ニュートン法の話: 近似法から微分方程式の解の存在証明まで

岡本 久
〒 606-8502　京都市左京区北白川追分町
京都大学数理解析研究所

平成 31 年 4 月 13 日

1　序

大学の数値解析の講義で習うニュートン法は通常次の形で与えられる．関数 $f : [a, b] \to \mathbb{R}$ が根 α を持つと仮定する．x_0 を α に近い値とする．そして $n = 0, 1, 2, \cdots$ に対して次の操作を繰り返す:

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}. \tag{1}$$

f と α に適当な仮定をおくと, $\{x_n\}$ は極めて速く α に収束する．また, このプロセスはしばしば図 1 のように接線を逐次構成することと解釈できることも周知のこととなっており, 高校の教科書にも書かれている．

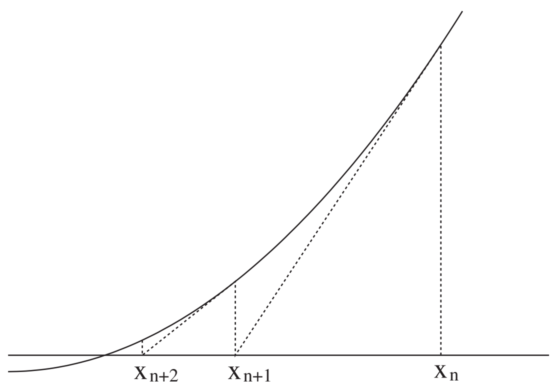

図 1: ニュートン法のグラフ的な解釈

ニュートン法はきわめて精度がよい．一例として $f(x) = x^2 - 2$ の場合に $x_0 = 2.0$ として計算してみよう．$x_n - \sqrt{2}$ の値を列挙すると,

```
n=0   0.585786437626905
n=1   0.085786437626905
n=2   0.002453104293571
n=3   0.000002123901415
n=4   0.000000000001595
n=5   0.000000000000000
```

となる. x_4 で 10 桁以上正しくなるわけである. ゼロへの収束の仕方がきわめて急速であることも理解できるであろう.

ニュートン法にはまた, 思わぬ応用もある. 次の問題はクイズのようなものであるが, 頭の体操にはいいかもしれないと思い, ここに掲示しておく. 解答は本稿最後に記すので参照していただきたい.

練習問題 コンピュータの計算言語には通常 $f(x) = 1/x$ という関数が組み込まれており, $a \neq 0$ が与えられたとき $1/a$ は簡単に計算できる. しかし, もし加減乗除のうち除算がない計算言語が与えられたときには割り算はできなくなるのであろうか？ 実は, ニュートン法を使えば (加減乗算のみを使って) 除算を高速かつ高精度にプログラムすることができる. どうすればよいか？

ニュートン法は数値解析では必須項目であるが, その歴史を語られることは少ない. これは不幸なことである. 古典的な文献を読んでみると意外に思われることが多々発見される. それが次の発見につながることも多いから歴史をきちんと認識することは現在の研究を進める上でも有用である. 本稿はニュートン法の歴史について筆者が学んだことをまとめたものである.

ニュートン法は元来, 精度の高い数値計算法にすぎなかった. しかし２０世紀には, 微分方程式の解の存在証明のためにニュートン法を使うようになり, 様々な応用が見いだされるようになった. こうした事情を時の流れに沿って見てみよう.

2 ニュートンの方法とニュートン法

ニュートン[1] は流率法 (すなわち微分法) の発見者として有名である ([2]). 通常, ニュートン法が歴史上最初に現れるのは彼の De Analysi $\cdots$ [25] で, これは 1711 年まで出版されることはなかったのであるが, 元の原稿は 1669 年に書かれたというのが定説である[2]. この論文こそ彼の流率法の概要が世に示された最初の論文なのであるが, これはニュートン法が現れた論文としても画期的なものである. 彼はそのアイデアを次の例で示している: 3 次方程式

$$x^3 - 2x - 5 = 0$$

の根を近似するのに, 近似値 2 から出発する. 根を $x = 2 + p$ とおいて方程式に代入すると,

$$p^3 + 6p^2 + 10p - 1 = 0 \tag{2}$$

[1] Isaac Newton, 1643–1727(1642 生まれとすることも多い)

[2] 文献 [25] の Whiteside による解説参照

となる．視察によって根は２に近いということを彼は知っていたから, p は小さい．したがって２次以上のべきは無視できる．すると $10p-1=0$ すなわち, $p=0.1$ を得る．さらに良い近似を得たいならば $p=0.1+q$ とおいて (2) に代入すると

$$q^3+6.3q^2+11.23q+0.061=0 \tag{3}$$

を得る． 同じ規則に従えば $q \approx -0.061/11.23 = -0.005431878\cdots$． 近似値は $2.0+0.1-0.005431878=2.09456812$. これは５桁正しい.

この規則が前節の方法と数学的に同等であることを見ることは比較的易しい．実際, $f(x)=0$ の近似値 x_0 が与えられたとき, (2) の左辺は

$$0=f(x_0+p)=f(x_0)+f'(x_0)p+\frac{1}{2}f''(x_0)p^2+\cdots$$

に他ならない．２次以上のべきを無視すれば $p=-f(x_0)/f'(x_0)$ であるから, 近似値を

$$x_0-\frac{f(x_0)}{f'(x_0)}$$

と取り直していることになるが, これはすなわち (1) に他ならない.

この理由により, (1) はニュートン法と呼ばれているわけであるが, この名前に抵抗を感じている人も多い．文献 [8][23][28] などがそうである．その理由は, ニュートンの方法が逐次代入形式になっていないということと微分の概念を用いていないという２点にあるようである ([23]). 逐次近似になっていないというのは次のことを意味している．なるほど数学的には (1) とニュートンの方法は同値であると解釈できるが, ニュートンの方法では p に対する多項式から q に対する多項式を新たに計算し直さねばならない．これは計算をさらに進めてみても同様である．計算の手間という点では大違いである．文献 [28] によれば, ニュートンは後にプリンキピアの中で逐次近似的なニュートン法を使っているという．これはプリンキピアの命題 31 の補足説明 (Scholium) であるというのだが, ここで彼は $x-e\sin x=a$ (a,e は与えられた定数) の根の近似を逐次的に計算しているという．しかし, これを読んでみても幾何学的な論証があるのみであり, これが (1) と同等である, ということを見抜くことは素人には至難の業である．結論として, 筆者も [8][23][28] といった人々と意見を共有する: ニュートンの方法とニュートン法は同じ物ではないのである．こうした判断は個人的な好悪も絡むものでもあり, 非常に難しい．しかし, 昨今のように先取権争いが激しくなっている現代ではもっと議論すべき問題であろう.

