VI. Iterationsverfahren

To infinity and beyond

Falls eine direkte Lösung des Problems nicht möglich oder ineffizient ist.

6.1. Fixpunktgleichungen

6.1.1. Problemstellung:

Iterationsfunktion $\Phi(x)$

Iteration: $x_0 \in R$ Startwert,

$x_{k+1} := \Phi(x_k) \in \Re$, k=0,1,...

Dadurch ist eine Folge x_k definiert.

Ist diese Folge konvergent:

$$x_k \to \bar{x} \quad \text{für } k \to \infty,$$

so folgt für eine stetige Funktion Φ:

$$\bar{x} = \lim_{k\to\infty} x_k = \lim_{k\to\infty} \Phi(x_{k-1}) = \Phi(\lim_{k\to\infty} x_{k-1}) = \Phi(\bar{x})$$

$\bar{x}$ heißt Fixpunkt von Φ in **R**.

Beispiel: $\Phi(x) = x^2$, also $\quad x_{k+1} := \Phi(x_k) = (x_k)^2$
oder

$$x_k = x_{k-1}^2 = x_{k-2}^4 = \cdots = x_0^{\left(2^k\right)}$$

Das Konvergenzverhalten hängt vom Startwert x_0 ab :

$$
\begin{array}{lllll}
x_0 = 0 & \Rightarrow & x_k = 0, & k = 1,2,\ldots, & \bar{x}_1 = 0 \\
x_0 = \pm 1 & \Rightarrow & x_k = 1, & k = 1,2,\ldots, & \bar{x}_2 = 1 \\
0 < |x_0| < 1 & \Rightarrow & x_k = x_0^{2^k} \to 0 & k \to \infty, & \bar{x}_1 = 0 \\
|x_0| > 1 & \Rightarrow & x_k = x_0^{2^k} \to \infty & k \to \infty, & (\bar{x}_3 = \infty)
\end{array}
$$

Φ hat daher den Fixpunkte 0, bzw. ist divergent (könnte man als Fixpunkt ∞ bezeichnen).

Die Folgen sind jeweils monoton für $k > 0$.

1 ist ebenfalls Fixpunkt, kann aber nur vorkommen, wenn man mit ± 1 startet; 1 heißt daher <u>abstoßender</u> Fixpunkt.

6.1.2. Beispiel: Ausbreitung eines Grippevirus in einem Kindergarten

Zu Zeitpunkt t_i bezeichnen wir mit k_i die relative Anzahl erkrankter Kinder, also

k_i = # kranke Kinder / # Kinder.

t_i ist diskrete Folge von Zeitpunkten.

Infektionsrate sei $\alpha > 1$, (= Übertragungswahrscheinlichkeit)

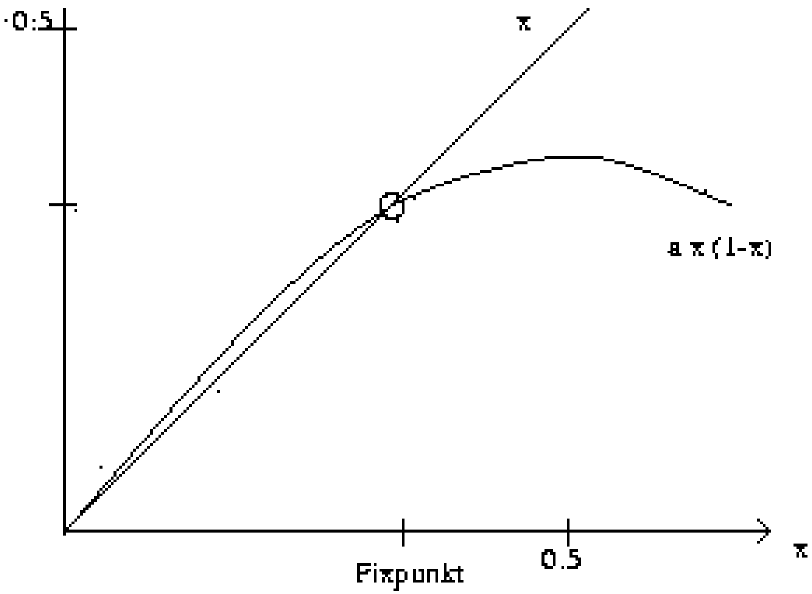
Bei jeder Kontaktaufnahme zweier Kinder kann daher eine Virusübertragung stattfinden; daher ist die Zahl neu Erkrankter zum nächsten Zeitpunkt direkt proportional zur Zahl der möglichen Begegnungen zwischen einem kranken und einem gesunden Kind. Ein Kind, das zum Zeitpunkt t_i krank ist, ist zum Zeitpunkt t_{i+1} wieder gesund, kann sich aber wieder anstecken (Diese Annahme dient der Vereinfachung des Modells).

$$k_{i+1} = \alpha\, k_i (1 - k_i)$$

= α* *#Kranke * #Gesunde*

Dazugehörige Iterationsfunktion: $\Phi(x) = \alpha\ x(1-x)$

beschreibt eine konkave Parabel, die logistische Parabel.



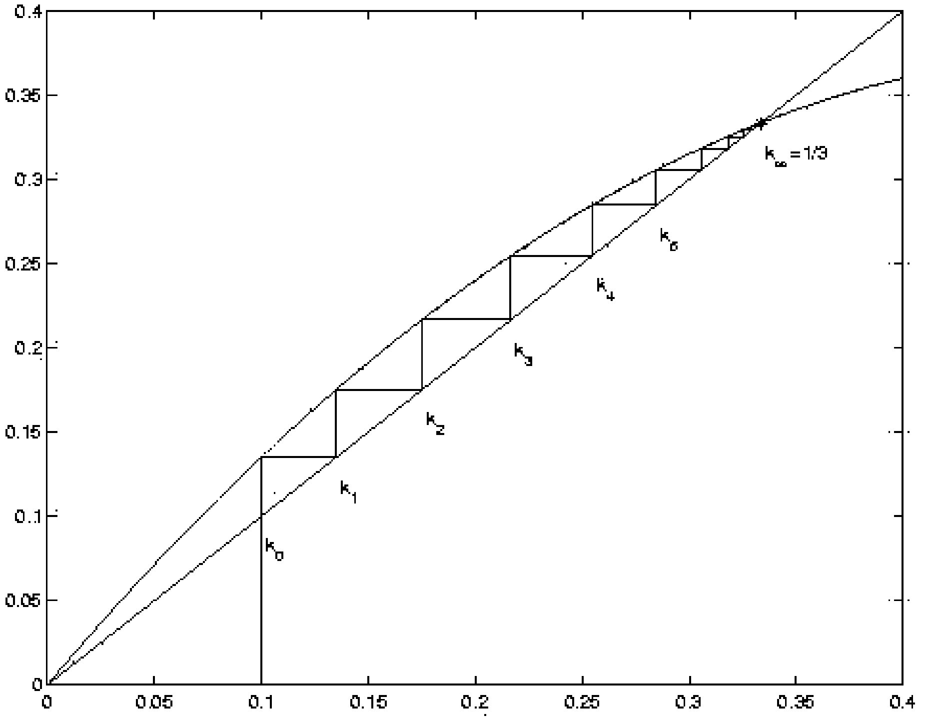
Relatives Maximum bei x = 0.5 mit Wert Φ(0.5) = 0.25α;

Nullstellen bei x = 0 und x = 1;

Fixpunkt als Schnittpunkt von Φ mit der Funktion $g(x) \equiv x$:

$$\bar{x} = \Phi(\bar{x}) = \alpha\bar{x}(1-\bar{x}) \quad \Rightarrow \quad \bar{x} = \frac{\alpha - 1}{\alpha}$$

Für $1 < \alpha < 2$ existiert genau ein eindeutiger Fixpunkt dieser Iteration zwischen 0 und 0.5.



z.B. $x_0=0.1$ und $\alpha=1.5$

Konvergiert monoton wachsend gegen 1/3

Fixpunkt geometrisch:

Schnittpunkt zwischen Iterationsfunktion $\Phi(x)$ und Gerade $g(x) \equiv x$.

6.1.3. Banach'scher Fixpunktsatz

Frage:

Wann konvergiert die so erzeugte Folge und wann nicht?

Welche Eigenschaften müssen Φ und x_0 haben, damit Konvergenz gegen einen (ev. eindeutigen) Fixpunkt vorliegt?

