中心化ニュートン法

田辺 國士

1. 序

次のような2次計画問題（Dが0行列のときは線形計画問題）を考えよう．

（1） 条件 $Ax \leqq b, x \geqq 0$，の下で $c^t x - x^t Dx/2$ を最大化せよ．

ただしAは$m \times n$行列，Dは$n \times n$正定値行列とする．この双対問題は

（2） 条件 $A^t y + Dx \geqq c, y \geqq 0$，の下で $b^t y + x^t Dx/2$ を最小化せよ．

となる．これらの問題は非線形等式・不等式系

（3） $r(z) \equiv \begin{pmatrix} [x](A^t y + Dx - c) \\ [y](b - Ax) \end{pmatrix} = \begin{pmatrix} [A^t y + Dx - c]x \\ [b - Ax]y \end{pmatrix} = 0,$

（4） $x \geqq 0, y \geqq 0,\ b - Ax \geqq 0,\ A^t y + Dx - c \geqq 0,$

を満たすベクトル $x \in R^n$，$y \in R^m$ を求める問題と同値である．ただしR^kはk次元ユークリッド空間を表わし，$z \equiv (x, y) \in R^{m+n}$となる．また$v$を$k$次元ベクトル$v$とするとき，$v_i$は$v$の第$i$要素を表わし，$|v|$は$|v_i|$を第$i$要素とする非負ベクトルを表わし，$[v]$は$v_i$を第$i$対角要素とする対角行列を表わし，$\Pi(v)$は$v$のすべての要素の積を表わすとする．したがってこれは，条件（4）を満たしながら非線形方程式（3）を解けば問題の解が与えられることを意味する．この立場から解法アルゴリズムを設計することができる．

非線形方程式の数値解法としてはニュートン法が広く用いられている．解に十分近い初期値から出発するときこの方法は速い収束性をもっている．しかし，解から遠く離れた初期値から出発させると，振動したり発散することがある．振動や発散をおさえて安定して解に導くには，解くべき方程式の残差のノルムが単調減少するようにステップ幅を制御する制動型ニュートン法や常微分方程式の解曲線を追跡する連続型ニュートン法が知られている．しかしこれらの方法においてはステップ幅が小さくなりがちで，一般に計算効率がよくない．方程式（3）のように比較的低次の連立代数方程式におけるニュートン法の修正ベクトルは，解から遠く離れた点においても十分意味あるベクトルであることが多い．したがって，ステップ幅を縮小しないでできるかぎり大きな幅（すなわち1に近いステップ幅）で反復解法を実行できるならば，そのほうが望ましい．このような動機から設計された解法が以下に述べる中心化ニュートン法（Centered Newton Method）である．

たなべ くにお 統計数理研究所
〒106 港区南麻布4－6－7

2. 連立非線形方程式を解く中心化ニュートン法

連立非線形方程式

（5） $r_1(z) = 0,$
$r_2(z) = 0,$
$\cdots\cdots$
$r_k(z) = 0,$

の解法を考える．ただし，$z \in R^k$とする．r_iを第i要素とする列ベクトルを$r(z)$とし，そのヤコビ行列を$J(z)$とする．このときニュートン法は

（6） $z^+ = z + d_n,$
$J(z) d_n = -r(z),$

と書ける．ニュートン法の修正ベクトルd_nを補助するベクトルd_cを

（7） $J(z) d_c = -r(z) + \left(\sum_{i=1}^{k} |r_i(z)| / \|s(z)\|_2^2\right) s(z)$

で定義する．ただし，$s(z)$は

（8） $s_i(z) \equiv \begin{cases} 1, (r_i(z) > 0) \\ 0, (r_i(z) = 0) \\ -1, (r_i(z) < 0) \end{cases}$

を第i要素とする列ベクトルとし，$\|*\|_2$はユークリッド・ノルムを表わすとする．d_cは中心曲線，

（9） $C \equiv \{z : |r_1(z)| = |r_2(z)| = \ldots = |r_k(z)|\}$

に向かうベクトルとなる．以下d_n, d_cをそれぞれニュートン・ベクトル，中心化ベクトルと呼ぶ．関数$\rho(z)$を

（10） $\rho(z) \equiv \left(\sum_{i=1}^{k} |r_i(z)| / k\right) \Big/ \left(\sqrt[k]{\prod_{i=1}^{k} |r_i(z)|}\right),$

と定義すると，分母が0でない範囲で$\rho(z) \geqq 1$，かつ$\rho(z)$は中心曲線上で最小値1をとる．したがって，$\log \rho(z)$は点zの中心曲線からの離れ具合いの尺度となる．このとき，式に現われる各要素が存在する範囲で次式が

