ニュートン法

非線形方程式を解くときには計算時間などのコストを考えると「ニュートン法」[1])を用いるのが良い. ニュートン法は「非線形の世界も極く小さな領域だけ見ればほぼ線形の世界である[2])」という「一般原理」に基づいているので, 理論的には適用可能範囲が大変広く, 実用的である.

公式

ニュートン法の公式は反復公式である. 今, 方程式 $f(x)=0$ の一つの解 α の「近くに」ある初期近似値 $x^{(0)}$ から出発し, 近次解 $x^{(1)}$ を求める.

$$f(x^{(0)}) = f'(x^{(0)}) \times (x^{(0)} - x^{(1)})$$

なので

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

よって,

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

1)ニュートン・ラフソン (Newton-Raphson) 法ともいう. ラフソンってどんな人かを調べておくこと.

2)非線形の関数 $f(x)$ があるとする. そのとき, $f(x)$ を $x=a$ のまわりでテイラー展開を行う. ただし, $x-a=h$ とする.

$$f(a+h) = f(a) + f'(a)h + \frac{f''(a)h^2}{2!} + \cdots$$

ここで $h \ll 1$ のときは h が 2 次以上の項は無視できる. よって

$$f(a+h) \sim f(a) + f'(a)h.$$

右辺は線形なので $h \ll 1$ という極く小さな領域だと線形とみなせる. この近似は非線形の関数がデルタ関数や不連続な関数以外の性質の良い関数の場合に成り立つ. 例えばべき級数などである.

テイラー展開の説明は正しいか要確認.

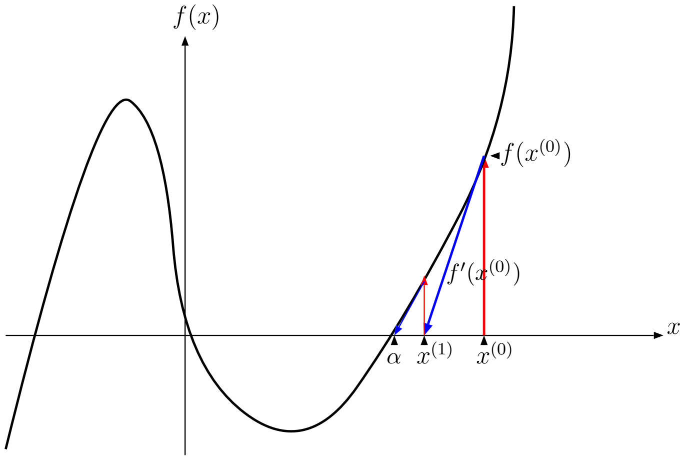

Figure 1: ニュートン法概略図. $x^{(0)}$ は初期値, $x^{(1)}$ は次の値. 青線は f' の勾配をもちかつ $f(x)$ を通る一次関数であり, 実際の解は α である.

これを図にすると図 1 にのようになる.

この式の $\nu+1$ 回目は

$$x^{(\nu+1)} = x^{(\nu)} - \frac{f(x^{(\nu)})}{f'(x^{(\nu)})}. \tag{1}$$

これを用いて, 次々と望むらくは改良された近似解 $x^{(1)}, x^{(2)}, \cdots$ を計算する [3].

さらに多変数の非線形連立方程式

$$f_i(x_1, \cdots, x_n) = 0 \qquad (i = 1, \cdots, n)$$

の場合にも (1) 式を自然に拡張することができる. $\alpha_i (i = 1, \cdots, n)$ を $f_i(x_i)(i = 1, \cdots, n)$ のとある解とする. $x_i^{(0)}(i = 1, \cdots, n)$ をその初期近似解としたとき更に精度の良い近似解 $x_i^{(1)}(i = 1, \cdots, n)$ があるとすると, $f_i(x_i^{(1)})(i = 1, \cdots, n)$ は

$$f_i(x_1^{(1)}, x_2^{(1)}, \cdots, x_n^{(1)}) \sim f_i(x_1^{(0)}, x_2^{(0)}, \cdots, x_n^{(0)}) + \sum_{j=1}^{n} \frac{\partial f_i(x_1^{(0)}, x_2^{(0)}, \cdots, x_n^{(0)})}{\partial x_j}(x_j^{(1)} - x_j^{(0)})$$

と表される. この式の $\nu+1$ 回目は

$$f_i(x_i^{(\nu+1)}) \sim f_i(x_i^{(\nu)}) + \sum_{j=1}^{n} \frac{\partial f_i(x_i^{(\nu)})}{\partial x_j}(x_j^{(\nu+1)} - x_j^{(\nu)}). \tag{2}$$

[3] x_0 が α ととても離れていると期待される解 α の近似値を求められない場合があるので, α の見当が付いていないと使えない.

