

## Newton 法

Newton 法は最も基本的な非線形方程式の反復解法であり，Newton-Raphson 法とも呼ばれる．

### 全般的注意

$x = (x_1, \dots, x_n) \in \mathbf{R}^n$  に関する非線形連立方程式  $f(x) = 0$ ，すなわち

$$f_i(x) = 0 \quad (i = 1, \dots, n) \quad (1)$$

を考える<sup>1</sup>．解  $x = \alpha$  を有限回の四則演算で求めるのは一般に不可能であり，通常は，適当な初期値  $x^{(0)}$  から出発して，なんらかの反復法

$$x^{(\nu+1)} = x^{(\nu)} + \Delta x^{(\nu)} \quad (\nu = 0, 1, \dots) \quad (2)$$

によって  $\alpha$  に収束する近似解列  $\{x^{(\nu)}\}$  を数値的に生成する．

近似解列  $\{x^{(\nu)}\}$  は，有限桁の演算によって計算される  $f_i(x)$  の値などに基づいて生成されるので，必然的に丸め誤差を含み，その結果，数学的な意味で  $\lim_{\nu \rightarrow \infty} x^{(\nu)} = \alpha$  となるようにはできない．使用する計算機などによって自然に定まる丸め誤差限界を  $\delta f_i$  とするとき， $|f_i(x)|$  が  $\delta f_i$  程度以下ならば  $f_i(x) = 0$  と見なさざるをえない．

すなわち，数値計算の立場からは，

方程式 (1) を解くというのは  $|f_i(x)| \leq \delta f_i$  ( $i = 1, \dots, n$ ) を満たす  $x$  を求めることであると考え，

$$|f_i(x^{(\nu)})| \leq \delta f_i \quad (i = 1, \dots, n) \quad (3)$$

を反復法の停止規則にするのが健全である．これとは対照的に，適当に定めたノルムと要求精度  $\varepsilon$  (例えば， $\varepsilon = 10^{-5}$  など) によって  $\|x^{(\nu+1)} - x^{(\nu)}\| \leq \varepsilon$  の形の規準が停止規則に用いられることもあるが， $\varepsilon$  の設定に関する恣意性もあり，上記のような立場からは推奨されない．

### Newton 法の算法 (1 変数の場合)

1 変数の場合の公式は

$$x^{(\nu+1)} = x^{(\nu)} - f(x^{(\nu)})/f'(x^{(\nu)}) \quad (4)$$

となる．これは，式 (2) における修正量  $\Delta x^{(\nu)}$  を， $f(x)$  の  $x = x^{(\nu)}$  付近での線形近似に基づいて定めたものである．実際， $f$  の線形近似式

$$f(x) \simeq f(x^{(\nu)}) + f'(x^{(\nu)})(x - x^{(\nu)}) \quad (5)$$

<sup>1</sup>本来は， $f_i$  が多項式の場合に代数方程式，そうでないとき超越方程式と呼ぶが， $f_i$  が多項式でない一般の場合にも微分方程式に対するものとして代数方程式と呼ぶことがある．

で  $x = \alpha$  とおいた式

$$0 \simeq f(x^{(\nu)}) + f'(x^{(\nu)})(\alpha - x^{(\nu)}) \quad (6)$$

を  $\alpha$  について解けば

$$\alpha \simeq x^{(\nu)} - f'(x^{(\nu)})^{-1} f(x^{(\nu)}) \quad (7)$$

となる．この右辺を新たな近似解としたものが式 (4) である．

幾何学的には，曲線  $y = f(x)$  の点  $(x^{(\nu)}, f(x^{(\nu)}))$  における接線と  $x$  軸との交点の  $x$  座標が  $x^{(\nu+1)}$  である．なお，式 (4) の右辺を変形して

$$x^{(\nu+1)} = [x^{(\nu)} f'(x^{(\nu)}) - f(x^{(\nu)})] / f'(x^{(\nu)}) \quad (8)$$

の形で計算するのは，丸め誤差の観点から不利である．

導関数の値は数値微分で代用することも多い．その場合，刻み幅  $h$  の前進差分  $[f(x+h) - f(x)]/h$  で近似する． $h$  の選び方は， $f(x)$  の計算における丸め誤差から決まる精度の限界を  $\delta f$  として，

$$h \simeq 2\sqrt{\delta f / |f''(x)|} \quad (9)$$

とするとよい．

実際に方程式を解くときには初期値の選び方が重要となる<sup>2</sup>．また，近似解の振舞いを安定化する工夫として減速の手法（後述）がある．

### Newton 法の算法 (多変数の場合)

多変数の場合の公式は

$$\mathbf{x}^{(\nu+1)} = \mathbf{x}^{(\nu)} - J(\mathbf{x}^{(\nu)})^{-1} \mathbf{f}(\mathbf{x}^{(\nu)}) \quad (10)$$