***Banach'scher Fixpunktsatz*:**

Sei I ein abgeschlossenes Intervall, Startwert $x_0 \in I$,

Φ eine Abbildung $\Phi: I \rightarrow I$, d.h. $\Phi(I) \subseteq I$,

Φ sei in I eine kontrahierende Abbildung, d.h. es gibt eine

Konstante 0 < L < 1 mit $x, y \in I \Rightarrow \left|\Phi(x) - \Phi(y)\right| \leq L \cdot \left|x - y\right|.$

Dann konvergiert die Folge $x_{k+1} := \Phi(x_k)$

gegen den eindeutigen Fixpunkt $\bar{x} \in I$,

also $x_k \rightarrow \bar{x} = \Phi(\bar{x})$ für $k \rightarrow \infty$.

Beweis: Offensichtlich gilt stets $x_k \in I$; daher folgt

$$\left|x_{k+1} - x_k\right| = \left|\Phi(x_k) - \Phi(x_{k-1})\right| \leq L\left|x_k - x_{k-1}\right| \leq$$

$$\leq L\left(L\left|x_{k-1} - x_{k-2}\right|\right) \leq \cdots \leq L^k\left|x_1 - x_0\right|.$$

Damit ergibt sich für den Abstand zweier Iterierter x_k und x_m für $m>k$:

$$\begin{aligned}
|x_m - x_k| &= |(x_m - x_{m-1}) + (x_{m-1} - x_{m-2}) + \cdots + (x_{k+1} - x_k)| \le \\
&\le |x_m - x_{m-1}| + |x_{m-1} - x_{m-2}| + \cdots + |x_{k+1} - x_k| \le \\
&\le \left(L^{m-1} + L^{m-2} + \cdots + L^k\right) \cdot |x_1 - x_0| \le \\
&\le \sum_{j=k}^{\infty} L^j \cdot |x_1 - x_0| = \frac{L^k}{1-L} \cdot |x_1 - x_0|.
\end{aligned}$$

Daher wird der Abstand zwischen zwei Iterierten beliebig klein, wenn sie beide großen Index haben:

$$|x_m - x_k| \le \varepsilon \qquad \mathit{für} \qquad m,k \ge N(\varepsilon)$$

Die Zahlen x_k bilden daher eine Cauchy-Folge!

Cauchy-Folgen sind konvergent in I, da I abgeschlossen ist!

(dies ist genau die mathematische Definition von Abgeschlossenheit).

Daher existiert eine Zahl $\bar{x} \in I$ mit $x_k \rightarrow \bar{x}$ für $k \rightarrow \infty$.

$\overline{x}$ ist der gesuchte Fixpunkt in I, denn

$$0 \le \left|\overline{x} - \Phi(\overline{x})\right| \le \left|\overline{x} - x_k\right| + \left|x_k - \Phi(\overline{x})\right| \le$$
$$\le \left|\overline{x} - x_k\right| + \left|\Phi(x_{k-1}) - \Phi(\overline{x})\right| \le$$
$$\le \left|\overline{x} - x_k\right| + L\left|x_{k-1} - \overline{x}\right| \to 0 \qquad \textit{für} \quad k \to \infty$$

Daher muss gelten $\quad \overline{x} - \Phi(\overline{x}) = 0.$

Dieser Fixpunkt in I ist eindeutig, denn

$$\overline{x} = \Phi(\overline{x}) \in I \qquad und \quad \hat{x} = \Phi(\hat{x}) \in I \qquad mit \quad \overline{x} \ne \hat{x} \quad \Rightarrow$$
$$\left|\overline{x} - \hat{x}\right| = \left|\Phi(\overline{x}) - \Phi(\hat{x})\right| \le L \cdot \left|\overline{x} - \hat{x}\right| < \left|\overline{x} - \hat{x}\right| \qquad\qquad falsch\ !$$

Widerspruch, daher Annahme $\overline{x} \ne \hat{x}$ falsch, also $\overline{x} = \hat{x}$.

Außerdem gilt: x_k konvergiert *linear* gegen $\overline{x}$, denn

$$\underline{\left|x_{k+1}-\overline{x}\right|} = \left|\Phi(x_k)-\Phi(\overline{x})\right| \le \underline{L\left|x_k-\overline{x}\right|} \le$$

$$\le L^{k+1}\left|x_0-\overline{x}\right| \to 0, \quad L<1$$

#

Um den Satz anwenden zu können benötigt man also,
dass Φ ein Intervall in sich abbildet, und zwar so,
dass die Funktionswerte dabei näher zusammenrücken.

L < 1 heißt Lipschitz-Konstante.

Ist Φ diff'bar, so lassen sich diese Bedingungen mit Hilfe des Mittelwertsatzes vereinfachen:

Für x,y$\in$I ist dann nämlich

$$\Phi(y) = \Phi(x) + \Phi'(z) \cdot (y - x)$$

mit einer Zahl z zwischen x und y.

Daher gilt: $$\left|\Phi(y) - \Phi(x)\right| \le \max_{z \in I} \left|\Phi'(z)\right| \cdot \left|y - x\right|$$

Ist nun $|\Phi'(z)| < 1$ in I, so kann man

$$L := \max_{z \in I} \left|\Phi'(z)\right|$$ setzen.

Wir betrachten den Mittelwertsatz jetzt speziell in der Nähe eines Fixpunktes x . Dann gilt:

6.1.4 Sei $|\Phi'(\overline{x})| < 1$; dann kann man eine Umgebung U von $\overline{x}$ angeben, in der Φ kontrahierend ist und $\Phi(U) \subseteq U$ gilt (es liegt dann also lokale Konvergenz vor).

Beweis: Setze U:= [$\bar{x}$-h , $\bar{x}$+h] und wähle dabei h > 0 so klein,

dass immer noch $L := \max_{z \in U} |\Phi'(z)| < 1$ gilt (Φ' ist ja stetig!).

Für ein x∈U folgt dann $|\Phi(x) - \bar{x}| = |\Phi(x) - \Phi(\bar{x})| \leq L \cdot |x - \bar{x}| < h$

und daher ist auch Φ(x)∈U.

Also insgesamt: Φ(U) ⊆ U und außerdem ist Φ kontrahierend in U.
Daher ist der Banach'sche Fixpunktsatz anwendbar.

#

Einen Fixpunkt $\bar{x}$ mit |Φ'($\bar{x}$)| < 1 nennt man
‚anziehenden Fixpunkt'.

Im Beispiel x^2 ist 0 anziehender Fixpunkt, da $\Phi'(0)=2*0 = 0<1$

Daher gilt für $L = 0.5$ und $U = [-0.25 , 0.25]$, dass $\Phi(x) = x^2$ in U eine kontrahierende Selbstabbildung von U ist → mit dem Fixpunktsatz von Banach: Konvergenz in U gegen Fixpunkt 0!

Andererseits heißt ein Fixpunkt $\bar{x}$ mit $|\Phi'(\bar{x})| > 1$

abstoßender Fixpunkt,

da keine kontrahierende Umgebung für Φ existiert.

Im Beispiel x^2 ist 1 abstoßender Fixpunkt,

da $\Phi'(1) = 2*1 = 2 > 1$.

Beispiel Grippevirus: $\Phi(x) = \alpha x(1-x)$

mit Fixpunkt $\bar{x} = (\alpha - 1)/\alpha$

Dann gilt $$\Phi'(\bar{x}) = \alpha(1-2\bar{x}) = \alpha\left(1 - 2\frac{\alpha - 1}{\alpha}\right) = 2 - \alpha$$

$\bar{x}$ =(1-α)/α ist anziehender Fixpunkt für 1 < α < 3.

Für α>=3 oder α<= 1 ergeben sich abstoßende Fixpunkte!
Keine Konvergenz in diesen Fällen!

