

Newton - Verfahren

$x_0 \in \mathbb{R}^n$ gegeben, $\bar{F}(x) = 0$ gesucht

x besser als x_0 : $\|\bar{F}(x)\| < \|\bar{F}(x_0)\|$

Welche Punkte sind ganz sicher
besser als x_0 ?

hängt von der Wahl
der Norm ab!

$$M = \{x \mid \|\bar{F}(x)\| < \|\bar{F}(x_0)\| \quad \forall \| \cdot \| \}$$

Sei: $\bar{F}(x) = \lambda \bar{F}(x_0) = \|\bar{F}(x)\| = |\lambda| \|\bar{F}(x_0)\| \quad \forall \lambda |$
 $|\lambda| < 1 \Rightarrow x$ ist sicher besser als x_0

$$M = \{x \mid \bar{F}(x) = (\lambda) \bar{F}(x_0), \lambda \in [0, 2]\}$$

nur ein Parameter

→ M ist ein Pfad

$$x(\lambda) : [0, 2] \rightarrow \mathbb{R}^n$$

Newton-Pfad

$$x(\lambda)$$

$x(x_0)$

$x(x_1)$

$x(x_2)$

$x(x_3)$

$x(x_4)$

$x(x_5)$

$x(x_6)$

$x(x_7)$

$x(x_8)$

$x(x_9)$

$x(x_{10})$

$x(x_{11})$

$x(x_{12})$

$x(x_{13})$

$x(x_{14})$

$x(x_{15})$

$x(x_{16})$

$x(x_{17})$

$x(x_{18})$

$x(x_{19})$

$x(x_{20})$

$x(x_{21})$

$x(x_{22})$

$x(x_{23})$

$x(x_{24})$

$x(x_{25})$

$x(x_{26})$

$x(x_{27})$

$x(x_{28})$

$x(x_{29})$

$x(x_{30})$

$x(x_{31})$

$x(x_{32})$

$x(x_{33})$

$x(x_{34})$

$x(x_{35})$

$x(x_{36})$

$x(x_{37})$

$x(x_{38})$

$x(x_{39})$

$x(x_{40})$

$x(x_{41})$

$x(x_{42})$

$x(x_{43})$

$x(x_{44})$

$x(x_{45})$

$x(x_{46})$

$x(x_{47})$

$x(x_{48})$

$x(x_{49})$

$x(x_{50})$

$x(x_{51})$

$x(x_{52})$

$x(x_{53})$

$x(x_{54})$

$x(x_{55})$

$x(x_{56})$

$x(x_{57})$

$x(x_{58})$

$x(x_{59})$

$x(x_{60})$

$x(x_{61})$

$x(x_{62})$

$x(x_{63})$

$x(x_{64})$

$x(x_{65})$

$x(x_{66})$

$x(x_{67})$

$x(x_{68})$

$x(x_{69})$

$x(x_{70})$

$x(x_{71})$

$x(x_{72})$

$x(x_{73})$

$x(x_{74})$

$x(x_{75})$

$x(x_{76})$

$x(x_{77})$

$x(x_{78})$

$x(x_{79})$

$x(x_{80})$

$x(x_{81})$

$x(x_{82})$

$x(x_{83})$

$x(x_{84})$

$x(x_{85})$

$x(x_{86})$

$x(x_{87})$

$x(x_{88})$

$x(x_{89})$

$x(x_{90})$

$x(x_{91})$

$x(x_{92})$

$x(x_{93})$

$x(x_{94})$

$x(x_{95})$

$x(x_{96})$

$x(x_{97})$

$x(x_{98})$

$x(x_{99})$

$x(x_{100})$

$x(x_{101})$

$x(x_{102})$

$x(x_{103})$

$x(x_{104})$

$x(x_{105})$

$x(x_{106})$

$x(x_{107})$

$x(x_{108})$

$x(x_{109})$

$x(x_{110})$

$x(x_{111})$

$x(x_{112})$

$x(x_{113})$

$x(x_{114})$

$x(x_{115})$

$x(x_{116})$

$x(x_{117})$

$x(x_{118})$

$x(x_{119})$

$x(x_{120})$

$x(x_{121})$

$x(x_{122})$

$x(x_{123})$

$x(x_{124})$

$x(x_{125})$

$x(x_{126})$

$x(x_{127})$

$x(x_{128})$

$x(x_{129})$

$x(x_{130})$

$x(x_{131})$

$x(x_{132})$

$x(x_{133})$

$x(x_{134})$

$x(x_{135})$

$x(x_{136})$

$x(x_{137})$

$x(x_{138})$

$x(x_{139})$

$x(x_{140})$

$x(x_{141})$

$x(x_{142})$

$x(x_{143})$

$x(x_{144})$

$x(x_{145})$

$x(x_{146})$

$x(x_{147})$

$x(x_{148})$

$x(x_{149})$

$x(x_{150})$

$x(x_{151})$

$x(x_{152})$

$x(x_{153})$

$x(x_{154})$

$x(x_{155})$

$x(x_{156})$

$x(x_{157})$

$x(x_{158})$

$x(x_{159})$

$x(x_{160})$

$x(x_{161})$

$x(x_{162})$

$x(x_{163})$

$x(x_{164})$

$x(x_{165})$

$x(x_{166})$

$x(x_{167})$

$x(x_{168})$

$x(x_{169})$

$x(x_{170})$

$x(x_{171})$

$x(x_{172})$

$x(x_{173})$

$x(x_{174})$

$x(x_{175})$

$x(x_{176})$

$x(x_{177})$

$x(x_{178})$