De Analysi は 1711 年まで出版されることはなかったが, 原稿のコピーはイギリスの数学者の間で回覧されていたということなので, 彼の近似方法は公然のものであったということができる．さらにウォリス[3]が 1685 年に刊行した代数学の教科書 Algebra において彼の方法を紹介しているので, 少なくともこの時点で世に現れていたということになる.

なお, De Analysi はもともと流率法の解説をしたものであり, なぜその中に数値計算法が現れるのか, 不思議に思われる方も多いと思う．彼の意図は De Analysi の中のニュートン法の部分の直後を読まないとわからない．彼は様々な関数をべき級数で展開することができることを

[3] John Wallis, 1616–1703

例示しているのであるが, その中で $y^3+axy+aay-x^3-2a^3=0$ で[4]定まる代数関数 $y=y(x)$ をべき級数に展開する方法を示すために, まずは数に関する方程式でウォーミングアップをして, これから上の代数関数の展開をほぼ同じ方法で導いているのである. 少なくともこの時点で彼の念頭にあったのは数値計算というよりはべき級数展開だったのである.

微分を使っていないという点は, 扱っている方程式が多項式である限り大した欠点にはならないが, 超越方程式の根を計算しようとすれば導関数の使用は不可欠であるといってもよい. これについては, 「本質的に微分を使っているわけであるからけちをつけなくてもいいではないか」という言い方も成り立つかもしれない. しかし, 計算する立場から言えば, 数値計算というのはあくまで簡単な手続き (アルゴリズム) が確立していることが重要なのである. そういう形式をもっていないニュートンの方法は我々が現在教えられているニュートン法とは異質のものであると言わねばならない. またよく知られている接線を使った幾何学的解釈 (図 1) はニュートンの著作のどこにも見あたらない.

では, (1) は誰が考えついたものであろうか？これについて答えるためにはラフソンとシンプソンについて話をする必要がある.

3　ラフソン

ラフソン[5]は 1690 年に Analysis Aequationum Universalis という本を出版し, その中でニュートン法を提案した ([5],[23]). ニュートン自身の論文の出版はこれよりもずっと後のことであるので, いくつかの解説にはこの本こそ印刷公表されたニュートン法の最初のものであると書いてあるものもあるが, 上で見たようにウォリスの本の中に書かれているのでこれは間違いである. 現在でもニュートン法と言わず, ニュートン・ラフソン法と呼ばれることも少なくないが, これは決して理由のないことではない. ラフソンも自分の方法はニュートンの方法とは別物であると思っていたふしがあるという (コラーストロム [23])[6].

コラーストロム [23] によれば, ラフソンもニュートンと同じく $x^3-2x-5=0$ を考えるが, それだけではなく彼はもう少し一般に $x^3-bx=c$ を考えている. そして

$$x \mapsto x - \frac{x^3-bx-c}{3x^2-b}$$

を考えている点で逐次近似の形を完成させているという. ただ, 同時に, この分母が分子の導関数であることについては一言もふれていないという. したがって彼のニュートン法は不完全なものである. これがコラーストロムの主張するところである. カジョリ [8] によれば, ラフソンは一般の多項式に対して下の表のように「ルール」を定めている. ここで, g は変数であり, b,c,d,f は定数である. 明らかに微分の操作であるが, ラフソンは微分（流率）という概念は全く使っておらず, これを完全に代数的なものであると理解していたふしがあるという. g^n というデカルトの記法も用いていらず, ラフソンの方法がそれほど画期的なものかどうかは怪しい.

[4]当時は a^2 と書かず, aa と書いた. 一方 $n \geq 3$ ならば a^n と記した.

[5]Joseph Raphson, 1648–1715

[6]19 世紀にはニュートン法と呼ばれることが多かった. ２０世紀中頃の文献ではニュートン・ラフソン法と呼ばれることが多かった. これは [8] の影響が大きいという ([23]). 現在では再びニュートン法と呼ばれることが多い.

$ggggg$	$5gggg$
$bgggg$	$4bggg$
$cggg$	$3cgg$
dgg	$2dg$
fg	f

いずれにせよ逐次近似としての形が出来ているわけであるから, 計算の手間はニュートンの方法に比べて格段に便利になっている. 文献 [5, 28] によれば, 18 世紀後半の数学者はニュートンの方法とラフソンの方法と比べ, ラフソンの方法が優れているというコンセンサスを持っていたようである.

4 シンプソン

シンプソン[7]の名前を最も有名にしているのは数値積分におけるシンプソン公式であろう. これは精度が比較的高いので現在でも使われる積分近似式である. しかし, 彼が公式 (1) を 1740 年に公表したことはあまり知られていない. 我々が知っているニュートン法の公式は実はシンプソンのものだったのである. これをもってコラーストロム [23] はニュートン法でもなくニュートン・ラフソン法でもなくニュートン・ラフソン・シンプソン法という名前こそ正しい名前であると言っている. なぜシンプソンの貢献が無視されてきたのか私は知らない.

シンプソンの貢献は単に形式だけのものではない. 実際, 彼は多変数の問題にもニュートン法を定式化しているが, 彼以前は 1 変数の方程式しか考えられていなかった. 多変数関数への拡張は, 現在の我々の目から見ればそれほど大きな拡張とは見えないかもしれないが, 当時としては画期的なことと言える [23].

歴史というのは不思議なもので, シンプソンは現代版ニュートン法の提唱者であるにもかかわらずその名を残すことは出来なかったが, 一方で数値積分のシンプソン公式はすでにニュートンが発見したものであった (このことはシンプソン自身も認めていることらしい [23]) のに彼の名前で呼ばれている.

ニュートン法は 18 世紀を通してさまざまな人々によって使われてきた. たとえばオイラーは, 1770 年に書いた代数学の教科書 [12] の第 16 章においてニュートン法を解説しているが, これを純粋に代数的なプロセスとしているように思える. 微分を使った形跡はない. これから見てもシンプソンの先見の明は賞賛に値する.