Wegen $$x_{k+1} = \Phi(x_k) = \bar{x} + \Phi'(z)(x_k - \bar{x})$$

gilt außerdem $$x_{k+1} - \bar{x} = \Phi'(z)(x_k - \bar{x})$$

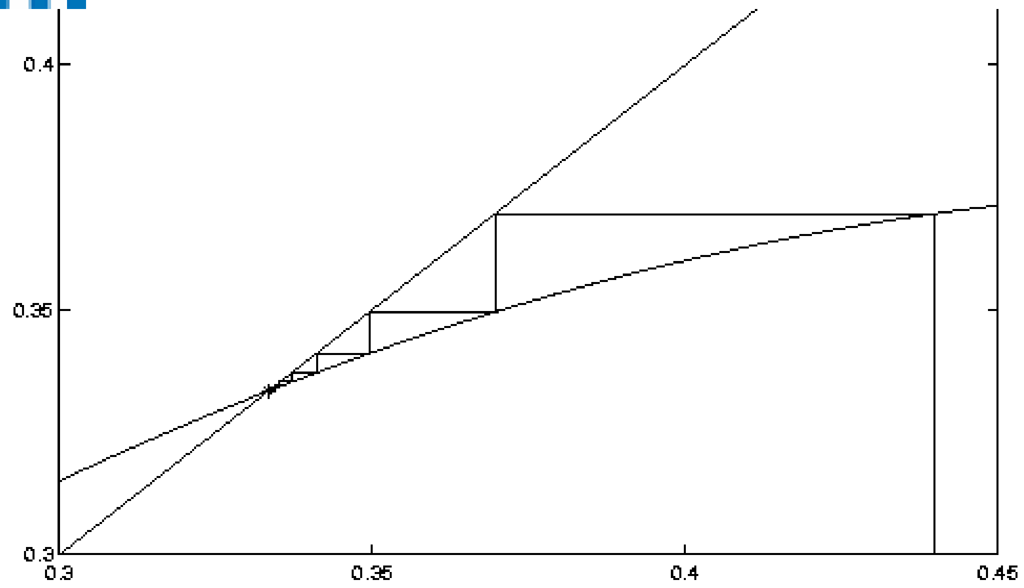
Daher ist für $1 < \alpha < 2$ $\Phi'(\bar{x}) > 0$, und die Folge x_k ist lokal monoton wachsend oder fallend, je nach Startwert rechts oder links vom Fixpunkt.

$$x_{k+1} - \bar{x} = \Phi'(z)(x_k - \bar{x})$$

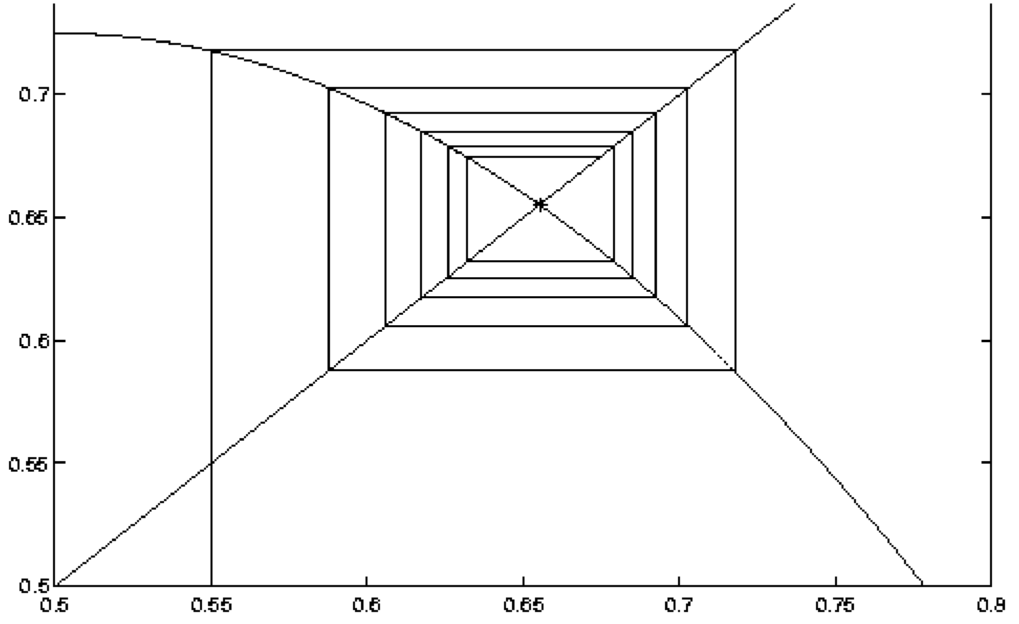
(die x_k liegen stets auf derselben Seite, rechts oder links vom Fixpunkt)

Dagegen konvergiert für $2 < \alpha < 3$ die Folge alternierend, da $\Phi'(\bar{x}) < 0$ ist.

Beispiele:



α = 1.5 und
Startwert 0.44 →
monoton fallende
Konvergenz



α = 2.9 und
Startwert 0.55 →
alternierende
Konvergenz

6.2. Das Newtonverfahren zur Nullstellenbestimmung

Gesucht sind Nullstellen einer nichtlinearen (stetig diff'baren) Funktion f:R→R, also $\overline{x} \in R$ mit $f(\overline{x}) = 0$!

Zurückführung des Nullstellenproblems auf das inzwischen bekannte Fixpunktproblem.

Also gesucht: $x_k \to \overline{x}$ für $k \to \infty$

Betrachte dazu die Taylorentwicklung in letzter Iterierten $\mathbf{x_k}$:

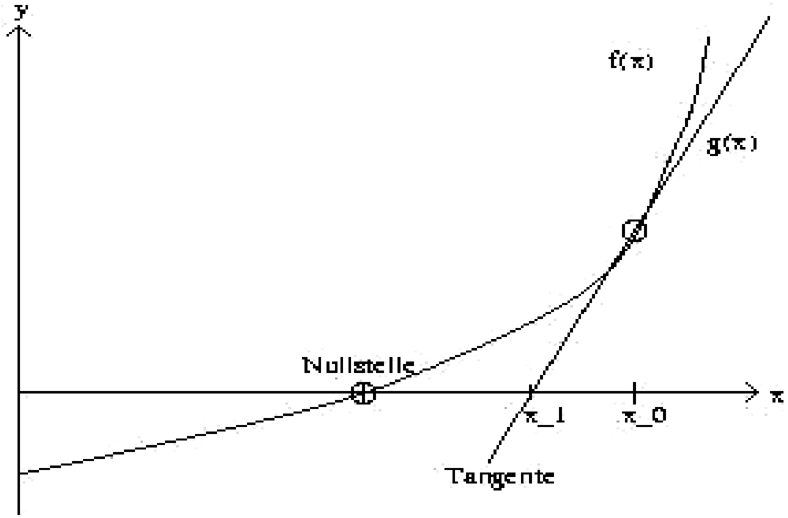
$$0 = f(\overline{x}) = f(x_k) + f'(x_k) \cdot (\underbrace{\overline{x}}_{x_{k+1}} - x_k) + \underbrace{O\left((\overline{x} - x_k)^2\right)}_{\text{wird ignoriert}}.$$

Auflösen nach $\overline{x} \to x_{k+1}$ liefert die

Newton-Iteration:

$$x_{k+1} := x_k - \frac{f(x_k)}{f'(x_k)} \qquad \textit{falls} \quad f'(x_k) \neq 0$$

6.2.1. Geometrische Interpretation des Newtonverfahrens



Ersetze f(x) lokal an der Stelle x_k durch bestmögliche Gerade = Tangente, entspricht linearem Anteil der Taylorreihe.

Die Nullstelle dieser Geraden

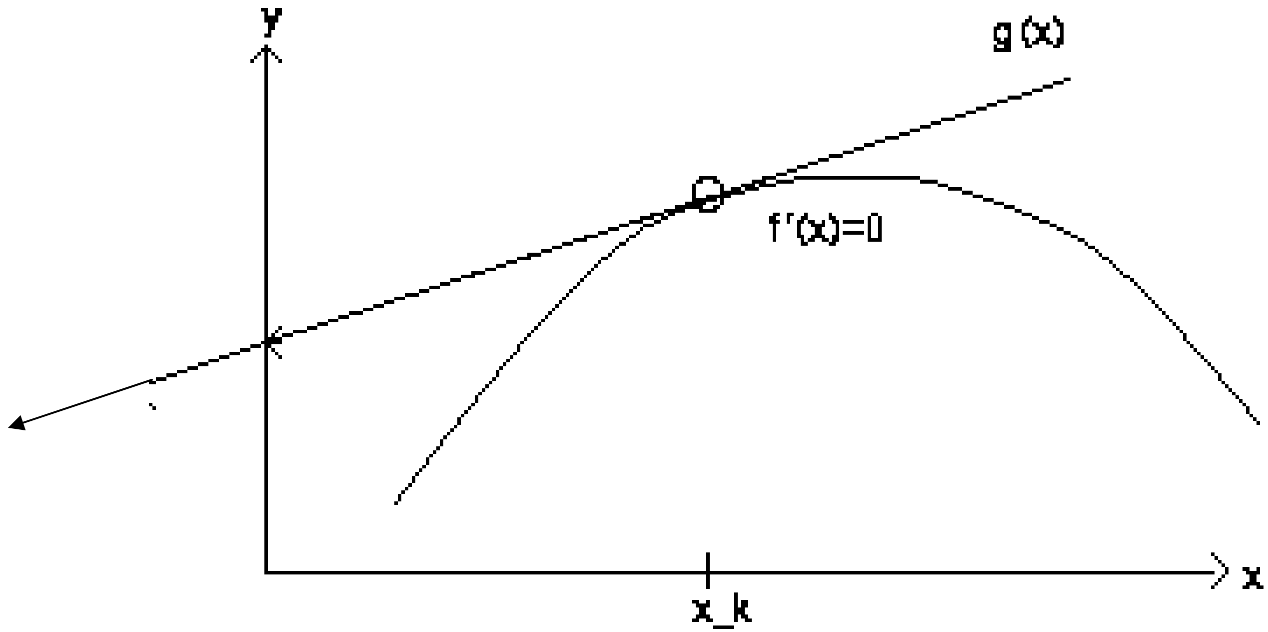
$$g(x) = f(x_k) + f'(x_k) \cdot (x - x_k)$$

ist genau $x_{k+1} = x_k - f(x_k)/f'(x_k)$, und wird als nächste Näherung für die ‚echte' Nullstelle von f gewählt.

