Newton-Verfahren

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

x besser als x_0 : $\|F(x)\| < \|F(x_0)\|$ ← hängt von der Wahl der Norm ab!

Welche Punkte sind ganz sicher besser als x_0?

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

Sei: $F(x) = \lambda F(x_0) \Rightarrow \|F(x)\| = |\lambda| \, \|F(x_0)\| \quad \forall \|\cdot\|$

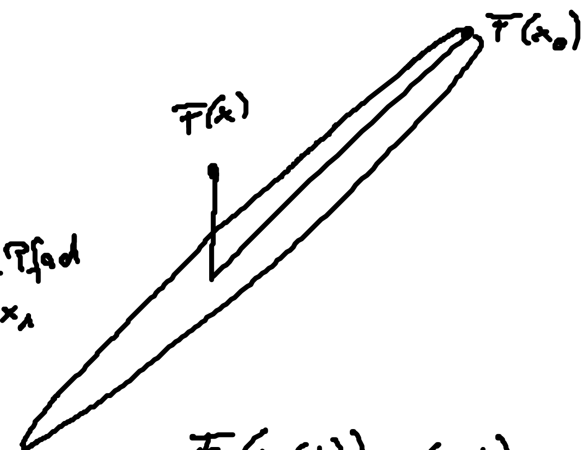
$|\lambda| < 1 \Rightarrow x$ ist sicher besser als x_0

$$M = \{ x \mid F(x) = (1-\lambda) F(x_0), \ \lambda \in [0,2] \}$$

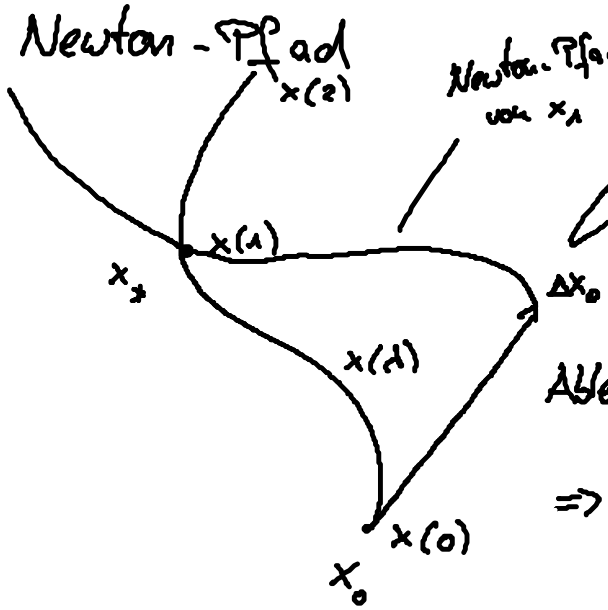
nur ein Parameter

$\longrightarrow M$ ist ein Pfad

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



Newton-Pfad



$$F(x(\lambda)) = (1-\lambda) F(x_0)$$

Ableiten: $F'(x(\lambda)) \cdot x'(\lambda) = -F(x_0)$

$$\Rightarrow F'(x(\lambda)) \cdot x'(\lambda) = -\frac{1}{1-\lambda} F(x(\lambda))$$

$\Rightarrow x'(\lambda) = \Delta x$ ist skalierte Newton-Korrektur in $x(\lambda)$

$\Rightarrow$ 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 $\longrightarrow$ Fehler ist groß $O(\lambda^2)$

$\longrightarrow$ Schrittweitensteuerung

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

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

<u>Adaptive Schrittweitenwahl</u>

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

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

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

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

Algorithmus

gegeben: ω (Vorschlag, Schätzer)

löse $F'(x)\Delta x = -F(x)$

do

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

$x_1 = x + \lambda \Delta x$

berechne $F(x_1)$

$\omega := \frac{2}{\lambda^2} \frac{\|F(x_1) - (1-\lambda)F(x)\|}{\|F(x)\|}$

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

$x := x_1$

ω verwenden als Vorschlag für nächsten Schritt

ganz zu Beginn: ω wird vom Benutze vorgeschlage

$\omega \approx 1$: mild nichtlinear

$\omega \gg 1$: stark nichtlinear

Inexakte Newtonverfahren

$F'(x)\,\delta x = -F(x) + r$

r: Residuum, $\|r\| \le TOL\,\|F(x)\|$

(iterativer Lösung, Diskretisierung, abweichender Ableitung (z.B. bei Verwendung voriger F'))

$$F(x+\lambda\delta x) = F(x) + \int_{s=0}^{\lambda} F'(x+s\delta x)\,\delta x\,ds$$

$$= F(x) + \int_{s=0}^{\lambda} \left(F'(x+s\delta x) - F'(x)\right)\delta x + \underbrace{F'(x)\,\delta x}_{=-F(x)+r}\,ds$$

$$= (1-\lambda)F(x) + \int_{s=0}^{\lambda} \left(F'(x+s\delta x) - F'(x)\right)\delta x\,ds + \lambda r$$