5 収束の吟味・誤差解析

さて, 18 世紀も後半になればニュートン法はよく知られたものとなっていたが, その近似解がどのように収束するのか, 誤差はどれくらいか, といった問題を理論的に解明しようとする動きはなかなか出なかったようである (オイラー [12] には誤差を理論的に見積もろうとする姿勢も

[7] Thomas Simpson, 1710–1761

見られない). ニュートン法の誤差解析を最初に行ったのが誰なのか筆者は知らないが, グラビナー [17] はラグランジュ[8]が最初であるとみているようだ. しかしながらこれはムレイユ[9]の業績を無視した話であり, 受け入れがたい. カジョリ [7] に従ってムレイユの仕事を紹介しよう. 彼は 1768 年に Traité de la résolution des équations en géneral という書物を発表し, そこでニュートン法についても論じている. へたな初期値を選べばニュートン法は収束しないということは当時常識になっていたようであるが, 彼の優れた点は, どのような条件があればニュートン点列が収束するかを考察したことである. たとえば, 適当な初期値から出発するとニュートン点列が $x_{n+2} = x_n$ を満たす周期列になることも指摘している (付録参照).

さらに, ニュートンもその後継者もこの方法を代数的あるいは解析的に考察したのに対し, ムレイユは幾何学的かつグラフ的に考察して簡単な収束の十分条件を導いている. ムレイユの主張によれば図 1 のような解釈は彼以前にはないという.

ラグランジュは彼の教科書 [24] においてニュートン法の収束を考察した. $f(x)$ が実数の根のみを持つと仮定し, それらを $\alpha_0, \alpha_1, \cdots, \alpha_{N-1}$ とする. つまり, $f(x) =$ 定数 $\times (x - \alpha_0)(x - \alpha_1) \cdots (x - \alpha_{N-1})$. そして, α_0 に近い値 β を初期値にとってニュートン法を適用する場合を考える. 彼はこのとき

$$\gamma = \beta - \frac{f(\beta)}{f'(\beta)}$$

で定義される γ が β よりもさらによい近似になるためにはどのような条件があればよいか？という問題を考察した. 彼の得た結果は,

$$R = \frac{1}{\alpha_1 - \beta} + \frac{1}{\alpha_2 - \beta} + \cdots + \frac{1}{\alpha_{N-1} - \beta}$$

と定義したときに,

$$2(\alpha_0 - \beta)R + 1 > 0 \tag{4}$$

であれば $|\gamma - \alpha_0| < |\beta - \alpha_0|$ になる, というものである. 本来計算し得ない量である R に依存する十分条件であるからそれほど印象的なものではないが, 近似が良くなることを保証する条件は何か？と疑問に思う姿勢はオイラーとは明らかに違っている. ムレイユを除けば彼以前の数学者は, 様々な具体例で近似の精度が優れていることを確かめただけであって, 一般論を展開しようとはしていないのである ([17]). 誤差を事前に見積もろうとした彼の研究姿勢がその後のフーリエやコーシーの研究を刺激したのは間違いない[10].

注意 5.1 条件 (4) は全く無意味なものではない. たとえば, $\{\alpha_j\}$ のうち α_0 が最小ならば $\beta < \alpha_0$ なる任意の β について (4) は満たされる. α_0 が最大で $\alpha_0 < \beta$ でも同様である.

しかしながら, グラビナーが [17] の 67 ページで指摘しているように, ラグランジュのやったことをほめすぎるのも良くないことである. ラグランジュは第 1 近似が第 0 近似よりも良くなる条件を求めただけであって, つぎつぎと得られる近似解が直前の近似解よりも常に良くなるような条件を導いたわけではないのである. 近似解が収束することを証明したのはコーシーが最初のようである.

[8] Joseph-Louis Lagrange, 1736–1813

[9] J. Raym. Mourraille

[10] ムレイユの著書はほとんど読まれることなく, 後世に与えた影響はほぼゼロであるようだ.

6　フーリエの貢献・コーシーの解析

コーシーの収束証明に入る前にフーリエについて述べておこう．フーリエと言えば熱の理論とフーリエ級数ということになるが, 彼はニュートン法でも重要な脇役としての働きをしている．フーリエは 1818 年の論文 [15] および彼の死後に出版された代数学の教科書 Analyse des Équations déterminées[16] においてニュートン法について考察している．彼は,「以前から知られているように, もし初期近似があまりよくないと, 逐次近似列は, 近似はおろか真の値からどんどん離れてゆく」と問題点を指摘し, 引き続いてニュートン法に対するラグランジュの解析を引用する．そして,「収束のための条件が未知の根に依存しているだけになおさら収束の問題は難しい」というラグランジュの言葉を引用する ([15])．そして, 収束のためのはっきりとした十分条件が必要であることをフーリエは強調し, うまく初期値を選べばニュートン列が収束することを主張する．次の定理はフーリエの定理と呼ばれている:

定理 6.1 区間 $[a, b]$ で C^2 級の関数 f がただ一つの根を $a < x < b$ に持つものと仮定する．さらに, すべての $a \leq x \leq b$ に対して $f'(x) \neq 0, f''(x) \neq 0$ であると仮定する[11]．このとき, もしも $f(a)f''(a) > 0$ ならば a を初期値とするニュートン数列は根に収束する．また, $f(b)f''(b) > 0$ ならば b を初期値とするニュートン数列は根に収束する．

グラフを描いてみればわかるように, この定理の仮定のもとで近似列は単調になり, その事実を使えば定理の証明はそれほど難しいものではない．

フーリエは根の上界と下界について次のような事実に言及している．今, ξ を根とし, $a < \xi < b$, $f(a)f''(a) > 0$ としよう．このとき $x_0 = a$ から出発するニュートン近似列は ξ に下から単調に収束する．$x_0 < x_1 < x_2 < \cdots < \xi$．一方, $f(b)f''(b) < 0$ ならば根の上限は b から出発するニュートン列からは必ずしも得られない．そこで,

$$y_0 = b, \qquad y_{n+1} = y_n - \frac{f(y_n)}{f'(x_n)}$$

を考える．ここで x_n は a から出発するニュートン点列である．こうすると $\xi < \cdots < y_2 < y_1 < y_0 = b$ がわかる．彼はこうして根の上限と下限を得ることに初めて成功した ([26])．しかし, 彼は誤差がどれくらい速く減衰するかは示せなかった．

ここまでラグランジュとフーリエの問題意識を述べてきたが, 収束に関する決定的な貢献はコーシーの 1823 年のエコールポリテクニクの教科書 [9] であって, 現在から見ればラグランジュの研究ももフーリエの研究も過渡的なものと言わざるを得ない．コーシーは [9] の 575 ページにおいて次のような結果を証明する．まず, a を与えられた数とする．そして

$$i = -\frac{f(a)}{f'(a)}$$

とおく．区間 $[a, a+2i]$ (i が負のときは $[a+2i, a]$) を K とおき,

$$B := \max_{x \in K} |f''(x)|$$

とおく．

[11] 条件 $f'(x) \neq 0$ は必要ないことが知られている．この拡張された定理は通常ダルブーの定理と呼ばれている．

定理 6.2 もし $|f'(a)| > 2B|i|$ ならば $f(x) = 0$ は区間 K の中にただ一つの根を持つ.

