x_{k+1} - x_k\| \leq q^k\,\|x_1 - x_0\|.$$

This result can be used to show that x_k is a Cauchy sequence, and since $D \subset \mathbb{R}^N$ is complete this implies that x_k converges to some limit $x \in D$. This limit is a fixed point, $x = g(x)$, since $g(x)$ is continuous.

Showing that x_k is indeed a Cauchy sequence requires a bit of work.

I omit the details since they are irrelevant for our purposes.

Banach's Fixed Point Theorem

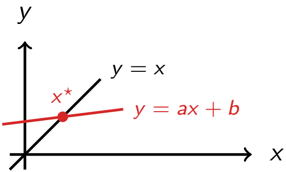
Example

Assume $g : \mathbb{R} \to \mathbb{R}, x \mapsto ax + b$ with $|a| < 1$. We have

$$|g(x_1) - g(x_2)| = |a|\,|x_1 - x_2|;$$

thus $g(x)$ is a contraction and has a unique fixed point $x^\star$.

This finding is easily confirmed by drawing the graphs of $y = x$ and $y = ax + b$ and finding their intersection $x^\star$.



Discussion

Showing that $g(x)$ is a contraction is easy in this example because $g(x) = ax + b$ is a linear function. For more general, nonlinear functions, we can use the result on the following slide.

Banach's Fixed Point Theorem

Thm: Lipschitz constants and derivatives

Assume $D \subset \mathbb{R}^N$ is convex and $g : D \to D$ has a bounded derivative. Then,

$$\|g(x_2) - g(x_1)\| \leq L \|x_2 - x_1\| \qquad \text{where} \qquad L = \sup_{x \in D} \|\nabla g(x)\|$$

Proof. By the chain rule, we have that

$$\frac{d}{dt}\Big(g\left(x_1 + t\left(x_2 - x_1\right)\right)\Big) = \nabla g\left(x_1 + t\left(x_2 - x_1\right)\right)\left(x_2 - x_1\right)$$

and hence we conclude using the fundamental theorem of calculus that

$$\begin{aligned}
\|g(x_2) - g(x_1)\| &= \left\| \int_0^1 \nabla g\left(x_1 + t\left(x_2 - x_1\right)\right)\left(x_2 - x_1\right) dt \right\| \\
&\leq \int_0^1 \left\|\nabla g\left(x_1 + t\left(x_2 - x_1\right)\right)\right\| \left\|x_2 - x_1\right\| dt \\
&\leq \left(\sup_{x \in D} \|\nabla g(x)\|\right) \left\|x_2 - x_1\right\|.
\end{aligned}$$

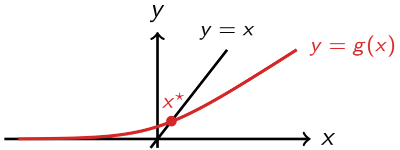
Banach's Fixed Point Theorem

Example

We have

$$g(x) = \log\bigl(1 + \exp(\tfrac{x}{2})\bigr) \quad \Longrightarrow \quad g'(x) = \frac{1}{2}\,\frac{1}{1 + \exp(-\frac{x}{2})} \in \bigl(0, \tfrac{1}{2}\bigr).$$

Thus, $g(x)$ is a contraction and has a unique fixed point $x^\star$.



Banach's Fixed Point Theorem

Corollary: Local convergence of Newton's method

We have seen that Newton's method $x_{k+1} = x_k - \nabla f(x_k)^{-1} f(x_k)$ may fail to converge to a root $x^\star$ if the initial guess x_0 is too far from $x^\star$.

Using Banach's fixed-point theorem, we can show the following converse:

Assume $f : \mathbb{R}^N \to \mathbb{R}^N$ is a twice continuously differentiable function with root $x^\star \in \mathbb{R}^N$ such that $\nabla f(x^\star)$ is invertible. Then, there exists $\delta > 0$ such that $\|x_0 - x^\star\| < \delta$ guarantees that $x_k \to x^\star$.

Proof. We only consider the case $N = 1$ for simplicity. The proof for $N > 1$ is analogous but requires special notation since $\nabla^2 f \in \mathbb{R}^{N \times N \times N}$.

We observe that Newton's method is the fixed-point iteration

$$x_{k+1} = g(x_k) \qquad \text{where} \qquad g(x) = x - \frac{f(x)}{f'(x)}.$$

In order to apply Banach's fixed-point theorem, we thus need to determine an interval D such that $x \in D \implies g(x) \in D$ and $g(x)$ is a contraction on D.

Banach's Fixed Point Theorem

Proof (continued).

We compute

$$g'(x) = 1 - \frac{f'(x)}{f'(x)} + \frac{f(x)\, f''(x)}{(f'(x))^2} = \frac{f(x)\, f''(x)}{(f'(x))^2}.$$

Since $g'(x^\star) \propto f(x^\star) = 0$ and $g'(x)$ is continuous, we conclude that for every $q < 1$ there is a $\delta > 0$ such that $|x - x^\star| \le \delta \implies |g'(x)| \le q$ and hence $g(x)$ is a contraction on $D = x^\star + [-\delta, \delta]$.

It remains to show that $x \in D \implies g(x) \in D$, but this follows immediately from the fixed-point and contraction properties: we have

$$|g(x) - x^\star| = |g(x) - g(x^\star)| \le q\, |x - x^\star|$$

and thus

$$\Big(x \in D \iff |x - x^\star| \le \delta\Big) \quad \implies \quad \Big(|g(x) - x^\star| \le \delta \iff g(x) \in D\Big).$$

Banach's Fixed Point Theorem

Remark: Definition of contractions

It is tempting to abbreviate the contraction condition

$$\|g(x_1) - g(x_2)\| \le q\,\|x_1 - x_2\| \qquad \text{for some } q < 1$$

as

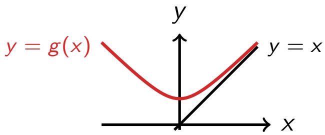
$$\|g(x_1) - g(x_2)\| < \|x_1 - x_2\|.$$

However, these two statements are not equivalent, and the second one is not strong enough for the Banach fixed-point theorem.

As a counterexample, consider $g(x) = \sqrt{1 + x^2}$. We have

$$|g'(x)| = \left|\frac{x}{\sqrt{1+x^2}}\right| < 1 \qquad \Longrightarrow \qquad |g(x_1) - g(x_2)| < |x_1 - x_2|,$$

but $g(x)$ does not have a fixed point.



Banach's Fixed Point Theorem

Summary

- Banach's fixed-point theorem:

$$\text{g(x) is a contraction} \quad \Longrightarrow \quad g(x) \text{ has a unique fixed point.}$$

- Lipschitz constants and derivatives:

$$\|g(x_2) - g(x_1)\| \leq L \|x_2 - x_1\| \qquad \text{where} \qquad L = \sup_{x \in D} \|\nabla g(x)\|.$$