となる．ここで，

$$J = J(\mathbf{x}) = (J_{ij}(\mathbf{x})) = (\partial f_i / \partial x_j) \quad (11)$$

は  $f(\mathbf{x})$  の Jacobi 行列であり，その存在と正則性は仮定している．実際の計算は， $\mathbf{d}^{(\nu)}$  に関する線形方程式

$$J(\mathbf{x}^{(\nu)}) \mathbf{d}^{(\nu)} = -\mathbf{f}(\mathbf{x}^{(\nu)}) \quad (12)$$

を解いて， $\mathbf{x}^{(\nu+1)} = \mathbf{x}^{(\nu)} + \mathbf{d}^{(\nu)}$  とする．

式 (10) は，式 (2) における修正量  $\Delta \mathbf{x}^{(\nu)}$  を， $f(\mathbf{x})$  の  $\mathbf{x} = \mathbf{x}^{(\nu)}$  付近での線形近似に基づいて定めたものである．線形近似式

$$\mathbf{f}(\mathbf{x}) \simeq \mathbf{f}(\mathbf{x}^{(\nu)}) + J(\mathbf{x}^{(\nu)})(\mathbf{x} - \mathbf{x}^{(\nu)}) \quad (13)$$

で  $\mathbf{x} = \alpha$  とおいた式

$$\mathbf{0} \simeq \mathbf{f}(\mathbf{x}^{(\nu)}) + J(\mathbf{x}^{(\nu)})(\alpha - \mathbf{x}^{(\nu)}) \quad (14)$$

を  $\alpha$  について解けば

$$\alpha \simeq \mathbf{x}^{(\nu)} - J(\mathbf{x}^{(\nu)})^{-1} \mathbf{f}(\mathbf{x}^{(\nu)}) \quad (15)$$

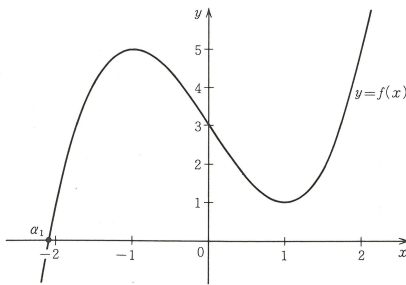
<sup>2</sup>例えば，連続変形法と呼ばれる方法で粗い近似解を求めて，これを初期値とすることができる．

表 1: Newton 法による方程式の解法

$$f(x) = x^3 - 3x + 3; \quad \alpha_1 = -2.1038034\dots$$

$\nu$	$x^{(\nu)}$	$f(x^{(\nu)})$
0	-3.00000	-1.50 E+1
1	-2.37500	-3.27 E+0
2	-2.14001	-3.80 E-1
3	-2.10458	-8.01 E-3
4	-2.10380	-9.54 E-7

$\nu$	$x^{(\nu)}$	$f(x^{(\nu)})$
0	2.50000	1.11 E+1
1	1.79365	3.39 E+0
2	1.28406	1.26 E+0
3	0.634163	1.35 E+0
4	1.38830	1.51 E+0
5	0.845229	1.07 E+0
6	2.09196	5.88 E+0
7	1.51152	1.92 E+0
8	1.01366	1.00 E+0
9	-11.1105	-1.34 E+3
10	-7.47567	-3.92 E+2
11	-5.09281	-1.14 E+2
12	-3.57146	-3.18 E+1
13	-2.66858	-8.00 E+0
14	-2.23305	-1.44 E+0
15	-2.11298	-9.48 E-2
16	-2.10385	-5.20 E-4
17	-2.10380	-9.54 E-7



となる．この右辺を新たな近似解としたものが式 (10) である．このような Newton 法の考え方は， $x$  が複素数ベクトルのときにも適用できる．さらに，抽象的な関数空間における作用素を含んだ方程式に対しても Newton 法を拡張することができる．

多変数の場合も，導関数の値を数値微分で代用することが多い．各方向の刻み幅  $h_j$  ( $j = 1, \dots, n$ ) を適当に定めて，

$$J_{ij}(x) \simeq [f_i(x + h_j e_j) - f_i(x)]/h_j \quad (16)$$

(ただし  $e_j$  は第  $j$  単位ベクトル) と近似することにすれば，1 反復あたり関数  $f$  を  $n+1$  回計算すれば済む．刻み幅  $h$  の選び方は，1 変数の場合と同様である．

### 例題

3 次方程式  $f(x) = x^3 - 3x + 3 = 0$  は唯一の実根  $\alpha_1 = -2.1038034\dots$  をもつ． $x^{(0)} = -3$  として Newton 法を適用すると，表 1(a) のように速やかに  $x^{(\nu)} \rightarrow \alpha_1$  となる． $\delta f = 10^{-5}$  とすると， $x^{(4)} = -2.10380$ ， $f(x^{(4)}) = -9.5 \times 10^{-7}$  で反復を停止する． $x^{(0)} = 2.5$  とすると， $x^{(\nu)}$  は表 1(b) のような激しい挙動を示す．結局は  $x^{(\nu)} \rightarrow \alpha_1$  となるが安心して使えない．初期値の取り方が大切である．

### Newton 法の収束性

Newton 法の解の近傍での振舞い (局所的収束性) は理論上も実用上も十分満足できるものであり，よい初期値から出発すれば十分速く解に収束することが保証される．一方，解

から遠い所での振舞い(大域的収束性)については, 近似解が大きく変動したり, 周期的に振動したりするなど, 問題点もある.