彼の証明を読めばわかるのだが, この定理は単なる中間値の定理の応用にすぎない[12]. しかし, 次の定理はニュートン法の収束に関する重要な発見である.

定理 6.3 ([9],576 ページ) $b = a + i$ とし,

$$j = -\frac{f(b)}{f'(b)}$$

とおく. さらに,

$$A := \min_{x \in K} |f'(x)|$$

とおく. このとき, もし $2B|i| < A$ ならば

$$|j| < \frac{B}{2A}|i|^2$$

であり, b も $b + 2j$ も a と $a + 2i$ の内部に属する. さらに, 根は b と $b + 2j$ の間に存在する.

彼はこの定理に引き続いて次の系の証明に移る.

系 6.1 $\rho = |i|$, $\epsilon = B\rho/(2A)$ とおく. ニュートン点列を $a, b, c, d, \cdots$ とするとき[13] それに対応する区間の長さ, すなわち

$$i = -\frac{f(a)}{f'(a)}, \quad j = -\frac{f(b)}{f'(b)}, \quad k = -\frac{f(c)}{f'(c)}, \quad \ell = -\frac{f(d)}{f'(d)}, \quad \cdots$$

とおいたときの $2|i|, 2|j|, 2|k|, 2|\ell|, \cdots$ は, それぞれ

$$2\rho, \quad 2\rho\epsilon, \quad 2\rho\epsilon^3, \quad 2\rho\epsilon^7, \cdots$$

で上から押さえられる.

彼の結果を現代の用語をつかって説明するために次のような用語を用いることにする.

定義 6.1 数列 $\{x_n\}_{n=0}^{\infty}$ が $x_n \neq \xi$ を満たしながら ξ に収束するものとする. このとき $\{x_n\}$ が s 次収束するとは

$$\lim_{n \to \infty} \frac{|x_{n+1} - \xi|}{|x_n - \xi|^s} = C \in (0, \infty)$$

が成り立つこととする.

たとえば数列 $x_n = 2^{-n}$ はゼロに 1 次収束するし, $x_n = 2^{-2^n}$ はゼロに 2 次収束する. 一般に s が大きいほど速く収束する.

彼はニュートン数列が収束するとは明言していないけれども, 彼の命題をまとめれば, 根の存在とニュートン数列の収束を証明したと言っても言い過ぎではない. 系 6.1 は現代的な用語で言えば $|x_{n+1} - x_n| = O(\epsilon^{2^n})$ になっている訳であるから, ほとんど 2 次収束の証明に達している. しかし, 厳密に言えば, 2 次収束を理解していたわけではない. もちろん, 彼の証明をまねすれば, c を正定数として $|x_{n+1} - \xi| \leq c|x_n - \xi|^2$ となることは簡単に証明できるけれど.

[12]だから彼もその意義をよく理解していなかったのかもしれない. コーシーは非常に多くの論文を書いたので彼の導いた定理がどういう意義を持っているのか, 自分自身で良く理解していなかったことが多い. たとえばコーシー・シュワルツの不等式など, その重要性に気づいていなかったふしがある.

[13]コーシーは a_n といった表記を用いていない.

7　高次収束のアルゴリズム

上で述べたようにニュートン法は 2 次収束である．a が $0 < a < 1$ を満たすとき, a^n はゼロに 1 次収束する．a^{2^n} は 2 次収束である．では, 3 次収束するアルゴリズムはあるのであろうか, と疑問に思うことは当然である．そういうアルゴリズムは確かに存在する．実際上はニュートン法が十分早いので 3 次以上のアルゴリズムを考える御利益はないのであるが, 数学的な問題としては面白いものなのでここでその一部を見ておこう．

こうした拡張がいったいいつ頃から考えられるようになったのか, 筆者はよく知らない．しかし, [18] の 96 ページには興味のある記述がある．1694 年にハリー[14]が線形近似ではなく 2 次多項式による近似を用いて計算したというのである． たとえば, 第 1 節の例 (3) において $11.23q + 0.061 = 0$ の代わりに $6.3q^2 + 11.23q + 0.061 = 0$ を用いることにすれば

$$q = \frac{-11.23 + \sqrt{11.23^2 - 4 \times 6.3 \times 0.061}}{12.6} = -0.005448532949\cdots$$

近似解は $2.1 + q = 2.0945514670\cdots$ これは 7 桁正しい．もともとのニュートン法よりもさらに精密な近似値が得られることがわかる．

a を近似値としたときに $f(a+z) = f(a) + f'(a)z + \frac{1}{2}f''(a)z^2 + \cdots$ である．もし $a+z$ が真の根であるならばこの右辺がゼロになる．そこで,

$$f(a) + f'(a)z + \frac{1}{2}f''(a)z^2 = 0$$

を解いてこれを近似値に使うことを考えれば良い．$|z| \approx 0$ であるから, 2 根のうち

$$z = \frac{-f'(a) + f'(a)\sqrt{1 - \frac{2f(a)f''(a)}{(f'(a))^2}}}{f''(a)}$$

を採用する．つまり,

$$x_{n+1} = x_n + \frac{-f'(x_n) + f'(a)\sqrt{1 - \frac{2f(a)f''(a)}{(f'(a))^2}}}{f''(a)} \tag{5}$$

