FIXED POINT ITERATIONS

EIRIK HOEL HØISETH

1 Introduction

Let $D \subset \mathbb{R}^n$ and $\mathbf{F}\colon D \mapsto \mathbb{R}^n$ be a continuous vector valued mapping in n variables[1]. Our goal is the solution of an equation

$$\mathbf{F}(\mathbf{x}) = \mathbf{0}. \tag{1}$$

Assume the root finding problem (1) has the equivalent form

$$\mathbf{G}(\mathbf{x}) = \mathbf{x}, \tag{2}$$

for $\mathbf{G}\colon D \mapsto \mathbb{R}^n$, i.e. $\mathbf{x}^* \in D$ is a solution of (1) if and only if it is a solution of (2). We say that $\mathbf{x}^*$ is a *fixed point* of $\mathbf{G}$. A simple example of this is

$$\mathbf{G}(\mathbf{x}) = \mathbf{x} + \mathbf{F}(\mathbf{x}).$$

A straightforward idea for the solution of fixed point equations (2) is that of *fixed point iterations*. Starting with some point $\mathbf{x}^{(0)} \in \mathbb{R}^n$ (preferably an approximation to a solution of (2)), we define the sequence $\{\mathbf{x}^{(k)}\}_{k=0}^{\infty} \subset \mathbb{R}^n$ by the recursive relation

$$\mathbf{x}^{(k+1)} = \mathbf{G}(\mathbf{x}^{(k)}). \tag{3}$$

This idea is based on the observation that this sequence will become stationary after some index k if and only if $\mathbf{x}^{(k)}$ is a fixed point of $\mathbf{G}$. Moreover it seems reasonable that $\mathbf{x}^{(k)}$ is close to a fixed point if $\mathbf{x}^{(k)} \approx \mathbf{x}^{(k+1)}$. However we still need to develop sufficient conditions on $\mathbf{G}$ for when this function has a fixed point, and when the iteration (3) actually works.

2 Main results

The first result gives a sufficient condition for $\mathbf{G}$ to have a fixed point. We first define a convex domain/set

[1]Notation: For two sets A, B we write $A \subset B$ if and and only if $x \in A$ implies $x \in B$. So $A \subset A$ is true. Some people use the notation $\subseteq$ instead.

Definition 1. *A domain D is* convex *if for any $\mathbf{x}, \mathbf{y} \in D$,*

$$(1-t)\mathbf{x} + t\mathbf{y} \in D \quad \text{for all} \quad t \in [0,1].$$

Theorem 2 (Brouwer's fixed-point theorem)**.** *Let $D \subset \mathbb{R}^n$ be closed, bounded and convex. If $\mathbf{G}\colon D \mapsto D$ is continuous it has a fixed point.*

The general proof of this is well outside the scope of this course. Note that for the special case $n = 1$, i.e. scalar functions, this theorem states that if $G\colon [a,b] \mapsto [a,b]$ where $a \leq b \in \mathbb{R}$ then G has a fixed point if it is continuous.

The next theorem is very important. It gives sufficient conditions for both the existence and uniqueness of the fixed point, as well as for the fixed point iterations to converge towards this fixed point. First, recall the definition of a Cauchy sequence

Definition 3. $\{\mathbf{x}^{(k)}\}_{k=0}^{\infty}$ *in some subset of $\mathbb{R}^n$ is* Cauchy *in some norm $\|\cdot\|$ if elements of the sequence become arbitrarily close together as the sequence progresses. To be specific, for any $\epsilon > 0$ there exists a positive integer n_0 such that for all $q, r \in \mathbb{N}$ such that $q, r > n_0$*

$$\|\mathbf{x}^{(q)} - \mathbf{x}^{(r)}\| < \epsilon$$

Theorem 4 (Banach's fixed-point theorem)**.** *Assume that $D \subset \mathbb{R}^n$ is closed, and that $\mathbf{G}\colon D \mapsto D$ is a contraction. That is, there exists $0 \leq C < 1$ such that*

$$\|\mathbf{G}(\mathbf{x}) - \mathbf{G}(\mathbf{y})\| \leq C\|\mathbf{x} - \mathbf{y}\|,$$

for all $\mathbf{x}, \mathbf{y} \in D$, where $\|\cdot\|$ is any norm on $\mathbb{R}^n$. Then the function $\mathbf{G}$ has a unique fixed point $\mathbf{x}^* \in D$.