Probleme des Newtonverfahrens, wenn $f'(x_k) \cong 0$:

Division durch Null!

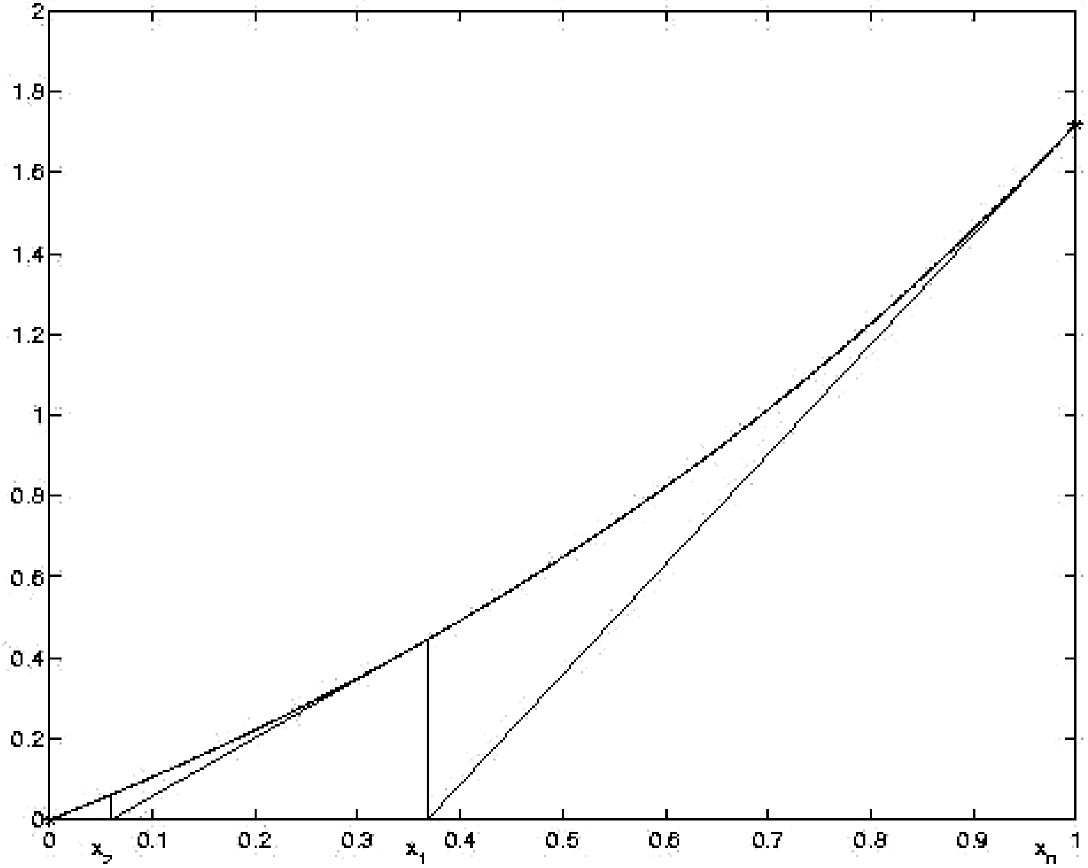
Geometrisch:



Ist x_k nahe einem Punkt mit waagrechter Tangente, so ist die Gerade g(x) fast parallel zur x-Achse, und die nächste Iterierte liegt weit entfernt von x_k →

Kein konvergentes Verhalten!

Beispiel für Iterationen beim Newtonverfahren:



Funktion: f(x) = exp(x) - 1

Nullstelle: $\overline{x} = 0$

Startwert: $x_0 = 1$

Newton-Verfahren als Fixpunktiteration:

Dazugehörige Iterationsfunktion ist

$$\Phi(x) = x - \frac{f(x)}{f'(x)}.$$

Gesuchte Nullstelle $\bar{x}$ von f(x) ist gleichzeitig Fixpunkt von Φ(x) (falls $f'(\bar{x}) \neq 0$):

$$x = \Phi(x) = x - \frac{f(x)}{f'(x)} \Leftrightarrow \frac{f(x)}{f'(x)} = 0$$

Damit sind die Resultate über Iterationsfunktionen und deren Fixpunkte anwendbar (Fixpunktsatz, L<1, usw.):

$$\Phi'(x) = 1 - \frac{f'^2(x) - f(x) f''(x)}{f'^2(x)} = \frac{f(x) f''(x)}{f'^2(x)}$$

und (falls $\bar{x}$ einfache Nullstelle von f) gilt:

$$|\Phi'(\bar{x})| = 0 < 1 \,.$$

Daher ist die gesuchte Nullstelle $\bar{x}$ von f dann ein anziehender Fixpunkt von Φ !

6.2.2
Satz: Das Newtonverfahren für eine stetig diff'bare Funktion f mit einfacher Nullstelle $\bar{x}$ ist lokal quadratisch konvergent.

Beweis:

Lokal konvergent nach Banach'schem Fixpunktsatz!

Startwert nahe genug bei Fixpunkt → lineare Konvergenz im Intervall $\mathbf{U = [\,\bar{x}-h\,,\,\bar{x}+h\,]}$

Zum Beweis der quadratische Konvergenz:

$$0 = f(\bar{x}) = f(x_k) + f'(x_k)(\bar{x} - x_k) + \frac{1}{2} f''(z)(\bar{x} - x_k)^2$$

mit Zwischenstelle z (Taylorentwicklung).

Umformung:

$$\bar{x} - \boxed{x_k + \frac{f(x_k)}{f'(x_k)}} = -\frac{1}{2} \cdot \frac{f''(z)}{f'(x_k)} (\bar{x} - x_k)^2$$

$$\left| \bar{x} - \boxed{x_{k+1}} \right| = \left| \frac{f''(z)}{2 f'(x_k)} \right| \cdot \left| \bar{x} - x_k \right|^2 = C(x_k, \bar{x}) \cdot \left| \bar{x} - x_k \right|^2$$

Ist $f'(\bar{x}) \neq 0$, so kann C lokal durch eine Konstante L nach oben beschränkt werden.

Daher folgt: $$\left| \bar{x} - x_{k+1} \right| \leq L \cdot \left| \bar{x} - x_k \right|^2$$

Der Abstand von der gesuchten Lösung verringert sich in jedem Schritt quadratisch (wenn man nahe genug an der Lösung ist, so dass $|\bar{x}-x_k| << 1$).

z.B. Abstand von der Lösung in jedem Schritt $|\bar{x}-x_k|$ z.B. wie 10^{-1}, 10^{-2}, 10^{-4}, 10^{-8}, 10^{-16} ....

Also schnelle Konvergenz!

