NTNU

Norwegian University of Science and Technology

TMA4125 Matematikk 4N

#4 – Nonlinear Equations & Newton's method

Ronny Bergmann

Department of Mathematical Sciences, NTNU.

February 2nd, 2023

Introduction / Recap.

Recap 1. For a given function f we want to consider the equation

$$f(x) = 0 \tag{1}$$

and find its solutions x^*, which we also call roots to the equation.

Challenge. In many applications:
Often these (again) might be complicated/impossible to solve analytically
$\Rightarrow$ Develop numerical techniques to solve (1).

Scalar and systems of equations

We consider scalar equations first, i. e. $f\colon \mathbb{R} \to \mathbb{R}$, with just one equation and one variable, for example

$$x^3 + x^2 - 3x - 3 = 0.$$

After that we will consider systems of equations, for example

$$\begin{aligned} x\mathrm{e}^y &= 1, \\ -x^2 + y &= 1. \end{aligned}$$

We can write this also as a functions $f\colon \mathbb{R}^n \to \mathbb{R}^n$, which in this example would be $n = 2$ and

$$f(x,y) = \begin{pmatrix} x\mathrm{e}^y - 1 \\ -x^2 + y - 1 \end{pmatrix} \quad \text{and we want to solve} \quad f(x,y) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Recap 2. Fix point iteration.

Idea. reformulate $f(x) = 0$ into a fix point equation $g(x) = x$.

Algorithm. Fix point iteration

Input Given a function g and a starting value $x^{(0)}$

1. For $k = 0, 1, 2, \ldots$ compute

$$x^{(k+1)} = g(x^{(k)}).$$

and we discussed its convergence:

Theorem. If there is an interval $[a, b]$ such that $g \in C^1[a, b]$, $g([a, b]) \subset (a, b)$ and there exist a positive constant $L < 1$ such that $|g'(x)| \leq L < 1$ for all $x \in [a, b]$, then

- g has a unique fixed point r in (a, b).
- The fix point iterations $x^{(k+1)} = g(x^{(k)})$ converges towards x^* for all starting values $x^{(0)} \in [a, b]$.

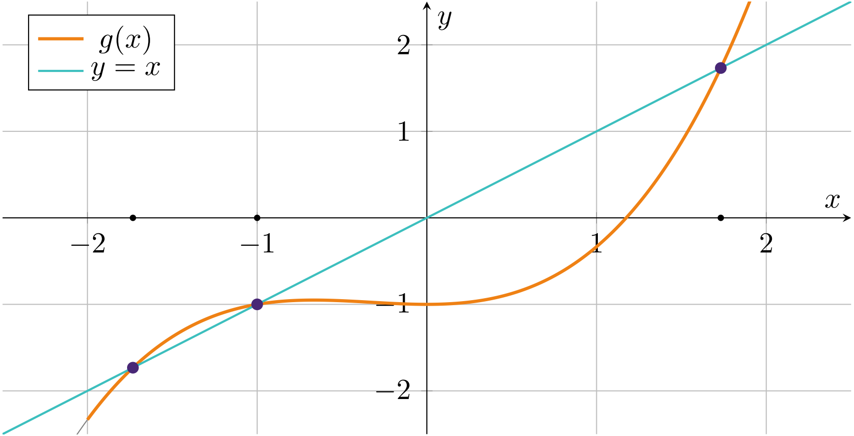
We obtained linear convergence: $|x^* - x^{(k+1)}| \leq |x^* - x^{(k)}|L$ for all k.

Fix point equation: Rearrange f to find g

We rewrite

$$f(x) = x^3 + x^2 - 3x - 3 = 0 \quad \text{to} \quad x = \frac{x^3 + x^2 + 3}{3} = g(x).$$

Interpretation. Fixed points are intersections of $g(x)$ with $y = x$.