Now let $\mathbf{x}^{(0)} \in D$ be arbitrary and define the sequence $\{\mathbf{x}^{(k)}\}_{k=0}^{\infty} \subset D$ by $\mathbf{x}^{(k+1)} := \mathbf{G}(\mathbf{x}^{(k)})$. Then we have the estimates

$$\begin{aligned}
\|\mathbf{x}^{(k+1)} - \mathbf{x}^*\| &\leq C\|\mathbf{x}^{(k)} - \mathbf{x}^*\|, \\
\|\mathbf{x}^{(k)} - \mathbf{x}^*\| &\leq \frac{C^k}{1-C}\|\mathbf{x}^{(1)} - \mathbf{x}^{(0)}\|, \\
\|\mathbf{x}^{(k)} - \mathbf{x}^*\| &\leq \frac{C}{1-C}\|\mathbf{x}^{(k)} - \mathbf{x}^{(k-1)}\|.
\end{aligned}$$

In particular the sequence $\{\mathbf{x}^{(k)}\}_{k=0}^{\infty}$ converges to $\mathbf{x}^*$.

Proof. Assume that $\mathbf{x}^*$ and $\mathbf{y}^*$ are two fixed points of $\mathbf{G}$ in D. Then

$$\|\mathbf{x}^* - \mathbf{y}^*\| = \|\mathbf{G}(\mathbf{x}^*) - \mathbf{G}(\mathbf{y}^*)\| \leq C\|\mathbf{x}^* - \mathbf{y}^*\|.$$

From $0 \leq C < 1$, it follows that $\|\mathbf{x}^* - \mathbf{y}^*\| = 0$ and thus $\mathbf{x}^* = \mathbf{y}^*$. Thus the fixed-point (if it exists is unique).

Now let let $\mathbf{x}^{(0)} \in D$ be arbitrary and define inductively $\mathbf{x}^{(k+1)} := \mathbf{G}(\mathbf{x}^{(k)})$. Then we have for all $k, m \in \mathbb{N}$ the estimate

$$\begin{aligned}
\|\mathbf{x}^{(k+m)} - \mathbf{x}^{(k)}\| &\leq \sum_{j=1}^{m} \|\mathbf{x}^{(k+j)} - \mathbf{x}^{(k+j-1)}\| \\
&\leq \sum_{j=1}^{m} C^{j-1}\|\mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}\| \\
&\leq \sum_{j=1}^{\infty} C^{j-1}\|\mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}\| \qquad (4)\\
&= \frac{1}{1-C}\|\mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}\| \\
&\leq \frac{C}{1-C}\|\mathbf{x}^{(k)} - \mathbf{x}^{(k-1)}\| \\
&\leq \frac{C^k}{1-C}\|\mathbf{x}^{(1)} - \mathbf{x}^{(0)}\|.
\end{aligned}$$

It is clear from (4) that $\{\mathbf{x}^{(k)}\}_{k=0}^{\infty}$ is a Cauchy sequence by letting $q = k + m$, and $r = k$ in the definition above. Thus this sequence has a limit $\mathbf{x}^* \in \mathbb{R}^n$. In fact $\mathbf{x}^* \in D$ since D is closed, which implies that any Cauchy sequence in D has a limit in D. Now note that the function $\mathbf{G}$ is by definition Lipschitz-continuous with Lipschitz-constant C, which again shows that $\mathbf{G}$ is continuous. Thus

$$\mathbf{x}^* = \lim_{k\to\infty} \mathbf{x}^{(k)} = \lim_{k\to\infty} \mathbf{x}^{(k+1)} = \lim_{k\to\infty} \mathbf{G}(\mathbf{x}^{(k)}) = \mathbf{G}\left(\lim_{k\to\infty} \mathbf{x}^{(k)}\right) = \mathbf{G}(\mathbf{x}^*).$$

This shows that $\mathbf{x}^*$ is a fixed point of $\mathbf{G}$. The different estimates claimed in the theorem now follow easily from the fact that $\mathbf{G}$ is a contraction and from (4) by letting $m \to \infty$. (The norm is a continuous mapping, so the left hand side of this equation tends to $\|\mathbf{x}^* - \mathbf{x}^{(k)}\|$). □

