Technische Universität Braunschweig

Institute for Numerical Analysis

Ramp Up Mathematics — Numerical Analysis Ramp Up for Data Science

Matthias Bollhöfer, SS 2024

Contents I

Contents II

Nonlinear Systems of Equations

Objective: for $D, W \subset \mathbb{R}^m$ we consider a function

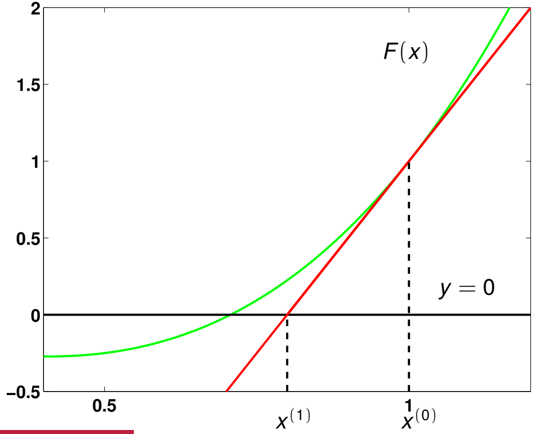
$$\begin{array}{rcl} \vec{F} : D & \longrightarrow & W \\ \begin{bmatrix} x_1 \\ \vdots \\ x_m \end{bmatrix} & \longmapsto & \begin{bmatrix} F_1(x_1, \ldots, x_m) \\ \vdots \\ F_m(x_1, \ldots, x_m) \end{bmatrix} \end{array}$$

and we want to numerically compute a solution vector $\vec{x} = [x_1, \ldots, x_m]^T$ such that $F_i(x_1, \ldots, x_m) = 0$ for all $i = 1, \cdots, m$, i.e., solve

$$\boxed{\vec{F}(\vec{x}) = \vec{0}}$$

Newton's Method

- basic idea of Newton's method: linearization
- we start with the scalar case



Newton: approximate function F by its tangent

$$F'(x^{(0)}) = \frac{F(x^{(0)}) - 0}{x^{(0)} - x^{(1)}}$$

Technische Universität Braunschweig

Newton' Method

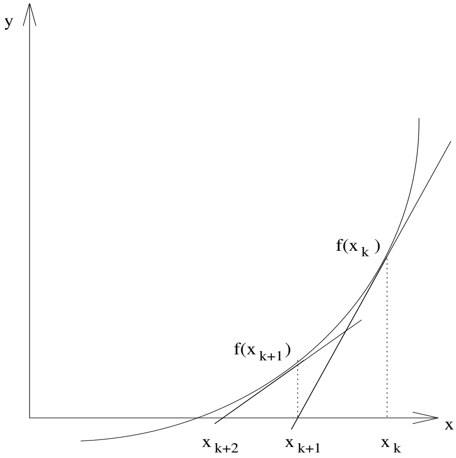
- Determine $x^{(1)}$ from tangent equation $F'(x^{(0)}) = \frac{F(x^{(0)})-0}{x^{(0)}-x^{(1)}}$:

$$\Rightarrow \quad F'(x^{(0)})(x^{(0)} - x^{(1)}) = F(x^{(0)})$$

$$\Rightarrow$$

- repeated application of this ideas yields *Newton's method* in the scalar case

$$x^{(k+1)} := x^{(k)} - \frac{F(x^{(k)})}{F'(x^{(k)})}, \; k = 0, 1, 2, \ldots$$



Newton Raphson Method in $\mathbb{R}^m$

- Newton's method for multidimensional systems in $\mathbb{R}^m$
- Jacobian

$$\nabla \vec{F}(\vec{x}) = \left(\frac{\partial F_i(\vec{x})}{\partial x_j} \right)_{i,j=1,\dots,m}$$