$x(x_{179})$

$x(x_{180})$

$x(x_{181})$

$x(x_{182})$

$x(x_{183})$

$x(x_{184})$

$x(x_{185})$

$x(x_{186})$

$x(x_{187})$

$x(x_{188})$

$x(x_{189})$

$x(x_{190})$

$x(x_{191})$

$x(x_{192})$

$x(x_{193})$

$x(x_{194})$

$x(x_{195})$

$x(x_{196})$

$x(x_{197})$

$x(x_{198})$

$x(x_{199})$

$x(x_{200})$

$x(x_{201})$

$x(x_{202})$

$x(x_{203})$

$x(x_{204})$

$x(x_{205})$

$x(x_{206})$

$x(x_{207})$

$x(x_{208})$

$x(x_{209})$

$x(x_{210})$

$x(x_{211})$

$x(x_{212})$

$x(x_{213})$

$x(x_{214})$

$x(x_{215})$

$x(x_{216})$

$x(x_{217})$

$x(x_{218})$

$x(x_{219})$

$x(x_{220})$

$x(x_{221})$

$x(x_{222})$

$x(x_{223})$

$x(x_{224})$

$x(x_{225})$

$x(x_{226})$

$x(x_{227})$

$x(x_{228})$

$x(x_{229})$

$x(x_{230})$

$x(x_{231})$

$x(x_{232})$

\Rightarrow Newton-Schritte sind tangential
zum Newton-Pfad

- $x' = -F'(x)^{-1} F(x)$ ist die Davidenko-Differentialgleichung
- Newton-Verfahren ist das expl. Euler-Verfahren für diese ODE.

Schrittweite im Euler-Verfahren zu groß \rightarrow Fehler ist groß $O(\Delta^2)$
 \rightarrow Schrittweitensteuerung

Es gilt $F(x(\lambda)) = (1-\lambda)F(x_0) + O(\lambda^2)$ und daher
für jede feste Norm:

$$\exists \bar{\lambda} > 0 : \forall \lambda \in]0, \bar{\lambda}]: \|F(x(\lambda))\| \leq \|F(x_0)\|$$

Adaptive Schrittweitenwahl

Modell: $\|F(x + \lambda \Delta x) - (1-\lambda)F(x)\| \leq \frac{\omega}{2} \lambda^2 \|F(x)\|$

$$\begin{aligned}\Rightarrow \|F(x + \lambda \Delta x)\| &= \|F(x + \lambda \Delta x) - (1-\lambda)F(x) + (1-\lambda)F(x)\| \\ &\leq \|F(x + \lambda \Delta x) - (1-\lambda)F(x)\| + |1-\lambda| \|F(x)\| \\ &= \left(|1-\lambda| + \frac{\omega}{2} \lambda^2 \right) \|F(x)\|\end{aligned}$$

theoretisch optimale Schrittweite minimiert $(|1-\lambda| + \frac{\omega}{2} \lambda^2)$

$$\text{Minimierer ist } \lambda_{\text{opt}} = \min(1, \frac{1}{\omega})$$

Algorithmus

gegeben: ω (Vorschlag, Schätzer)

löse $F'(x) \Delta x = -F(x)$

do

$$\lambda = \min(1, \frac{1}{\omega})$$

$$x_1 = x + \lambda \Delta x$$

berechne $F(x_1)$

$$\omega := \frac{\omega}{e^{\frac{\|F(x_1) - (1-\lambda)F(x)\|}{\|F(x)\|}}}$$

until $\|F(x_1)\| < \|F(x)\|$

$$x := x_1$$

ω verwenden als Vorschlag für nächsten Schritt

ganz zu Beginn: ω wird vom Benutzer vorgeschlage

$\omega \approx 1$: mild nichtlinear

$\omega \gg 1$: stark nichtlinear

Inexakte Newtonverfahren

$$F'(x) \Delta x = -F(x) + r$$

r : Residuum, $\|r\| \leq \text{TOL} \|F(x)\|$

(iterativer Lösung, Diskretisierung, abweichende Ableitung
(z.B. bei Verwendung variger F'))

$$\begin{aligned} F(x + \lambda \Delta x) &= F(x) + \int_{s=0}^{\lambda} F'(x + s \Delta x) \Delta x \, ds \\ &= F(x) + \int_{s=0}^{\lambda} (F'(x + s \Delta x) - F'(x)) \Delta x + \underbrace{F'(x) \Delta x}_{= -F(x) + r} \, ds \\ &= (1-\lambda) F(x) + \int_{s=0}^{\lambda} (F'(x + s \Delta x) - F'(x)) \Delta x \, ds + \lambda r \end{aligned}$$

$$\text{Wegen } \|F'(x + s\delta_x) - F'(x)\| \leq \beta s \|\delta_x\|$$

$$\text{und } \|\delta_x\| \leq \|F'(x)^{-1}\| \cdot \| -F(x)_{+r} \| \leq \gamma (1+\text{TOL}) \|F(x)\|$$