がハリーの考えたアルゴリズムなのであろう[15]．$f(x_n)f''(x_n)$ が小さいならば

$$\begin{aligned}
&\frac{-f'(x_n) + f'(x_n)\sqrt{1 - 2f(x_n)f''(x_n)/f'(x_n)^2}}{f''(x_n)} \\
&\quad = \frac{-f'(x_n) + f'(x_n)\left\{1 - f(x_n)f''(x_n)/f'(x_n)^2 + O(f(x_n)^2)\right\}}{f''(x_n)} \\
&\quad \approx -\frac{f(x_n)}{f'(x_n)}
\end{aligned}$$

であるから, $f(x_n)f''(x_n)$ が小さいときにはハリーのアルゴリズムはニュートンのアルゴリズムになる．

[14]Edmond Halley, 1656–1742, ハリー彗星で有名なハリーである．
[15]彼の論文を読んでいないので確認はできない

なお, (5) は

$$x_{n+1} = x_n - \frac{2f(x_n)}{f'(x_n)} \frac{1}{1 + \sqrt{1 - 2\frac{f(x_n)}{f'(x_n)}\frac{f''(x_n)}{f'(x_n)}}} \tag{6}$$

と書くこともできる. コーシーはこのアルゴリズムを考察し, $|x_{n+1} - x_n| = O(\epsilon^{3^n})$ を証明している.

ハリーやコーシーの方法では平方根をとっているが, これを使わずに済ませる方法もある. コンピュータがなかった昔, 手計算するしかなかった時代には平方根を求めることは手間のかかることであったから, 平方根を使わずに済ませることができるという事実は大きな利点であったはずである. これは次に述べるシュレーダーのアルゴリズムによって可能になる.

1860 年のシュレーダーの論文 [27] は様々な意味で画期的なものである. 任意次数の近似公式が作れることを示したのは彼が最初であろう. 彼の公式 ([27] の 327 ページ) は次のようなものである.

$$F(z) = z + \sum_{n=1}^{N-1} (-1)^n \frac{f(z)^n}{n!} \left(\frac{1}{f'(z)} \frac{\mathrm{d}}{\mathrm{d}z} \right)^{n-1} \frac{1}{f'(z)}$$

とおいて, $z_{n+1} = F(z_n)$ とすると $\{z_n\}$ は N 次収束する. 簡単なものではあるが, この証明は省略する.

例として $N = 3$ の時を考えてみよう. シュレーダーのアルゴリズムは

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} - \frac{f(x_n)^2 f''(x_n)}{2\left(f'(x_n)\right)^3} \tag{7}$$

となる. 通常のニュートン法と比較してみよう. $f(x) = x^2 - 2, x_0 = 2.0$ で計算し, $x_n - \sqrt{2}$ を書き下してみると

0.585786437626905	0.585786437626905
0.085786437626905	0.023286437626905
0.002453104293572	0.000003042993881
0.000002123901415	0.000000000000000
0.000000000001595	0.000000000000000
0.000000000000000	0.000000000000000

ここで, 左側が通常のニュートン法であり, 右側が (7) である. 3 次収束の早さは明白である.

だが, シュレーダーの方法では 1 ステップ進むために通常のニュートン法よりも多くの四則演算が必要とされていることに注意して欲しい. ニュートン法で x_5 までかかるものがシュレーダー法では x_3 ですむのであるが, 合計の計算の手間としては減っていない. また, 2 階導関数 f'' が必要となることにも注意されたい. ある種の問題では 2 階導関数の計算に著しく時間をとられるが, そのような場合にはニュートン法で x_5 まで行った方が速い. 急がば回れである.

シュレーダーの論文はあまり知られていない. ２０世紀の中頃になって, シュレーダーの論文を知らずに「高次精度のアルゴリズムができた」と報告している論文もある. 世の無常を感ずるけれども, こういうことはしばしば起きる. 科学上の評価を行うにはきちんとした歴史的知識を持たねばならないということを教示しているようにも思える.

8　存在定理とニュートン法

コーシーの解析が現れるまでは, ニュートン法は長い間近似法であった．つまり, 解 ξ の存在を仮定し, それにどれくらい速く近づけるかが問題であった．これに対し, コーシーの論法では解の存在自体も証明されてしまうことが決定的に重要である．解の存在自体をニュートン法を使って証明するという発想はコーシーによってそれほど強く主張されたわけではないので, この点はその後の多くの研究者によって見過ごされてきたのかもしれない．しかし, ファーベル [13] では解の存在は仮定されるものではなく証明されるものであることが明瞭に述べられている．

ファーベルの 1910 年の論文 [13] は最近はほとんど引用されていないと思われるが, 大変興味のあるものである．この中で彼は複素数値関数 f に対して

$$z_{n+1} = z_n - \frac{f(z_n)}{f'(z_n)} \tag{8}$$

によって複素数列 $\{z_n\}_{n=0}^{\infty}$ を定義したとき「どのような条件があればこれが収束し, その収束先が $f(z)=0$ の根になるか？」という問題を提起し, それについて明快な解答を与えている．ここで重要なことは, 彼は f の零点の存在を仮定することなく, 収束先が零点となることを証明している点である．根の存在は仮定し, 近似解がどれくらい精度良く計算できるか, という問題意識とは全く違う問題意識に基づいている点が重要である．もちろん彼はコーシーを読んでいるからこれは彼のオリジナルではないが, 存在定理の重要性を彼はコーシーよりも深く理解していたように思える．

彼は様々な問題設定をしているが, その中のひとつとして次の定理を証明することに成功している:

定理 8.1 $0 < \alpha < 1$ とする．複素平面の点 z_1 を中心とする半径

$$R \equiv \frac{1}{1-\alpha}\left|\frac{f(z_0)}{f'(z_0)}\right|$$

の閉じた円のすべての点で

$$\left|\frac{f(z)f''(z)}{\left(f'(z)\right)^2}\right| < \alpha$$

であることを仮定する．このとき z_0 を初期値とするニュートン法の点列 z_n はすべてこの円の内部に属し, $n \to \infty$ のときに f のゼロ点に収束する．（ゼロ点の存在も同時に証明される）

ファーベルはまた, 微分が出来なくなる点の近傍でどうなるかについても詳しく論じており, これも興味深いものである．たとえば彼は, 実関数 $f(x) = |x|^p$ を考えてみる．ここで $0 < p$ とする．このとき $f(x)/f'(x) = \frac{1}{p}x$ となるから,

$$x_{n+1} = \left(1 - \frac{1}{p}\right)x_n$$

となる．したがって $p > 1/2$ ならばニュートン法の点列は根 $x = 0$ に収束するけれども $p \le 1/2$ では収束しない．

このように画期的な問題意識を持っているファーベルの論文であるが, 現在ではほとんど忘れ去られている．これは残念なことである．

注意 8.1 (8) において収束のための必要十分条件は書き下すことが難しいことが多い. (8) が収束しない z_0 は何か？と考えると複素力学系の世界に入り込んでゆくことになる. 複素力学系についてはたとえば [11] を見よ.