ここで $x_i^{(\nu+1)}(i=1,\cdots,n)$ が $\alpha_i(i=1,\cdots,n)$ に十分近いとすれば $f_i(x_i^{(\nu+1)})\sim f_i(\alpha_i)=0$ と見なせる. このとき (2) 式は

$$0\sim f_i(x_i^{(\nu)})+\sum_{j=1}^{n}\frac{\partial f_i(x_i^{(\nu)})}{\partial x_j}(x_j^{(\nu+1)}-x_j^{(\nu)}).$$

よって,

$$f_i(x_i^{(\nu)})\sim\sum_{j=1}^{n}\frac{\partial f_i(x_i^{(\nu)})}{\partial x_j}(x_j^{(\nu)}-x_j^{(\nu+1)}).$$

ここで $\Delta x_j^{(\nu)}\equiv x_j^{(\nu)}-x_j^{(\nu+1)}$ とすると,

$$x_j^{(\nu+1)}=x_j^{(\nu)}-\Delta x_j^{(\nu)}. \tag{3}$$

$$f_i(x_i^{(\nu)})\sim\sum_{j=1}^{n}\frac{\partial f_i(x_i^{(\nu)})}{\partial x_j}\Delta x_j^{(\nu)}. \tag{4}$$

となって (1) 式を多変数に拡張した式となる. (4) 式を $\Delta x_1^{(\nu)},\cdots,\Delta x_n^{(\nu)}$ を未知数とする連立 1 次方程式として解いて $\Delta x_j^{(\nu)}$ を定め, それを (3) 式に入れることで解を求められる. $f_i(x)$ が複素変数 x の複素関数であっても (4) 式は形式的にそのまま使える[4]).

収束性

ニュートン法をつかって誤差がどのような振る舞いをするか考える. 簡単のために 1 変数のときを考える. 解 α が単純孤立した解[5])で, α の近傍で f が素直な関数でかつ

$$f'(\alpha)\neq 0 \qquad \left(\text{連立方程式のときは, 行列}\left[\frac{\partial f_i(\alpha)}{\partial x_j}\right]\text{が正則}\right).$$

そして出発値 $x^{(0)}$ が α に「十分」 近ければ, つまり ν を ∞ までもっていくと

$$x^{(\nu)}\to\alpha \qquad (\nu\to\infty)$$

であるとする. このとき誤差を

$$\epsilon_\nu=x^{(\nu)}-\alpha \tag{5}$$

4)複素数 z を $z\equiv x+iy$ (i は虚数単位) とすると $f(z)=u(x,y)+iv(x,y)$ と置ける. よって, $f(z)=f_2(u,v)$ と表せて 2 変数連立方程式と書ける. したがって, (4) 式を $n=2$ として解けばよい

5)「単純独立した解」がどういう解なのかよくわからない.

とする. ある N より先の $\nu\ (> N)$ に対して ϵ_ν がとても小さくなると考える. 反復公式 (1) 式より $\nu+1$ 回目の x の値は

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

$f(x^{(\nu)})$ をテイラー展開して

$$f(x^{(\nu)}) = f(\alpha) + f'(\alpha)\epsilon_\nu + \frac{1}{2!}f''(\alpha)\epsilon_\nu^2 + \cdots .$$

$f'(x^{(\nu)})$ をテイラー展開して

$$f'(x^{(\nu)}) = f'(\alpha) + f''(\alpha)\epsilon_\nu + \frac{1}{2!}f'''(\alpha)\epsilon_\nu^2 + \cdots .$$

これを (1) 式の右辺第 2 項に代入する.

$$\begin{aligned}
\frac{f(x^{(\nu)})}{f'(x^{(\nu)})} &= \frac{f(\alpha) + f'(\alpha)\epsilon_\nu + f''(\alpha)\frac{\epsilon_\nu^2}{2}}{f'(\alpha) + f''(\alpha)\epsilon_\nu} + O(\epsilon_\nu^3) \\
&= \frac{1}{f'(\alpha)} \frac{f(\alpha) + f'(\alpha)\epsilon_\nu + f''(\alpha)\frac{{\epsilon_\nu}^2}{2}}{1 + \frac{f''(\alpha)}{f'(\alpha)}\epsilon_\nu} + O(\epsilon_\nu^3).
\end{aligned}$$