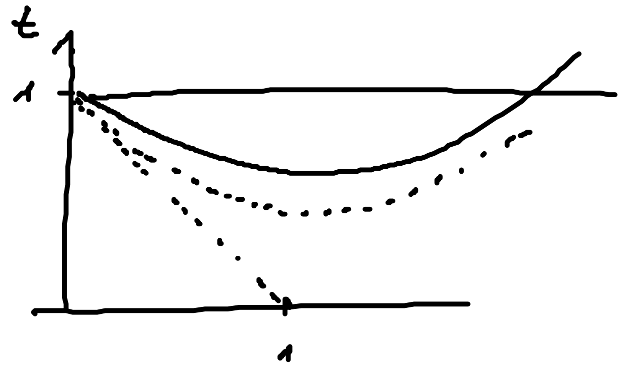
Wegen $\|F'(x+s\delta x) - F'(x)\| \le \beta s \|\delta x\|$

und $\|\delta x\| \le \|F'(x)^{-1}\| \cdot \|-F(x)+r\| \le \gamma\,(1+TOL)\|F(x)\|$

gilt

$$\|F(x+\lambda\delta x)\| \le |1-\lambda|\, \|F(x)\| + \int_{s=0}^{\lambda} \beta s \gamma^2 (1+TOL)^2 \|F(x)\|^2\, ds + \lambda\, TOL\, \|F(x)\|$$

$$= \underbrace{\left(1-(1-TOL)\lambda + \beta\gamma^2(1+TOL)^2\|F(x)\|\frac{\lambda^2}{2}\right)}_{t(\lambda)} \|F(x)\|$$



Interpretation:

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

$\to$ Konvergenz

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

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

$TOL = O(\|F\|) \Rightarrow \|F(x+\delta x)\| = O(\|F(x)\|^2)$ quadratische Konvergenz

QR-Zerlegung mit Givens (1953)

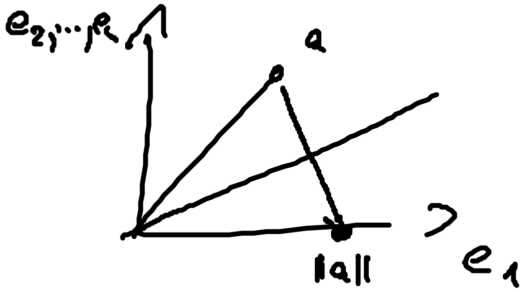
Prinzip QR:

$Q = Q_1 \cdots Q_k$ mit $QR = A$ $R=$

Q_i : elementare orthonormale Matrizen

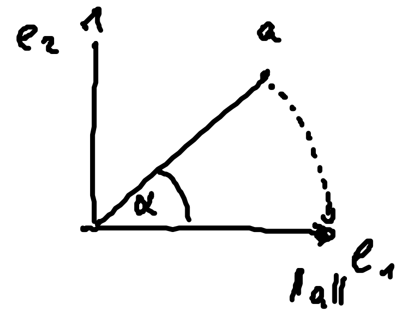
Elementare Orthonormalmatrizen

(i) Householder-Reflexionen



Spiegelungen

(ii)



Drehungen

Givens-Rotationen

$$Q \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} r \\ 0 \end{bmatrix} \qquad r = \left\| \begin{matrix} a_1 \\ a_2 \end{matrix} \right\|$$

$$Q = \begin{bmatrix} c & s \\ -s & c \end{bmatrix} \qquad \begin{matrix} c = \cos\alpha \\ s = \sin\alpha \end{matrix}$$

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

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

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

QR-Zerlegung

$$\begin{bmatrix} * & * & * \\ & * & * \\ & & * \\ \end{bmatrix}$$