分解型準ニュートン更新公式の逐次2次計画法への応用†*

矢部 博†† 高橋 悟†† 八巻直一†††

制約条件付き最小化問題に対する数値解法の中で，最近特に逐次2次計画法が注目されている．この方法は有効ではあるけれども，大域的収束性を保証するためには2次計画部分問題を狭義凸に保つ必要がある．本稿ではセカント条件のある種の緩和により，部分問題がつねに狭義凸2次計画問題になるようなアルゴリズムを提案する．アルゴリズムの構築過程で分解型準ニュートン更新公式が統一的に議論されるが，この結果は例えば Powell の修正 BFGS 公式を含むことが示される．さらに分解型公式が2次計画部分問題を解くための Goldfarb and Idnani 法ともうまく適合することを数値実験を行って確かめる．

1. はじめに

本稿では次の制約条件付き非線形最小化問題に対する数値解法を提案する：

(問題 NLP) 不等号制約条件 $g(x)\leqq 0$, $g(x)=(g_1(x), g_2(x), \cdots, g_m(x))^T$ の下で目的関数 $f(x)$ を最小にせよ．ただし x は n 次元実ベクトル，$f, g_i\ (i=1, \cdots, m)$ はそれぞれ微分可能な実数値関数である．

問題 NLP を解くことは次の Kuhn-Tucker 条件

$$\nabla_x L(x, \lambda)=0,\quad g(x)\leqq 0, \tag{1.1a}$$

$$\lambda^T g(x)=0,\quad \lambda\geqq 0,\quad \lambda\in R^m \tag{1.1b}$$

を満たす点 (x, λ) を見つけることにほかならない．ただし $L(x, \lambda)$ は問題 NLP のラグランジュ関数で

$$L(x, \lambda)=f(x)+\lambda^T g(x) \tag{1.2}$$

で与えられる．λ は不等号制約条件に対するラグランジュ乗数ベクトルである．なお等号制約条件 $h(x)=0$ は二つの不等号制約条件 $h(x)\leqq 0$, $-h(x)\leqq 0$ として扱えるので，不等号制約条件に限定しても一般性は失われない．

問題 NLP に対する数値解法として，ペナルティ関数法，拡張ラグランジュ関数法，射影法等いくつかの有力な方法が知られているが，近年特に逐次2次計画法 (SQP : Successive Quadratic Programming method) が注目を浴びている (Han[6),7)], Powell[10),11)], Yamashita[16)])．

逐次2次計画法の特徴は，毎回方向ベクトルが凸2次計画問題を解くことによって得られることにある．Han らはこの方法の大域的収束性や局所的超1次収束性を示している．逐次2次計画法は有望な数値解法ではあるけれども，同時にいくつかの課題も抱えている．代表的な課題として，(1)大域的収束性を実現するために2次計画部分問題の目的関数が狭義凸関数であることが要請されること，(2)問題によっては超1次収束性が得られない場合があること，(3)問題 NLP が実行可能解をもつにもかかわらず2次計画部分問題に実行可能解が存在しなくなる場合が有り得ること，などがあげられる．

課題(1)に関しては Powell[10),11)] の研究がよく知られている．課題(2)は Maratos effect と呼ばれる現象で，これについては Mayne and Polak[8)] らの研究がある．また課題(3)は非線形制約関数を線形化することに原因があり，これについては Powell[10)] らの研究がある．

本稿では課題(1)を解決するために，狭義凸性を保つアルゴリズムを構築することを考える．これは Powell[10)] の研究を特別な場合として含む．さらに我我の提案したアルゴリズムと狭義凸2次計画部分問題に対する Goldfarb and Idnani 法[5),12)] を併用すれば，計算の効率が高まることを示す．なお以下では必要に応じて添字 k は省略し，$k+1$ を "+" で表す．またノルム $\|\cdot\|$ はユークリッドノルムである．

2. 逐次2次計画法のアルゴリズムおよび収束性

この章では逐次2次計画法のアルゴリズムと2次計画部分問題に対する Goldfarb-Idnani 法を紹介する．

2.1 逐次2次計画法

逐次2次計画法のアルゴリズムは以下のとおりである．

† Application of Factorized Quasi-Newton Updates to Successive QP Method for Constrained Optimization by HIROSHI YABE, SATORU TAKAHASHI (Department of Applied Mathematics, Faculty of Science, Science University of Tokyo) and NAOKAZU YAMAKI (Research Institute of Systems Planning Ltd.).
†† 東京理科大学理学部応用数学科
††† (株)システム計画研究所
* 本研究の一部は，文部省科学研究費補助金・奨励研究 (A) 61740133 によって補助された．