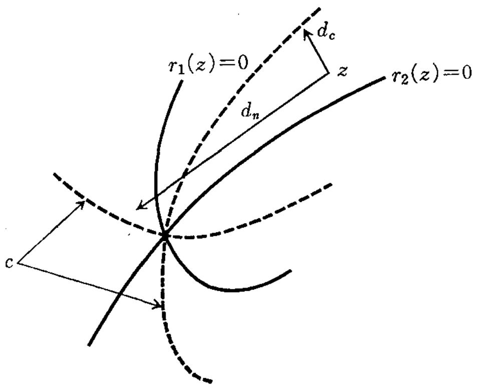

図 1 ニュートン・ベクトル，中心曲線，中心化ベクトル

成り立つ．

(11) $(\nabla\rho(z), d_n(z))=0$,

$(\nabla\|r(z)\|_1, d_n(z))\leqq 0.$

$(\nabla\rho(z), d_c(z))\leqq 0,$

$(\nabla\|r(z)\|_1, d_c(z))=0,$

ただし，∇はグラディエント・ベクトル，$(*,*)$ は内積（要素の積和）を表わす．すなわち，ニュートン・ベクトル d_n は，案内錐

(12) $\mathrm{Cone}(r,\lambda)\equiv\{z\in R^k:(\sum_{i=1}^{k}|r_i(z)|/k)$

$\leqq\lambda\sqrt[k]{\prod_{i=1}^{k}|r_i(z)|}\}$

の境界に接し，$\|r(z)\|_1\equiv\sum_{i=1}^{k}|r_i(z)|$ を減少させる方向ベクトルであり，中心化ベクトルd_cは，集合

(13) $\{z\in R^k:\|r(z)\|_1=\nu\}$

に接し，$\rho(z)$を減少させる方向ベクトルである．ただし λは1より大きな実数，ν は正数とする．

これらのベクトル d_n と d_c を組み合せて用いると，制動型ニュートン法などに比べてより柔軟で大きなステップ幅を許容する反復解法が得られる．

中心化ニュートン法：

(14) $z^+=z+\alpha d_n+\beta d_c$

ただし，ステップ幅 $\alpha,\beta\geqq 0$ は，条件

(15) $0\leqq\alpha\leqq 1$,

$z+\alpha d_n(z)+\beta d_c(z)\in\mathrm{Cone}(r,\lambda)$,

$\|r(z+\alpha d_n(z)+\beta d_c(z))\|_1\leqq(1-\delta)\|r(z)\|_1$,

を満たす範囲でニュートン・ステップ幅αができるだけ1に近くなるように中心化ステップ幅βを定める．

ステップ幅は評価関数（罰金つきノルム）を

(16) $\mu(\omega,z)=(\rho(z))^{\omega}\|r(z)\|_1$,

と定義し，条件(15)の代わりに

(17) $0\leqq\alpha\leqq 1$,

$\mu(\omega, z+\alpha d_n(z)+\beta d_c(z))\leqq(1-\delta)\mu(\omega,z)$,

を用いてもよい．ただし，ωは正数とする．実際のアルゴリズムにおいては $\log\ \mu(\omega,z)$ を計算する．

ニュートン法と中心化ニュートン法（ただし，$\alpha=\beta$に固定）が作るベクトル場，$dz/dt=d_n$および$dz/dt=d_n+d_c$の様子を見るために，連立方程式

$$r(z)\equiv\begin{pmatrix}z_1-z_2^2\\ z_2-z_1^2\end{pmatrix}=0,$$

を例にとり図2，3に示した．

$J(z)$ が特異になる場合には，レヴェンバーグ・マーカート型の修正を加えて，(6)(7)で定められるd_nとd_cの代わりに

(18) $(J^t(z)J(z)+\delta I)d_n^*(z)=-J^t(z)r(z)$,

(19) $(J^t(z)J(z)+\delta I)d_c^*(z)=-J^t(z)P(z)r(z)$,

で定まるd_n^*とd_c^*を用いる．ただし，$P(z)$は射影行列

(20) $P(z)\equiv I-s(z)(s(z))^t/\|s(z)\|_2^2$

とし，δは小さな正数とする．これを修正中心化ニュートン法と呼ぶことにする．

一般の連立非線形方程式に中心化ニュートン法を適用する場合には，予め次のような局所的スケール変換

(21) $r_i(z):=\dfrac{r_i(z)}{\|\nabla r_i(z)\|_2}$, $(i=1,\cdots,k)$

を各方程式にほどこした方が良い．これにより中心曲線はスケール不変になる利点がある．ただし，こうするとヤコビ行列 $J(z)$ に元の関数の2次微係数が現われる．しかし，この2次微係数の部分を1次微係数の数値微分によって計算することができ，そうすることにより性能が向上する．