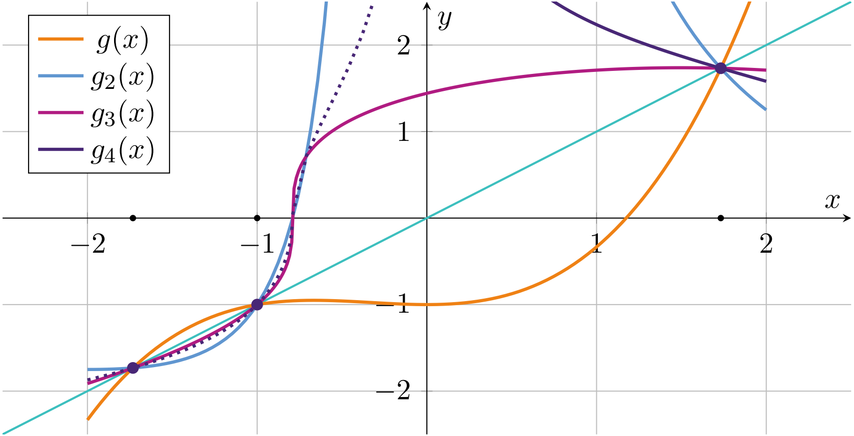


Fix point equation: g is not unique

Other possibilities:

$$g_2(x) = \frac{-x^2 + 3x + 3}{x^2}, \quad g_3(x) = \sqrt[3]{3 + 3x - x^2}, \quad g_4(x) = \sqrt{\frac{3 + 3x - x^2}{x}}$$

Challenge. all of them have different g' and hence different L.



Newton's method – a first idea.

Observation.
Locally around a fix point x^* a small $g'(x^*)$ is preferable (small L).
$\Rightarrow$ What happens if we could (can) choose g such that $g'(x^*) = 0$?

We use $e^{(k)} = x^* - x^{(k)}$ for the iterates of the fix point algorithm and a Taylor expansion (around x^*)

$$
\begin{aligned}
e^{(k+1)} &= x^* - x^{(k+1)} = g(x^*) - g(x^{(k)}) = g(x^*) - g(x^* - e^{(k)}) \\
&= -g'(x^*)e^{(k)} + \frac{1}{2}g''(\xi_k)\big(e^{(k)}\big)^2, \quad (\xi_k \text{ between } x^* \text{ and } x^{(k)})
\end{aligned}
$$

Assume we choose $g'(x^*) = 0$ and that we have a constant M such that $|g''(x)|/2 \leq M$. Then

$$
|e^{(k+1)}| \leq M|e^{(k)}|^2
$$

Also known as quadratic convergence.

How to find our favourite fix point equation.

Question. Given an equation

$$f(x) = 0$$

with an unknown solution x^*.
Can we find a g with fixed point $x^* = g(x^*)$ and $g'(x^*) = 0$?
Idea. For any $h(x)$ we have due to $f(x^*) = 0$ that

$$g(x) = x - h(x)f(x)$$

has a fixed point $x^* = g(x^*)$. The derivate reads using the product rule

$$g'(x) = 1 - h'(x)f(x) - h(x)f'(x)$$

At the fixed point we get

$$g'(x^*) = 1 - h(x^*)f'(x^*).$$

$\Rightarrow$ Choose $h(x) = 1/f'(x)$ we achieve $g'(x^*) = 0$.

Newton's method.

Algorithm.

Input a function f, its derivative f', and a start point $x^{(0)}$.

1. For $k = 0, 1, 2, \ldots$ compute

$$x^{(k+1)} = x^{(k)} - \frac{f(x^{(k)})}{f'(x^{(k)})}$$

Error analysis

We constructed the method to give quadratic convergence

$$|e^{(k+1)}| \leq M\left|e^{(k)}\right|^2,$$

where $e^{(k)} = x^* - x^{(k)}$.
But under which conditions can we say something about the size of M?

Let's compare the Taylor series around x^* with Newton's method

$$0 = f(x^{(k)}) + f'(x^{(k)})(x^* - x^{(k)}) + \frac{1}{2}f''(\xi_k)(x^* - x^{(k)})^2 \quad \text{(Taylor series)}$$

$$0 = f(x^{(k)}) + f'(x^{(k)})(x^{(k+1)} - x^{(k)}), \quad \text{(Newton's method)}$$

where ξ_k is between x^* and $x^{(k)}$. Subtracting both yields

$$f'(x^{(k)})(x^* - x^{(k+1)}) + \frac{1}{2}f''(\xi_k)(x^* - x^{(k)})^2 = 0 \quad \Rightarrow \quad e^{(k+1)} = -\frac{1}{2}\frac{f''(\xi_k)}{f'(x^{(k)})}\left(e^{(k)}\right)^2$$

$\Rightarrow$ we need $f'(x^{(k)}) \neq 0$, $f \in C^2[a,b]$ and $x^{(0)}$ sufficiently close to x^*.