近似解  $x^{(\nu)}$  が真の解  $\alpha$  に十分近いとき, 1 回の Newton 反復でどの程度近似が改善されるかを調べよう. まず 1 変数 ( $n = 1$ ) の場合を考え,  $\varepsilon = x^{(\nu)} - \alpha$ ,  $\tilde{\varepsilon} = x^{(\nu+1)} - \alpha$  とおく.  $x = \alpha$  が  $f(x) = 0$  の  $M$  重根とすると,  $x$  が  $\alpha$  に近いとき

$$f(x) \simeq a(x - \alpha)^M [1 + b(x - \alpha)] \quad (a \neq 0) \quad (17)$$

とかけるので, 式 (4) より

$$\tilde{\varepsilon} \simeq \varepsilon - \frac{a\varepsilon^M [1 + b\varepsilon]}{a\varepsilon^{M-1} [M + (M+1)b\varepsilon]} \simeq \left(1 - \frac{1}{M}\right) \varepsilon + \frac{b}{M^2} \varepsilon^2 \quad (18)$$

である. したがって,  $\alpha$  が単根 ( $M = 1$ ) なら 2 次収束, 重根 ( $M \geq 2$ ) なら 1 次収束である. なお,  $M$  が既知ならば, 修正量を  $M$  倍して

$$x^{(\nu+1)} = x^{(\nu)} - Mf(x^{(\nu)})/f'(x^{(\nu)}) \quad (19)$$

とすると,  $\{x^{(\nu)}\}$  は  $\alpha$  に 2 次収束する.

多変数 ( $n \geq 2$ ) の場合も,  $J(\alpha)$  が正則ならば ( $n = 1$  の単根の場合に対応して) 2 次収束する. 次の定理では,  $J(x)$  の Lipschitz 連続性

$$\|J(x) - J(y)\| \leq L\|x - y\| \quad (x, y \in D) \quad (20)$$

が仮定されるが, これは“ $f(x)$  が適当に滑らかならば”と読みかえても実用上差し支えない.  $f(x)$  が  $C^2$  級ならば  $J(x)$  は Lipschitz 連続である.

定理 1  $J(x)$  が (1) の解  $\alpha$  の近傍で Lipschitz 連続 (20) で,  $J(\alpha)$  が正則ならば,  $\alpha$  に十分近い初期値から出発する Newton 法の近似解列は  $\alpha$  に 2 次収束する.

補足: 十分大きい  $\nu$  に対して  $\|x^{(\nu+1)} - \alpha\| \leq C\|x^{(\nu)} - \alpha\|^p$  ( $C$  は定数) のとき,  $\{x^{(\nu)}\}$  は  $\alpha$  に  $p$  次収束するといい,  $p$  を収束次数と呼ぶ. とくに,  $p = 1$  のとき,  $C < 1$  を収束率という.

## 減速

Newton 法の大域的収束性を改善するために, 減速と呼ばれる手法が有効な場合がある. これは, (12) で定めた修正方向ベクトル  $d^{(\nu)}$  を用いながら, (適当なノルムで測った) 関数値の大きさ  $\|f(x^{(\nu)})\|$  を単調に減らそうとするものである.

適当にパラメタ  $0 < \beta < 1$ ,  $\lambda > 1$  を定めておいて,

$$\|f(x^{(\nu)} + \mu d^{(\nu)})\| \leq (1 - (1 - \beta)\mu)\|f(x^{(\nu)})\| \quad (21)$$

を満たす  $\mu (> 0)$  を  $\mu = 1, \lambda^{-1}, \lambda^{-2}, \dots$  の順に探していき, そのような  $\mu$  が見つかったら  $\Delta x^{(\nu)} = \mu d^{(\nu)}$  とする.  $J(x^{(\nu)})$  が正則ならば, このような正の数  $\mu = \mu^{(\nu)}$  が存在することが知られている. このとき  $\{\|f(x^{(\nu)})\|\}_\nu$  は単調減少列になる (しかし, 0 に収束するとはかぎらない).

以上 (2013-08-05)

[この資料は, 杉原正顯, 室田一雄, 数値計算法の数理, 4 章 (岩波書店, 1994) の記述に基づいて編集したものである.]