Introduction to Scientific Computing

(Lecture 5: Stability)

Bojana Rosić

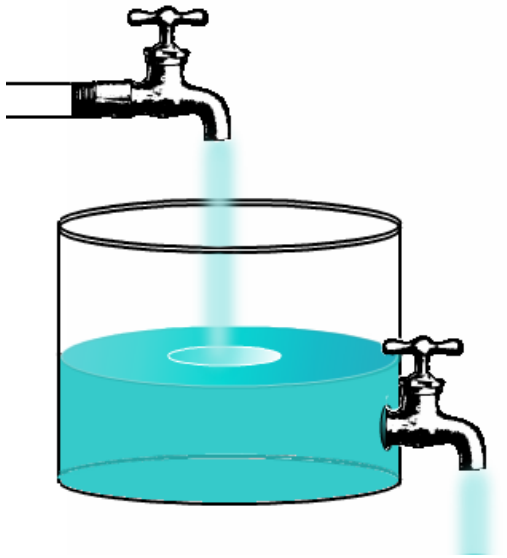
Institute of Scientific Computing

November 22, 2016

Equilibrium (stationary/fixed points)

The equilibrium of the dynamical system is a steady state which does not change in time:

$$\mathbf{x}_0 = \mathbf{x}_1 = \mathbf{x}_2 = \ldots = \mathbf{x}_n = \mathbf{x}_*.$$



Example:

The water level x_n in reservoir does not change as the amount of water that runs in is equal to the amount of water that comes out.

Equilibrium (stationary/fixed points)

Equilibrium 1

Equilibrium 2

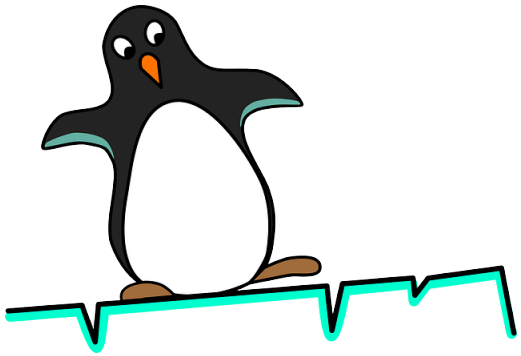
Equilibrium 3

Equilibrium (stationary/fixed points)

If our system is described by the difference equation

$$\Delta \mathbf{x}_n = \mathbf{x}_{n+1} - \mathbf{x}_n = G(n, \mathbf{x}_n, \mathbf{x}_{n-1}, \ldots, \mathbf{x}_{n-k+1})$$

in which G can be linear or nonlinear, then the equilibrium point $\mathbf{x}_*$ can be obtained by requiring



$$G(n, \mathbf{x}_*) = 0, \quad \forall n \in \mathbb{N}$$

Note that if we add $\mathbf{x}_*$ to the both sides of the previous equation we obtain

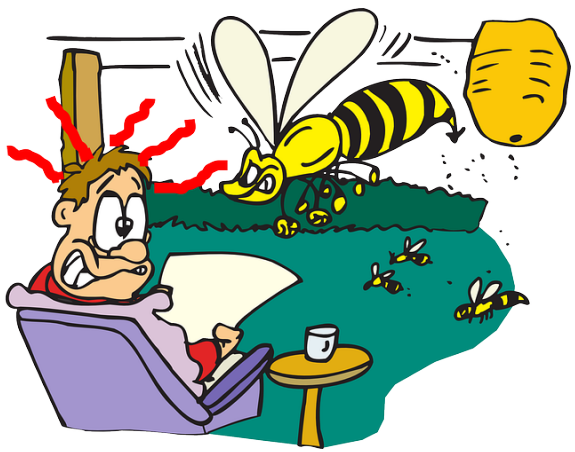
$$\mathbf{x}_* = G(n, \mathbf{x}_*) \Rightarrow \mathbf{x}_* = F(n, \mathbf{x}_*)$$

The last equation is known as **the fixed point iteration**.

How to judge if equilibrium is stable or not?

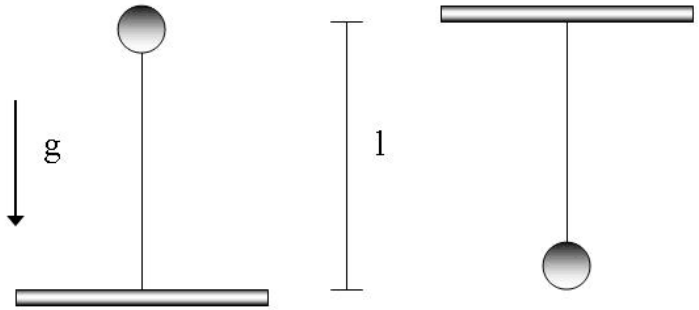
To see whether $\mathbf{x}_*$ is stable or not, use as initial condition $\mathbf{x}_0 = \mathbf{x}_* + \delta$ (**a little perturbation**) where $|\delta|$ is small. Then, evaluate $\mathbf{x}_n$ and see what is the behavior of the solution for $n \to \infty$? Does it converge towards some value, does it cycle periodically around some value, or does encounter some other behavior?

- if for $n \to \infty$ the solution $x_n \to \infty$, then the system is **unstable**
- if for $n \to \infty$ the solution $x_n \to x_*$ or near to it, then the system is **stable**



Example

Consider a pendulum. When the pendulum is pointing straight up then the system can have a steady state. However, minor changes will result in the pendulum swinging to either side. This steady state is an unstable steady state. Now consider the pendulum hanging down. When the pendulum is in is vertical position then the system does not change and this is also a steady state of a system. However, any perturbation of the pendulum will not have an effect in the long run as the pendulum will ultimately settle down to its steady state. This steady state is a stable steady state.

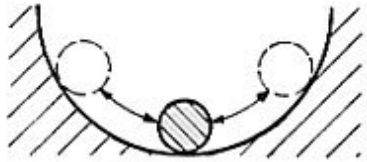


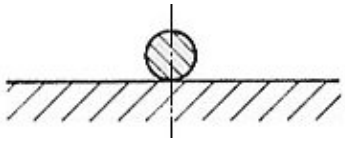
http://alcheme.tamu.edu

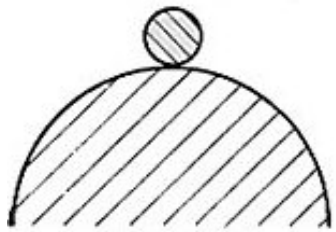
Stability

Assume that a system is in equilibrium. If the system is disturbed a little, one of the following may happen:

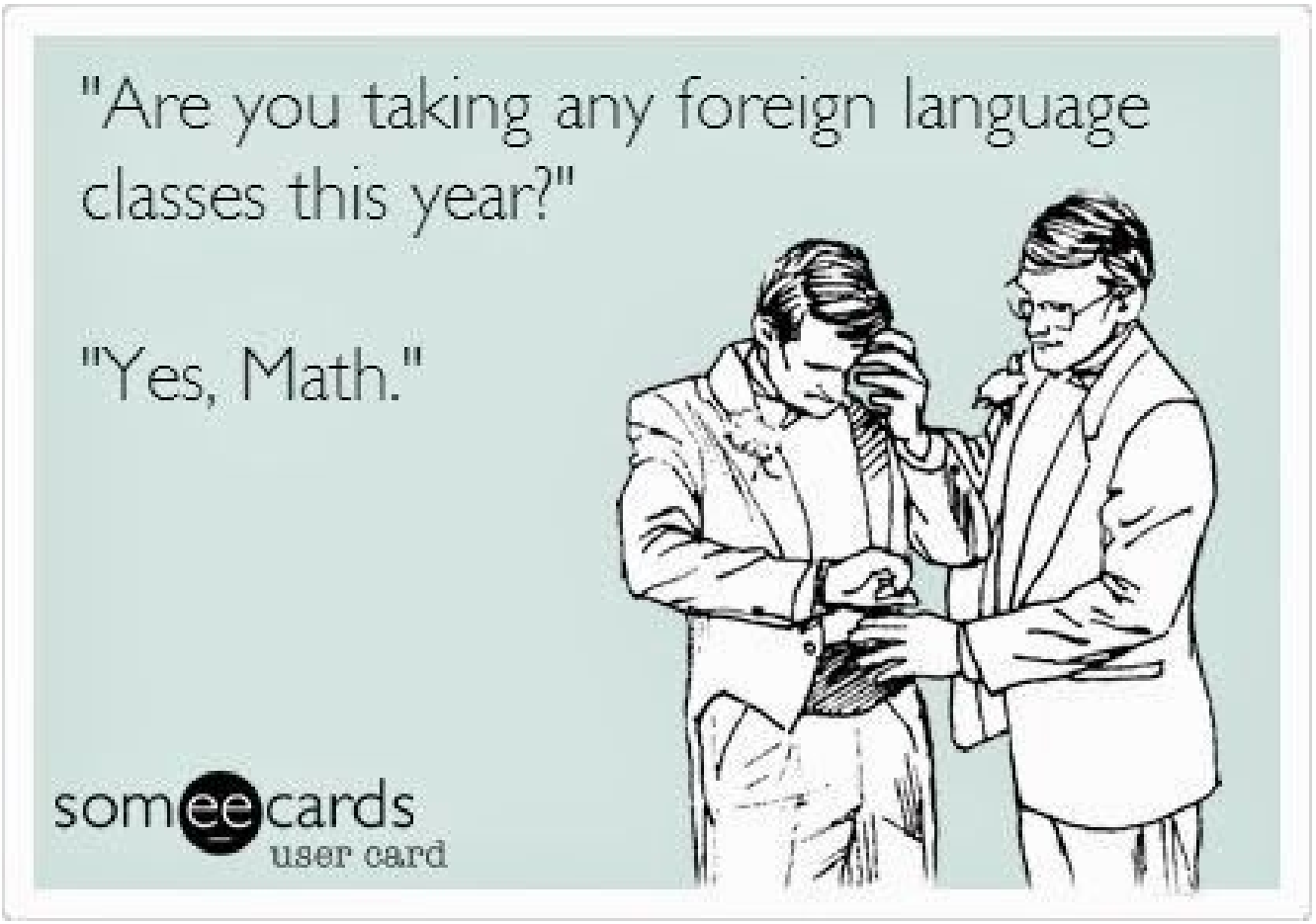
- The system might immediately return to the equilibrium. In this case, the equilibrium is called asymptotically stable.
- Or the system might move around in the neighborhood of the old equilibrium without returning to it. However, the system will not move far away. In this case, the equilibrium is stable but not asymptotically stable.
- The system might move far away from the old equilibrium. In this case, the equilibrium is instable.







Let us now transform this to math language



Stability of FODE

Consider: general difference equation of order 1 and dimension d

$$\mathbf{x}_{n+1} = F(\mathbf{x}_n), \qquad \mathbf{x}_n \in \mathbb{R}^d,\ n \in \mathbb{N}.$$

Definition: An equilibrium point of the dynamical system

$$\mathbf{x}_{n+1} = F(\mathbf{x}_n), \qquad \mathbf{x}_n \in \mathbb{R}^d,\ n \in \mathbb{N}$$

is a state–vector $\mathbf{x}_* \in \mathbb{R}^d$ such that

$$F(\mathbf{x}_*) = \mathbf{x}_*$$

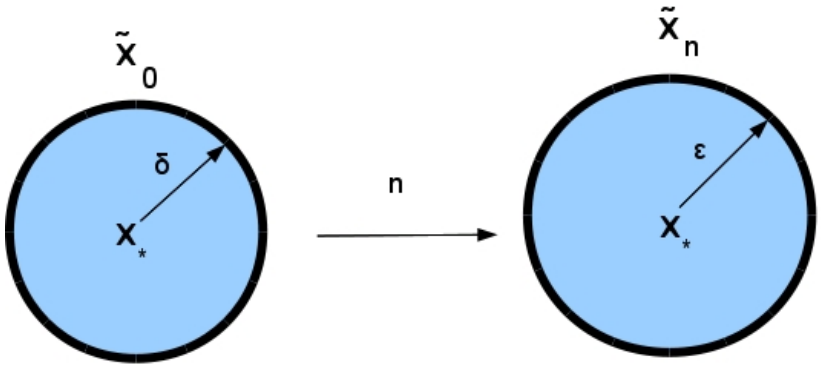
holds.

Stability of FODE

Let $\mathbf{x}_* = F(\mathbf{x}_*) \in \mathbb{R}^d$ be an equilibrium point of the dynamical system $\mathbf{x}_{n+1} = F(\mathbf{x}_n)$ where $\mathbf{x}_n \in \mathbb{R}^d$, $n \in \mathbb{N}$. Furthermore, let

$$\tilde{\mathbf{x}}_{n+1} = F(\tilde{\mathbf{x}}_n), \qquad \tilde{\mathbf{x}}_n \in \mathbb{R}^d,\ n \in \mathbb{N}$$

be the **pertubed** solution.



Definition (1)

The equilibrium point $\mathbf{x}_*$ is called stable if for all $\epsilon > 0$ exist a $\delta > 0$ such that for all $\tilde{\mathbf{x}}_0$ coming from $\|\mathbf{x}_* - \tilde{\mathbf{x}}_0\| \leq \delta$ holds

$$\|\mathbf{x}_* - \tilde{\mathbf{x}}_n\| < \epsilon, \quad \forall n > 0.$$

Here, $\|\cdot\|$ stands for **the norm**.

Stability of FODE

Definition (2)

The equilibrium point $\mathbf{x}_*$ *is called attractive if there exists a* $\delta > 0$ *such that for all* $\tilde{\mathbf{x}}_0$ *coming from* $\|\mathbf{x}_* - \tilde{\mathbf{x}}_0\| < \delta$ *holds*

$$\lim_{n\to\infty} \|\mathbf{x}_* - \tilde{\mathbf{x}}_n\| = 0.$$

Definition (3)

The equilibrium point $\mathbf{x}_*$ *is called asymptotically stable if* $\mathbf{x}_*$ *is stable and attractive.*

Definition (4)

The equilibrium point $\mathbf{x}_*$ *is called unstable if it is not stable.*

Exercise: one equation

Let us observe

$$x_{n+1} = ax_n$$

where a is a scalar. Then the equilibrium point is

$$x_* = ax_* \Rightarrow x_* = 0.$$

The solution of the difference equation is something we have learned before and reads