9　多変数関数に対するニュートン法

$f : \mathbb{R}^N \to \mathbb{R}^N$ のヤコビ行列を $f'(x)$ で表そう. f のゼロ点 $\alpha = (\alpha_1, \alpha_2, \cdots, \alpha_N)$ の近傍で $f'(x)$ が正則行列であると仮定する. このとき

$$x_{n+1} = x_n - [f'(x_n)]^{-1} f(x_n)$$

がニュートン法である. 形式的には 1 変数の場合と同じであるが, 実際に計算するときは逆行列を掛けるわけではない. そんなことをすれば無駄な時間がかかるだけである. 数学的には同値であるが, 実際の数値計算には次の手続きを繰り返す:

1. 近似値 $x_n \in \mathbb{R}^N$ を選ぶ.
2. $f(x_n)$ を計算する.
3. 連立方程式 $f'(x_n)y = f(x_n)$ を解く.
4. $x_{n+1} = x_n - y$ とおく.

多変数の場合のニュートン法は 1740 年にシンプソンによって定式化されたが, その収束証明は 1910 年のブリュテル (Blutel[6]) まで待たねばならない. ここで 2 次収束が証明されている.

彼の論文に遅れること数年, アメリカに興味ある研究が現れた. ファイン [14] は $\mathbb{R}^N$ でニュートン法を考え, それが収束することを証明した. ここで彼は根の存在は仮定せず, 定理の結論として導くことができることをはっきりと認識している (Blutel[6] にはその認識がない).

ファインの論文に触発されてベネット [4] が彼の定理を無限次元空間に拡張した. $C[0,1], L^p(0,1)$ における非線形方程式についてもファインのアイデアが使えることを示しており, 次に述べるカントロビッチに３０年ほど先だって, 彼と同様の理論を展開しているのであるが, なぜか現在ではほとんど忘れ去られている. ベネットが論文を書いたときにはまだバナッハ空間の理論も現れていない時代であり, 彼の先見性は賞賛に値する.

10　カントロビッチの定理

カントロビッチの定理は無限次元空間におけるニュートン法の収束のための十分条件を与えるもので, たいへん有名なものである. コーシーもファーベルもファインもある領域全体での $f'(x)$ と $f''(x)$ の評価式を使っていたが, カントロビッチの条件はもっと緩い. [21] の定理 6 (532 頁) を紹介しよう.

定理 10.1 $\|\ \ \|$ をノルムとするバナッハ空間 X の開集合 Ω で定義された写像 $f : \Omega \to X$ が次の条件を満たすものとする.

i) ある $x_0 \in \Omega$ があって, $f'(x_0)^{-1}$ が存在して有界作用素となる[16].

ii) $\|f'(x_0)^{-1}(f(x_0))\| \leq \eta$,

iii) $\|f'(x_0)^{-1}f''(x)\| \leq K$ が $\|x - x_0\| \leq r$ なるすべての x について成立する.

ここで,

$$h = K\eta \leq \frac{1}{2}, \qquad r = \frac{1-\sqrt{1-2h}}{h}\eta$$

と仮定する. このとき x_0 を初期値とするニュートン列は収束し, その収束先は f のゼロ点となる. さらに, $x^* = \lim_{n\to\infty} x_n$ とすると

$$\|x_n - x^*\| \leq \frac{1}{2^n}(2h)^{2^n}\frac{\eta}{h}$$

も成り立つ.

カントロビッチの定理を関数空間で用いると微分方程式の境界値問題の解の存在が証明できる. ここでは次の境界値問題を例にそのアイデアを示そう.

$$u'' + f(u) = 0 \qquad (0 < x < 1) \tag{9}$$

$$u(0) = 0, \qquad u(1) = 0. \tag{10}$$

ここで f は u の非線形関数である. f がべき関数の場合, つまり $f(u) = u^m$ の場合は様々なモデルに現れる. $m = 2$ として微分方程式を差分法で離散化し, それをニュートン法を用いて解くと図 2 のように収束してゆくことがわかる. ($n = 3$ で真の解とほとんど違わない.)

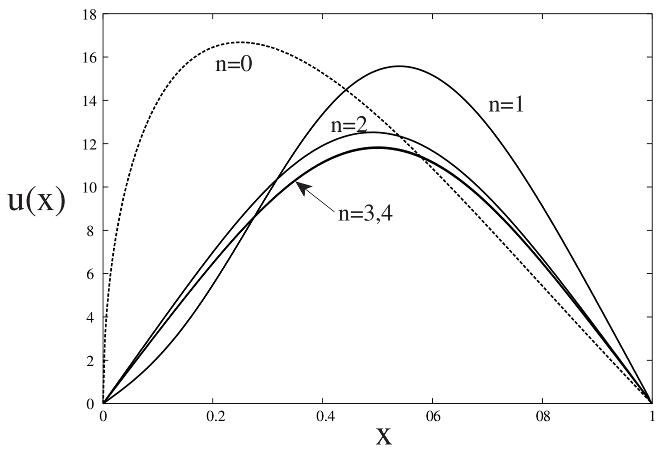

図 2: $u'' + u^2 = 0$ のニュートン近似列. 初期値は $u_0(x) = 1.69\pi^2 \sin(\pi\sqrt{x})$.

この例は比較的簡単なものであるが, アイデアはどのような微分方程式にも適用可能である. たとえば流体力学に現れる複雑な非線形偏微分方程式の解も計算可能となる (図 3).