Voraussetzungen für quadratische Konvergenz:

- **$\bar{x}$ einfache Nullstelle, d.h.**
  $f(\bar{x}) = 0$, aber $f'(\bar{x}) \neq 0$, oder
  $f(x)=(\bar{x}-x)g(x)$ mit $g(\bar{x})\neq 0$
- **Startwert x_0 nahe bei $\bar{x}$**

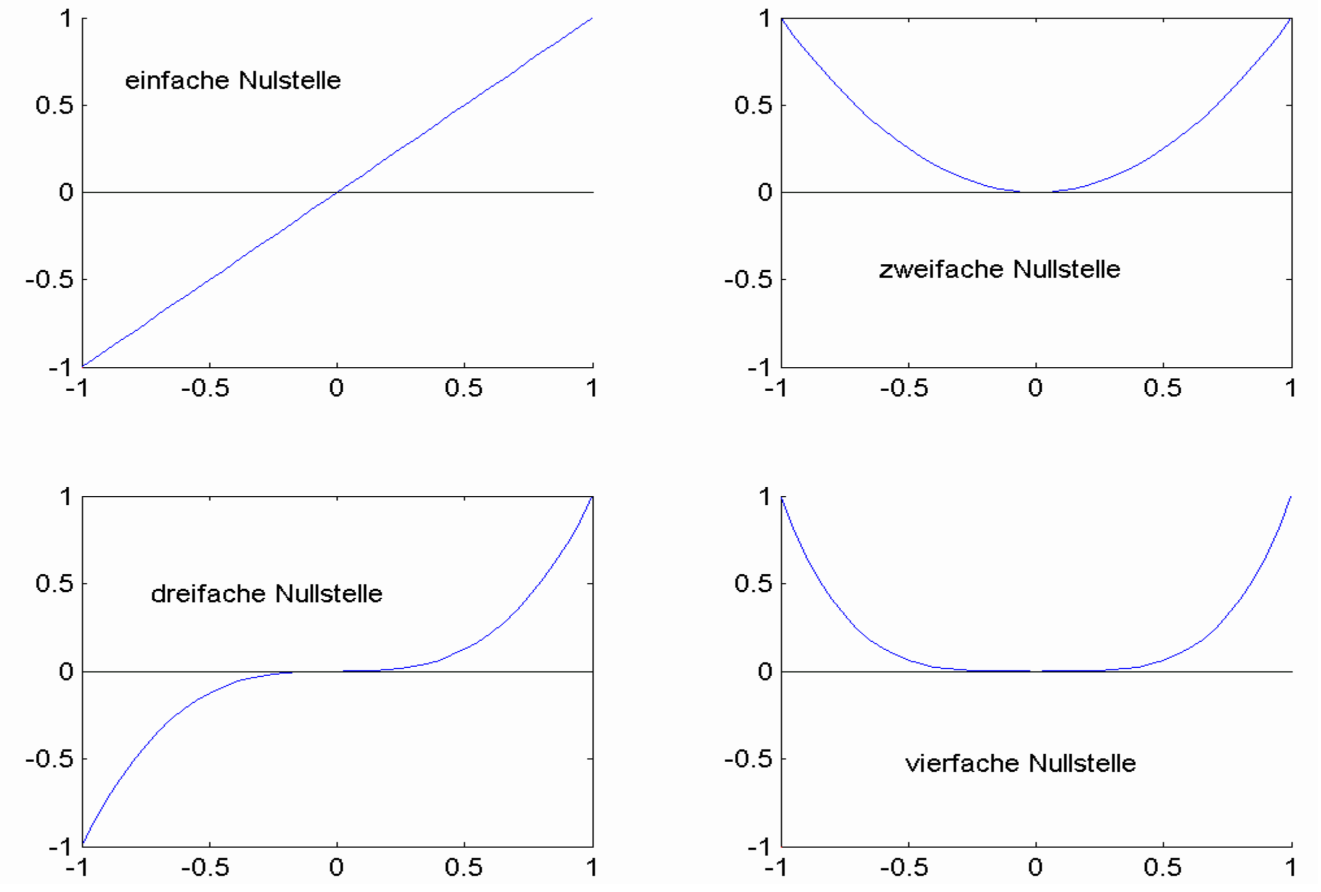
a heißt m-fache Nullstelle, wenn:

$$f^{(j)}(a) = 0 \quad \text{für } j=0,\dots,m-1 \text{ , aber} \quad f^{(m)}(a) \neq 0$$

oder

$$f(x) = (x-a)^m \cdot g(x) \quad \text{mit } g(a) \neq 0$$

Graphisch:



6.2.3. Definition der Konvergenzordnung:

Linear konvergent: $\left|\bar{x} - x_{k+1}\right| \leq L \cdot \left|\bar{x} - x_k\right| \quad \textit{und} \quad L < 1$

Konvergent von Ordnung p > 1 : $\left|\bar{x} - x_{k+1}\right| \leq L \cdot \left|\bar{x} - x_k\right|^p$

Falls f an der Stelle $\bar{x}$ eine m-fache Nullstelle hat mit m > 1 , so ist das Newtonverfahren nur noch lokal linear konvergent:

Ist nämlich $\bar{x}$ eine m-fache Nullstelle, so folgt

$$f(x) = (x - \bar{x})^m g(x) \qquad mit \quad g(\bar{x}) \neq 0$$

und die Iterationsfunktion lautet

$$\Phi(x) = x - \frac{f(x)}{f'(x)} = x - \frac{(x - \bar{x})^m g(x)}{m(x - \bar{x})^{m-1} g(x) + (x - \bar{x})^m g'(x)}$$

Die Ableitung der Iterationsfunktion ist daher

$$\Phi'(\bar{x}) = 1 - \frac{g(\bar{x})}{m \cdot g(\bar{x})} = 1 - \frac{1}{m} < 1$$

Nach dem Banach'schen Fixpunksatz folgt:

Φ ist lokal kontrahierend und es liegt *lineare* Konvergenz vor!

Variationen bei m-facher Nullstelle, um quadratische Konvergenz zu erreichen:

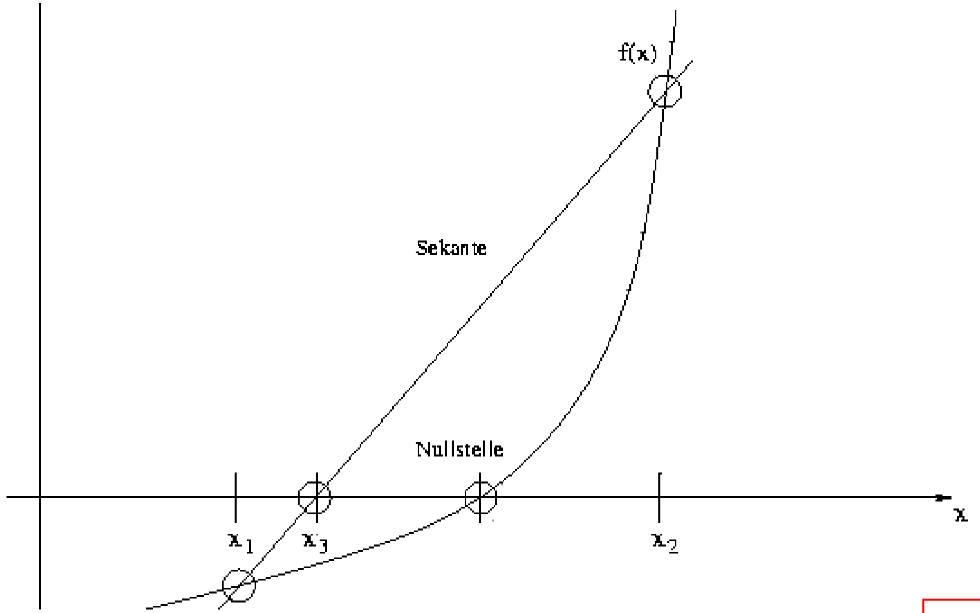
Wende Newtonverfahren an auf

- (m-1) – te Ableitung von f(x)
- f(x)/f'(x)
- m-te Wurzel von |f(x)|
- oder modifiziere Newtonformel (bei bekanntem m) zu

$$\Phi(x) = x - m \cdot \frac{f(x)}{f'(x)}.$$

6.2.4. Verwandte Verfahren

Sekantenverfahren:
Ersetze Tangente in letztem Punkt durch die Sekante, die die beiden letzten Punkte verbindet; verwende deren Nullstelle!



$$\approx f'(x_k)$$

$$g(x) = f(x_{k-1}) + \boxed{\frac{f(x_k) - f(x_{k-1})}{x_k - x_{k-1}}} \cdot (x - x_{k-1})$$