$$x_n = a^n x_0.$$

To check if $x_* = 0$ is stable we need to perturb the initial condition

$$\tilde{x}_0 = x_* + \delta, \quad \delta > 0$$

Exercise: one equation

The new solution after perturbation reads

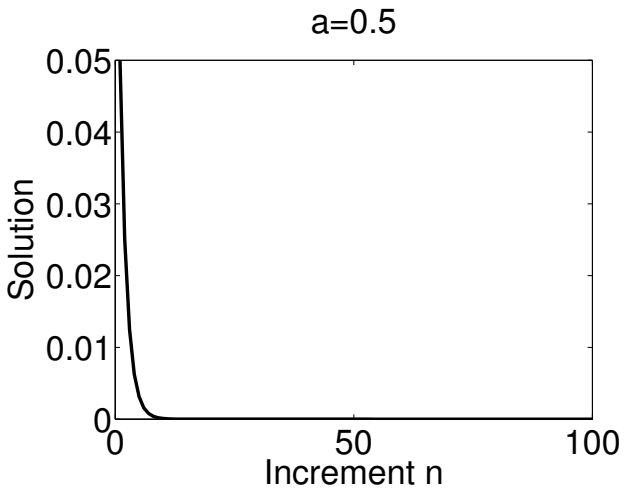
$$\tilde{x}_n = a^n(x_* + \delta).$$

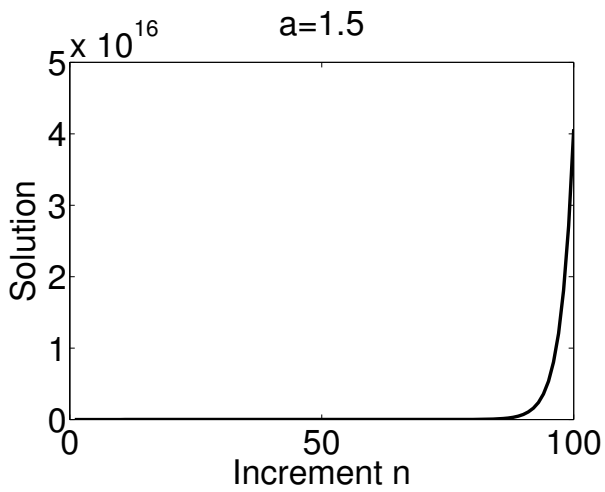
Having that $x_* = 0$ the last equation becomes

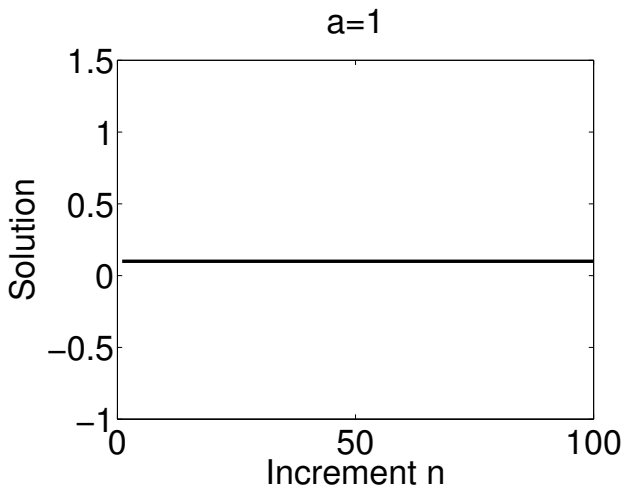
$$\tilde{x}_n = a^n(\delta).$$

Therefore, whether $x_* = 0$ is stable (i.e. $\tilde{x}_n \to x_*$ when $n \to \infty$) or not depends on a.

Exercise: one equation







Exercise: one equation

- If $|a| > 1$ x_n will grow without bound. Therefore x_* is unstable.
- If $|a| < 1$, then x_n will converge to $x_* = 0$, i.e. for small perturbations δ, the system won't move away from the equilibrium but return to it. In this case, $x_* = 0$ is a stable equilibrium.
- if $|a| = 1$ then $x_n = \delta$, and hence the system moves around a neighborhood of the old equilibrium without returning to it. Hence, $x_* = 0$ is stable but not asymptotically

Exercise: system of equations

For a system of FODE

$$\mathbf{x}_{n+1} = A\mathbf{x}_n = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 2 & 1 \\ 1 & 2 & 1 \end{pmatrix} \mathbf{x}_n$$

we may find the equilibrium point

$$\mathbf{x}_* = A\mathbf{x}_* = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 2 & 1 \\ 1 & 2 & 1 \end{pmatrix} \mathbf{x}_* \Rightarrow \mathbf{x}_* = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

On the other side, we know that the solution of the previous system is

$$\mathbf{x}_n = A^n \mathbf{x}_0 = \sum_{j=1}^{3} c_j \lambda_j^n \mathbf{v}_j$$

Exercise: system of equations

To check if $\mathbf{x}_*$ is stable let us perturb the initial condition

$$\tilde{\mathbf{x}}_0 = \mathbf{x}_* + \delta, \quad \delta > 0$$

such that

$$\tilde{\mathbf{x}}_n = A^n \tilde{\mathbf{x}}_0 = A^n(\mathbf{x}_* + \delta)$$

holds. Having that $\mathbf{x}_* = \mathbf{0}$ one obtains

$$\tilde{\mathbf{x}}_n = A^n \delta = \sum_{j=1}^{3} m_j \lambda_j^n \mathbf{v}_j$$

Hence, $A^n \delta \to 0$ when $|\lambda_j^n| \to 0$.

Exercise: system of equations

In our case

$$\lambda_{1,2} = 0.3652 \pm 0.6916i, \quad \lambda_3 = 3.2695$$

which further means

$$|\lambda_{1,2}| = \sqrt{0.3652^2 + 0.6916^2} = 0.7821 < 1$$

and

$$|\lambda_3| = 3.2695 > 1$$

Thus, the system is **unstable**.

Exercise: system of equations II

For autonomous system of FODE

$$\mathbf{x}_{n+1} = A\mathbf{x}_n + \mathbf{b}$$

we may find the equilibrium point

$$\mathbf{x}_* = A\mathbf{x}_* + \mathbf{b} \Rightarrow \mathbf{x}_* = (I - A)^{-1}\mathbf{b}$$

The solution is

$$\mathbf{x}_n = A^n\mathbf{x}_0 + \mathbf{x}^{(p)}$$

where $\mathbf{x}^{(p)}$ is assumed to be in the form of $\mathbf{c}$ since $\mathbf{b}$ does not depend on n. Substituting $\mathbf{x}^{(p)}$ to the original difference equation

$$\mathbf{c} = A\mathbf{c} + \mathbf{b} \Rightarrow \mathbf{x}^{(p)} = \mathbf{x}_*$$

Exercise: system of equations II

Hence, the solution is given by

$$\mathbf{x}_n = A^n \mathbf{x}_0 + \mathbf{x}_* = \sum_{i=1}^{d} c_i \lambda_i^n \mathbf{v}_i + \mathbf{x}_*$$

and no matter how we choose $\mathbf{x}_0$ the solution will converge to $\mathbf{x}_*$ (when $n \to \infty$)only if $|A^n| \to 0$.

This further means that the eigenvalues of A must be less than 1, i.e.

$$|\lambda_i| < 1$$

Exercise: Stability of system of FODE

Stability criteria:

- If all λ_i of A have absolute value smaller than one: $(\forall i = 1, \ldots, d : |\lambda_i| < 1)$, then for every $\mathbf{x}_0 \in \mathbb{R}^d$ the sequence $\mathbf{x}_n \overset{n\to\infty}{\longrightarrow} 0$, and $\mathbf{x}_*$ is asymptotically stable.
- If any λ_i of A has absolute value greater than one: $(\exists i : |\lambda_i| > 1)$ then there exist $\mathbf{x}_0 \in \mathbb{R}^d$ such that the sequence $\mathbf{x}_n \overset{n\to\infty}{\longrightarrow} \infty$, and $\mathbf{x}_*$ is unstable.
- If all λ_i of A have absolute value smaller or equal than one: $(\forall i = 1, \ldots, d : |\lambda_i| \leq 1)$ and if there are λ_j's with $|\lambda_j| = 1$, then we cannot decide whether or not $\mathbf{x}_*$ is stable.

Exercise: Application problem ??