Example 3.1

Consider the nonlinear system

$$\vec{F}(x,y,z) = \begin{bmatrix} F_1(x,y,z) \\ F_2(x,y,z) \\ F_3(x,y,z) \end{bmatrix} = \begin{bmatrix} x+y-zx \\ 2y-zy \\ \frac{1}{2}-\frac{1}{2}x^2-\frac{1}{2}y^2 \end{bmatrix}$$

determine its Jacobian:

$$\nabla \vec{F}(x,y,z) = \begin{bmatrix} \frac{\partial F_1(x,y,z)}{\partial x} & \frac{\partial F_1(x,y,z)}{\partial y} & \frac{\partial F_1(x,y,z)}{\partial z} \\ \frac{\partial F_2(x,y,z)}{\partial x} & \frac{\partial F_2(x,y,z)}{\partial y} & \frac{\partial F_2(x,y,z)}{\partial z} \\ \frac{\partial F_3(x,y,z)}{\partial x} & \frac{\partial F_3(x,y,z)}{\partial y} & \frac{\partial F_3(x,y,z)}{\partial z} \end{bmatrix} = \begin{bmatrix} \quad & \quad & \quad \\ & & \\ & & \end{bmatrix}$$

Newton Raphson Method in $\mathbb{R}^m$

- use Taylor expansion for the scalar case $m = 1$ to approximate $\hat{x}$ such that $F(\hat{x}) = 0$, where we assume that our initial guess $x^{(0)}$ is close enough to $\hat{x}$

$$0 = F(\hat{x}) = F(x^{(0)}) + F'(x^{(0)}) \cdot (\hat{x} - x^{(0)}) + \underbrace{\mathcal{O}(\|\hat{x} - x^{(0)}\|^2)}_{\text{omit}}$$

- use Taylor expansion to approximate $\hat{\vec{x}}$ such that $\vec{F}(\hat{\vec{x}}) = \vec{0}$, again assuming that $\vec{x}^{(0)}$ is close to $\hat{\vec{x}}$

$$\underbrace{\vec{0}}_{\in\mathbb{R}^m} = \underbrace{\vec{F}(\hat{\vec{x}})}_{\in\mathbb{R}^m} = \underbrace{\vec{F}(\vec{x}^{(0)})}_{\in\mathbb{R}^m} + \underbrace{\underbrace{\nabla\vec{F}(\vec{x}^{(0)})}_{\in\mathbb{R}^{m,m}} \cdot \underbrace{(\hat{\vec{x}} - \vec{x}^{(0)})}_{\in\mathbb{R}^m}}_{\text{matrix-vector product}} + \underbrace{\mathcal{O}(\|\hat{\vec{x}} - \vec{x}^{(0)}\|^2)}_{\text{omit}}$$

- skipping the term of second order gives

$$0 \approx \vec{F}(\vec{x}^{(0)}) + \nabla\vec{F}(\vec{x}^{(0)}) \cdot \underbrace{(\vec{x}^{(1)} - \vec{x}^{(0)})}_{-\vec{d}}$$

Newton Raphson Method in $\mathbb{R}^m$

- Suppose that $\nabla \vec{F}(\vec{x}^{(0)})$ is nonsingular, then we can uniquely solve the system

$$\nabla \vec{F}(\vec{x}^{(0)})\ \vec{d} = \vec{F}(\vec{x}^{(0)})$$

  to obtain the solution $\vec{d} \equiv \vec{x}^{(0)} - \vec{x}^{(1)}$
- From this solution $\vec{d}$ we compute the next iterate $\vec{x}^{(1)} := \vec{x}^{(0)} - \vec{d}$

Algorithm 3.1 (Blue Print of the Newton Raphson Method)

Input: $\vec{F} : \mathbb{R}^m \to \mathbb{R}^m$ *with Jacobian* $\nabla \vec{F}(\vec{x})$, *initial guess* $\vec{x} \in \mathbb{R}^m$, *tolerance* ϵ
Output: *compute approximate solution* $\vec{x} \in \mathbb{R}^m$ *of* $\vec{F}(\vec{x}) = \vec{0}$

1: $\vec{b} \leftarrow \vec{F}(\vec{x})$
2: While $\|\vec{b}\| > \epsilon$
3: *determine Jacobian* $A = \nabla \vec{F}(\vec{x})$
4: *solve* $A\vec{d} = \vec{b}$ *using LU decomposition*
5: $\vec{x} \leftarrow \vec{x} - \vec{d}$
6: $\vec{b} \leftarrow \vec{F}(\vec{x})$
7: [End] While

