第6章 Newton 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[2)]も極く小さな領域だけ見ればほぼ線形の世界である[3)]」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.

6.1 公式

ニュートン法の公式は反復公式である. 方程式 $f(x) = 0$ の一つの解 α の "近くに" ある初期近似値 $x^{(0)}$ から出発したとき,

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

なので

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

よって,

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

これを一般化すると

$$x^{(\nu+1)} = x^{(\nu)} - \frac{f(x^{(\nu)})}{f'(x^{(\nu)})}. \tag{6.1}$$

[1)] ニュートン・ラフトン (Newton-Raphson) 法ともいう
[2)] 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. 例えばべき級数など.
[3)] 非線形の関数 $f(x+h)$ があるの時 $x = a$ のまわりでのテイラー展開を行う.

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

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

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

右辺は線形なので極く小さな領域だと線形とみなせる.

これを用いて, 次々と望むらく[4)]は改良された近似解 $x^{(1)}, x^{(2)}\cdots$を計算する. さらに多変数の非線形連立方程式

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

の場合にも (6.1) 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. これを一般化すると

$$f_i(x_i^{(\mu+1)})\sim f_i(x_i^{(\mu)})+\sum_{j=1}^{n}\frac{\partial f_i(x_i^{(\mu)})}{\partial x_j}(x_j^{(\mu)}-x_j^{(\mu)}). \tag{6.3'}$$

ここで $x_i^{(\mu+1)}(i=1,\cdots,n)$ が $\alpha_i(i=1,\cdots,n)$ に十分近いとすれば $f_i(x_i^{(\mu+1)})\sim f_i(\alpha_i)=0$ と見なせる. このとき (6.3') は

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

なので

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

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

$$x_i^{(\mu+1)}=x_j^{(\mu)}-\Delta x_j^{(\mu)}. \tag{6.2}$$

$$f_i(x_i^{(\mu)})\sim\sum_{j=1}^{n}\frac{\partial f_i(x_i^{(\mu)})}{\partial x_j}\Delta x_j^{(\mu)}. \tag{6.3}$$

となって (6.1) を多変数に拡張した式となる. (6.3) 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 (6.2) に入れることで解を求められる. $f_i(x)$ が複素変数 x の複素数関数であっても式 (6.3) は形式的にそのまま使える[5)].

[4)] x_o 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 が見当が付いていないと使えない.

[5)] 複素数 z を $z\equiv x+iy$(i は虚数単位) とすると $f(z)=u+iv$ と置ける. ここで $u=u(x,y), v=v(x,y)$ なので $f(z)=f_2(u,v)$ と表せて 2 変数連立方程式と書ける. よって (6.3) を $n=2$ として解けばよい

6.2　収束性

今, 簡単のために連立でない時を考える. 解 α が単純孤立した解で, α の近傍で f が素直な関数でかつ

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

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

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

であり, かつある N より先の ν ($< N$) に対して

$$x^{(\nu+1)} = x^{(\nu)} - \frac{f(x^{(\nu)})}{f'(x^{(\nu)})}. \tag{6.4'}$$

このとき $\epsilon_\nu = x^{(\nu)} - \alpha$ とおく. $f(x^{(\nu)})$ 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 (6.4') の右辺第 2 項に代入する.

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

$\left|\frac{f''(\alpha)}{f'(\alpha)}\epsilon_\nu\right| \ll 1$ より[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

$$
\frac{f(x^k)}{f'(x^k)} \approx -\frac{f''(\alpha)}{2f'(\alpha)}\epsilon_\nu^2+\epsilon_\nu+O(\epsilon_\nu^3). \tag{6.4''}
$$

今 (6.4') 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 (6.4") 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 $O(\epsilon_\nu^3)$ を無視して ϵ の定義を思い出すと

$$
\| x^{(\nu+1)}-\alpha \| \approx \frac{f''(\alpha)}{2f'(\alpha)} \| x^{(\nu)}-\alpha \|^2 . \tag{6.4}
$$

(6.4) 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」. この諸条件は, 満足されているかどうかを前もって具体的に調べる手段がないようなものである. (6.4) 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. $x^{(\mu)}-x^{(\mu+1)}$ が (6.4) の様に振舞いをするようになった場合は

$$
x^{(\nu)}-x^{(\nu+1)} \approx x^{(\nu)}-\alpha
$$

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

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

と近似できることを用いた

とみなしてよい[7].

6.3 計算例

実際に以下の二つの方程式を計算してみる.

(1)$f(x) = \tanh(x) + 0.2x + 0.3 = 0$

(2)$f(x) = x^5 - 7x^4 + 16x^3 - 8x^2 - 16x + 16 = 0$

解答:

(1) のプログラム詳細は付録 1 を参照されたい. 結果は付録 1′ を参照のこと.

(2) のプログラム詳細は付録 2 を参照されたい. 結果は付録 2′ を参照のこと.

7)

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

なので (6.4”) より

$$x^{(\nu)} - x^{(\nu+1)} \approx \frac{f''(\alpha)}{2f'(\alpha)}\epsilon_\nu^2 + \epsilon_\nu + O(\epsilon_\nu^3).$$

$\epsilon_\nu \ll 1$ より

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