Let us now look at the difference equation

$$x_{n+1} = \sin x_n.$$

The equilibrium point is satisfying

$$x_* = \sin x_* \Rightarrow x_* - \sin x_* = 0$$

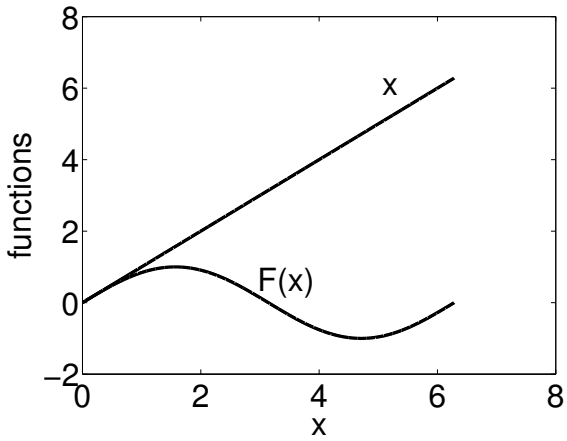
which is the **nonlinear** equation.

Exercise: Application problem ??

The equilibrium point is easy to calculate

$$x_* = \sin x_* \Rightarrow x_* - \sin x_* = 0$$

as we know that there is one solution $x_* = 0$ (check the graph on the right).



How to check if this point is stable now?? We didn't study how to solve the nonlinear difference equation!!

Stability of FODE

For stability investigation we usually observe the normal and perturbed systems

System	Perturbed system
$\mathbf{x}_{n+1} = F(\mathbf{x}_n)$ with equilibrium $\mathbf{x}_* = F(\mathbf{x}_*)$	$\tilde{\mathbf{x}}_{n+1} = F(\tilde{\mathbf{x}}_n)$ with initial condition $\tilde{\mathbf{x}}_0 = \mathbf{x}_* + \delta$

which gives us the difference

$$\mathbf{y}_n := \tilde{\mathbf{x}}_n - \mathbf{x}_*$$

Hence, $\mathbf{x}_*$ is asymptotically stable if $\mathbf{y}_n \to 0$ for $n \to \infty$.

Stability of FODE

In case of general FODE we may compute the previously defined difference by using the Taylor expansion

$$\begin{aligned} \mathbf{y}_n &= \tilde{\mathbf{x}}_n - \mathbf{x}_* = F(\tilde{\mathbf{x}}_{n-1}) - \mathbf{x}_* \\ &= \underbrace{F(\mathbf{x}_*)}_{=\mathbf{x}_*} + F'(\mathbf{x}_*)(\tilde{\mathbf{x}}_{n-1} - \mathbf{x}_*) + \mathcal{O}(|\tilde{\mathbf{x}}_{n-1} - \mathbf{x}_*|^2) - \mathbf{x}_* \\ &= F'(\mathbf{x}_*)\mathbf{y}_{n-1} + \mathcal{O}(|\mathbf{y}_{n-1}|^2). \end{aligned}$$

Assume, that $|\mathbf{y}_{n-1}|$ is small. Then

$$\mathbf{y}_n = F'(\mathbf{x}_*)\mathbf{y}_{n-1},$$

in which $F'(\mathbf{x}_*)$ denotes the Jacobi–matrix of F in the equilibrium point $\mathbf{x}_*$.

Stability of FODE

The system

$$\mathbf{y}_n = F'(\mathbf{x}_*)\mathbf{y}_{n-1}.$$

is now linear FODE and we may diagonalise the matrix $F'(\mathbf{x}_*)$ such that it has d linearly independent eigenvectors $\mathbf{v}_1, \ldots, \mathbf{v}_d \in \mathbb{R}^d$ and eigenvalues $\lambda_1, \ldots, \lambda_d \in \mathbb{C}$. Then there exist coefficients $c_1, \ldots, c_d$ such that

$$\mathbf{y}_0 = \sum_{j=1}^{d} c_j \mathbf{v}_j \quad \text{and} \quad \mathbf{y}_n = \sum_{j=1}^{d} c_j \lambda_j^n \mathbf{v}_j.$$

Now judging on λ values one may define the stability criteria.

Stability of FODE

Stability criteria:

- If all λ_i of $F'(\mathbf{x}_*)$ have absolute value smaller than one: $(\forall i = 1, \ldots, d : |\lambda_i| < 1)$, then for every $\mathbf{y}_0 \in \mathbb{R}^d$ the sequence $\mathbf{y}_n \overset{n\to\infty}{\longrightarrow} 0$, and $\mathbf{x}_*$ is asymptotically stable.
- If any λ_i of $F'(\mathbf{x}_*)$ has absolute value greater than one: $(\exists i : |\lambda_i| > 1)$ then there exist $\mathbf{y}_0 \in \mathbb{R}^d$ such that the sequence $\mathbf{y}_n \overset{n\to\infty}{\longrightarrow} \infty$, and $\mathbf{x}_*$ is unstable.
- If all λ_i of $F'(\mathbf{x}_*)$ have absolute value smaller or equal than one: $(\forall i = 1, \ldots, d : |\lambda_i| \leq 1)$ and if there are λ_j's with $|\lambda_j| = 1$, then higher order terms in the Taylor–series of F are required to decide whether or not $\mathbf{x}_*$ is stable.

Exercise: Application problem ??

To check stability of equilibrium point $x_* = 0$ of

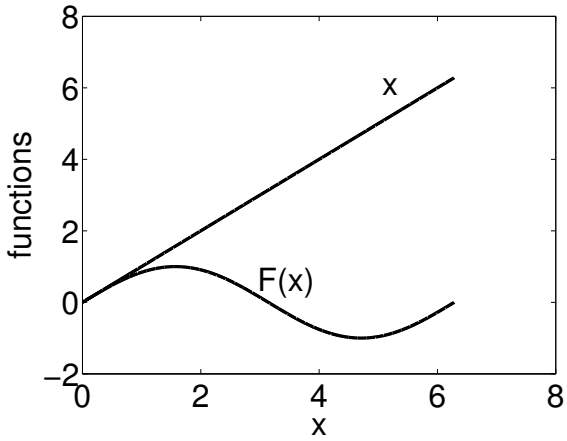
$$x_{n+1} = F(x_n) = \sin x_n$$

we may find the Jacobian of the function

$$F(x_*) = \sin x_* \Rightarrow F'(x_*) = \cos x_*$$

Now we substitute the point $x_* = 0$ to obtain

$$F'(x_*) = \cos 0 = 1$$



Since we have obtained 1 and we have no other eigenvalues we may conclude that the system is stable.

Stability of HODE

The difference equation

$$x_{n+2} + 2x_{n+1} + x_n = 2$$

has the equilibrium point

$$x_* + 2x_* + x_* = 2 \Rightarrow x_* = \frac{1}{2}.$$

To check if this point is stable one may transform HODE to the FODE system

$$\mathbf{y}_{n+1} = \begin{pmatrix} 0 & 1 \\ -1 & -2 \end{pmatrix} \mathbf{y}_n + \begin{pmatrix} 0 \\ 2 \end{pmatrix} = A\mathbf{y}_n + \mathbf{b}$$

And we have already learned that the point x_* will be stable if the eigenvalues of A are smaller than 1.

Stability of HODE

Since A has eigenvalues:

$$\lambda_{1,2} = -1$$

and both by absolute value are equal to 1 one may conclude that the point is **unstable**.