(アルゴリズム SQP)

初期点，x_1，$n\times n$ 正定値対称行列 B_1 および三つの定数 $r>0$，$\omega\in(0,\ 0.5)$，$\tau\in(0,\ 1)$ が与えられたとき，以下の手順を $k=1,\ 2\cdots$に対して繰り返す；

ステップ1．x_k と B_k に対する次の2次計画問題

(2次計画部分問題)

線形制約条件 $g(x_k)+\nabla g(x_k)d\leqq 0$ の下で，2次関数 $(1/2)\ d^{\mathrm{T}}B_kd+\nabla f(x_k)^{\mathrm{T}}d$ を最小にせよ

を解いて，方向ベクトル d_k とラグランジュ乗数ベクトル λ_{k+1} を求める．$\nabla g(x)$ は g のヤコビ行列である．

ステップ2．もし x_k，λ_{k+1} が問題 NLP の Kuhn-Tucker 条件(1.1)を満足していれば停止し，さもなければステップ3へ行く．

ステップ3．以下に示す直線探索を行って刻み幅 α_k を決定する；

ステップ 3.1 $\beta_{k,1}=1$ および $j=1$ とおく．

ステップ 3.2 もし

$$\theta_r(x_k+\beta_{k,j}d_k)\leqq\theta_r(x_k)-\omega\beta_{k,j}d_k^{\mathrm{T}}B_kd_k \tag{2.1}$$

が成り立てば $\alpha_k=\beta_{k,j}$ とおきステップ4へ行く．さもなければステップ3.3へ行く．ただし $\theta_r(x)$ は直線探索に用いるコスト関数で

$$\theta_r(x)=f(x)+r\max(0,g_1(x),\cdots,g_m(x)) \tag{2.2}$$

で与えられる．

ステップ 3.3 $\beta_{k,j+1}=\tau\beta_{k,j}$, $j=j+1$ とおいてステップ3.2へ行く．

ステップ4．$x_{k+1}=x_k+\alpha_kd_k$ とおく．

ステップ5．行列 B_k を準ニュートン公式に従って更新し B_{k+1} を作成する．$k=k+1$ とおいてステップ1へ行く．

アルゴリズム SQP において，(2.2)で与えられた $\theta_r(x)$ は正確なペナルティ関数[16]であり，直線探索で使用された基準(2.1)は微分不可関数への Armijo 規則の拡張[16]である．2次計画部分問題の B_k は $\nabla_{xx}L(x_k,\ \lambda_{k+1})$ の近似行列である．このとき Han[7]は，Bが正定値行列ならばステップ3の直線探索が有限回の手順で終了すること，およびこのアルゴリズムが大域的収束することを示している．したがって行列 B_k の正定値性の保存が重要になる．しかしながら普通の準ニュートン更新公式を単に逐次2次計画法に適用しただけでは，必ずしも正定値性は保存されない．このことについては第3.1節で触れる．

また等号制約条件付き最小化問題では，2次計画部分問題を解くことは非線形方程式 $\nabla_xL(x,\ \lambda)=0$, $h(x)=0$ をニュートン法に基づいた方法で解くことに相当している．後者に関しては Tanabe[15]の研究がある．

2.2 Goldfarb and Idnani 法

逐次2次計画法の効率を高める重要な要素の一つは，2次計画部分問題に対する数値解法の選択である．近年それぞれに特徴のある数値解法が研究されているが，我々は Goldfarb and Idnani[5]によるアルゴリズムを使用した．この方法は狭義凸2次目的関数の無制約最小点から出発して，満足していない制約条件を順々に考慮しながら有限回の手順で最適解を得るアルゴリズムであり，逐次2次計画法に大変よく馴染む方法であると思われる[12]．2次計画部分問題に対する具体的なアルゴリズムは以下のとおりである．ただし繁雑になるので反復回数の添字は省略する．

(アルゴリズム GI)

ステップ0．効いている制約条件の個数を q，条件式の番号の集合をA，対応する双対変数からなるq次元ベクトルを u とすれば，初回は $q=0$, $A=\phi$ となる．行列 B を $B=FF^{\mathrm{T}}$ と Cholesky 分解して無制約最小点 $d=-B^{-1}\nabla f(x)$ を求め，かつ $J_2=F^{-\mathrm{T}}$ とおく．