$\left|\frac{f''(\alpha)}{f'(\alpha)}\epsilon_\nu\right| \ll 1$ とすると[6]

$$\begin{aligned}
\frac{f(x^{(\nu)})}{f'(x^{(\nu)})} &\approx \frac{1}{f'(\alpha)}\left(f(\alpha) + f'(\alpha)\epsilon_\nu + f''(\alpha)\frac{\epsilon_\nu^2}{2}\right)\left(1 - \frac{f''(\alpha)}{f'(\alpha)}\epsilon_\nu\right) \quad + O(\epsilon_\nu^3) \\
&= \frac{1}{f'(\alpha)}\left(f(\alpha) + f'(\alpha)\epsilon_\nu + f''(\alpha)\frac{\epsilon_\nu^2}{2} - \frac{f(\alpha)f''(\alpha)}{f'(\alpha)}\epsilon_\nu - f''\epsilon_\nu^2 - \frac{(f''(\alpha))^2}{f'(\alpha)}\frac{\epsilon_\nu^3}{2}\right) + O(\epsilon_\nu^3) \\
&= \frac{f(\alpha)}{f'(\alpha)} + \frac{f'(\alpha)\epsilon_\nu}{f'(\alpha)} + \frac{f''(\alpha)}{f'(\alpha)}\frac{\epsilon_\nu^2}{2} - \frac{f(\alpha)f''(\alpha)}{(f'(\alpha))^2}\epsilon_\nu - \frac{f''(\alpha)}{f'(\alpha)}\epsilon_\nu^2 + O(\epsilon_\nu^3).
\end{aligned}$$

ここで $f(\alpha) = 0$ なので

$$\frac{f(x^{(\nu)})}{f'(x^{(\nu)})} \approx -\frac{f''(\alpha)}{2f'(\alpha)}\epsilon_\nu^2 + \epsilon_\nu + O(\epsilon_\nu^3). \tag{6}$$

[6] $|\delta| \ll 1$ のとき $\frac{1}{1+\delta}$ をマクローリン展開して

$$\begin{aligned}
\frac{1}{1+\delta} &= \frac{1}{1-0} + \frac{d}{d\delta}\left(\frac{1}{1+\delta}\right)\bigg|_{\delta=0} \cdot \delta + \frac{1}{2!}\frac{d^2}{d\delta^2}\left(\frac{1}{1+\delta}\right)\bigg|_{\delta=0} \cdot \delta^2 + \cdots \\
&= 1 + \left(-\frac{1}{(1+0)^2}\cdot\delta\right) + \frac{1}{(1+0)^3}\cdot\delta^2 + \cdots \\
&\sim 1 - \delta
\end{aligned}$$

と近似できることを用いた

今 (1) 式に戻り, 両辺から α を引く

$$x^{(\nu+1)}-\alpha = x^{(\nu)}-\alpha-\frac{f(x^{(\nu)})}{f'(x^{(\nu)})}.$$

ここで (6) 式を代入すると

$$\begin{aligned}\epsilon_{(\nu+1)} &= \epsilon_\nu-\epsilon_\nu+\frac{f''(\alpha)}{2f'(\alpha)}{\epsilon_\nu}^2+O(\epsilon_\nu^3)\\ &= \frac{f''(\alpha)}{2f'(\alpha)}{\epsilon_\nu}^2+O(\epsilon_\nu^3).\end{aligned}$$

ここで $O(\epsilon_\nu^3)$ を無視する. さらに (5) 式より,

$$\parallel x^{(\nu+1)}-\alpha \parallel \approx \frac{f''(\alpha)}{2f'(\alpha)} \parallel x^{(\nu)}-\alpha \parallel^2 . \tag{7}$$

(7) 式の形の収束状況は 2 次収束と呼ばれる. 大雑把にいって「誤差が毎回 2 乗で減る」あるいは「正しい桁数が反復 1 回ごとに倍に増える」. N を事前に知ることはできない. そのため (7) 式になるための条件 ν が N を超えたかどうかは前もって具体的に調べる手段はない. (7) 式のような収束状況に到達したかどうかは, $x^{(\nu)}-x^{(\nu+1)}$ を観察していればわかる. 誤差 $x^{(\nu)}-x^{(\nu+1)}$ が (7) 式の様に振舞いをするようになった場合は誤差 $x^{(\nu)}-x^{(\nu+1)}$ が真の解からの誤差とほぼ同じになった場合である. つまり,

$$\begin{aligned}x^{(\nu)}-x^{(\nu+1)} &\approx \epsilon_\nu\\ &= x^{(\nu)}-\alpha\end{aligned} \tag{8}$$

となったとき (7) 式に到達したとみなしてよいということである.