Stability of HODE: another way

Also, note that the difference equation

$$x_{n+2} + 2x_{n+1} + x_n = 2$$

has roots: $\rho_{1,2} = -1$ and the solution is

$$x_n = c_1(-1)^n + c_2 n(-1)^n$$

Hence, whatever we choose for initial conditions we get $n \to \infty \Rightarrow x_n \to \infty$

Evaluating equilibrium point

To find the equilibrium of a dynamical system

$$\mathbf{x}_* = F(\mathbf{x}_*),$$

a set of linear or nonlinear equations

$$G(\mathbf{x}_*) = 0,$$

has to be solved, where

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

http://www.vitamin-ha.com/

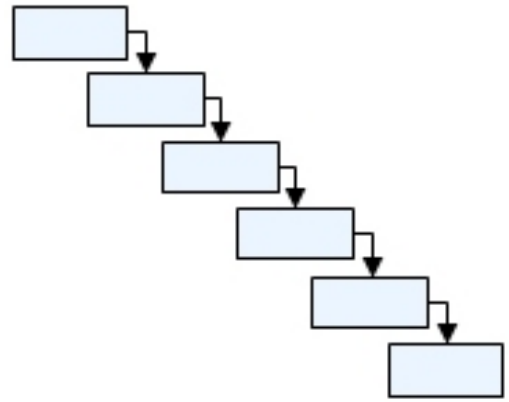
Solving set of equations: Fixed point iterations

Similarly, the system of equations

$$G(\mathbf{x}) = 0$$

can be solved by constructing a dynamical system with

$$F(\mathbf{x}_*) = \mathbf{x}_*$$



copyright@fcsl

in which $\mathbf{x}_*$ denotes the equilibrium point. Once the appropriate dynamical system is found, one may apply the evolution rule by starting from some initial point $\mathbf{x}_0$

$$\mathbf{x}_{n+1} = F(\mathbf{x}_n)$$

until $\mathbf{x}_{n+1}$ "**almost matches**" $\mathbf{x}_n$ (convergence).

Problem

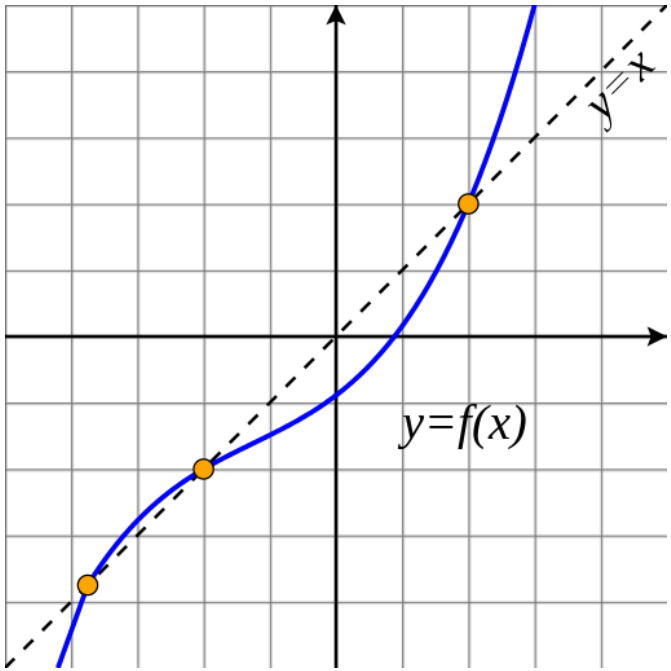
We want to solve

$$F(\mathbf{x}_*) = \mathbf{x}_*$$

such that

- the mapping F has the fixed point (solution)
- the fixed point $\mathbf{x}_*$ is unique
- can be obtained by iterative process

$$\mathbf{x}_{n+1} = F(\mathbf{x}_n)$$



copyright@wiki

Exercise

Problem: Find the root 1.8556 *of the polynomial*

$$x^4 - x - 10 = 0$$

in an iterative manner.

First, one has to find the proper dynamical system $x_* = F(x_*)$. Let us try with

$$\text{I case} \quad F(x) = \frac{10}{x^3 - 1}$$

and the fixed point iterative scheme

$$x_{n+1} = \frac{10}{x_n^3 - 1}, \quad x_0 = 2$$

Exercise

Let us now try another one

$$\text{II case} \quad F(x) = (x + 10)^{1/4}$$

and the scheme

$$x_{n+1} = (x_n + 10)^{1/4}, \quad x_0 = 2.0$$

Exercise

Finally, let us try

$$\text{III case} \quad F(x) = \frac{(x+10)^{1/2}}{x}$$

and the scheme

$$x_{n+1} = \frac{(x_n+10)^{1/2}}{x_n}, \quad x_0 = 2.0$$

Comparison

For the same accuracy $abs(x_n - 1.8556) < 1e\text{-}3$ we get

Term	I case	II case	III case
No. of iterations	maxIter	4	55
Final value	1e-2	1.855	1.855

Hence, the first iterative procedure is not good. The other two are giving correct result. However, the second one is much faster and hence more suitable.

Thus, the question: *How to choose the proper scheme?*

How to choose

In order to choose the proper scheme we need to study convergence. This means that we need to study the error in the solution and its behaviour with respect to the number of iterations.

blog.allstream.com

The question is now how we will know this without doing calculations (apriori)? The answer is that the function

$$\mathbf{x} = F(\mathbf{x})$$

has to fulfill some conditions given by **Banach fixed point theorem**.

Problem

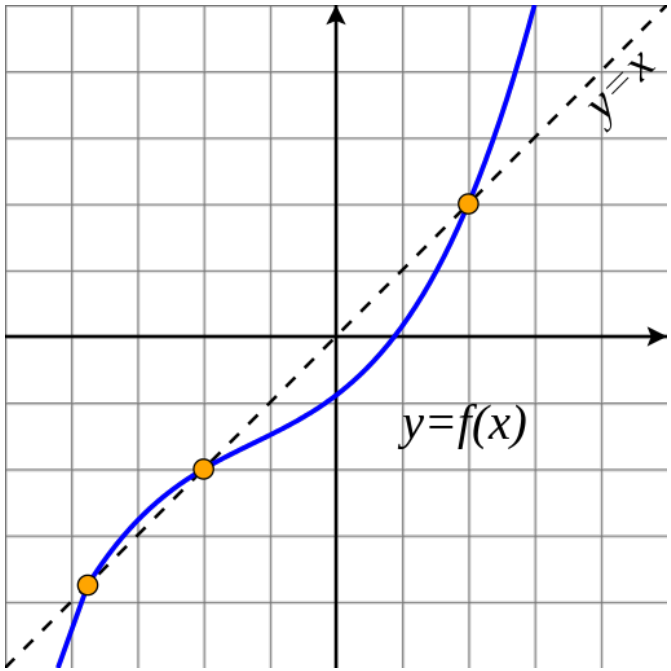
We want to solve

$$F(\mathbf{x}_*) = \mathbf{x}_*$$

such that

- the mapping F has the fixed point (solution)
- the fixed point $\mathbf{x}_*$ is unique
- can be obtained by iterative process

$$\mathbf{x}_{n+1} = F(\mathbf{x}_n)$$



copyright@wiki

Fixed point iteration

The question is now how we will know this without doing calculations if the previous iteration will give us the result? The answer is that the function

$$\mathbf{x} = F(\mathbf{x})$$

has to fulfill some conditions given by **Banach fixed point theorem**.

Banach fixed point theorem

Definition

Let V be a Banach space, let $K \subset V$ be a closed subset, and let $F : K \longrightarrow K$ be a contraction ($0 \leq q < 1$) such that

$$\|F(\mathbf{x}) - F(\mathbf{y})\| \leq q\|\mathbf{x} - \mathbf{y}\| \qquad \text{for all } \mathbf{x}, \mathbf{y} \in K.$$

holds. Then the following conclusions hold:

1.) *F has a unique fixed point $\mathbf{x}_* \in K$, i. e. $F(\mathbf{x}_*) = \mathbf{x}_*$.*

2.) *For any $\mathbf{x}_0 \in K$ the sequence $\mathbf{x}_{n+1} = F(\mathbf{x}_n)$ converges to $\mathbf{x}_*$.*

Error estimates

Additionally, one may define

- A posteriori error estimate:

$$\|\mathbf{x}_* - \mathbf{x}_n\| \leq \frac{q}{1-q}\|\mathbf{x}_n - \mathbf{x}_{n-1}\|$$

- A priori error estimate:

$$\|\mathbf{x}_* - \mathbf{x}_n\| \leq \frac{q^n}{1-q}\|\mathbf{x}_1 - \mathbf{x}_0\|.$$

Lipschitz continuity (LC)

Definition

The function $F : K \to K$ is called Lipschitz continuous if:

$$\|F(x) - F(y)\| \leq q\|x - y\|$$

holds for any $x, y \in K$. Here, q is some positive constant.