Newton Raphson Method in $\mathbb{R}^m$

Example 3.2 (Newton Raphson Method)

Consider the nonlinear system

$$\vec{F}(x,y,z) = \begin{bmatrix} x + y - zx \\ 2y - zy \\ \frac{1}{2} - \frac{1}{2}x^2 - \frac{1}{2}y^2 \end{bmatrix}$$

its associated Jacobian is given by

$$\nabla \vec{F}(x,y,z) = \begin{bmatrix} \\ \\ \\ \end{bmatrix}$$

compute zero of $\vec{F}(x,y,z) = \vec{0}$*: suppose our initial guess is*

$$\vec{x}^{(0)} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

Newton Raphson Method in $\mathbb{R}^m$

Example 3.3 (Newton Raphson Method (continued))

we obtain

$$\vec{F}(1,0,0) = \begin{bmatrix} \\ \\ \\ \end{bmatrix}$$

solve $\nabla \vec{F}(1,0,0)d = \vec{F}(1,0,0)$, *i.e., solve*

$$\begin{bmatrix} & & \\ & & \\ & & \end{bmatrix} \begin{bmatrix} d_1 \\ d_2 \\ d_3 \end{bmatrix} = \begin{bmatrix} \\ \\ \\ \end{bmatrix}$$

thus our next iterates becomes

$$\vec{x}^{(1)} := \vec{x}^{(0)} - \vec{d} = \begin{bmatrix} \\ \\ \\ \end{bmatrix} - \begin{bmatrix} \\ \\ \\ \end{bmatrix} = \begin{bmatrix} \\ \\ \\ \end{bmatrix}$$

observe that $\vec{F}(1,0,1) = \vec{0}$, *our method terminates and* $\vec{x}^{(1)}$ *is the desired solution*

Convergence of Newton's Method

Newton Raphson Method is a fix point method

$$
\begin{aligned}
\vec{x}^{(k+1)} &= \vec{x}^{(k)} - \vec{d}, \text{ where } \nabla\vec{F}(\vec{x}^{(k)})\vec{d} = \vec{F}(\vec{x}^{(k)}) \\
&= \vec{x}^{(k)} - \nabla\vec{F}(\vec{x}^{(k)})^{-1}\vec{F}(\vec{x}^{(k)}) \\
&\equiv \vec{\Phi}(\vec{x}^{(k)})
\end{aligned}
$$

Theorem 3.1 (Banach)

Let $D \subseteq \mathbb{R}^m$ be a closed area and suppose that $\vec{\Phi} : D \longrightarrow D$ is contracting, i.e., there exists $0 < \alpha < 1$ such that

$$\|\vec{\Phi}(\vec{x}) - \vec{\Phi}(\vec{y})\| \leqslant \alpha\|\vec{x} - \vec{y}\|, \qquad \textit{for all } \vec{x}, \vec{y} \in D.$$

i) *Then there exists a unique fix point $\hat{\vec{x}} = \vec{\Phi}(\hat{\vec{x}})$ in D.*

ii) *The fix point iteration $\vec{x}^{(k+1)} = \vec{\Phi}(\vec{x}^{(k)})$ converges to $\hat{\vec{x}}$ for any initial guess $\vec{x}^{(0)} \in D$.*

Asymptotic Quadratic Convergence of Newton's Method

Theorem 3.2 (Quadratic Convergence Speed)

Let $D \subseteq \mathbb{R}^m$ be open and convex. Suppose that $\vec{F} : D \to \mathbb{R}^m$ is still continuous on the closure $\bar{D}$ of D and satisfies the following conditions:

1. *$\vec{F}$ is continuously differentiable and $\nabla \vec{F}^{-1}$ exists and is uniformly bounded, i.e.,*

$$\|\nabla \vec{F}(\vec{x})^{-1}\| \leqslant \beta.$$

2. *$\nabla \vec{F}(\vec{x})$ itself is Lipschitz continuous, i.e.,*

$$\|\nabla \vec{F}(\vec{x}) - \nabla \vec{F}(\vec{y})\| \leqslant \gamma \|\vec{x} - \vec{y}\|.$$