α が m 重解のとき, つまり $f'(x^{(\nu)})=f''(x^{(\nu)})=\cdots=f^{(m-1)}(x^{(\nu)})=0$ であるとき,

$$f(x)=(x-\alpha)^m g(x) \quad (g(\alpha)\neq 0) \tag{9}$$

となるとする. このときの収束性を (7) 式と同様に求めていく. 先ほどと同じ様に $\epsilon_\nu = x^{(\nu)}-\alpha$ とおく. 更に $f(x^{(\nu)})$ をテイラー展開して

$$\begin{aligned}f(x^{(\nu)}) &= f'(\alpha)+\cdots+\frac{1}{(m)!}f^{(m)}\epsilon_\nu^m+\frac{1}{(m+1)!}f^{(m+1)}\epsilon_\nu^{m+1}+\cdots\\ &= \frac{1}{(m)!}f^{(m)}\epsilon_\nu^m+\frac{1}{(m+1)!}f^{(m+1)}\epsilon_\nu^{m+1}+\cdots .\end{aligned}$$

$f'(x^{(\nu)})$ をテイラー展開して

$$\begin{aligned}f'(x^{(\nu)}) &= f^{(2)}(\alpha)+\cdots+\frac{1}{(m-1)!}f^{(m)}\epsilon_\nu^{m-1}+\frac{1}{(m)!}f^{(m+1)}\epsilon_\nu^m+\cdots\\ &= \frac{1}{(m-1)!}f^{(m)}\epsilon_\nu^{m-1}+\frac{1}{(m)!}f^{(m+1)}\epsilon_\nu^m+\cdots .\end{aligned}$$

これを (1) 式の右辺第 2 項に代入する.

$$
\begin{aligned}
\frac{f(x^{(\nu)})}{f'(x^{(\nu)})} &= \frac{\frac{1}{(m)!}f^{(m)}\epsilon_\nu^m + \frac{1}{(m+1)!}f^{(m+1)}\epsilon_\nu^{m+1} + \cdots}{\frac{1}{(m-1)!}f^{(m)}\epsilon_\nu^{m-1} + \frac{1}{(m)!}f^{(m+1)}\epsilon_\nu^{m} + \cdots} \\
&= \frac{(m-1)!}{f^{(m)}\epsilon_\nu^{m-1}} \frac{\frac{1}{(m)!}f^{(m)}\epsilon_\nu^m + \frac{1}{(m+1)!}f^{(m+1)}\epsilon_\nu^{m+1} + \cdots}{1 + \frac{f^{m+1}(\alpha)}{mf^m(\alpha)}\epsilon_\nu + \cdots} \\
&= \frac{1}{f^{(m)}} \frac{\frac{1}{m}f^{(m)}\epsilon_\nu + \frac{1}{m(m+1)}f^{(m+1)}\epsilon_\nu^{2} + \cdots}{1 + \frac{f^{m+1}(\alpha)}{mf^m(\alpha)}\epsilon_\nu + \cdots}
\end{aligned}
$$

$\left|\dfrac{f^{(m+1)}(\alpha)}{f^{(m)}(\alpha)}\epsilon_\nu\right| \ll 1$ とすると

$$
\begin{aligned}
\frac{f(x^{(\nu)})}{f'(x^{(\nu)})} &\approx \frac{1}{f^{(m)}}\left(\frac{1}{m}f^{(m)}\epsilon_\nu + \frac{1}{m(m+1)}f^{(m+1)}\epsilon_\nu^2\right)\left(1 + \frac{f^{m+1}(\alpha)}{mf^m(\alpha)}\epsilon_\nu\right) \quad + O(\epsilon_\nu^3) \\
&= \left(\frac{1}{m}\epsilon_\nu + \frac{1}{m(m+1)}\frac{f^{(m+1)}}{f^{(m)}}\epsilon_\nu^2\right)\left(1 + \frac{f^{m+1}(\alpha)}{mf^m(\alpha)}\epsilon_\nu\right) \quad + O(\epsilon_\nu^3) \\
&= \frac{1}{m}\epsilon_\nu + O(\epsilon_\nu^2).
\end{aligned} \tag{10}
$$