The next result gives a local condition on $\mathbf{G}$ for there to exist a set where Banach's fixed point theorem is applicable. First, recall that

Definition 5. *A* neighbourhood *of a point* $\mathbf{x}$ *is a set* D *that contains an open set containing* $\mathbf{x}$*.*

Theorem 6. *Assume that* $\mathbf{G}\colon \mathbb{R}^n \mapsto \mathbb{R}^n$ *is continuously differentiable in a neighbourhood of a fixed point* $\mathbf{x}^*$ *of* $\mathbf{G}$*, and that there exists a norm* $\|\cdot\|$ *on* $\mathbb{R}^n$ *with subordinate matrix norm* $\|\cdot\|$ *on* $\mathbb{R}^{n\times n}$ *such that*

$$\|\mathbf{J_G}(\mathbf{x}^*)\| < 1$$

where $\mathbf{J_G}$ *is the Jacobian of* $\mathbf{G}$*. Then there exists a closed neighbourhood* D *of* $\mathbf{x}^*$ *such that* $\mathbf{G}$ *is a contraction on* D*. In particular, the fixed point iteration* $\mathbf{x}^{(k+1)} = \mathbf{G}(\mathbf{x}^{(k)})$ *converges for every* $\mathbf{x}^{(0)} \in D$ *to* $\mathbf{x}^*$*.*

Proof. Because of the continuous differentiability of $\mathbf{G}$ there exists $\rho > 0$ and a constant $0 < C < 1$ such that

$$\|\mathbf{J_G}(\mathbf{x})\| \le C \quad \text{for all} \quad \mathbf{x} \in B_\rho(\mathbf{x}^*) := \{\mathbf{y} \in \mathbb{R}^n \colon \|\mathbf{y} - \mathbf{x}^*\| \le \rho\}$$

Let now $\mathbf{x}, \mathbf{y} \in B_\rho(\mathbf{x}^*)$. Then

$$\mathbf{G}(\mathbf{y}) - \mathbf{G}(\mathbf{x}) = \int_0^1 \mathbf{J_G}(\mathbf{x} + t(\mathbf{y} - \mathbf{x}))(\mathbf{y} - \mathbf{x})dt.$$

Thus

$$\begin{aligned}
\|\mathbf{G}(\mathbf{y}) - \mathbf{G}(\mathbf{x})\| &\le \int_0^1 \|\mathbf{J_G}(\mathbf{x} + t(\mathbf{y} - \mathbf{x}))(\mathbf{y} - \mathbf{x})\|\, dt \\
&\le \int_0^1 \|\mathbf{J_G}(\mathbf{x} + t(\mathbf{y} - \mathbf{x}))\|\|\mathbf{y} - \mathbf{x}\|\, dt \\
&= \int_0^1 \|\mathbf{J_G}(\mathbf{x} + t(\mathbf{y} - \mathbf{x}))\|\, dt\, \|\mathbf{y} - \mathbf{x}\| \qquad (5) \\
&\le C\|\mathbf{y} - \mathbf{x}\|
\end{aligned}$$

In particular we obtain with $\mathbf{y} = \mathbf{x}^*$ the inequality

$$\|\mathbf{x}^* - \mathbf{G}(\mathbf{x})\| = \|\mathbf{G}(\mathbf{x}^*) - \mathbf{G}(\mathbf{x})\| \le C\|\mathbf{x}^* - \mathbf{x}\| \le C\rho < \rho.$$

Thus $\mathbf{x} \in B_\rho(\mathbf{x}^*)$ implies $\mathbf{G}(\mathbf{x}) \in B_\rho(\mathbf{x}^*)$. This shows that $\mathbf{G}$ is in fact a contraction on $B_\rho(\mathbf{x}^*)$. The final claim of the theorem is a simple application of Banach's fixed point theorem. □