ステップ1．満足されない制約条件を一つ選んでその番号を p とおく．もしそのような p が存在しなければ，最適解が得られたので終了する．もし $q>0$ ならば $\hat{u}=[u^{\mathrm{T}}\ \ 0]^{\mathrm{T}}$ と拡張する．

ステップ2．$h_2=-J_2^{\mathrm{T}}\nabla g_p(x)$, $z=J_2h_2$ とおく．もし $q>0$ ならば $h_1=-J_1^{\mathrm{T}}\nabla g_p(x)$, $r=U^{-1}h_1$ を求める．ここでzは方向ベクトル，$-r$ は効いている制約条件に関する双対変数空間の方向ベクトルである．

ステップ3．$q=0$ もしくは $r\leqq 0$ ならば $t_1=\infty$ とおき，さもなければ $t_1=\min\{e_j^{\mathrm{T}}\hat{u}/e_j^{\mathrm{T}}r\,|\,e_j^{\mathrm{T}}r>0,\ j=1,\cdots,q\}=e_N^{\mathrm{T}}\hat{u}/e_N^{\mathrm{T}}r$ とおく．ただし e_j は単位行列の第j列ベクトルである．さらに $\|z\|=0$ ならば $t_2=\infty$ とおき，さもなければ $t_2=(g_p(x)+\nabla g_p(x)^{\mathrm{T}}d)/z^{\mathrm{T}}\nabla g_p(x)$ とおく．ここで t_1 は双対変数空間での r 方向の刻み幅，t_2 は z 方向の刻み幅である．

ステップ4．$t=\min(t_1,t_2)$ としたとき，$t=\infty$ ならば2次計画部分問題には実行可能解が存在しないので終了する．$\hat{u}=\hat{u}+t[-r^{\mathrm{T}}\ 1]^{\mathrm{T}}$ ($q=0$ のとき $\hat{u}=t$) とおき，$t_2=\infty$ のとき手順(a)を実行し，さもなければ手順(b)を実行する；

(a) $A=A\backslash\{A$の第N番目の要素$\}$ とおき，$\hat{u}$ の

第 N 要素を取り除いたベクトルを改めて $\hat{u}$ とおく．行列 U の第 N 列目を取り除き（不要な制約条件の除去），右下の $(q-N)\times(q-N)$ 小行列を Givens 法を用いて上三角行列にする．すなわち

$$U=\begin{bmatrix} U_1 & \cdot & S \\ \cdots & \cdot & \cdots \\ 0 & \cdot & C \end{bmatrix}$$

↑——第N列目

のとき $U_2=QC$ を計算して

$$U=\begin{bmatrix} U_1 & S \\ 0 & U_2 \end{bmatrix}$$

を求める．ただし

$$U_1\in R^{(N-1)\times(N-1)},\quad U_2\in R^{(q-N)\times(q-N)}$$

は上三角行列，Q は Givens 法によって得られる直交行列である．このとき

$$\left[J_1\begin{bmatrix} I & 0 \\ 0 & Q^{\mathrm{T}} \end{bmatrix} \middle| J_2\right]$$

の最初の $n\times(q-1)$ 小行列を J_1，残りの部分を J_2 とする．$q=q-1$ とおいてステップ2へ行く．

（b） $d=d+tz$ とおく．もし $t=t_1$ ならば手順（a）へ行く．$t=t_2$ ならば

$$A=AU\{p\},\quad U=\begin{bmatrix} U & h_1 \\ 0 & \|h_2\| \end{bmatrix}$$

（$q=0$ のとき $U=[\|h_2\|]$），$u=\hat{u}$ とおく．また $[J_1 : J_2Q^{\mathrm{T}}]$（$q=0$ のとき $[J_2 \ \ Q^{\mathrm{T}}]$）の最初の $n\times(q+1)$ 小行列を J_1，残りの部分を J_2 とする（効いている制約条件式の追加）．ただし Q は $Qh_2=\pm\|h_2\|e_1$ を満たすように Givens 法で求めた直交行列である．$q=q+1$ とおいてステップ1へ行く．

3. 分解型準ニュートン更新公式

3.1 正定値対称性を保存する準ニュートン更新公式

逐次2次計画法では B は $\nabla_{xx}L(x,\lambda)$ の近似行列であるから，新しい行列 B_+ が満足すべき条件は

$$B_+s=y,\quad s=x_+-x, \tag{3.1}$$

$$y=\nabla_xL(x_+,\lambda_+)-\nabla_xL(x,\lambda_+) \tag{3.2}$$

となる．ここで(3.1)はセカント条件と呼ばれ，この条件を満足するような B から B_+ への更新公式を総称して準ニュートン更新公式と呼ぶ．代表的な更新公式としては BFGS (Broyden, Fletcher, Goldfarb and Shanno) 公式と DFP (Davidon, Fletcher and Powell) 公式があり，それぞれ