3. 線形・2次計画問題を解く中心化ニュートン法

問題(1)，(2)の解法を説明する前に，記号を導入しよう．2つのベクトルu,vに対して，$[u]/[v]$は対角行列$[u][v]^{-1}=[v]^{-1}[u]$を表わす．$\mathbf{1}$はすべての要素が1の列ベクトルを表わす．

以下，条件(4)で定義される許容集合の内点$z^0\equiv(x^0, y^0)$が与えられていると仮定する．すなわち，x^0, y^0は不等式

(22) $x>0, y>0, b-Ax>0, A^ty+Dx-c>0$,

を満たしているとする．

初期値z^0から出発して，この不等式を常に満足させながら前節で述べた中心化ニュートン法を連立非線形方程式(3)に適用することができる．条件(22)の下で

(23) $\gamma=\|r(z)\|_1/(m+n)$

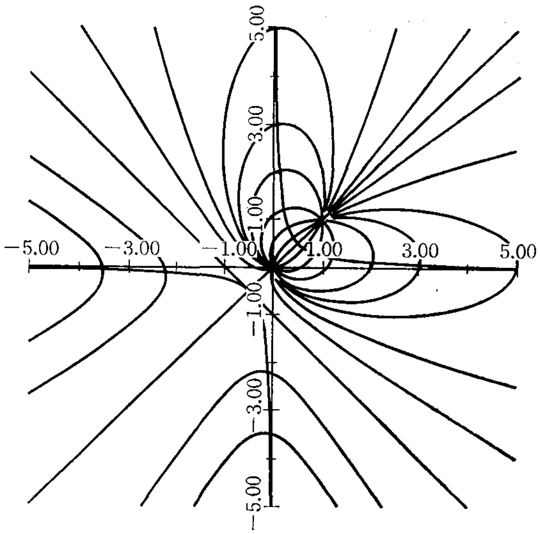

図 2 ニュートン法のベクトル場

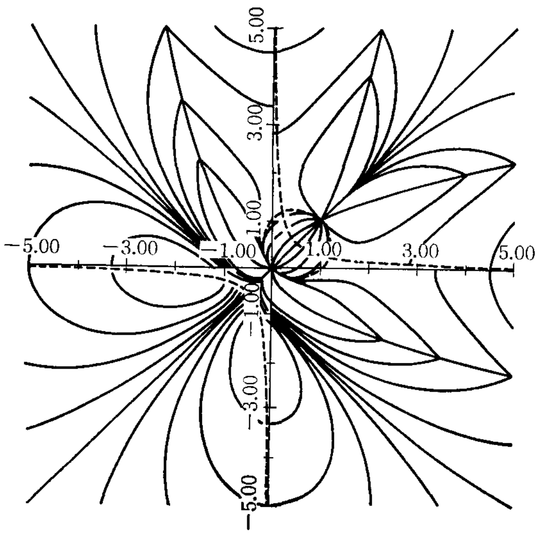

図 3 中心化ニュートン法のベクトル場

$$=(b^ty-c^tx+x^tDx)/(m+n)>0$$

が成り立ち，ニュートン・ベクトルと中心化ベクトルをそれぞれ$d_n\equiv(x_n,y_n)$, $d_c\equiv(x_c,y_c)$とおくと，これらは連立1次方程式

(24) $$\begin{bmatrix}[A^ty+Dx-c]/[x]+D & A^t\\ -A & [b-Ax]/[y]\end{bmatrix}\begin{pmatrix}x_n\\ y_n\end{pmatrix}=-\begin{pmatrix}A^ty+Dx-c\\ b-Ax\end{pmatrix},$$

(25) $$\begin{bmatrix}[A^ty+Dx-c]/[x]+D & A^t\\ -A & [b-Ax]/[y]\end{bmatrix}\begin{pmatrix}x_c\\ y_c\end{pmatrix}=-\begin{pmatrix}A^ty+Dx-c\\ b-Ax\end{pmatrix}+\gamma\begin{pmatrix}[x]^{-1}\mathbf{1}\\ [y]^{-1}\mathbf{1}\end{pmatrix},$$

の解である．許容集合に含まれる中心曲線は

(26) $$C^+=\{z\in R^{m+n}:[x](A^ty+Dx-c)=\gamma\mathbf{1},\ [y](b-Ax)=\gamma\mathbf{1},$$

