Numerical Methods and Optimization

Chapter 3: Systems of linear Equations - Iterative Methods

10.10.2022

Content

1. The Banach Fixed Point Theorem and its applications

Theorem (Banach fixed point theorem (BFT))

Let $\Phi\colon \mathcal{K} \to \mathcal{K}$ a contracting self-map (a contraction) of a closed subset $\mathcal{K} \subset \mathbb{R}^n$ with contraction factor q, i.e.

$$\|\Phi(x) - \Phi(y)\| \leq q\|x - y\| \tag{1}$$

for some $q < 1$ and all $x, y \in \mathcal{K}$. Then the fixed point equation

$$x = \Phi(x) \tag{2}$$

has exactly one solution $\hat{x} \in \mathcal{K}$ ($\hat{x}$ is called the fixed point of Φ), and the fixed point iteration

$$x^{(k+1)} := \Phi(x^{(k)}), \quad k = 0, 1, 2, \dots \tag{3}$$

converges for each initial vector $x^{(0)} \in \mathcal{K}$ to $\hat{x}$ when $k \to \infty$. Furthermore, for $k \geq 1$, we have

$$\|x^{(k)} - \hat{x}\| \leq \frac{q^k}{1-q}\|x^{(1)} - x^{(0)}\| \qquad \text{(a priori estimate)}$$

and

$$\|x^{(k)} - \hat{x}\| \leq \frac{q}{1-q}\|x^{(k)} - x^{(k-1)}\|. \qquad \text{(a posteriori estimate)}$$

PROOF: Choose an arbitrary $x^{(0)} \in \mathcal{K}$ an let $(x^{(k)})$ be defined as in (3).

Step 1: $x^{(k)}$ is a Cauchy sequence:

Due to the contraction property (1), for any $k \in \mathbb{N}$ it holds

$$\|x^{(k+1)} - x^{(k)}\| = \|\Phi(x^{(k)}) - \Phi(x^{(k-1)})\| \leq q\|x^{(k)} - x^{(k-1)}\|.$$

Therefore, we obtain by induction (over $p \in \mathbb{N}_0$) that for all $k \geq 1$ and $p \geq 0$

$$\|x^{(k+p)} - x^{(k+p-1)}\| \leq q^p\|x^{(k)} - x^{(k-1)}\|. \tag{4}$$

From this one can deduce (using the triangle inequality for the telescoping sum)

$$\begin{aligned}
\|x^{(k+p)} - x^{(k)}\| &\leq \|x^{(k+p)} - x^{(k+p-1)}\| + \ldots + \|x^{(k+1)} - x^{(k)}\| \\
&\overset{(4)}{\leq} (q^p + q^{p-1} + \ldots + q)\|x^{(k)} - x^{(k-1)}\| \\
&\leq \frac{q}{1-q}\|x^{(k)} - x^{(k-1)}\| \\
&\leq \frac{q^k}{1-q}\|x^{(1)} - x^{(0)}\|.
\end{aligned}$$

$$\begin{aligned}\|x^{(k+p)} - x^{(k)}\| &\leq \|x^{(k+p)} - x^{(k+p-1)}\| + \ldots + \|x^{(k+1)} - x^{(k)}\| \\ &\overset{(4)}{\leq} (q^p + q^{p-1} + \ldots + q)\|x^{(k)} - x^{(k-1)}\| \\ &\leq \frac{q}{1-q}\|x^{(k)} - x^{(k-1)}\| \\ &\leq \frac{q^k}{1-q}\|x^{(1)} - x^{(0)}\|.\end{aligned}$$

Since $0 < q < 1$ we can conclude that for all k large enough with $m := k + p$ it is

$$\|x^{(m)} - x^{(k)}\| \leq \frac{q^k}{1-q}\|x^{(1)} - x^{(0)}\| < \varepsilon,$$

meaning that $(x^{(k)})$ is a Cauchy sequence. Letting $\hat{x}$ denote the unique limit (it may depend on $x^{(0)}$), the closedness of $\mathcal{K}$ then implies that $\hat{x} \in \mathcal{K}$.