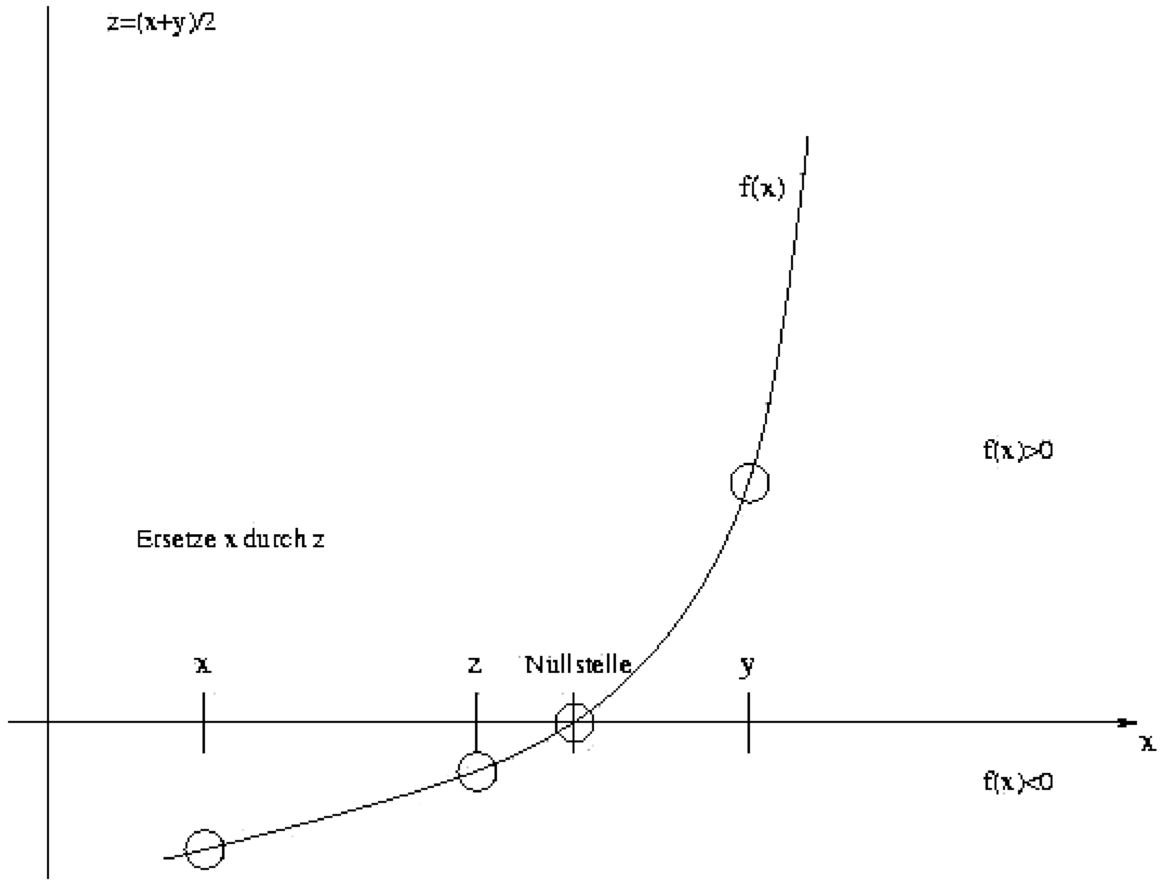
Nullstelle der Sekante als nächste Iterierte:

$$x_{k+1} = x_{k-1} - \frac{f(x_{k-1})(x_k - x_{k-1})}{f(x_k) - f(x_{k-1})} = \frac{x_{k-1} f(x_k) - x_k f(x_{k-1})}{f(x_k) - f(x_{k-1})}$$

Liefert lokale Konvergenz wie beim Newtonverfahren.
Vorteil: keine Ableitungen benötigt, jeweils nur die letzten beiden Funktionswerte! Billiger!
Aber Problem, falls Nenner gleich Null ist!

Konvergenzordnung p nur mit $1<p<2$, aber dafür pro Schritt nur eine neue Funktionsauswertung nötig
(bei Newton f und f ’ pro Schritt)

Bisektionsverfahren:



Starte mit zwei Werten x und y, für die gilt, dass $f(x)*f(y) < 0$ ist, für die also f verschiedene Vorzeichen hat.
Daher liegt für stetiges f zwischen x und y garantiert eine Nullstelle!
Setze $z:=(x+y)/2$ den Wert genau in der Mitte zwischen x und y:
Berechne f(z).

Ist f(z) = 0: fertig
Ist f(x)*f(z) > 0 , so setze x:=z, y bleibt
Ist f(y)*f(z) > 0 , so setze y:=z, x bleibt

*Damit gilt wieder f(x)*f(y) < 0, aber der Abstand zwischen x und y hat sich halbiert.*

Daher liegt die gesuchte Nullstelle nach k Schritten garantiert in einem kleineren Intervall der Größe const/2^k , und es gilt:
(Bezeichnung: x_k, y_k → z_k usw.)

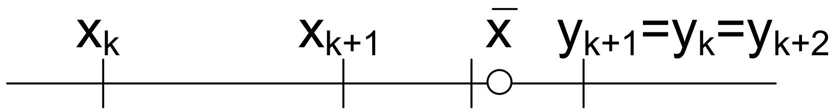
$$\underline{\left|\bar{x} - x_{k+1}\right| + \left|y_{k+1} - \bar{x}\right|} = \left|y_{k+1} - x_{k+1}\right| =$$

$$= 0.5 \cdot \left|y_k - x_k\right| = 0.5 \cdot \left(\underline{\left|\bar{x} - x_k\right| + \left|y_k - \bar{x}\right|}\right)$$

$$<= \max \{ |\bar{x} - x_k| , |\bar{x} - y_k| \}$$

Daher schrumpft der „Abstand“ der Nullstelle
zu den beiden Intervallgrenzen linear in jedem Schritt
um den Faktor 0.5

$$\max\{|x_{k+2}-\bar{x}|,|y_{k+2}-\bar{x}|\} \le |x_{k+2}-y_{k+2}| \le |x_{k+1}-y_{k+1}|/2 \le \max\{|x_k-\bar{x}|,|y_k-\bar{x}|\}/2$$



Regula falsi durch Verbindung von Bisektion und Sekantenverf.:

**Iteriere wieder Intervall-Einschließungsgrenzen.
Neuer Kandidat für Ober/Untergrenze ist jetzt nicht der Punkt in der Mitte, sondern die Nullstelle der Sekante.
Ersetze eine der beiden alten Grenzen durch diesen neuen Kandidaten, so dass die gesuchte Nullstelle wieder garantiert in dem neuen Intervall liegt!**

In der Praxis wird häufig eine Kombination von Bisektion und Newton eingesetzt:

Starte mit ‚sicherer' Bisektion, bis der ‚Einzugsbereich' der quadratischen Konvergenz des Newtonverfahrens erreicht ist.

Dazu nötig sind Werte x und y mit $f(x)f(y) < 0$;
solche Werte erhält man z.B. durch Auswertung von f(x) an Zufallszahlen oder an äquidistanten Stellen.

Falls solche x und y nicht auffindbar sind, liegt ev. keine Nullstelle vor oder eine Nullstelle gerader Ordnung!
Im letzteren Fall kann das Newtonverfahren auf $f'(x)$ angewendet werden.

6.2.5. Newtonverfahren und Minimierung:

Nullstelle einer Funktion f: $R^n \rightarrow R^n$,

$$f(x) = f(x_1,\dots,x_n) = \begin{pmatrix} f_1(x_1,\dots,x_n) \\ \vdots \\ f_n(x_1,\dots,x_n) \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} = 0$$

Jacobi-matrix der Ableitungen von f:

$$J = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial x_1} & \cdots & \frac{\partial f_n}{\partial x_n} \end{pmatrix}$$

Newtonverfahren im R^n :

$$\boxed{x^{k+1} := x^k - inv(J) \cdot f(x^k)}$$

Vektor = Vektor – Matrix * Vektor

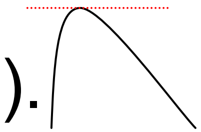
In jedem Schritt ist ein neues Gleichungss. in $J(x^k)$ zu lösen!

6.2.6. Minimierungsproblem:

Gegeben: F: $R^n \rightarrow R$, $F(x_1,\ldots,x_n) \in R$

Gesucht: (relatives) Minimum / Maximum

Bestimme dazu Nullstelle der Ableitung
(entspricht waagrechter Tangente, also Extremwert).



Gradientenvektor von F:

$$\nabla F = f(x_1,\ldots,x_n) = \begin{pmatrix} f_1(x_1,\ldots,x_n) \\ \vdots \\ f_n(x_1,\ldots,x_n) \end{pmatrix} = \begin{pmatrix} \dfrac{\partial F}{\partial x_1} \\ \vdots \\ \dfrac{\partial F}{\partial x_n} \end{pmatrix} \overset{!}{=} 0$$

Jacobi-matrix J der ersten Ableitungen von f entspricht der sog. Hesse-matrix H der zweiten Ableitungen von F:

$$H(F) = J(f) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial x_1} & \cdots & \frac{\partial f_n}{\partial x_n} \end{pmatrix} = \left(\frac{\partial^2 F}{\partial x_i \partial x_j} \right)_{i,j=1}^{n}$$

Newtonverfahren zur Bestimmung der Nullstelle des Gradienten ergibt:

$$\boxed{x^{k+1} = x^k - inv(H) \cdot \nabla F(x^k)}$$

Löse dazu in jedem Schritt ein lineares Gleichungssystem mit sog. Hesse-matrix $H(x^k)$!