[16] ここで f' は f のフレッシェ微分を表す

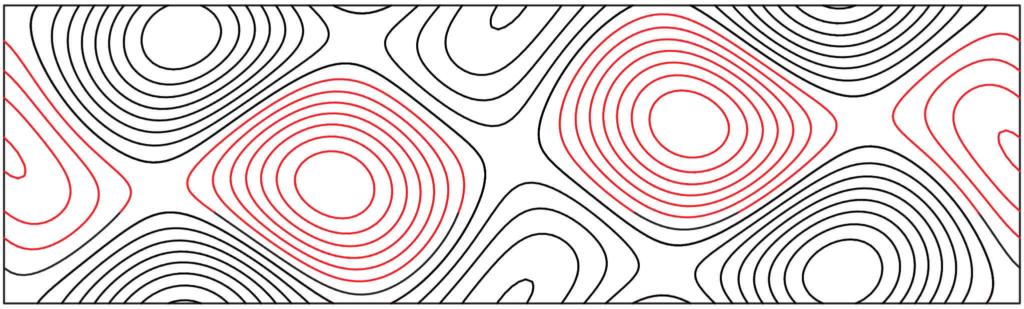

図 3: Navier-Stokes 方程式の解の流線.

11　その他の話題

KAM (Kolmogorov-Arnold-Moser) の理論や Nash-Moser の陰関数定理にもニュートン法は深く関わってくる．こうした理論は保存型微分方程式の解の存在を示すときには強力な武器となる．しかし, 同時にたいへん複雑な理論であるため, 本稿での解説には適さない．文献 [22] を参考としていただきたい．

ニュートン法の変形に割線法[17]と呼ばれる物がある．ふたつの異なる近似値 x_0 と x_1 を用意する．そして

$$x_{n+1} = x_n - \frac{f(x_n)(x_n - x_{n-1})}{f(x_n) - f(x_{n-1})} \qquad (n = 1, 2, \cdots)$$

で逐次近似を構成する．すなわち, 微分を差分で置き換えるわけである．ある種の関数では導関数の計算が困難あるいは不可能な場合がある．このようなときには割線法が威力を発揮する．このときこの近似列は, $\lambda = (1+\sqrt{5})/2$ として λ 次収束することがわかっている．なぜこのようなところに黄金比が出てくるのか不思議であるが, 説明は止める．証明は [10, 20, 26] などをみよ．付録の例からわかるようにニュートン法ほどには速く収束しないが, それでも実用的な範囲では十分役立つものであると言うことができよう．

これ以外にもおもしろい話題はつきないが, 文献 [1][3][19] を参照していただくことにし, ここで終わることにする．

参考文献

[1] 杉原正顯, 室田一雄, 数値計算法の数理, 岩波書店 (1994).

[2] 高橋 秀裕, ニュートン—流率法の変容, 東大出版会, (2003).

[3] 山本哲朗, Newton 法とその周辺, 数学, 第 37 巻, (1985), 1-15.

[4] A.A. Bennet, Newton's method in general analysis, Proc. Nat. Acad. Sci., vol. 2 (1916), 592–598.

[17] the secant method. secant(割線) とは曲線と 2 個以上の点で交わる直線のこと．

[5] N. Bićanić and K.H. Johnson, Who was '-Raphson'? Inter. J. Numer. Meth. Eng., vol. 14 (1979), 148–152.

[6] E. Blutel, Sur l'approximation de la méthod d'approximation de Newton à la résolution approach'ee des équations à plusieur inconnues, Comp. R. Acad. Sci., vol. 151 (1910), 1109–1112.

[7] F. Cajori, Fourier's improvement of the Newton-Raphson method of approximation anticipated by Mourraille, Bibliotheca Mathematica, vol. 11 (1910-11), 132–137.

[8] F. Cajori, Historical note on the Newton-Raphson method of approximation, Amer. Math. Month., vol. 18 (1911), 29–32.

[9] A.-L. Cauchy, Résumé des Leçons données a L'École Royale Polytechnique sur le Calcule Infinitésimal, (1823); Œuvres Compléte d'Augustin Cauchy, II Série Tome IV, (1899), 573–609.

[10] G. Dahlquist nd A. Björck, Numerical Methods, Prentice Hall (1974).

[11] ロバート・L. ドゥヴェイニー, カオス力学系入門 第 2 版, 国府 寛司 他訳（原著は英語）, 共立出版.

[12] L. Euler, Elements of Algebra, (translated by J. Hewitt), Longman, Orme and Co. (1840). Springer (1972). (原著はドイツ語. Opera Omnia の第 1 巻に相当する.)

[13] G. Faber, Über die Newtonsche Näherungsformel, J. Reine Angew. Math., vol. 138 (1910), 1–21.

[14] H.B. Fine, On Newton's method of approximation, Proc. Nat. Acad. Sci., vol. 2 (1916), 546–552.

[15] J.B.J. Fourier, Question d'analyse algébrique, Bull. Soc. Philomathique, (1818), 61–67; Œuvres de Fourier, vol. II 243–253.

[16] J.B.J. Fourier, Analyse des équations déterminées, Paris (1831); ドイツ語翻訳, Die Auflösung der bestimmeten Gleichungen, A. Loewy, Leipzig, (1912) (Ostwald's Klassiker der Exakten Wissenschaften)

[17] J.V. Grabiner, The Origin of Cauchy's Rigorous Calculus, MIT Press (1981), Dover (2005)

[18] E. Hairer and G. Wanner, Analysis by Its History, Springer, (1996).

[19] A.S. Householder, The Numerical Treatment of a Single Nonlinear Equation, McGraw-Hill (1970).

[20] E. Isaacson and H.B. Keller, Analysis of Numerical Mathods, John Wiley & Sons (1966).

[21] L.V. Kantorovich and G.P. Akilov, Functional Analysis, 2nd Ed., Pergamon Press (1982). 第 18 章

[22] S.G. Krantz and H.R. Parks, The Implicit Function Theorem, History, Theory, and Applications, Birkhäuser (2002).

[23] N. Kollerstrom, Newton's method of approximation, British J. Hist. Sci., vol. 25 (1992), 347–354.

[24] J.-L. Lagrange, Traité de la Résolution des Équations numériques de tous lrs degrés, 2nd ed., (1808), Œuvres de Lagrange, vol. 8 (1879). 初版は 1798 年.

[25] I. Newton, De Analysi per Aequationes Numero Terminorum Infinitas, これには英訳がある. D.T. Whiteside, The Mathematical Papers of Isaac Newton, vol. II, Cambridge Univ. Press, (1968), 206–247.

[26] A.M. Ostrowski, Solution of Equations in Euclean and Banach Spaces, Academic Press (1973).

[27] E. Schröder, Ueber unendlich viele Algorithmen zur Auflösung der Gleichungen, Math. Ann., vol. 2 (1870), 317–365.