Step 2: $\hat{x}$ is a unique fixed point of Φ:

First, Φ is lipschitz continuous (and therefore continous) on $\mathcal{K}$. Hence we can pass to the limit in the iteration equation $x^{(k+1)} = \Phi(x^{(k)})$ to show that $\hat{x}$ is a fixed point of Φ:

$$\hat{x} = \lim_{k\to\infty} x^{(k+1)} = \lim_{k\to\infty} \Phi(x^{(k)}) = \Phi(\lim_{k\to\infty} x^{(k)}) = \Phi(\hat{x}).$$

Uniqueness now follows again with inequality (1): Suppose there are two fixed points $\hat{x}$ and $\tilde{x}$. Then

$$\|\hat{x} - \tilde{x}\| = \|\Phi(\hat{x}) - \Phi(\tilde{x})\| \leq q\|\hat{x} - \tilde{x}\|$$

follows and because of $q < 1$ this can only hold if $\hat{x} = \tilde{x}$.

$$\begin{aligned}
\|x^{(k+p)} - x^{(k)}\| &\leq \|x^{(k+p)} - x^{(k+p-1)}\| + \ldots + \|x^{(k+1)} - x^{(k)}\| \\
&\overset{(4)}{\leq} (q^p + q^{p-1} + \ldots + q)\|x^{(k)} - x^{(k-1)}\| \\
&\leq \frac{q}{1-q}\|x^{(k)} - x^{(k-1)}\| \qquad (5) \\
&\leq \frac{q^k}{1-q}\|x^{(1)} - x^{(0)}\|. \qquad (6)
\end{aligned}$$

Step 3: Proof of the a priori/a posteriori estimates:

Letting $p \to \infty$ in (6) and (5) the a priori and a posteriori estimates follow. ■

Examples:

1. The Newton method (in 1d) for determining a zero of a differentiable function $f\colon \mathbb{R} \to \mathbb{R}$ is given by the iteration formula

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}$$

with given initial guess $x^{(0)} \in \mathbb{R}$.
For the function $f(x) = x^2 - 2$, this results in the iteration formula

$$x_{k+1} = \frac{1}{2}\Big(x_k + \frac{2}{x_k}\Big).$$

This defines the fixed point iteration

$$x_{k+1} = \Phi(x_k)$$

for the approximation of $\sqrt{2}$, where

$$\Phi(x) \coloneqq \frac{1}{2}\Big(x + \frac{2}{x}\Big).$$

For $I \coloneqq [1;3]$ we have that $\Phi\colon I \to I$ is a contraction with contraction factor $q = \frac{1}{2}$ and thus the BFT is applicable. That is, there is exactly one fixed point and Newton's iteration converges to this.

2) alternative fixed point iterations to determine a zero of a function: Consider

$$f(x) := x^4 - x + \frac{1}{4}$$

in the intervall $D = [0; 1/2]$, rewrite problem in fixed point form:

$$x = \Phi(x),$$

where $\Phi(x) = x^4 + \frac{1}{4}$ describes a contraction on D so that the BFT is applicable:

- Φ is contracting: For $x, y \in D$ we have

$$|\Phi(x) - \Phi(y)| \leq \max_{\xi \in D} |\Phi'(\xi)||x - y|,$$

  and therefore Φ is contracting with contraction factor $\frac{1}{2}$.
- Φ is a self-map: For $x \in D$ it follows that

$$0 \leq \Phi(x) \leq \frac{1}{16} + \frac{1}{4} = \frac{5}{16} \leq \frac{1}{2}.$$

  Hence, using the iteration

$$x_{k+1} = \Phi(x_k),$$

  one can find the unique zero of f in D for all initial data $x^{(0)} \in D$.

3. The BFT also reveals an interesting property of maps/city plans:
   - X: urban area of Rosenheim
   - put the city map on the floor (not on the boundary of the city area!).
   - obtain mapping $\Phi\colon X \to X$ as follows:
   - each $x \in X$ is uniquely assigned to a point y on the city map.
   - pierce with a needle through this point
   - needle tip meets point $\Phi(x)$ in the urban area!
   - Φ is a self map: ✓
   - Φ is contracting (i.e. Lischitz with $L < 1$):
     Since the map scale is $1 : M > 1$ it follows that

$$|\Phi(x) - \Phi(y)| = \frac{1}{M}|x - y|.$$

   - BFT implies: $\exists! x \in X$ with $\Phi(x) = x$. //.

- For given $A \in \mathrm{Gl}(n)$ and $b \in \mathbb{R}^n$, the BFT can be used to construct convergent iteration methods for the numerical solution of SLE $Ax = b$.
- choose decomposition $A = M - N$, where $M \in \mathrm{Gl}(n)$.
- transform $Ax = b$ into a fixed point form:

$$\begin{aligned} Ax = b \quad &\Leftrightarrow \quad Mx = Nx + b \\ &\Leftrightarrow \quad x = M^{-1}Nx + M^{-1}b \end{aligned}$$

  or

$$x = Tx + c,$$

  where $T = M^{-1}N$ and $c = M^{-1}b$.
- $\rightsquigarrow$ fixed point operator is an affine function

$$\Phi(x) := Tx + c.$$

- M should be simple concerning calc. of M^{-1} and that the fixed point iteration converges
- the following theorems provide conditions for the convergence

Proposition

Let $\|\cdot\|_*$ be a Norm in $\mathbb{R}^{n\times n}$ that is compatible with a vector norm $\|\cdot\|$. Suppose $A = M - N$ with $A, M \in \mathrm{Gl}(n)$. Further let $T = M^{-1}N$ with

$$\|T\|_* = \|M^{-1}N\|_* < 1.$$

Then the fixed point iteration with fixed point operator $\Phi(x) = Tx + c$ and $c = M^{-1}b$ converges according to the iteration

$$x^{(k+1)} = \Phi(x^{(k)}) = Tx^{(k)} + c \tag{7}$$

for each initial guess $x^{(0)}$ to the solution $\hat{x}$ of $Ax = b$, i.e. $y = A^{-1}b$.