Problem, falls H an der Stelle x^k singulär.
Billiger: Quasi-Newton-Verfahren, ersetze H durch billiges B.

6.2.7. Nichtlineare Ausgleichsrechnung: Gauss-Newton-Verfahren

$$\min_x \left\| \begin{pmatrix} f_1(x_1,\ldots,x_n) \\ \vdots \\ f_m(x_1,\ldots,x_n) \end{pmatrix} - b \right\|_2^2 ,$$

(vgl. min ||Ax-b||)

Erster Schritt:

Linearisierung mittels Jacobi-matrix und Taylorreihe

$$f(x) = f(x^k) + J(x - x^k) + O(|h|^2)$$

Damit ergibt sich neue Iterierte x^{k+1} aus der Lösung des <u>linearen</u> Ausgleichsproblems

$$\min_x \left\| \left(f(x^k) + J^k(x - x^k) \right) - b \right\|_2^2$$

Neue Linearisierung an der Stelle x^{k+1} ergibt neues lineares Ausgleichsproblem mit neuer Matrix J^{k+1}.

6.3. Iterative Lösung linearer Gleichungssysteme

Großes lineares dünnbesetztes Gleichungssystem $A x = b$

Gauss-Elimination nutzt in der Regel die Dünnbesetztheit nicht aus und führt meist auf Kosten $O(n^3)$;
Im Gegensatz dazu ist oft nur $O(n)$ Speicherbedarf für Matrix A

Idee: Formuliere iteratives Verfahren, das in jedem Schritt nur Matrix*Vektor berechnet, so dass die Iterierten gegen die Lösung von $Ax = b$ konvergieren.

Gesamtkosten: #Iterationen * $O(n)$.
Effizient, falls schnelle Konvergenz vorliegt!
Einfach parallelisierbar!

6.3.1. Stationäre Methoden:

Richardson-Verfahren:
Formulierung eines passenden Fixpunktproblems

$$b = Ax = (A - I)x + x \Leftrightarrow \boxed{x = b + (I - A)x =: \Phi(x)}$$

Daraus ergibt sich die Iteration:

$$x_{k+1} = b + (I - A)x_k = x_k + (b - Ax_k) = x_k + r_k = \Phi(x_k)$$

mit Residuum r_k .

Frage:
Wann konvergiert die ausgehend von einem x_0 definierte Folge x_k gegen die gesuchte Lösung $\bar{x} = A^{-1}b$?

Betrachte dazu den Fehler im k-ten Schritt:

$$\bar{x} - x_{k+1} = \bar{x} - x_k - (b - Ax_k) = \bar{x} - x_k - (A\bar{x} - Ax_k) = (I - A)(\bar{x} - x_k)$$

Ergibt in der Norm

$$\left\|\bar{x} - x_{k+1}\right\|_2 = \left\|(I - A)(\bar{x} - x_k)\right\|_2 \leq \left\|I - A\right\|_2 \cdot \left\|\bar{x} - x_k\right\|_2$$

und daher $\left\|\bar{x} - x_k\right\|_2 \leq \left\|I - A\right\|_2^k \cdot \left\|\bar{x} - x_0\right\|_2$

Daher liegt Konvergenz vor für $\|I - A\|_2 < 1$.

Dies entspricht der Kontraktionsbedingung für die Iterations-funktion $\Phi(x)$:

$$\left\|\Phi(x) - \Phi(y)\right\|_2 \leq \left\|I - A\right\|_2 \cdot \left\|x - y\right\|_2$$

Richardsonverfahren ist daher nur sinnvoll, wenn $A \approx I$!

Geschickter:

Wende das Verfahren auf das modifizierte Problem an:

$$D = diag(A); \qquad \boxed{\left(D^{-1}A\right)x = \left(D^{-1}b\right)} \quad \Leftrightarrow \quad Ax = b$$

Die Bedingung $D^{-1}A \approx I$ ist besser erfüllt!

Bezeichnung: $\tilde{A} = D^{-1}A$ und $\tilde{b} = D^{-1}b$ ergibt neues Gleichungssystem $\tilde{A}x = \tilde{b}$.

Richardson für das tilde-System liefert dann:

$$x_{k+1} = x_k + (\tilde{b} - \tilde{A}x_k) = x_k + D^{-1}(b - Ax_k)$$

oder

$$Dx_{k+1} = Dx_k + (b - Ax_k) = b + (D - A)x_k$$

Dies ist das Jacobiverfahren zur iterativen Lösung von Ax = b
Wesentliche Kosten in jedem Schritt: Ax_k

Konvergent, falls $\left\| I - D^{-1}A \right\|_2 < 1$

Notwendig: Diagonalmatrix D regulär!

Allgemeinere Idee zur Formulierung einer Iterationsfunktion:

Matrix-splitting

$$L: \quad Unterdiagonalteil \quad von \quad A$$

$$U: \quad Oberdiagonalteil \quad von \quad A$$

$$b = Ax = (L + D + U)x = (L + D)x + Ux$$

führt auf Iteration

$$(L + D)x_{k+1} = b - Ux_k = b - (A - L - D)x_k$$

$$= (L + D)x_k + (b - Ax_k)$$

Dies ist das Gauss-Seidel-Verfahren

$$\begin{aligned} x_{k+1} &= x_k + (L+D)^{-1}(b - Ax_k) \\ &= (L+D)^{-1}b + \left(I - (L+D)^{-1}A\right)x_k \end{aligned}$$

In jedem Schritt ist dabei nur ein Dreiecksgleichungssystem zu lösen.

Das Verfahren ist konvergent, falls $\left\| I - (L+D)^{-1}A \right\|_2 < 1$

Gauss-Seidel ist äquivalent zu Richardson angewendet auf

$$\left((L+D)^{-1}A\right)x = \left((L+D)^{-1}b\right) \quad \Leftrightarrow \quad Ax = b$$

NEU: Komponentenweise Formel für Jacobi und GS!

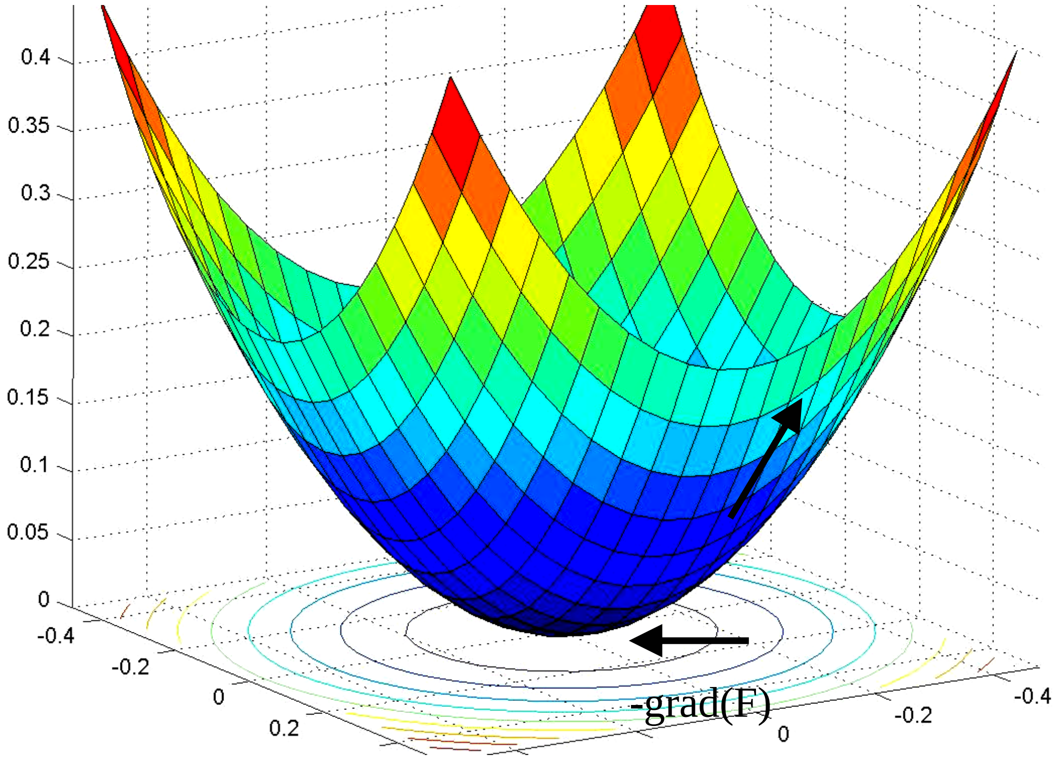
6.3.2. Nichtstationäre Verfahren:

Das Gradientenverfahren für symmetrisch positiv definite Matrix A:

Betrachte Funktion F(x): $R^n \rightarrow R^n$,

$$F(x) = \frac{1}{2} x^T A x - b^T x$$

F beschreibt einen Paraboloid im R^n:



x-Ebene

Minimum dieser Funktion ist wieder der Punkt mit waagrechter Tangente, also die Stelle mit Gradient gleich Null:

$$\boxed{\nabla F(x) = Ax - b = 0 \Leftrightarrow Ax = b}$$

Stelle x, an der der Paraboloid sein Minimum annimmt ist gleich gesuchte Lösung des Gleichungssystems!