（i） **BFGS 公式**

$$B_+=B-Bss^{\mathrm{T}}B/s^{\mathrm{T}}Bs+yy^{\mathrm{T}}/s^{\mathrm{T}}y, \tag{3.3}$$

（ii） **DFP 公式**

$$B_+=B+(1+s^{\mathrm{T}}Bs/s^{\mathrm{T}}y)(yy^{\mathrm{T}}/s^{\mathrm{T}}y) -(Bsy^{\mathrm{T}}+ys^{\mathrm{T}}B)/s^{\mathrm{T}}y \tag{3.4}$$

で与えられる．

これらの公式では，B が正定値のとき B_+ が正定値になる（正定値性の保存）ための必要十分条件は $s^{\mathrm{T}}y>0$ が成り立つことである．しかしながら制約条件付き最小化問題ではラグランジュ関数は Kuhn-Tucker 点で凸ではないから $s^{\mathrm{T}}y>0$ は必ずしも成り立たない．したがって正定値性の保存が保証されないので，このままでは逐次2次計画法の大域的収束性は期待できない．この難点を解決するために Powell[10] は(3.1)の代わりに

$$B_+s=\eta,\quad \eta=\phi y+(1-\phi)Bs \tag{3.5}$$

を導入し，更新公式として**修正 BFGS 公式**

$$B_+=B-Bss^{\mathrm{T}}B/s^{\mathrm{T}}Bs+\eta\eta^{\mathrm{T}}/s^{\mathrm{T}}\eta \tag{3.6}$$

を提案した．ただし ϕ は $s^{\mathrm{T}}y\geqq 0.2\,s^{\mathrm{T}}Bs$ のとき 1，さもなければ $0.8\,s^{\mathrm{T}}Bs/s^{\mathrm{T}}(Bs-y)$ とおかれる．このとき B の正定値性が保存され，さらに $\phi=1$ のとき(3.6)は従来の BFGS 公式(3.3)に帰着する．

3.2 分解型準ニュートン更新公式

B の正定値性を保存する別の方法として，行列 B を $B=LL^{\mathrm{T}}$ と分解表現することが考えられている．ここで L は正則な n 次行列（三角行列とは限らない）である．この立場での研究としては Brodlie, Gourlay and Greenstadt[1], Dennis and Schnabel[2], Gill, Murray and Saunders[4] らの結果がある．本節ではセカント条件(3.1)に相当する条件

$$L_+L_+^{\mathrm{T}}s=y \tag{3.7}$$

を満足するような L から L_+ への更新公式（**分解型準ニュートン更新公式**）を考える．この結果は Brodlie らや Dennis らの結果を特別な場合として含む．さらにセカント条件をゆるめた条件を提案し，Powell の修正 BFGS 公式(3.6)の分解型公式を導出する．

まず次の集合を定義しておく．

$$V(y,s)=\{v\in R^n \mid v^{\mathrm{T}}v=s^{\mathrm{T}}y\}, \tag{3.8}$$

$$Q(y,s)=\{M\in R^{n\times n} \mid MM^{\mathrm{T}}s=y\}. \tag{3.9}$$

ここで上の条件(3.7)を満足する L_+ を見つけることと $v\in V(y,s)$ なる v に対する行列方程式

$$L_+v=y,\quad L_+^{\mathrm{T}}s=v \tag{3.10}$$

を解くことが同値であることに注目すれば，$s^{\mathrm{T}}y>0$ のとき集合 $Q(y,s)$ の要素が

$$L_+=yv^{\mathrm{T}}/s^{\mathrm{T}}y+[I-ps^{\mathrm{T}}/s^{\mathrm{T}}p]\Omega[I-vq^{\mathrm{T}}/q^{\mathrm{T}}v] \tag{3.11}$$

で表されることがわかる．ただし v は $v\in V(y,s)$ な

る任意ベクトル，p と q は $s^Tp\neq 0$, $q^Tv\neq 0$ なる任意ベクトル，Ω は $n\times n$ 任意行列である．

この結果を利用すれば，以下の手順によって BFGS 公式や DFP 公式の分解型公式（それぞれ分解型 BFGS 公式，分解型 DFP 公式と呼ぶ）を求めることができる[13]．ただし L の正則性と $s^Ty>0$ を仮定しておく．