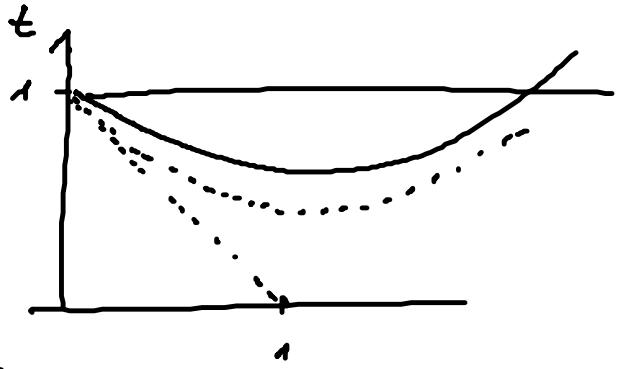
gilt

$$\begin{aligned} \|F(x + \lambda \delta_x)\| &\leq |1-\lambda| \|F(x)\| + \int_0^\lambda \beta s \gamma^2 (1+\text{TOL})^2 \|F(x)\|^2 ds \\ &= \underbrace{\left(1 - (1-\text{TOL})\lambda + \beta \gamma^2 (1+\text{TOL})^2 \|F(x)\|^2 \frac{\lambda^2}{2}\right)}_{t(\lambda)} \|F(x)\| \end{aligned}$$

Interpretation:

$$\text{TOL} < 1 \Rightarrow \exists \lambda \in [0, 1]: t(\lambda) < 1$$

\rightarrow Konvergenz



gewöhnliches iteratives Newton-Verfahren: $\lambda = 1$

$$\Rightarrow \|F(x + \delta_x)\| \leq \left[\text{TOL} + \frac{\beta \gamma^2 (1+\text{TOL})^2}{2} \|F(x)\|^2 \right] \|F(x)\|$$

$$\text{TOL} = O(\|F\|) \Rightarrow \|F(x + \delta_x)\| = O(\|F(x)\|^2)$$

quadratische Konvergenz

QR-Zerlegung mit Givens (1953)

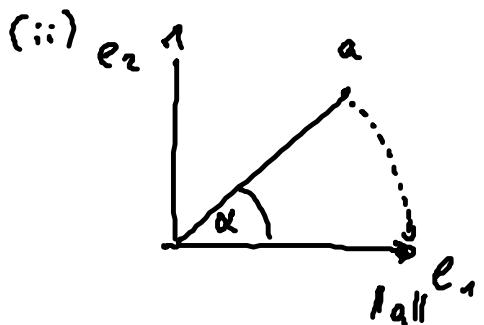
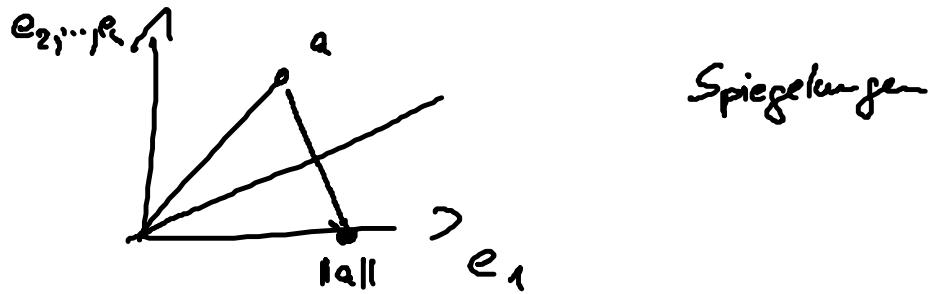
Prinzip QR:

$$Q = Q_1 \dots Q_k \text{ mit } QR = A \quad R = \begin{pmatrix} * & * & \dots & * \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix}$$

Q_i : elementare orthonormale Matrizen

Elementare Orthonormalmatrizen

(i) Householder-Reflektionen



Drehungen
Givens-Rotationen

$$Q \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} r \\ 0 \end{pmatrix} \quad r = \| \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \|$$

$$Q = \begin{pmatrix} c & s \\ -s & c \end{pmatrix} \quad c = \cos \alpha \quad s = \sin \alpha$$

aus $\begin{pmatrix} c & s \\ -s & c \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} r \\ 0 \end{pmatrix}$ und $c^2 + s^2 = 1$

folgt $c \frac{a}{r} + s \frac{b}{r} = 1 = c^2 + s^2$

$$\Rightarrow c = \frac{a}{r}, s = \frac{b}{r}, r = \sqrt{a^2 + b^2}$$

QR-Zerlegung

$$= \begin{pmatrix} * & * & * \\ * & * & * \\ 1 & & * \end{pmatrix}$$