今 (1) 式に戻り, 両辺から α を引く

$$
x^{(\nu+1)} - \alpha = x^{(\nu)} - \alpha - \frac{f(x^{(\nu)})}{f'(x^{(\nu)})}.
$$

ここで (10) 式を代入すると

$$
\begin{aligned}
\epsilon_{(\nu+1)} &= \epsilon_\nu - \frac{1}{m}\epsilon_\nu + O(\epsilon_\nu^2) \\
&= \left(1 - \frac{1}{m}\right)\epsilon_\nu + O(\epsilon_\nu^2).
\end{aligned}
$$

ここで $O(\epsilon_\nu^2)$ を無視する. 更に, (5) 式より

$$
|x^{\nu+1} - \alpha| \approx \left(1 - \frac{1}{m}\right)|x^\nu - \alpha| \tag{11}
$$

となる. これは $x^\nu \to \alpha$ となったとしても, 縮小率 $\left(1 - \dfrac{1}{m}\right)$ の「のんびりとした 1 次収束」である.

計算の終了判定

数値計算でどれだけ精度が良くできるか δ で見積もってみる.

$x = \alpha$ の近くで関数値 $f(x)$ を計算する際に, 予想される丸め誤差などの計算誤差の大体の大きさ δ の見当を, 何らかの方法でつけておけるとする. このとき,

$$|f(x^{(\nu)})| < \delta \tag{12}$$

となったら計算を終えることができる. 今, $f(x) = 0$ となる $x = \alpha$ を求めようとしているので計算誤差の範囲内で $f(x^{(\nu)}) = 0$ となったらもうそれ以上は精度良くは求めることができないのである.

(12) 式で, 重解を持たないとき, つまり $f'(x^{(\nu)}) \neq 0$ とはっきり分かっていれば, (8) 式より

$$\begin{aligned} |x^{(\nu)} - \alpha| &\approx |x^{(\nu)} - x^{(\nu+1)}| \\ &= \left| \frac{f(x^{(\nu)})}{f'(x^{(\nu)})} \right| \\ &\leq \frac{\delta}{|f'(x^{(\nu)})|} \end{aligned}$$

である. α が m 重解のときの計算終了規則で停止したときは (9) 式より

$$\begin{aligned} (x^\nu - \alpha)^m &= \frac{f(x^\nu)}{g(x^\nu)} \\ &\leq \frac{\delta}{g(x^\nu)}, \\ |x^\nu - \alpha| &\leq \left| \frac{\delta}{g(x^\nu)} \right|^{\frac{1}{m}} \end{aligned}$$

となる. すなわち m 重解の精度は, 桁数にして, 計算精度の $\dfrac{1}{m}$ くらいしかないということである.

減速

出発値 $x^{(0)}$ を決めることも難しく, 見当外れの場所から始めると収束どころか発散や振動を起こすこともある. よって, 収束性向上の工夫として「減速」をおこなう. この手法は反復公式 (1) 式の代わりに

$$x^{(\nu+1)} = x^{(\nu)} - \mu^{(\nu)} \frac{f(x^{(\nu)})}{f'(x^{(\nu)})} \tag{13}$$

を用いる.

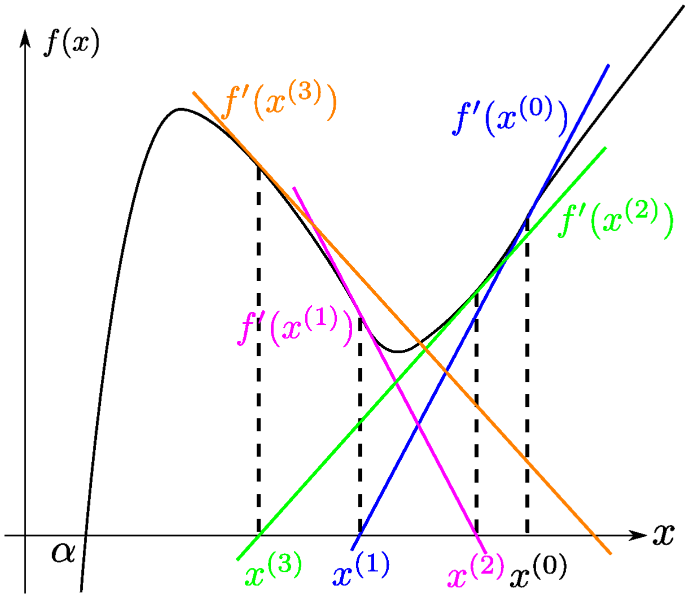

Figure 2: ニュートン法が収束しない例. α が求めたい解. $x^{(0)}$ から始めて, 解を求めようとすると, 途中で勾配 f' が負になってしまい, 収束しない.

