Broyden 法

春日 悠

2016 年 4 月 17 日

目 次

1 ニュートン法

1.1 ニュートン法

次の方程式を満たす x を数値的に求めることを考える．

$$f(x) = 0 \tag{1}$$

繰り返し計算により x を求めるとしよう．k 回目の計算により求まる x を x_k と表す．$k+1$ 回目の x である x_{k+1} は，次式で求められるものとする．

$$x_{k+1} = x_k + \Delta x_k \tag{2}$$

関数 $f(x)$ を x_{k+1} についてテーラー展開すると，次のようになる．

$$f(x_{k+1}) = f(x_k + \Delta x_k) = f(x_k) + f'(x_k)\Delta x_k + O(\Delta x_k^2) \tag{3}$$

$k+1$ 回目で x が方程式を満たすとすると

$$f(x_k) + f'(x_k)\Delta x_k + O(\Delta x_k^2) = 0 \tag{4}$$

左辺第 3 項を無視して

$$f(x_k) + f'(x_k)\Delta x_k \approx 0 \tag{5}$$

これより次式を得る．

$$\Delta x_k = -\frac{f(x_k)}{f'(x_k)} \tag{6}$$

以上より，x の更新式として次式が得られる．

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)} \tag{7}$$

上式で方程式を解く方法をニュートン法という．

1.2　多変数関数の方程式系の場合

次の方程式系をニュートン法で解くことを考える．

$$\begin{aligned} f_1(x_1, x_2) &= 0 \\ f_2(x_1, x_2) &= 0 \end{aligned} \tag{8}$$

k 回目の x_1，x_2 をそれぞれ x_1^k，x_2^k と表す．各変数の更新式は次のように表される．

$$\begin{aligned} x_1^{k+1} &= x_1^k + \Delta x_1^k \\ x_2^{k+1} &= x_2^k + \Delta x_2^k \end{aligned} \tag{9}$$

$k+1$ 回目の変数で各関数をテーラー展開すると，次のようになる．

$$\begin{aligned} f_1(x_1^{k+1}, x_2^{k+1}) &= f_1(x_1^k, x_2^k) + \frac{\partial}{\partial x_1} f_1(x_1^k, x_2^k)\Delta x_1^k + \frac{\partial}{\partial x_2} f_1(x_1^k, x_2^k)\Delta x_2^k \cdots \\ f_2(x_1^{k+1}, x_2^{k+1}) &= f_2(x_1^k, x_2^k) + \frac{\partial}{\partial x_1} f_2(x_1^k, x_2^k)\Delta x_1 + \frac{\partial}{\partial x_2} f_2(x_1^k, x_2^k)\Delta x_2^k \cdots \end{aligned} \tag{10}$$

右辺第 4 項以降を無視するとして，上式をベクトルと行列の形で表すと

$$\begin{pmatrix} f_1(x_1^{k+1}, x_2^{k+1}) \\ f_2(x_1^{k+1}, x_2^{k+1}) \end{pmatrix} = \begin{pmatrix} f_1(x_1^k, x_2^k) \\ f_2(x_1^k, x_2^k) \end{pmatrix} + \begin{bmatrix} \frac{\partial}{\partial x_1} f_1(x_1^k, x_2^k) & \frac{\partial}{\partial x_2} f_1(x_1^k, x_2^k) \\ \frac{\partial}{\partial x_1} f_2(x_1^k, x_2^k) & \frac{\partial}{\partial x_2} f_2(x_1^k, x_2^k) \end{bmatrix} \begin{pmatrix} \Delta x_1^k \\ \Delta x_2^k \end{pmatrix} \tag{11}$$

ベクトル・行列表記では次のように表せる．

$$\boldsymbol{f}(\boldsymbol{x}_{k+1}) = \boldsymbol{f}(\boldsymbol{x}_k) + J(\boldsymbol{x}_k)\Delta \boldsymbol{x}_k \tag{12}$$

ここで $J(\boldsymbol{x}_k)$ をヤコビ行列という．もとの方程式系は次式のように表される．

$$\boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{0} \tag{13}$$

これを $\boldsymbol{x}_{k+1}$ が満たすとすると，次式が得られる．

$$J(\boldsymbol{x}_k)\Delta \boldsymbol{x}_k = -\boldsymbol{f}(\boldsymbol{x}_k) \tag{14}$$

上式を $\Delta \boldsymbol{x}_k$ について解くことにより，$\boldsymbol{x}$ を更新することができる．3 変数以上の方程式系でも同様である．