Betrachte Minimierungsaufgabe!

Von aktueller Stelle x_k aus soll die nächste Iterierte x_{k+1} so gewählt werden, dass sie näher am Minimum liegt.

$$x_{k+1} = x_k + \alpha_k d_k$$

mit Suchrichtung d_k und Schrittweite α_k.
Finde Suchrichtung so, dass Funktionswerte kleiner werden:

Abstiegsrichung ist gegeben durch Richtung des negativen Gradienten!

Denn Richtungsableitung in Richtung *n* ist gleich $\nabla F \cdot n$, und wird am betragsgrößten für $n = \nabla F$, nämlich $\left\| \nabla F \right\|^2$!

Daher ist $n = -\nabla F$ lokal die Richtung des steilsten Abstiegs zum Minimum.

Daher verkleinern sich die Funktionswerte auf jeden Fall, wenn man in diese Abstiegsrichtung geht.

Also wähle: $$\boxed{x_{k+1} = x_k - \alpha_k \nabla F(x_k) = x_k + \alpha_k (b - Ax_k)}$$

Noch zu bestimmten ist optimale Schrittweite α_k, die am nähesten ans Minimum führt!

Betrachte dazu ein-dimensionale Minimierung:

$$\min_\alpha g(\alpha) := \min_\alpha \left(F(x_k + \alpha(b - Ax_k)) \right) =$$

$$\min_\alpha \left(\frac{1}{2}(x_k + \alpha d_k)^T A(x_k + \alpha d_k) - b^T (x_k + \alpha d_k) \right) =$$

$$\min_\alpha \left(\frac{1}{2}\alpha^2 d_k^T A d_k - \alpha d_k^T d_k + \frac{1}{2} x_k^T A x_k - x_k^T b \right)$$

mit Lösung $\alpha_k = d_k^T d_k / d_k^T A d_k$, $d_k := b - Ax_k$

Insgesamt:

$$\boxed{x_{k+1} = x_k + \frac{\|b - Ax_k\|_2^2}{(b - Ax_k)^T A(b - Ax_k)} \cdot (b - Ax_k)}$$

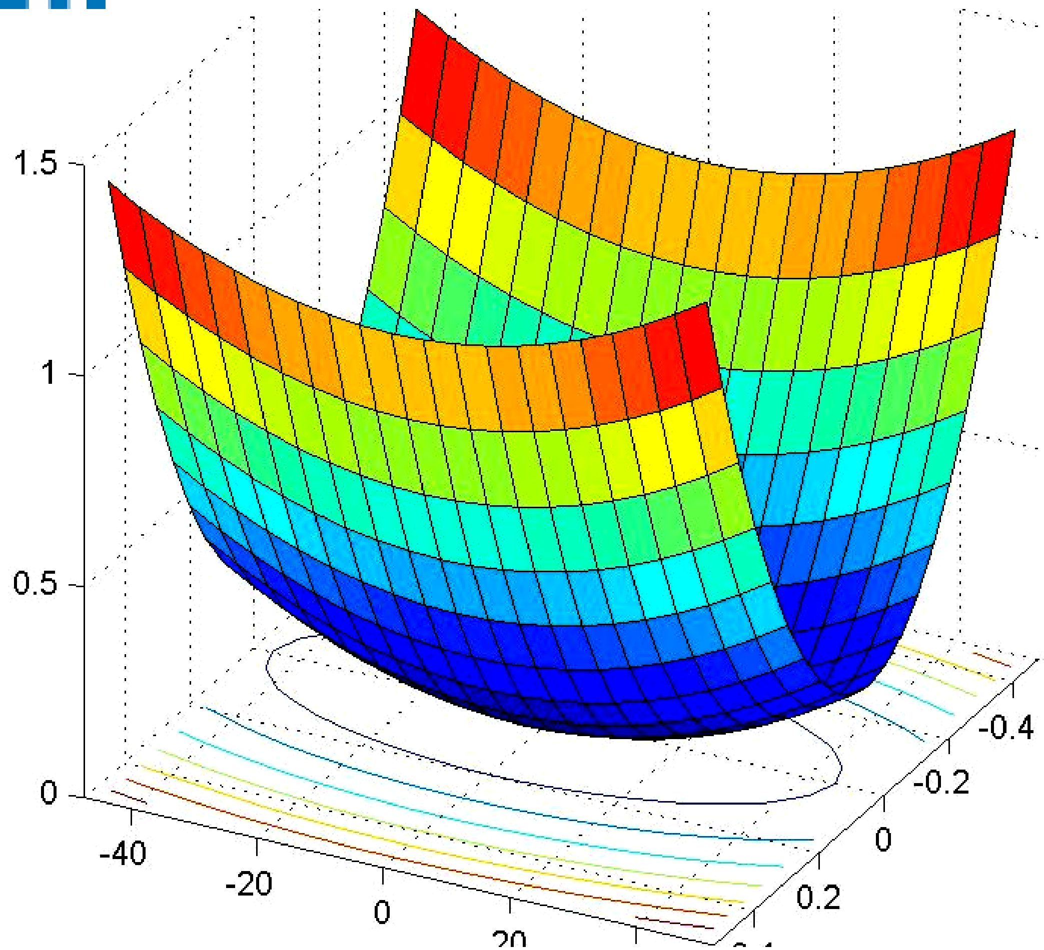
Dazugehörige Fixpunktgleichung ist

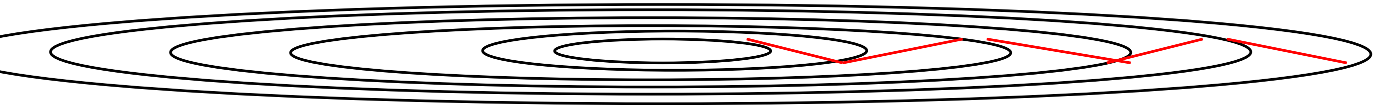
$$x = \Phi(x) := x + \frac{\left\| b - Ax \right\|_2^2}{(b - Ax)^T A (b - Ax)} \cdot (b - Ax)$$

mit Fixpunkt $\bar{x} = A^{-1} b$.

Ergibt Verfahren des steilsten Abstiegs (‚steepest descent') oder Gradientenverfahren

Nachteil:
Bei stark verzerrtem Paraboloiden ergibt sich sehr langsame Konvergenz.





Für A ≈ I ist der Paraboloid unverzerrt, die Höhenlinien fast kreisförmig → schnelle Konvergenz!

Daher versucht man, Ax=b zu präkonditionieren:
Ersetze Ax=b durch $M^{-1}Ax = M^{-1}b$ mit $M^{-1}A \approx I$

Bessere Variante eines Gradientenverfahren:
Verfahren der konjugierten Gradienten, kurz cg-Verfahren.

Suchrichtung nicht der negative Gradient selbst, sondern die Projektion des Gradienten, so dass alle Suchrichtungen in gewisser Weise orthogonal zueinander sind
Genauer: Suchrichtungen seien A-konjugiert, d.h.

$$d_k^T A d_j = 0 \quad \text{für} \quad j \neq k$$

Damit ergibt sich iteratives Verfahren, das nach k Schritten jeweils die beste Näherung an die Lösung in einem k-dimensionalen Unterraum liefert, und daher nach n Schritten fertig ist (in exakter Arithmetik).

NEU: Norm-Minimierung, Unterraum Krylov