ただし，$\gamma>0$ でかつ x と y は(22)を満足する$\}$,

で定義される．また許容集合に含まれる中心曲線は

(27) $$\mathrm{Cone}^+(r,\lambda)=\{z\in R^{m+n}:(b^ty-c^tx+x^tDx)/(m+n)\leqq\lambda\Pi(x)\Pi(y)\Pi(b-Ax)\Pi(A^ty+Dx-c)$$

ただし，x と y は(22)を満足$\}$,

で定義される．ただし，$\lambda\geqq1$ とする．罰金つきノルムは

(28) $$\mu^+(\omega,z)=(m+n)\{(b^ty-c^tx+x^tDx)/(m+n)\}^{1+\omega}/\{\Pi(x)\Pi(y)\Pi(b-Ax)\Pi(A^ty+Dx-c)\}^{\frac{\omega}{m+n}}$$

で定義される．

以下，$D=0$，すなわち線形計画問題の場合を考えよう．このとき，連立1次方程式(24)，(25)を解くと

(29) $$x_n=\{[A^ty-c]/[x]+A^t([y]/[b-Ax])A\}^{-1}c,$$

$$y_n=-\{[b-Ax]/[y]+A([x]/[A^ty-c])A\}^{-1}b,$$

$$x_c=\{[A^ty-c]/[x]+A^t([y]/[b-Ax])A\}^{-1}\{c+\gamma([x]^{-1}\mathbf{1}-A^t[b-Ax]^{-1}\mathbf{1})\},$$

$$y_c=-\{[b-Ax]/[y]+A([x]/[A^ty-c])A^t\}^{-1}\{b-\gamma([y]^{-1}\mathbf{1}+A[A^ty-c]^{-1}\mathbf{1})\}.$$

が得られる．この公式は実際計算にはなじまない．行列Aが疎な場合には，直接(24)や(25)の係数行列をスパース技法を用いて分解する方が得策である．

条件(22)の下では

(30) $c^tx_n>0$ かつ $b^ty_n<0$.

が成り立つ．また等式

(31) $$b^ty_n-c^tx_n=-(b^ty-c^tx),$$
$$b^ty_c-c^tx_c=0.$$

が成り立つ．したがって

(32) $$\|r(z+\alpha d_n+\beta d_c)\|_1=b^t(y+\alpha y_n+\beta y_c)-c^t(x+\alpha x_n+\beta x_c)=(1-\alpha)(b^ty-c^tx)=(1-\alpha)\|r(z)\|_1$$

が成り立つ．

4. 結　　び

中心化ニュートン法は主・双対問題が非対称の線形計画問題，非線形計画問題，相補性問題にも適用できる．[4,6] また，ニュートン・ベクトルや中心化ベクトルを計算する実用的アルゴリズムは，たとえば[5]を参照されたい．

参 考 文 献

[1] K. Tanabe, Continuous Newton-Raphson

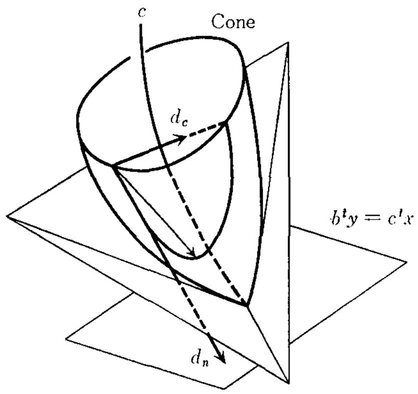

図 4 線形計画問題におけるニュートン・ベクトル，中心曲線，中心化ベクトル，案内錐

method for solving an underdetermined system of nonlinear equations, Journal of nonlinear analysis 3 (1979), 495-503.

[2] K. Tanabe, A geometric method in nonlinear programming, Journal of Optimization Theory and Applications 30 (1980), 181-210.

[3] K. Tanabe, Global analysis of continuous analogues of the Levenberg-Marguardt and the Newton-Raphson methods for solving nonlinear equations, Annals of the Institute of Statistical Mathematics 37 (1985), 189-203.

[4] K. Tanabe, Complementality-enforcing centered Newton method for Mathematical Programming : Global Method, The Institute of Statistical Mathematics Cooperative Research Report 5, "New Methods for Linear Programming", (1987).

[5] K. Tanabe, Algorithm for Computing Search Directions of the Interior Point Method for Linear Programming, The Institute of Statistical Mathematics Cooperative Research Report 5, "New Methods for Linear Programming 2", (1988)

[6] K. Tanabe, Centered Newton Method for Mathematical Programming, Proceedings of the 12th IFIP Conference on System Modelling and Optimization : Lecture Notes in Control and Information Sciences, Springer, (1988)