2　Broyden 法

2.1　セカント法

ニュートン法の問題点は，関数の微分が必要なことである．そこで，関数の微分を近似することを考える．関数の微分に近似を用いるニュートン法を準ニュートン法 (quasi-Newton method) という．

次の方程式の解を求めることを考える．

$$f(x) = 0 \tag{15}$$

ニュートン法では次式で x を更新する．

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)} \tag{16}$$

ここで，$f'(x_k)$ を次式で近似する．

$$f'(x_k) \approx \frac{f(x_k) - f(x_{k-1})}{x_k - x_{k-1}} \tag{17}$$

これにより，x の更新式は次式のようになる．

$$x_{k+1} = x_k - \frac{x_k - x_{k-1}}{f(x_k) - f(x_{k-1})} f(x_k) \tag{18}$$

この方法をセカント法 (secant method) という．

セカント法は関数の微分が不要であるが，初期値を 2 つ必要とする．

2.2　Broyden 法

次の方程式系をセカント法で解くことを考える．

$$\boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{0} \tag{19}$$

変数 $\boldsymbol{x}$ は次式で更新される．

$$\boldsymbol{x}_{k+1} = \boldsymbol{x}_k + \Delta\boldsymbol{x}_k \tag{20}$$

ニュートン法の場合，$\Delta\boldsymbol{x}_k$ は次式で得られる．

$$J(\boldsymbol{x}_k)\Delta\boldsymbol{x}_k = -\boldsymbol{f}(\boldsymbol{x}_k) \tag{21}$$

ヤコビ行列 $J(\boldsymbol{x}_k)$ の近似行列を B_k とすると

$$B_k\Delta\boldsymbol{x}_k = \boldsymbol{f}(\boldsymbol{x}_k) \tag{22}$$

ここで，右辺の負号は邪魔なので B_k の中に入れてしまっている．次の繰り返し計算におけるヤコビ行列 $J(\boldsymbol{x}_{k+1})$ に対する近似行列 B_{k+1} をセカント法同様に近似することを考えると，次式のように表される．

$$B_{k+1}\Delta\boldsymbol{x}_k = \Delta\boldsymbol{f}_k \tag{23}$$

ここで，$\Delta \boldsymbol{f}_k = \boldsymbol{f}(\boldsymbol{x}_k) - \boldsymbol{f}(\boldsymbol{x}_{k-1})$ である．上式では B_{k+1} は一意に決まらないので，条件を適当に与えてやる必要がある．

上式の両辺から $B_k \Delta \boldsymbol{x}_k$ を引くと

$$(B_{k+1} - B_k)\Delta \boldsymbol{x}_k = \Delta \boldsymbol{f}_k - B_k \Delta \boldsymbol{x}_k \tag{24}$$

右辺を $A\boldsymbol{x}_k$ の形にしてやれば，$B_{k+1} = B_k + A$ となって B_{k+1} を B_k の修正によって求めることができる．単純に考えると，両辺に $\Delta \boldsymbol{x}_k^T \Delta \boldsymbol{x}_k$ をかけてやればよい．

$$\begin{aligned} (B_{k+1} - B_k)\Delta \boldsymbol{x}_k \Delta \boldsymbol{x}_k^T \Delta \boldsymbol{x}_k &= (\Delta \boldsymbol{f}_k - B_k \Delta \boldsymbol{x}_k)\Delta \boldsymbol{x}_k^T \Delta \boldsymbol{x}_k \\ (B_{k+1} - B_k)\Delta \boldsymbol{x}_k &= \frac{(\Delta \boldsymbol{f}_k - B_k \Delta \boldsymbol{x}_k)\Delta \boldsymbol{x}_k^T \Delta \boldsymbol{x}_k}{\Delta \boldsymbol{x}_k^T \Delta \boldsymbol{x}_k} \end{aligned} \tag{25}$$

以上から，次式が得られる．

$$B_{k+1} = B_k + \frac{(\Delta \boldsymbol{f}_k - B_k \Delta \boldsymbol{x}_k)\Delta \boldsymbol{x}_k^T}{\Delta \boldsymbol{x}_k^T \Delta \boldsymbol{x}_k} \tag{26}$$

ヤコビ行列の近似式を上式で計算する準ニュートン法を Broyden 法という [2]．上の式は，$\Delta \boldsymbol{x}$ に直交するベクトル $\boldsymbol{q}$ を両辺にかけると $B_{k+1}\boldsymbol{q} = B_k \boldsymbol{q}$ となるような条件を課していることになっている．