Exercise

The function $F = x^2$ satsifies

$$\|x^2 - y^2\| = \|(x-y)(x+y)\| \leq \|(x+y)\|\|x-y\|$$

Hence, comapring to

$$\|x^2 - y^2\| \leq q\|x-y\|$$

one has that

$$q = \|(x+y)\|$$

Thus function is Lipschitz continuous only **locally** because for $x \to \infty \Rightarrow q \to \infty$.

LC of differentiable functions

Every differentiable function at x is Lipschitz continuous at x. Opposite does not hold.

LC of differentiable functions

Let us observe the equation of straight line connecting $\mathbf{y}$ and $\mathbf{x}$ in K

$$\xi = \mathbf{y} + \eta(\mathbf{x} - \mathbf{y}), \quad y \leq \xi \leq x$$

in which $\eta \in [0, 1]$. Then one may observe that

$$F(\xi) = F(\mathbf{y} + \eta(\mathbf{x} - \mathbf{y})) = G(\eta)$$

and

$$\frac{dG}{d\eta} = \frac{\partial F}{\partial \xi}\frac{\partial \xi}{\partial \eta}, \quad \frac{\partial \xi}{\partial \eta} = (\mathbf{x} - \mathbf{y})$$

Observing that

$$F(\mathbf{x}) = G(1), \quad F(\mathbf{y}) = G(0)$$

one may write

$$F(\mathbf{x}) - F(\mathbf{y}) = G(1) - G(0) = \int_0^1 \frac{dG}{d\eta} d\eta = \int_0^1 \frac{\partial F}{\partial \xi}(\mathbf{x} - \mathbf{y}) d\eta$$

LC of differentiable functions

Let F be differentiable, then the fundamental theorem of calculus asserts that there is an $\eta \in [0,1]$ such that

$$F(\mathbf{x}) = F(\mathbf{y}) + \int_0^1 F'(\xi)(\mathbf{x}-\mathbf{y})d\eta \qquad \text{for all } \mathbf{x},\mathbf{y} \in \mathbb{R}^d,$$

$$\begin{aligned}
\|F(\mathbf{x}) - F(\mathbf{y})\| &\leq \int_0^1 \|F'(\xi)(\mathbf{x}-\mathbf{y})\| d\eta \\
&\leq \left(\int_0^1 \|F'(\xi))\| d\eta \right) \|\mathbf{x}-\mathbf{y}\| \\
&\leq \underbrace{\max_{\eta\in[0,1]} \|F'(\xi))\|}_{=:q} \|\mathbf{x}-\mathbf{y}\| \\
&= q\|\mathbf{x}-\mathbf{y}\|.
\end{aligned}$$

LC of differentiable functions

Accordingly, F has for a Lipschitz constant:

$$q := \sup_{\eta \in K} \|F'(\xi)\|$$

Exercise

For the function $F = x^2$ one has

$$\|F'(\xi)\| = 2\|\xi\|$$

i.e.

$$q := \sup_{\eta \in K} \|F'(\xi)\| = 2 \max \|\xi\|$$

Comapring to

$$q = \|(x + y)\|$$

one may conclude that we have obtained the same result havng in mind that for some $\xi \in K$

$$q = \|(x + y)\| \leq 2 \max \|\xi\|$$

LC of differentiable functions

In practical applications q is usually not known analytically. A lower bound can be computed while performing the iteration

$$\mathbf{x}_{k+1} = F(\mathbf{x}_k)$$

by setting

$$q_k := \frac{\|\mathbf{x}_{k+1} - \mathbf{x}_k\|}{\|\mathbf{x}_k - \mathbf{x}_{k-1}\|}, \qquad \hat{q} := \max_k q_k.$$

The value $\hat{q}$ obtained in this manner is often an acceptable estimate for q and it has the advantage that its computation is inexpensive.

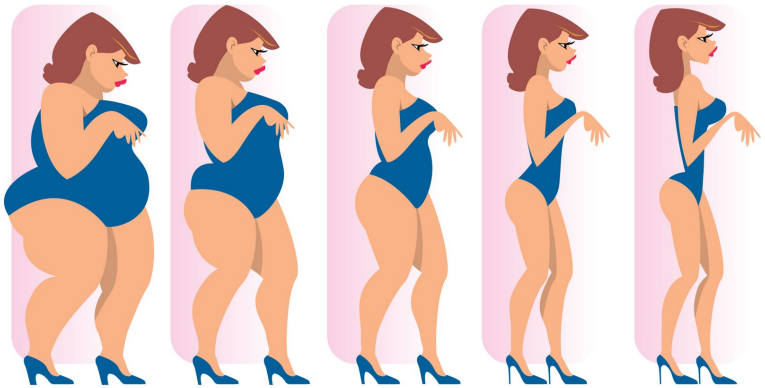
Contraction

Definition

A contraction mapping $F(x)$, or contraction or contractor is a function F from K to itself, with the property that there is some non-negative real number $0 \leq q < 1$ such that for all x and y in K

$$\|F(x) - F(y)\| \leq q\|x - y\|$$

holds.



www.proacaiberry.com

Exercise

For the function $F = x^2$ the Lipschitz constant

$$q := \sup_{\eta \in K} \|F'(\xi)\| = 2 \max \|\xi\|$$

is less than 1 only when $\|\xi\| < 0.5$. Otherwise, it is not. Hence, for $x \in [-0.5, 0.5]$ the function will be continuous and contractive.

Hence,

In order to solve

$$F(x) = x$$

we may apply fixed point iteration and obtain unique x_* from any $x_0 \in K$, only if $F(x)$ satisifes Banach fixed point theorem, i.e. if the mapping $F(x) : K \to K$ is Lipschitz continous and contractive.

Exercise

In our example we had that the first scheme reads

$$x_{n+1} = \frac{10}{x_n^3 - 1}$$

The mapping $F(x) = \frac{10}{x_n^3 - 1}$ is defined on $\mathbb{R}$ and returns values in $\mathbb{R}$. However, we will show that $F(x)$ is not Lipschitz continuous on $\mathbb{R}$, but only on some intervals K. Now we need to find interval K such that $F(x) : K \to K$, i.e. for each $x \in K$ we get some $F(x) \in K$.

Exercise

The mapping

$$F(x) = \frac{10}{x_n^3 - 1}$$

is differentiable and has Jacobian

$$F'(x) = \frac{-30x^2}{(x_n^3 - 1)^2}$$

The derivative goes to infinity when x_n approaches 1. Hence, we need to observe interval K which does not contain 1. Let us observe the interval (a, ∞) where $a > 1$.

Exercise

The absolute value of Jacobian is

$$F'(x) = 30 \left| \frac{x^2}{(x^3 - 1)^2} \right|$$

and it represents the decrising function with increasing x. This means that the supremum happens at the beginning of interval a:

$$q = \sup |F'(x)| \geq F'(a)$$

For interval $(1, 2]$ this value is larger than 1 and hence the scheme is not contractive. This means that we cannot find the root 1.856 in this interval.