3. *there exists $\hat{\vec{x}}$ s.t. $\vec{F}(\hat{\vec{x}}) = \vec{0}$ and the initial guess $\vec{x}^{(0)}$ satisfies $\|\vec{x}^{(0)} - \hat{\vec{x}}\| < \varepsilon$ for sufficient small $\varepsilon > 0$ such that $h := \alpha\beta\gamma/2 < 1$, where*

$$\alpha := \|\nabla \vec{F}(\vec{x}^{(0)})^{-1} \vec{F}(\vec{x}^{(0)})\|.$$

Asymptotic Quadratic Convergence of Newton's Method

Theorem 3.3 (Quadratic Convergence Speed (ctd.))

Then the iterates $\vec{x}^{(k)}$ of the Newton Raphson method satisfy:

1. $\|\vec{x}^{(k)} - \hat{\vec{x}}\| < \varepsilon$ *and* $\lim_{k\to\infty} \vec{x}^{(k)} = \hat{\vec{x}}$
2. $$\|\vec{x}^{(k)} - \hat{\vec{x}}\| \leqslant \frac{\alpha}{h(1-h^{2^k})} \cdot h^{2^k}.$$

Roughly spoken for small h we obtain an error bound of type

$$\|\vec{x}^{(k)} - \hat{\vec{x}}\| \leqslant \frac{C}{h} \cdot h^{2^k}$$

and

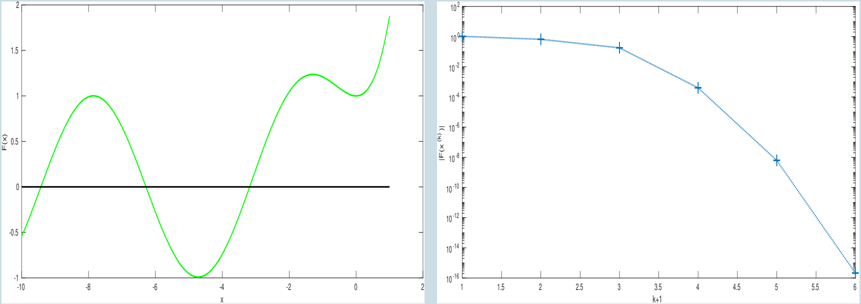
$$\|\vec{x}^{(k+1)} - \hat{\vec{x}}\| \leqslant \frac{C}{h} \cdot h^{2^{k+1}} = \frac{C}{h} \cdot h^{2^k \cdot 2} = \frac{C}{h} \cdot \left(h^{2^k}\right)^2.$$

In this sense the error bound for $\vec{x}^{(k+1)}$ is essentially the square of the error bound for $\vec{x}^{(k)}$, which can be read as *quadratic convergence*.

Asymptotic Quadratic Convergence of Newton's Method

Example 3.4 (Newton's Method (scalar case))

$F(x) = e^x - \sin x \overset{!}{=} 0$. *Use* $x^{(0)} = -2.0$ *and iterate!* $\rightarrow x^{(5)} = -3.183063011933363$



$F(x)$ *(left),* $|F(x^{(k)})|$ *(right, logarithmic scale in* y*-direction)*

Asymptotic Quadratic Convergence of Newton's Method

Example 3.5 (Newton Raphson Method)

$$\vec{F}(x,y,z) = \begin{bmatrix} F_1(x,y,z) \\ F_2(x,y,z) \\ F_3(x,y,z) \end{bmatrix} = \begin{bmatrix} x+y-zx \\ 2y-zy \\ \frac{1}{2}-\frac{1}{2}x^2-\frac{1}{2}y^2 \end{bmatrix}, \quad x^{(0)} = y^{(0)} = 1,\ z^{(0)} = 2,$$

$$x^{(5)} = y^{(5)} = 7.071067811865476e-01 \approx \frac{1}{\sqrt{2}},\ z^{(5)} = 2.0$$