（i） 分解型 BFGS 公式

$\Omega=L, v=\sqrt{s^Ty/s^TBs}L^Ts$, $q=L^Ts$ とおけば

$$L_+=L+[\sqrt{s^TBs/s^Ty}y-Bs](s/s^TBs)^TL \tag{3.12}$$

を得る．このとき L_+ の行列式は $\det L_+=\sqrt{s^Ty/s^TBs}\cdot(\det L)\neq 0$ であるから L_+ の正則性は保証され，したがって B_+ は正定値になる．

（ii） 分解型 DFP 公式

$\Omega=L$, $v=\sqrt{s^Ty/y^TB^{-1}y}L^{-1}y$, $q=L^Ts$ とおけば

$$L_+=L+(y/s^Ty)[\sqrt{s^Ty/y^TB^{-1}y}B^{-1}y-s]^TL \tag{3.13}$$

を得る．このとき $\det L_+=\sqrt{y^TB^{-1}y/s^Ty}(\det L)\neq 0$ であるから B_+ は正定値になる．

以上二つの公式は L についてのランク 1 公式であり，これらはまた Brodlie らや Dennis らの結果とも一致する．さらに(3.11)からいろいろな公式が導かれる．例えば，Broyden の 1 パラメータ公式は L についてのランク 2 公式として導かれ

$$L_+=L+(1-\sqrt{\phi})(\sqrt{\delta}\,y-Bs)(s/s^TBs)^TL +\sqrt{\phi}(y/s^Ty)(\sqrt{\delta}\,B^{-1}y-s)^TL, \tag{3.14}$$

$$B_+=B-Bss^TB/s^TBs+yy^T/s^Ty +\phi(s^TBs)(Bs/s^TBs-y/s^Ty) \times(Bs/s^TBs-y/s^Ty)^T \tag{3.15}$$

で与えられる[13]．ただし $\phi\geqq 0, \delta=1/[(1-\phi)s^Ty/s^TBs+\phi y^TB^{-1}y/s^Ty]$ である．このとき $\det L_+=(\det L)\times\sqrt{\delta}\,\{(1-\phi)(s^Ty)^2+\phi(s^TBs)(y^TB^{-1}y)\}/((s^Ty)(s^TBs))\neq 0$ であるから B_+ は正定値になる．

以上の議論は $s^Ty>0$ の仮定の下で行ってきたが，$s^Ty\leqq 0$ の場合にはもはやセカント条件を満足する L_+ を見つけることはできない．したがって第 3.1 節と同じ理由から，このままでは分解型公式を逐次 2 次計画法に適用するわけにはいかない．そこで行列方程式(3.10)のうちの一方の条件を捨てて，緩和された次の 2 種類の条件を考える．

（擬似セカント条件 1）

与えられたゼロでないベクトル $v\in R^n$ に対して $L_+^Ts=v$ を満足し，かつ $(1/2)\|L_+v-y\|_{W^{-1}}$ を最小にする L_+ を見つけよ．ただし $n\times n$ 正定値対称行列 W に対してノルムを $\|p\|_{W^{-1}}{}^2=p^TW^{-1}p$ と定義する．

（擬似セカント条件 2）

与えられたゼロでないベクトル $v\in R^n$ に対して $L_+v=y$ を満足し，かつ $(1/2)\|L_+^Ts-v\|_{W^{-1}}$ を最小にする L_+ を見つけよ．

以下の二節では，擬似セカント条件について議論し BFGS 公式や DFP 公式に対応する結果を導く．

3.3 擬似セカント条件 1

L_+ に関する最小化問題に対するラグランジュ関数は

$$\Phi(L_+,z)=(1/2)\|L_+v-y\|_{W^{-1}}{}^2+z^T(L_+^Ts-v) \tag{3.16}$$

である．ただし z はラグランジュ乗数ベクトルである．ここで Φ を L_+ と z について微分すれば $W^{-1}\times(L_+v-y)v^T+sz^T=0$, $L_+^Ts=v$ が得られ，さらに z を消去すれば行列方程式

$$L_+vv^T=(y+\{(v^Tv-s^Ty)/s^TWs\}Ws)v^T$$

を得る．このとき上の方程式と $L_+^Ts=v$ を満たす共通解（もとの問題の最適解）L_+ の一つは

$$L_+=L+(\hat{y}-Lv)((v-L^Ts)/v^T(v-L^Ts))^T \tag{3.17}$$

で与えられる．ただし $\hat{y}=y+\{(v^Tv-s^Ty)/s^TWs\}Ws$ である．このとき $L_+L_+^Ts$ と y との差は $\|L_+L_+^Ts-y\|=(|v^Tv-s^Ty|/s^TWs)\|Ws\|$ である．もし $s^Ty>0$ ならば $v\in V(y,s)$ なる v が選べるからセカント条件が満足される．