Convergence of Newton's method

Theorem. Assume that the function f has a root x^*, and let $I_\delta = [x^* - \delta, x^* + \delta]$ for some $\delta > 0$.
Assume further that

- $f \in C^2(I_\delta)$.
- There is a $M > 0$ such that $\left|\frac{f''(y)}{f'(x)}\right| \leq 2M, \quad$ for all $x, y \in I_\delta$.

In this case, Newton's method converges quadratically,

$$|e^{(k+1)}| \leq M|e^{(k)}|^2$$

for all starting values satisfying $|x^{(0)} - x^*| \leq \min\{1/M, \delta\}$.

Systems of equations

Instead of one function f with one variable x we now consider

Systems of nonlinear equations.

$$\begin{aligned} f_1(x_1, x_2, \dots, x_n) &= 0 \\ f_2(x_1, x_2, \dots, x_n) &= 0 \\ &\vdots \\ f_n(x_1, x_2, \dots, x_n) &= 0 \end{aligned}$$

We can write this in short as

$$\boldsymbol{f}(\boldsymbol{x}) = \mathbf{0},$$

where $\boldsymbol{f} : \mathbb{R}^n \to \mathbb{R}^n$.

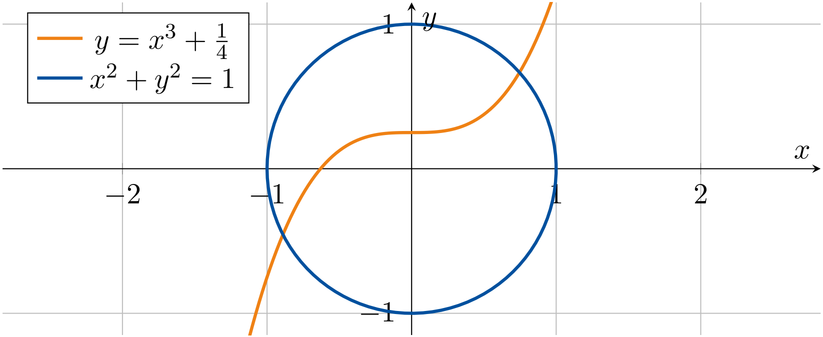
Example.

Example. Consider the two equations

$$x^3 - y + \frac{1}{4} = 0$$
$$x^2 + y^2 - 1 = 0$$

Interpretation.

Rewrite first equation to $y = x^3 + \frac{1}{4}$ $\Rightarrow$ solutions lie on this graph.
Second equation means $1 = x^2 + y^2$ $\Rightarrow$ point with distance 1 from origin



Towards Newton's method for systems of equations I

Idea. Extend fixed point iterations to systems of equations.
We concentrate on Newton's method and the case $n = 2$:

$$f(x, y) = 0$$
$$g(x, y) = 0$$

to avoid getting lost in indices.

Notation. We denote a solution (root) to these by $\boldsymbol{r} = (x^*, y^*)^{\mathrm{T}}$.
Let $\hat{\boldsymbol{x}} = (\hat{x}, \hat{y})^{\mathrm{T}}$ that approximates $\boldsymbol{r}$.

$\Rightarrow$ Search for a better approximation by linearizing $\boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{0}$.
In other words: use the multivariate Taylor expansion around $\hat{\boldsymbol{x}}$

$$f(x, y) = f(\hat{x}, \hat{y}) + \frac{\partial f}{\partial x}(\hat{x}, \hat{y})(x - \hat{x}) + \frac{\partial f}{\partial y}(\hat{x}, \hat{y})(y - \hat{y}) + \dots$$
$$g(x, y) = g(\hat{x}, \hat{y}) + \frac{\partial g}{\partial x}(\hat{x}, \hat{y})(x - \hat{x}) + \frac{\partial g}{\partial y}(\hat{x}, \hat{y})(y - \hat{y}) + \dots$$

where we omit higher order terms, which are small for $(x, y) = \boldsymbol{x} \approx \hat{\boldsymbol{x}}$.

Towards Newton's method for systems of equations II

Idea. Ignore the higher order terms and solve

$$
\begin{aligned}
f(\hat{x},\hat{y}) + \frac{\partial f}{\partial x}(\hat{x},\hat{y})(x-\hat{x}) + \frac{\partial f}{\partial y}(\hat{x},\hat{y})(y-\hat{y}) &= 0\\
g(\hat{x},\hat{y}) + \frac{\partial g}{\partial x}(\hat{x},\hat{y})(x-\hat{x}) + \frac{\partial g}{\partial y}(\hat{x},\hat{y})(y-\hat{y}) &= 0
\end{aligned}
$$

for x and y as a better approximation (or precise: next iterate).