ここで $\mu^{(\nu)}$ は各 ν ごとに, まず, $\mu^{(\nu)} = 1$ とおいて $x^{(\nu+1)}$ を計算し, もし

$$\left| f\left(x^{(\nu)} - \mu^{(\nu)} \frac{f(x^{(\nu)})}{f'(x^{(\nu)})} \right) \right| < \left(1 - \frac{\mu^{(\nu)}}{4} \right) |f(x^{(\nu)})| \tag{14}$$

ならばよく, (14) 式が成り立たないなら (14) 式が成立するまで $\mu^{(\nu)}$ を半分に減らしていくようにして定める. かなり緩やかな仮定のもとで, 「そのような $\mu^{(\nu)}$ は必ず存在する」ことが保証される[7]. つまり勾配 f' が負になることで起こる振動を防ごうとしている. 減速ニュートン法を図にまとめたのが図 3 である.

このようにして, 関数 $f(x^{(\nu)})$ の絶対値が次第に減少していけば, いつかは $x^{(\nu)}$ が解の近くにたどりつくであろうと思うのである.

[7] 図 3 の σ は条件 (14) 式が成り立つように $f\left(x^{(\nu)} - \mu^{(\nu)} \frac{f(x^{(\nu)})}{f'(x^{(\nu)})} \right)$ より少し上にするためのもの. $\sigma = \frac{1}{4}$ という数字には意味はない.

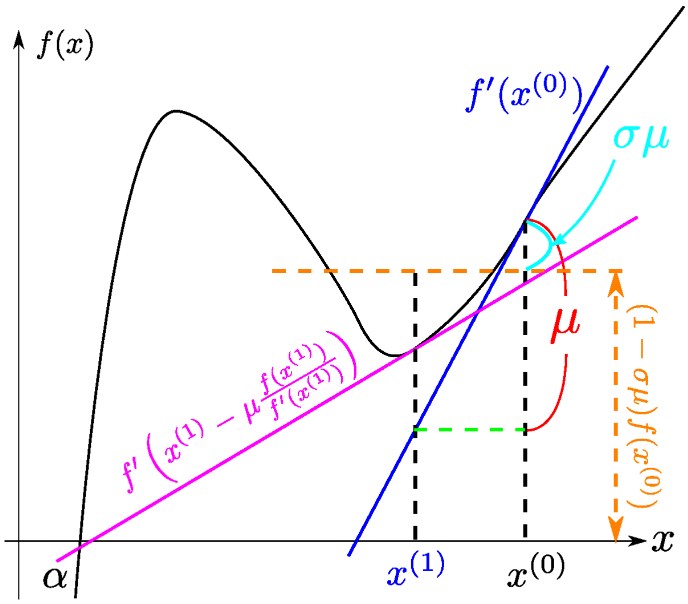

Figure 3: 減速ニュートン法の例. α が求めたい解. $f'(x^{(1)})$ が負にならないように $f\left(x^{(0)} - \mu^{(0)}\dfrac{f(x^{(0)})}{f'(x^{(0)})}\right) < (1-\sigma\mu^{(0)})f(x^{(0)})$ という制限をかけている.

例題

次の関数 $f(x)$ が $f(x) = 0$ となる解をニュートン法を用いて求めなさい.

(1)$f(x) = \tanh(x) + 0.2x + 0.3$

(2)$f(x) = x^5 - 7x^4 + 16x^3 - 8x^2 - 16x + 16$

(2) は解析的に求められて, $f(x) = (x+1)(x-2)^4$ となり, $x = -1$, $x = 2$ である.

参考文献

伊理正夫, 藤野和建, 1985,「数値計算の常識」 共立出版, ISBN 4320013433

川上一郎, 2009,「数値計算の基礎」URL: http://www7.ocn.ne.jp/ kawa1/

大石泰章, 2006, 「2006 年度 数値解析 資料その 4」
URL: http://www.misojiro.t.u-tokyo.ac.jp/ oishi/numeric/handout06-4.pdf