以下の議論では β を $s^T(y+\beta Ws)>0$ なるパラメータとする．式(3.17)で $v=\sqrt{s^T(y+\beta Ws)/s^TBs}\times L^Ts$ とおけば(3.12)に対応する公式

$$L_+=L+(\sqrt{s^TBs/s^T\hat{y}}\hat{y}-Bs)(s/s^TBs)^TL, \tag{3.18}$$

$$B_+=B-Bss^TB/s^TBs+\hat{y}\hat{y}^T/s^T\hat{y} \tag{3.19}$$

が得られる．ただし $\hat{y}=y+\beta Ws$ である．また(3.17)で $v=\sqrt{s^T\hat{y}/\hat{y}^TB^{-1}\hat{y}}L^{-1}\hat{y}$ とおけば(3.13)に対応する公式

$$L_+=L+(\hat{y}/s^T\hat{y})(\sqrt{s^T\hat{y}/\hat{y}^TB^{-1}\hat{y}}B^{-1}\hat{y}-s)^TL, \tag{3.20}$$

$$B_+=B+(1+s^TBs/s^T\hat{y})\hat{y}\hat{y}^T/s^T\hat{y} -(Bs\hat{y}^T+\hat{y}s^TB)/s^T\hat{y} \tag{3.21}$$

が得られる．ただし $\hat{y}=y+\beta Ws$ である．

特に $Ws=Bs-y$ なる正定値行列 W を選び，$\beta=1-\phi$, $\phi\in[0, 1]$ とすれば，公式(3.18)および(3.20)よりそれぞれ

$$L_+=L+(\sqrt{s^TBs/s^T\eta}\eta-Bs)(s/s^TBs)^TL, \quad (3.22)$$

$$L_+=L+(\eta/s^T\eta)(\sqrt{s^T\eta/\eta^TB^{-1}\eta}B^{-1}\eta-s)^TL \quad (3.23)$$

が得られる．ただし η は(3.5)で与えられる．ここで(3.22)は Powell の修正 BFGS 公式(3.6)の分解型公式である．したがって擬似セカント条件1の立場から Powell の修正 BFGS 公式を解釈することもできる．

3.4 擬似セカント条件 2

$Wv=v$ なる正定値行列 W を選んで，擬似セカント条件1の場合と同様にすれば L に関するランク1公式

$$L_+=L+((y-Lv)/s^T(y-Lv))((s^Ty/v^Tv)v-L^Ts)^T \quad (3.24)$$

が得られる．このとき $L_+L_+^Ts$ と y との差は $\|L_+\times L_+^Ts-y\|=(|v^Tv-s^Ty|/v^Tv)\|y\|$となる．もし$s^Ty>0$ ならば $v\in V(y,s)$ なる v が選べるからセカント条件が満足される．

以下の議論では γ を $(1-\gamma)s^Ty>0$ なるパラメータとする．式(3.24)で $v=\sqrt{s^Ty/((1-\gamma)s^TBs)}L^Ts$ とおけば(3.12)に対応する公式

$$L_+=L+(\sqrt{(1-\gamma)s^TBs/s^Ty}y-Bs)(s/s^TBs)^TL, \quad (3.25)$$

$$B_+=B-Bss^TB/s^TBs+(1-\gamma)yy^T/s^Ty \quad (3.26)$$

が得られる．また(3.24)で $v=\sqrt{s^Ty/((1-\gamma)y^TB^{-1}y)}\times L^{-1}y$ とおけば(3.13)に対応する公式

$$L_+=L+(y/s^Ty)(\sqrt{(1-\gamma)s^Ty/y^TB^{-1}y}B^{-1}y-s)^TL, \quad (3.27)$$

$$B_+=B+(1-\gamma+s^TBs/s^Ty)yy^T/s^Ty-(Bsy^T+ys^TB)/s^Ty \quad (3.28)$$

が得られる．

以上の二節で擬似セカント条件について議論してきた．簡単に言えば，擬似セカント条件1はyに補正ベクトルを付け加えて $B_+s=y+$(補正ベクトル) としてBの正定値性の保存を実現するものであり，他方擬似セカント条件2はyをスカラ倍してBの正定値性の保存を実現しようとしている．後者はy方向の情報しか取り入れておらず，特に $s^Ty=0$ の場合にはもはや正則な更新公式が定義できない．とりわけ $y=0$(線形計画問題では $y=0$ になる) の場合には条件 $L_+v=y$ そのものに特異性が含まれている．以上のことを考慮すれば，擬似セカント条件1から得られる更新公式の方が望ましいように思われる．