Exercise

On the other side, the second scheme

$$x_{n+1} = (x_n + 10)^{1/4}$$

defines the mapping characterised by a Jacobian

$$F'(x) = \frac{0.25}{(x_n + 10)^{3/4}}$$

The derivative goes to infinity when x_n approaches -10. Hence, we need to observe interval K which does not contain -10. Let us observe the interval (a, ∞) where $a > 0$ (in this interval lies 1.8556).

Exercise

The Jacobian is decreasing function with increasing x. Hence, the supremum becomes

$$q = \sup |F'(x)| \geq |F'(a)|$$

in which a is the beginning of interval. If we choose that the interval is $[1, 2]$. Then,

$$q = 4.6957e - 05$$

which tells us that the mapping is contractive. Thus, we may find the root 1.8566 using this scheme.

Exercise II

The difference equation

$$x_{n+1} - \frac{1}{6}(x_n^3 + x_{n-1}^2 + 1) = 0$$

can be rewritten to:

$$x_{n+1} = \frac{1}{6}(x_n^3 + x_{n-1}^2 + 1) =: F(x_n).$$

Thus, the fixed (equilibrium) point satisfies:

$$x_* = F(x_*)$$

and has values

$$x_{*1} = -3.0644, x_{*2} = 1.8920, x_{*3} = 0.1725$$

Exercise II

However, we would like to obtain them numerically by solving

$$x_*^{(k+1)} = F(x_*^{(k)}), \quad x_* \in K$$

The question is if we can do that? Yes, if

- $K = (-\infty, -1) \cup (-1, 1) \cup (1, \infty)$ is complete
- F is locally or globally Lipschitz continuous on K
- and Lipschitz constant is $0 \leq q < 1$

Exercise II

Having

$$F(x) = \frac{1}{6}(x^3 + x^2 + 1)$$

let us check for $a = -1.2$ in $K = (-\infty, -1)$ the value

$$F(a) = F(-1.2) = 0.1187$$

Hence, $F(a) \notin K$ and we may conclude that this interval is not complete. Similarly by taking $a = 1.2$ in $K = (1, \infty)$ we may show that

$$F(a) = 0.6947 \notin K$$

Hence, the interval $K = (1, \infty)$ is also not complete.

Exercise II

If we take $K = (-1, 1)$ we may show that for each a in K the value $F(a)$ falls into K (if you do not believe plot in Matlab). Hence, the interval is complete, and from three fixed points

$$x_{*1} = -3.0644, x_{*2} = 1.8920, x_{*3} = 0.1725$$

only the third one satisfies the first Banach fixed point theorem condition. But, we are not yet sure if we can compute it iteratively as we need to fullfill two more conditions.

Exercise II

We have proven that for differentiable functions F the Lipschitz constant q in

$$\|F(\mathbf{x}) - F(\mathbf{y})\| \leq q\|\mathbf{x} - \mathbf{y}\|$$

is given as derivative of F. In our case this reads as

$$\sup \|F'(x)\| = \|\frac{3x^2 + 2x}{6}\|$$

The maximum of $F'(x)$ in K is for $x = 1$. Hence,

$$\sup_{x \in [-1,1]} |F'(x)| = \frac{5}{6} < 1$$

Hence, the function is **locally** Lipschitz continuous with the Lipschitz constant $q = \frac{5}{6}$

Exercise II

Since,

$$q = \frac{5}{6} < 1$$

one may immediately conclude that the mapping $F(x)$ is contraction on interval $K = [-1, 1]$. This further means that we can compute the root $x_{*,3}$ inside the interval $[-1, 1]$ by using the fixed point iteration and starting from any x_0 in $[-1, 1]$.

Proof of fixed point theorem

Prove

- there exists a fixed point $\mathbf{x}_*$
- $\mathbf{x}_{n+1} = F(\mathbf{x}_n)$ converges to $\mathbf{x}_*$ for any $\mathbf{x}_0 \in K$

Idea: Show that $\mathbf{x}_k$ is a Cauchy sequence for any $\mathbf{x}_0 \in K$ and then use the fact that K as a closed subset of a complete space is itself a complete.

Proof of fixed point theorem I part

Show: that for all $\epsilon > 0$ there exists an $M \in \mathbb{N}$ such that

$$\|\mathbf{x}_{k+m} - \mathbf{x}_k\| < \epsilon \qquad \text{for all } k > M \text{ and all } m > 0.$$

To show this, choose some k and $m \in \mathbb{N}$ and observe that

$$\begin{aligned}\|\mathbf{x}_{k+m} - \mathbf{x}_k\| &= \left\|\mathbf{x}_{k+m} + \sum_{j=1}^{m-1}(\mathbf{x}_{k+j} - \mathbf{x}_{k+j}) - \mathbf{x}_k\right\| \\ &= \left\|\sum_{j=1}^{m}(\mathbf{x}_{k+j} - \mathbf{x}_{k+j-1})\right\| \leq \sum_{j=1}^{m}\|\mathbf{x}_{k+j} - \mathbf{x}_{k+j-1}\|\end{aligned}$$

Here is used the triangle inequality $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$,

Proof of fixed point theorem I part

Further from $\mathbf{x}_{k+j} = F(\mathbf{x}_{k+j-1})$ one has

$$\|\mathbf{x}_{k+j} - \mathbf{x}_{k+j-1}\| = \|F(\mathbf{x}_{k+j-1}) - F(\mathbf{x}_{k+j-2})\| \leq q\|\mathbf{x}_{k+j-1} - \mathbf{x}_{k+j-2}\|$$

Also,

$$\|\mathbf{x}_{k+j-1} - \mathbf{x}_{k+j-2}\| = \|F(\mathbf{x}_{k+j-2}) - F(\mathbf{x}_{k+j-3})\| \leq q\|\mathbf{x}_{k+j-3} - \mathbf{x}_{k+j-4}\|$$

such that

$$\|\mathbf{x}_{k+j} - \mathbf{x}_{k+j-1}\| = \|F(\mathbf{x}_{k+j-1}) - F(\mathbf{x}_{k+j-2})\| \leq q^2\|\mathbf{x}_{k+j-3} - \mathbf{x}_{k+j-4}\|$$

holds.

Proof of fixed point theorem I part

Hence, by induction one may conclude that

$$\|\mathbf{x}_{k+j} - \mathbf{x}_{k+j-1}\| = \|F(\mathbf{x}_{k+j-1}) - F(\mathbf{x}_{k+j-2})\| \leq \cdots \leq q^{j-1}\|\mathbf{x}_{k+1} - \mathbf{x}_k\|$$

holds. Substituting back to the inequality

$$\|\mathbf{x}_{k+m} - \mathbf{x}_k\| \leq \sum_{j=1}^{m} \|\mathbf{x}_{k+j} - \mathbf{x}_{k+j-1}\|$$

we get

$$\|\mathbf{x}_{k+m} - \mathbf{x}_k\| \leq \sum_{j=1}^{m} \|\mathbf{x}_{k+j} - \mathbf{x}_{k+j-1}\| \leq \sum_{j=1}^{m} q^{j-1}\|\mathbf{x}_{k+1} - \mathbf{x}_k\|$$

Proof of fixed point theorem I part

This further leads to

$$\|\mathbf{x}_{k+m} - \mathbf{x}_k\| \leq \sum_{j=1}^{m} q^{j-1} \|\mathbf{x}_{k+1} - \mathbf{x}_k\|$$
$$\leq \sum_{j=1}^{\infty} q^{j-1} \|\mathbf{x}_{k+1} - \mathbf{x}_k\| = \frac{1}{1-q} \|\mathbf{x}_{k+1} - \mathbf{x}_k\|.$$

Here is used the sum of geometric series:

$$\sum_{j=1}^{\infty} q^{j-1} = \frac{1}{1-q}$$

Proof of fixed point theorem I part

Also, we know that

$$\|\mathbf{x}_{k+1} - \mathbf{x}_k\| = \|F(\mathbf{x}_{k+1}) - F(\mathbf{x}_k)\| \leq q\|\mathbf{x}_k - \mathbf{x}_{k-1}\| \leq q^2\|\mathbf{x}_{k-1} - \mathbf{x}_{k-2}\|$$

which leads us to

$$\|\mathbf{x}_{k+1} - \mathbf{x}_k\| \leq q^k\|\mathbf{x}_1 - \mathbf{x}_0\|$$

With the previous slide this gives:

$$\|\mathbf{x}_{k+m} - \mathbf{x}_k\| \leq \frac{q^k}{1-q}\|\mathbf{x}_1 - \mathbf{x}_0\|$$

Because $q < 1$, q^k will decrease with increasing k, and hence one may conclude that $(\mathbf{x}_k)$ is a Cauchy sequence. Since K is complete there exists an $\mathbf{x}_* \in K$ such that $\mathbf{x}_k \longrightarrow \mathbf{x}_*$

Proof of fixed point theorem II part

Show that the fixed point is unique.