We now make some remarks regarding the spectral radius of a matrix $\mathbf{A} \in \mathbb{R}^{n\times n}$. The first is that every subordinate matrix norm on $\mathbb{R}^{n\times n}$ satisfies the inequality

$$\rho(\mathbf{A}) \leq \|\mathbf{A}\|.$$

This is a simple observation if the largest eigenvalue, i.e. λ such that $|\lambda| = \rho(\mathbf{A})$, is real. Then letting $\mathbf{x}$ be a corresponding normalized eigenvector such that $\|\mathbf{x}\| = 1$

$$\rho(\mathbf{A}) = |\lambda| = |\lambda|\|\mathbf{x}\| = \|\lambda\mathbf{x}\| = \|\mathbf{A}\mathbf{x}\| \leq \|\mathbf{A}\|\|\mathbf{x}\|.$$

Conversely you can show that for any $\epsilon > 0$ there exists a norm $\|\cdot\|_\epsilon$ on $\mathbb{R}^n$ such that the subordinate matrix norm $\|\cdot\|_\epsilon$ on $\mathbb{R}^{n\times n}$ satisfies

$$\|\mathbf{A}\|_\epsilon \leq \rho(\mathbf{A}) + \epsilon.$$

Together these results imply that there exists a subordinate matrix norm $\|\cdot\|$ on $\mathbb{R}^{n\times n}$ such that $\rho(\mathbf{A}) < 1$ if and only if $\|\mathbf{A}\| < 1$. Thus the main condition of the previous theorem can be equivalently reformulated as

$$\rho(\mathbf{J_G}(\mathbf{x}^*)) < 1.$$

The following lemma gives a practical way of determining when and where a differentiable function is a contraction, in order to apply Banach's fixed point theorem.

Lemma 7. *Assume that* $\mathbf{G}\colon D \mapsto D$*, with* $D \subset \mathbb{R}^n$ *convex, is continuously differentiable and let* $0 \leq C < 1$*. Let* $\|\cdot\|$ *be a norm on* $\mathbb{R}^n$ *with subordinate matrix norm* $\|\cdot\|$ *on* $\mathbb{R}^{n\times n}$*. If*

$$\|\mathbf{J_G}(\mathbf{x})\| \leq C,$$

for all $\mathbf{x} \in D$*, then* $\mathbf{G}$ *is a contraction using this norm.*

Proof. Choose any $\mathbf{x}, \mathbf{y} \in D$. The same computation as in (5) then shows that

$$\|\mathbf{G}(\mathbf{y}) - \mathbf{G}(\mathbf{x})\| \leq C\|(\mathbf{y} - \mathbf{x})\|$$

□

Banach's fixed point theorem then applies on D, provided D is closed. This is usually the easiest method to prove that $\mathbf{G}$ is a contraction. We remark that in $n = 1$ the criterion

$$\|\mathbf{J_G}(\mathbf{x})\| \leq C$$

becomes $|G'(x)| \leq C$ for some non-negative constant $C < 1$.

3 Convergence rate and order

We briefly mention two different ways to measure how fast an algorithm converges.

Definition 8. *Let the sequence* $\{\mathbf{x}^{(k)}\}_{k=0}^{\infty} \subset \mathbb{R}^n$ *converge to* $\mathbf{x}^* \in \mathbb{R}^n$*. If there exists a sequence* $\{\beta^{(k)}\}_{k=0}^{\infty} \subset \mathbb{R}$ *which converges to 0, and a positive constant* C *such that*

$$\|\mathbf{x}^{(k)} - \mathbf{x}^*\| \leq C|\beta^{(k)}| \quad \text{for sufficiently large} \quad k$$

Then $\{\mathbf{x}^{(k)}\}_{k=0}^{\infty}$ *is said to converge to* $\mathbf{x}^*$ *with* rate of convergence $\mathcal{O}(\beta^{(k)})$*.*

Note: $\beta^{(k)}$ is usually chosen as $\frac{1}{a^k}$ or $\frac{1}{k^a}$ for some positive constant a.

A central concept is the *order of convergence*.