4. 分解型準ニュートン公式を用いた逐次2次計画法

2次計画部分問題にアルゴリズム GI を適用した場合，行列 B の Cholesky 分解を求めなければならない．しかしながら Cholesky 分解そのものはアルゴリズム GI のステップ0における無制約最小点の計算には有効であるものの，途中の Givens 変換によって三角行列の形がくずされるから必ずしも B を Cholesky 分解しておく必然性はない．前章で論じてきた分解型公式は三角行列ではないけれども，上のような理由から利用できることが期待される．すなわち $B=LL^T$ に対して，L^{-T} がアルゴリズム GI の J_2 の初期行列としてそのまま使用できる．さらに $\hat{L}=L^{-T}$ に関する更新公式を構成すれば逆行列の計算は不要になる．そのために Sherman-Morrison 公式を使えば，(3.22)および(3.23)に関する $\hat{L}$ の更新公式はそれぞれ

$$\hat{L}_+=\hat{L}+(s/s^T\eta)[\sqrt{s^T\eta/s^TBs}Bs-\eta]^T\hat{L}, \quad (4.1)$$

$$\hat{L}_+=\hat{L}+(\sqrt{\eta^T\hat{L}\hat{L}^T\eta/s^T\eta}s-\hat{L}\hat{L}^T\eta)(\eta/\eta^T\hat{L}\hat{L}^T\eta)^T\hat{L} \quad (4.2)$$

で与えられる．このとき $\hat{L}$ は $O(n^2)$ の計算量で更新される．

以上のことを考慮すれば，分解型更新公式を用いたアルゴリズム SQP (FSQP) は次のようになる．

(アルゴリズム FSQP)

初期点 x_1, $n\times n$ 正則行列 $\hat{L}_1$ および三つの定数 $r>0$, $\omega\in(0,\ 0.5)$, $\tau\in(0,\ 1)$ が与えられたとき，以下の手順を $k=1,2,\cdots$に対して繰り返す；

ステップ1．x_k と $\hat{L}_k$ に対する次の2次計画問題

(2次計画部分問題)

線形制約条件 $g(x_k)+\nabla g(x_k)d\leqq 0$ の下で，2次関数 $(1/2)d^T(\hat{L}_k\hat{L}_k^T)^{-1}d+\nabla f(x_k)^Td$ を最小にせよ

を解いて，方向ベクトル d_k およびラグランジュ乗数ベクトル λ_{k+1} を求める．

ステップ2からステップ4まではアルゴリズム SQP と同じ手順である．

ステップ5．行列 $\hat{L}_k$ を分解型準ニュートン公式に従って更新し $\hat{L}_{k+1}$ を作成する．$k=k+1$ とおいてステップ1へ行く．

上のアルゴリズムの場合，Bs を直接計算することは不利であるから，実際には2次計画部分問題の Kuhn-Tucker 条件 $Bd+\nabla f(x)+\nabla g(x)^T\lambda_+=0$ から $Bs=\alpha Bd$ が求められる．

アルゴリズム FSQP の効率と SQP の効率を比較するために，以下に実行上の注意を述べておく．アルゴリズム SQP を実行する場合には行列 B 自身を更新するよりも Gill, Murray and Saunders[4]のテクニックを用いて，B の Cholesky 分解を更新する方が有効である．例えば Powell の修正 BFGS 公式(3.6)を用いる場合，まず(3.6)に Gill らのテクニックを2回適用して B の Cholesky 因子 F を求め ($O(n^2)$ の計算量ですむ)，次に F^{-T} を求める ($O(n^3)$ の計算量を必要とする)．そして2次計画問題を解く際に，$J_2=F^{-T}$ がアルゴリズム GI の初期行列として用いられる．このように Gill らのテクニックを組み込んだアルゴリズム SQP を**アルゴリズム GSQP** と呼ぶことにする．