PROOF:

- For $\Phi(x) = Tx + c$ we have $\Phi\colon \mathcal{K} \to \mathcal{K}$, where $\mathcal{K}\colon = \mathbb{R}^n$ is complete.Thus Φ is a self map ✓
- Φ is a contraction: By compatibility of the norms:

  $$\|\Phi(x) - \Phi(y)\| = \|T(x-y)\| \leq \|T\|_*\|x-y\|$$

  and because

  $$\|T\|_* = \|M^{-1}N\|_* < 1.$$

- The BFT implies: $(x^{(k)})$ defined by $x^{(k+1)} := \Phi(x^{(k)})$ converges for all $x^{(0)} \in \mathcal{K} = \mathbb{R}^n$ to the unique fixed point $\hat{x}$ with

$$\hat{x} = T\hat{x} + c.$$

- We have shown that there exists a unique $x = \hat{x}$ s.t. $x = Tx + c \quad \Leftrightarrow \quad Ax = b.$ ∎

There is another convergence criterion for iterative methods of the form discussed:

Proposition

Let A be regular and $T = M^{-1}N$ be as above. Then the method $x^{(k+1)} = Tx^{(k)} + c$ converges for every $x^{(0)}$ to $\hat{x} = A^{-1}b$ **if and only if** *for the spectral radius of T it holds*

$$\rho(T) < 1.$$

Furthermore,

$$\max_{x^{(0)}} \limsup_{k\to\infty} \|x^{(k)} - \hat{x}\|^{\frac{1}{k}} = \rho(T).$$

In the case of convergence, i.e. $\rho(T) < 1$, for every $\varepsilon > 0$ and every initial guess $x^{(0)}$ there exists an index $K \in \mathbb{N}$, such that

$$\|x^{(k)} - \hat{x}\| \le \big(\rho(T) + \varepsilon\big)^k \quad \forall k \ge K.$$

That is, the asymptotic convergence speed is $\rho(T)^k$. ■

Example (shows that speed of convergence really only applies asymptotically): Consider the invertible matrix

$$A = \begin{pmatrix} 1 & -1 & 0 & \cdots & \cdots & 0 \\ 0 & \ddots & \ddots & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & \ddots & \vdots \\ \vdots & & & \ddots & \ddots & 0 \\ \vdots & & & & \ddots & -1 \\ 0 & \cdots & \cdots & \cdots & 0 & 1 \end{pmatrix}.$$

Letting $M \coloneqq E$, we obtain from $A = M - N$ that

$$T = M^{-1}N = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ \vdots & & & \ddots & 1 \\ 0 & \cdots & \cdots & \cdots & 0 \end{pmatrix}.$$

- T: "shift matrix" because
  $x = (x_1, \ldots, x_n)^T \quad \Rightarrow \quad Tx = (x_2, \ldots, x_n, 0)^T$
- Clearly: $\sigma(T) = \{0\} \quad \Rightarrow \quad \rho(T) = 0$
- one would expect a very fast convergence: $\|x^{(k)} - \hat{x}\| \leq \big(\underbrace{\rho(T)}_{=0} + \varepsilon\big)^k$
- But:

- Let $x^{(0)} := \hat{x} + e_n$, where $\hat{x}$ is unique solution of

$$Ax = b \;\Leftrightarrow\; \Phi(x) = x, \text{ where } \Phi(x) = Tx + c.$$

- Then

$$\begin{aligned}
\|x^{(k)} - \hat{x}\| &= \|\Phi(x^{(k-1)}) - \Phi(\hat{x})\| \\
&= \|T(x^{(k-1)} - \hat{x})\| \\
&= \|T^k(x^{(0)} - \hat{x})\| \\
&= \|T^k e_n\| \\
&= \|e_{n-k}\| \\
&= 1 \\
&= \|x^{(0)} - \hat{x}\|.
\end{aligned}$$

- Therefore: in the first $n-1$ steps: no error reduction at all!
- $\rightarrow$ convergence rate in

$$\|x^{(k)} - \hat{x}\| \leq \big(\rho(T) + \varepsilon\big)^k$$

is asymptotic in nature