Definition 9. *We say that the sequence* $\{\mathbf{x}^{(k)}\}_{k=0}^{\infty} \subset \mathbb{R}^n$ *converges to* $\mathbf{x}^* \in \mathbb{R}^n$ *with* convergence order $p \geq 1$*, if* $\mathbf{x}^{(k)} \rightarrow \mathbf{x}^*$ *and there exists a positive constant* $0 < C < +\infty$ *(*$0 < C < 1$ *if* $p = 1$*) such that*

$$\|\mathbf{x}^{(k+1)} - \mathbf{x}^*\| \leq C\|\mathbf{x}^{(k)} - \mathbf{x}^*\|^p$$

We say that an iterative method on the form $\mathbf{x}^{(k+1)} = \mathbf{G}(\mathbf{x}^{(k)})$ *has order* p *if the sequence* $\{\mathbf{x}^{(k)}\}_{k=0}^{\infty}$ *it generates converges to the solution* $\mathbf{x}^* = \mathbf{G}(\mathbf{x}^*)$ *with convergence order* p*.*

We in particular mention the following cases:

- $p = 1$ is called *linear convergence*. Roughly the number of "correct digits" increases by a constant number.
- $p = 2$ is called *quadratic convergence*. Roughly the number of "correct digits" doubles in each iteration.
- $p = 3$ is called *cubic convergence*. Roughly the number of "correct digits" triples in each iteration.

If one aims for reasonable high accuracy, higher order methods are (all else being about equal) very desirable.

4 Convergence order of fixed point methods

From Banach's fixed point theorem, we are guaranteed (at least) linear convergence for the fixed point iteration. Now let us return to fixed point iterations for the case of $n = 1$. The following result tells us when we can expect higher convergence order.

Theorem 10. *Assume that $G\colon \mathbb{R} \mapsto \mathbb{R}$ is p-times continuously differentiable with $p \geq 1$ in a neighbourhood of a fixed point x^* of G. Furthermore assume that*

$$\begin{aligned} 0 = G'(x^*) = G''(x^*) = \ldots = G^{(p-1)}(x^*), \qquad & \text{if } \ p \geq 2, \\ G'(x^*) < 1, \qquad & \text{if } \ p = 1. \end{aligned}$$

Then the fixed point sequence $x^{(k+1)} = G(x^{(k)})$ converges to x^* with (at least) order p, provided that the starting point $x^{(0)}$ is sufficiently close to x^*. If in addition

$$G^{(p)}(x^*) \neq 0$$

this convergence order is precisely p.

Proof. First note that Theorem 6 shows that the fixed point sequence indeed converges to x^* for suitable starting points $x^{(0)}$. A Taylor expansion of G at the fixed point x^* now shows that

$$x^{(k+1)} - x^* = G(x^{(k)}) - G(x^*) = \sum_{s=1}^{p-1} \frac{G^{(s)}(x^*)}{s!} \left(x^{(k)} - x^*\right)^s + \frac{G^{(p)}(\xi)}{p!} \left(x^{(k)} - x^*\right)^p$$

for some ξ between $x^{(k)}$ and x^*. The sum on the left will be empty in the case $p = 1$. Since $G^{(s)}(x^*) = 0$ for $1 \leq s \leq p-1$, this further implies that

$$\left|x^{(k+1)} - x^*\right| = \frac{\left|G^{(p)}(\xi)\right|}{p!} \left|x^{(k)} - x^*\right|^p$$

Because $G^{(p)}$ is continuous, there exists $C > 0$ (with $C < 1$ for $p = 1$) such that

$$\frac{\left|G^{(p)}(\xi)\right|}{p!} \leq C$$

for ξ sufficiently close to x^*, Thus

$$\left|x^{(k+1)} - x^*\right| \leq C \left|x^{(k)} - x^*\right|^p ,$$

and therefore the convergence order of the sequence $\{x^{(k)}\}_{k=0}^{\infty}$ is (at least) p. If

$$G^{(p)}(x^*) \neq 0$$

then again because $G^{(p)}$ is continuous, there exist $K > 0$ such that

$$\frac{\left|G^{(p)}(\xi)\right|}{p!} \geq K$$

for ξ sufficiently close to x^*. Thus

$$\left|x^{(k+1)} - x^*\right| \geq K \left|x^{(k)} - x^*\right|^p,$$

which implies that the convergence order cannot be higher than p. Thus the convergence order is precisely p. □

Note: From the proof of Theorem 10 we expect that close to the fixed point x^*

$$\left|x^{(k+1)} - x^*\right| \approx \frac{\left|G^{(p)}(x^*)\right|}{p!} \left|x^{(k)} - x^*\right|^p,$$

when

$$0 = G'(x^*) = G''(x^*) = \ldots = G^{(p-1)}(x^*) \quad \text{but} \quad G^{(p)}(x^*) \neq 0.$$