More compact we can write

$$\mathbf{f}(\hat{\boldsymbol{x}}) + J(\hat{\boldsymbol{x}})(\boldsymbol{x} - \hat{\boldsymbol{x}}) = \mathbf{0},$$

where $J(\boldsymbol{x})$ denotes the Jacobian of $\boldsymbol{f}$ given by

$$J(\boldsymbol{x}) = \begin{pmatrix} \frac{\partial f}{\partial x}(x,y) & \frac{\partial f}{\partial y}(x,y) \\ \frac{\partial g}{\partial x}(x,y) & \frac{\partial g}{\partial y}(x,y) \end{pmatrix}$$

Newton's method for systems of equations

Algorithm.

Input A function $\boldsymbol{f}(\boldsymbol{x})$, its Jacobian $J(\boldsymbol{x})$ and a starting value $\boldsymbol{x}^{(0)}$.

1. For $k = 0, 1, 2, \ldots$
   - **1.1** Solve the linear system $J(\boldsymbol{x}^{(k)})\Delta^{(k)} = -\boldsymbol{f}(\boldsymbol{x}^{(k)})$
   - **1.2** Set $\boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)} + \Delta^{(k)}$

Note. This can be generalized to n equations with n unknowns, where the Jacobian reads

$$J(\boldsymbol{x}) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(\boldsymbol{x}) & \frac{\partial f_1}{\partial x_2}(\boldsymbol{x}) & \cdots & \frac{\partial f_1}{\partial x_n}(\boldsymbol{x}) \\ \frac{\partial f_2}{\partial x_1}(\boldsymbol{x}) & \frac{\partial f_2}{\partial x_2}(\boldsymbol{x}) & \cdots & \frac{\partial f_2}{\partial x_n}(\boldsymbol{x}) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial x_1}(\boldsymbol{x}) & \frac{\partial f_n}{\partial x_2}(\boldsymbol{x}) & \cdots & \frac{\partial f_n}{\partial x_n}(\boldsymbol{x}) \end{pmatrix}$$

Example.

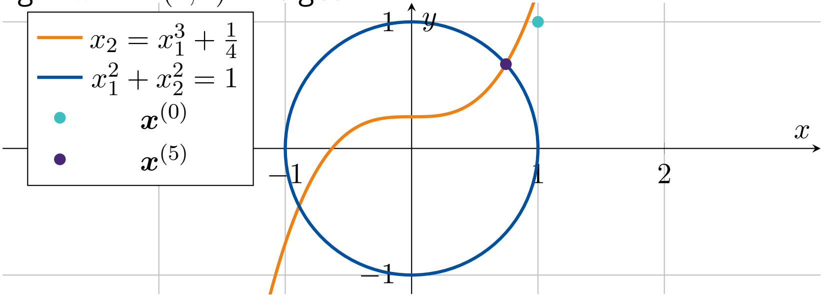
To solve our example from before

$$f(x,y) = x^3 - y + \frac{1}{4} = 0$$
$$g(x,y) = x^2 + y^2 - 1 = 0$$

we need its Jacobian

$$J(x,y) = \begin{pmatrix} 3x^2 & -1 \\ 2x & 2y \end{pmatrix}$$

Starting in $\boldsymbol{x}^{(0)} = (1,1)^{\mathrm{T}}$ we get



Error analysis for multivariate Newton's method

Error and convergence analysis for the multivariate case is beyond the scope of this lecture.

Summary. If $\boldsymbol{f}$ is sufficiently differentiable, and there is a solution $\boldsymbol{x}^*$ of the system $\boldsymbol{f}(\boldsymbol{x}) = \mathbf{0}$ with $J(\boldsymbol{x}^*)$ nonsingular, then the Newton iterations will converge quadratically towards $\boldsymbol{x}^*$ for all $\boldsymbol{x}^{(0)}$ sufficiently close to $\boldsymbol{x}^*$.

Exercise. Solve the other problem from the into

$$f(x, y) = \begin{pmatrix} x\mathrm{e}^y - 1 \\ -x^2 + y - 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$