[28] T.J. Ypma, Historical development of the Newton-Raphson method, SIAM Rev., vol. 37(1995), 531–551.

A 付録

年表

1669 年	ニュートンの De Analysi が書かれる
1685 年	ウォリスの教科書 Algabra にニュートンのアルゴリズム載る
1690 年	ラフソンの Analysis Aequationum $\cdots$ 出版される
1740 年	シンプソンのアルゴリズムが発表される
1768 年	ムレイユによる解析：収束しない例など
1822 年	コーシーの収束証明
1831 年	フーリエの本 Analyse des équations détérminée 出版される
1860 年	シュレーダーによる任意次数収束のアルゴリズム
1910 年	複素関数に対するファーベルによる収束証明:解の存在を仮定せず
1910 年	多変数の場合の収束証明 (ブリュテル):解の存在を仮定
1916 年	多変数の場合の収束証明 (ファイン):解の存在を仮定せず
1916 年	無限次元空間における収束証明 (ベネット):解の存在を仮定せず
1948 年	カントロビッチの定理

例 A.1 ニュートン数列が収束しない例

$f(x)=\frac{1}{3}x^3-\frac{3}{2}x^2+2x$ とする．この関数のグラフは図 4 のようになる．適当な $x_1\in(2,3)$ をとると $f'(x_2)=0$ となってニュートン法は破綻する．これを理解するためには，$f'(x)=(x-1)(x-2)$ であることに注意する．したがって x_1 が

$$2=x_1-\frac{f(x_1)}{f'(x_1)}$$

を満たせばよい．つまり x_1 が 3 次方程式 $f(x_1)=(x_1-2)f'(x_1)$ の根であればよい．

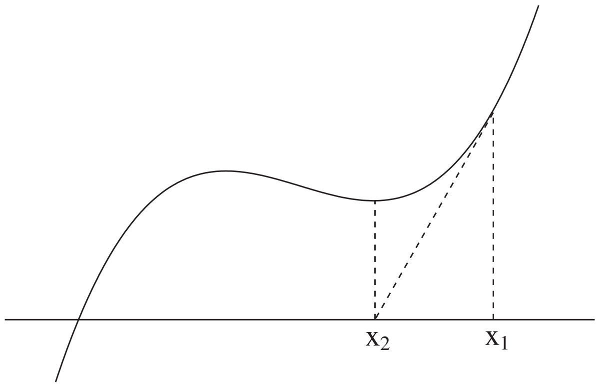

図 4: $x_1=2.802557\cdots$ のとき，$f'(x_2)=0$ となってニュートン法が破綻する

例 A.2 $f(x)=x/\sqrt{1+x^2}$ のとき，適当な ξ を取ると，$x_n=\xi$ ならば $x_{n+1}=-\xi$ となる（図 5）．関数 f は奇関数であるから $\xi,-\xi,\xi,-\xi,\cdots$ となって周期変動を繰り返すことになる．$f'(x)=(1+x^2)^{-3/2}$ であるから，このような ξ は

$$-\xi=\xi-\xi(1+\xi^2)$$

の根として特徴づけられるが，これが正根 $\xi=1$ を持つことは容易にわかる．さらに，$|x_1|>1$ ならばニュートン数列は発散し，$|x_1|<1$ ならばゼロに収束することも証明できる．同様のことは $f(x)=\arctan(x)$ でも成り立つし，様々な例を構成することはできる．

例 A.3 2 次収束する列は何もニュートン数列だけではなく，2 次収束する無限に多くの近似列を構成することは可能である．たとえばファーベル [13] の考えた数列

$$x_{n+1}=x_n-\frac{f(x_n)}{\alpha f(x_n)+f'(x_n)}$$

も 2 次収束である．ここで，α は定数である．これは直接証明することも容易ではあるが，この数列が関数 $e^{\alpha x}f(x)$ に対するニュートン数列であることと，$e^{\alpha x}f(x)=0\Longleftrightarrow f(x)=0$ であることからも証明できる．

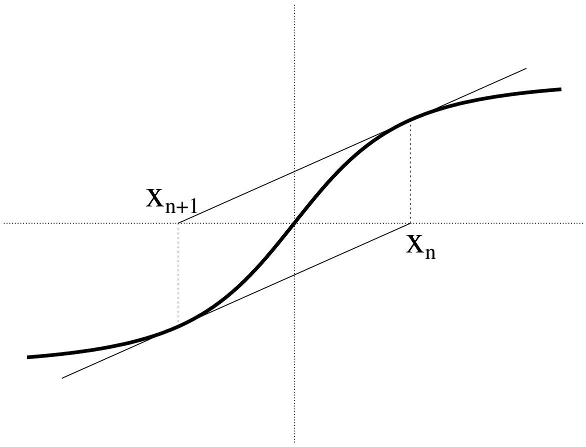

図 5: ニュートン点列が周期的になる場合

例 A.4 $x^2 - 2 = 0$ の正根を求めるために, $x_0 = 2.2, x_1 = 2.0$ とし,

$$x_{n+1} = x_n - \frac{f(x_n)(x_n - x_{n-1})}{f(x_n) - f(x_{n-1})} \qquad \text{(割線法)}$$

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} \qquad \text{(ニュートン法)}$$

で計算し, $x_n - \sqrt{2}$ がどれくらい小さくなって行くか, 調べてみるとつぎのようになる.

0.585786437626905	0.585786437626905
0.109595961436429	0.085786437626905
0.018218870059337	0.002453104293571
0.000675423260340	0.000002123901415
0.000004321763092	0.000000000001595
0.000000001031781	0.000000000000000

割線法（左側）はニュートン法（右側）ほど速くは収束しない.

例 A.5 根 ξ において $f'(\xi) = 0$ ならばニュートン法は 2 次収束しない. しかし, もしも $f(x) = (x - \xi)^m h(x)$ と書くことができて $h(\xi) \neq 0$ ならば

$$x_{n+1} = x_n - m\frac{f(x_n)}{f'(x_n)}$$

は 2 次収束する (文献 [13]).

本稿冒頭の問題の解答. $f(x) = a - 1/x$ とおく. $f'(x) = 1/x^2$ であるから

$$x_{n+1} = x_n - \frac{a - 1/x_n}{x_n^{-2}} = 2x_n - ax_n^2.$$

これは加減乗算だけで実行することができる.