Assume that two fixed points $\mathbf{x}_* \in K$ and $\mathbf{y}_* \in K$ exist

$$\mathbf{x}_* = F(\mathbf{x}_*) \quad \text{and} \quad \mathbf{y}_* = F(\mathbf{y}_*)$$

Then

$$\|\mathbf{x}_* - \mathbf{y}_*\| = \|F(\mathbf{x}_*) - F(\mathbf{y}_*)\| \leq q\|\mathbf{x}_* - \mathbf{y}_*\|$$

Since $0 \leq q < 1$, this requires $\|\mathbf{x}_* - \mathbf{y}_*\| = 0$. Therefore, $\mathbf{x}_* = \mathbf{y}_*$ and thus the fixed point is unique.

A posteriori estimate proof

Previously we have shown that

$$\|\mathbf{x}_{k+m} - \mathbf{x}_k\| \leq \frac{1}{1-q}\|\mathbf{x}_{k+1} - \mathbf{x}_k\|$$

but also according to Banach fixed point theorem

$$\|\mathbf{x}_{k+1} - \mathbf{x}_k\| \leq q\|\mathbf{x}_k - \mathbf{x}_{k-1}\|$$

Thus,

$$\|\mathbf{x}_{k+m} - \mathbf{x}_k\| \leq \frac{q}{1-q}\|\mathbf{x}_k - \mathbf{x}_{k-1}\|$$

and as $\mathbf{x}_{k+m} \longrightarrow \mathbf{x}_*$ for $m \longrightarrow \infty$ we obtain

$$\|\mathbf{x}_* - \mathbf{x}_k\| \leq \frac{q}{1-q}\|\mathbf{x}_k - \mathbf{x}_{k-1}\|.$$

A priori error estimate proof

Having

$$\|\mathbf{x}_k - \mathbf{x}_{k-1}\| \leq q\|\mathbf{x}_{k-1} - \mathbf{x}_{k-2}\|$$

and then

$$\|\mathbf{x}_{k-1} - \mathbf{x}_{k-2}\| \leq q\|\mathbf{x}_{k-2} - \mathbf{x}_{k-3}\|$$

one may sequentially obtain

$$\|\mathbf{x}_k - \mathbf{x}_{k-1}\| \leq q\|\mathbf{x}_{k-1} - \mathbf{x}_{k-2}\| \leq \cdots \leq q^{k-1}\|\mathbf{x}_1 - \mathbf{x}_0\|.$$

From this it follows

$$\|\mathbf{x}_* - \mathbf{x}_k\| \leq \frac{q^k}{1-q}\|\mathbf{x}_1 - \mathbf{x}_0\|.$$

Speed of convergence of fixed point iteration

Until now we have defined conditions under which the fixed point iteration

$$\mathbf{x}_{k+1} = F(\mathbf{x}_k)$$

converges to a unique solution $\mathbf{x}_*$. However, sometimes more than one scheme (think about our example) are converging to $\mathbf{x}_*$ To choose the faster one, we have to study the convergence of the scheme. To do this let us expand the fixed point mapping $F(\mathbf{x})$ into Taylor series around $\mathbf{x}_*$. Let $\mathbf{d}_k := \mathbf{x}_k - \mathbf{x}_*$ be the defect/residuum.

Speed of convergence of fixed point iteration

Then,

$$\begin{aligned}\mathbf{x}_{k+1} &= F(\mathbf{x}_k) = F(\mathbf{x}_* + \mathbf{d}_k) \\ &= F(\mathbf{x}_*) + F'(\mathbf{x}_*)\mathbf{d}_k + \frac{1}{2}F''(\mathbf{x}_*)\mathbf{d}_k^2 + \mathcal{O}(|\mathbf{d}_k|^3) \\ &= \mathbf{x}_* + F'(\mathbf{x}_*)\mathbf{d}_k + \frac{1}{2}F''(\mathbf{x}_*)\mathbf{d}_k^2 + \mathcal{O}(|\mathbf{d}_k|^3)\end{aligned}$$

and

$$\mathbf{d}_{k+1} = F'(\mathbf{x}_*)\mathbf{d}_k + \frac{1}{2}F''(\mathbf{x}_*)\mathbf{d}_k^2 + \mathcal{O}(|\mathbf{d}_k|^3).$$

Convergence speed

Thus

$$|\mathbf{d}_{k+1}| \leq |F'(\mathbf{x}_*)|\,|\mathbf{d}_k| + \frac{1}{2}|F''(\mathbf{x}_*)|\,|\mathbf{d}_k|^2 + \mathcal{O}(|\mathbf{d}_k|^3).$$

If $0 < q := |F'(\mathbf{x}_*)| < 1$, the scheme converges linearly with convergence factor q in a small region around $\mathbf{x}_*$. If $|F'(\mathbf{x}_*)| = 0$, the scheme converges at least quadratically in a region around $\mathbf{x}_*$. This analysis can be generalised for the multi-dimensional case.

Convergence

Definition

Let $\mathbf{x}_*, \mathbf{x}_1, \mathbf{x}_2, \ldots \in \mathbb{R}^d$ *with* $\mathbf{x}_n \longrightarrow \mathbf{x}_*$ *for* $n \longrightarrow \infty$.

1. $\{\mathbf{x}_n\}$ *converges linearly with convergence factor* $q \in (0,1)$ *if*

$$\|\mathbf{x}_{n+1} - \mathbf{x}_*\| \leq q\|\mathbf{x}_n - \mathbf{x}_*\|.$$

2. $\{\mathbf{x}_n\}$ *converges super-linearly with order* p *if there exists a* $p > 1$ *with*

$$\|\mathbf{x}_{n+1} - \mathbf{x}_*\| \leq q\|\mathbf{x}_n - \mathbf{x}_*\|^p.$$

3. *If* $q = 0$ *and* $p = 1$, *then* $(\mathbf{x}_n)$ *converges super-linearly.*
4. *When* $(\mathbf{x}_n)$ *converges with order* $p = 2$, *then* $(\mathbf{x}_n)$ *is said to converge quadratically.*

Fixed point iteration

Problems:

- The fixed point iteration is good when the function is Lipschitz continuous and contractive on the interval containing the equilibrium point (root). Otherwise, the method diverges.
- in case of functions with more than one root, the method can miss some of them
- requires complicated algebraic transformations to obtain $F(x)$ in

$$x = F(x)$$

- etc.

Other methods?

Is there **ANY OTHER** way to solve the system

$$\mathbf{f}(\mathbf{x}_*) = \mathbf{0}$$

Answer

There are other ways, but they strongly depend if f is

- linear
- or nonlinear.