欲しいのは $\Delta \boldsymbol{x}_k$ なので，必要なのはヤコビ行列よりはその逆行列である．そこで，B_{k+1} の逆行列を求めてみる．

ここで，次のような形の行列を考える．

$$B = A + \boldsymbol{u}\boldsymbol{v}^T \tag{27}$$

次の式を満たすベクトル $\boldsymbol{x}$ があるとする．

$$A\boldsymbol{x} = \boldsymbol{u} \tag{28}$$

このとき

$$\begin{aligned} B\boldsymbol{x} &= (A + \boldsymbol{u}\boldsymbol{v}^T)\boldsymbol{x} \\ &= A\boldsymbol{x} + \boldsymbol{u}\boldsymbol{v}^T\boldsymbol{x} \\ &= \boldsymbol{u} + (\boldsymbol{v}^T\boldsymbol{x})\boldsymbol{u} \\ &= (1 + \boldsymbol{v}^T\boldsymbol{x})\boldsymbol{u} \end{aligned} \tag{29}$$

これより，次式が得られる．

$$B^{-1}\boldsymbol{u} = \frac{1}{1 + \boldsymbol{v}^T\boldsymbol{x}} A^{-1}\boldsymbol{u} \tag{30}$$

上式の両辺から $A^{-1}\boldsymbol{u}$ を引くと

$$\begin{aligned}(B^{-1}-A^{-1})\boldsymbol{u} &= \left(\frac{1}{1+\boldsymbol{v}^T\boldsymbol{x}}-1\right)A^{-1}\boldsymbol{u} \\ &= -\frac{\boldsymbol{v}^T\boldsymbol{x}}{1+\boldsymbol{v}^T\boldsymbol{x}}A^{-1}\boldsymbol{u} \\ &= -\frac{A^{-1}\boldsymbol{u}\boldsymbol{v}^T\boldsymbol{x}}{1+\boldsymbol{v}^T\boldsymbol{x}} \\ &= -\frac{A^{-1}\boldsymbol{u}\boldsymbol{v}^T A^{-1}\boldsymbol{u}}{1+\boldsymbol{v}^T A^{-1}\boldsymbol{u}}\end{aligned} \tag{31}$$

これより次式を得る．

$$(A+\boldsymbol{u}\boldsymbol{v}^T)^{-1} = A^{-1} - \frac{A^{-1}\boldsymbol{u}\boldsymbol{v}^T A^{-1}}{1+\boldsymbol{v}^T A^{-1}\boldsymbol{u}} \tag{32}$$

これを Sherman-Morrison の公式という．

上の公式を用いると，B_{k+1} の逆行列は次式で表される．

$$B_{k+1}^{-1} = B_k^{-1} + \frac{(\Delta\boldsymbol{x}_k - B_k^{-1}\Delta\boldsymbol{f}_k)\Delta\boldsymbol{x}_k^T B_k^{-1}}{\Delta\boldsymbol{x}_k^T B_k^{-1}\Delta\boldsymbol{f}_k} \tag{33}$$

もう少し直接的に B_{k+1}^{-1} を求める方法もある．B_{k+1}^{-1} について，次式が成り立つ．

$$B_{k+1}^{-1}\Delta\boldsymbol{f}_k = \Delta\boldsymbol{x}_k \tag{34}$$

これについて，B_{k+1} の式を求めたのと同様の手順で B_{k+1}^{-1} の式を求めると，次式が得られる．

$$B_{k+1}^{-1} = B_k^{-1} + \frac{(\Delta\boldsymbol{x}_k - B_k^{-1}\Delta\boldsymbol{f}_k)\Delta\boldsymbol{f}_k}{\Delta\boldsymbol{f}_k^T\Delta\boldsymbol{f}_k} \tag{35}$$

ただし，Broyden によると，この式はあまりよろしくないらしい [2].

実際の問題に Broyden 法を適用する場合，B_k，B_k^{-1} のどちらを用いるにしても，それらの初期値を与える必要がある．初期値は単位行列にするか，何かしらの方法で求めたヤコビ行列の近似値を与える．また，最初だけヤコビ行列をきちんと計算するという方法もある．この場合は微分を避けるというより，それ自体手間のかかるヤコビ行列の計算を避けるために Broyden 法を用いる．

参考文献

[1] 戸川隼人：UNIX ワークステーションによる科学技術計算ハンドブック，サイエンス社 (1998).

[2] C.G.Broyden：A Class of Methods for Solving Nonlinear Simultaneous Equations，Mathematics of Computation，19 (92)，577-593 (1965).