それに対して FSQP では，アルゴリズム GI での J_2 の初期行列を毎回 $O(n^2)$ の計算量で作成できる．

5. 数値実験結果

公式(4.1)と(4.2)を用いたアルゴリズム FSQP と公式(3.6)を用いたアルゴリズム GSQP に関するいくつかの数値実験結果を報告する．アルゴリズムは FORTRAN 77 で記述され，すべての実変数は倍精度で計算された．また逆行列の計算，Givens 変換，ランク1更新公式に対する Gill, Murray, Saunders のテクニックなどはパッケージ LINPACK[3] を利用した．数値実験は NEC ACOS-350 で行われ，使用した収束判定条件は

$$\|\nabla_x L(x, \lambda_+)\|_\infty \leqq \varepsilon, \tag{5.1a}$$

$$g_i(x) \leqq \varepsilon, \quad e_i^T \lambda_+ \geqq -\varepsilon, \quad i=1, \cdots, m, \tag{5.1b}$$

$$|(e_i^T \lambda_+) g_i(x)| \leqq \varepsilon, \quad i=1, \cdots, m \tag{5.1c}$$

である．ただし ε はあらかじめ与えられた正定数である．等号制約条件 $h_j(x)=0$ $(j=1, \cdots, l)$ は二つの不等号制約条件 $h_j(x) \leqq 0$, $-h_j(x) \leqq 0$ として扱われた．またテスト問題は Pierre and Lowe[9] から引用した．それらを載せることは誌面の都合上割愛するとして，問題名と問題の規模を以下にあげておく；

問題 LP (Pierre) $n=4$, $m=7$, $l=1$.

問題 FM (Fiacco and McCormick) $n=2$, $m=2$, $l=0$.

問題 RS (Rosen and Suzuki) $n=4$, $m=3$, $l=0$.

問題 B (Beale) $n=3$, $m=4$, $l=0$.

問題 P (Powell) $n=5$, $m=0$, $l=3$.

問題 W 1 (Wong) $n=7$, $m=4$, $l=0$.

表 1 FSQP と GSQP の数値実験結果
Table 1 Summary of numerical experiments by algorithms FSQP and GSQP.

	公式(3.6)				(4.1)	公式(4.2)			
問題	IT	FE	QP	CPU	CPU	IT	FE	QP	CPU
LP	6	6	16	0.634	0.579	6	6	16	0.762
FM	2	2	4	0.492	0.309	2	2	4	0.281
RS	10	15	20	1.222	0.614	9	14	19	0.741
B	6	7	7	0.591	0.409	7	7	8	0.456
P	8	8	24	0.982	0.808	9	10	27	0.898
W1	15	33	32	1.878	1.415	14	29	29	1.367
W2	14	23	90	4.553	3.845	13	20	81	3.528
W3	24	47	281	35.448	31.505	28	53	332	38.104

問題 W 2 (Wong) $n=10$, $m=8$, $l=0$.

問題 W 3 (Wong) $n=20$, $m=17$, $l=0$.

公式(3.6), (4.1), (4.2)の初期行列は $B_1=I$, $L_1=I$ とし，直線探索ではパラメータ r, ω, τ をそれぞれ 10.0, 0.1, 0.5 とおいた．また問題 LP, FM, RS, B, P では $\varepsilon=10^{-6}$, 問題 W 1, W 2, W 3 では$\varepsilon=10^{-4}$ を用いた．**表 1** には反復回数 (IT)，目的関数評価回数 (FE)，CPU 時間 (CPU) および2次計画部分問題を解くのに必要とした総反復回数 (QP) を示す．

公式(3.6)と(4.1)は理論的に同じ2次計画部分問題を与えるから，表1の IT, FE, QP は同じ結果になる．しかしながら前章で述べたように両者の違いは演算回数にあるので，それが CPU 時間に現れている．問題の規模が大きくなればなるほど，この違いは大きくなることが予想される．なお我々は逆行列の分解型公式を用いているけれども，数値計算上の不安定現象は特に認められず，アルゴリズム GSQP と同様に安定に数値解が得られた．

6. おわりに

本稿では逐次2次計画法の大域的収束性を保証するために，擬似セカント条件の概念を導入した．その結果として Dennis ら，Brodlie らや Powell の結果を特別な場合として含むような分解型準ニュートン更新公式が統一的に構成された．さらに分解型公式は，2次計画部分問題を解くための Goldfarb-Idnani 法ともうまく適合することが数値実験でも確かめられた．

また我々は青山学院大学付属情報科学研究センターにおいて，制約条件付き非線形最小化問題のためのパッケージ ASNOP (An Application System for Nonlinear Optimization Problems) を開発しており，1985 年版に拡張ラグランジュ関数法に加えて逐

次2次計画法を組み込んだ[14]．このパッケージでは，公式(4.1)を用いたアルゴリズム FSQP が用意されている．本稿で取り上げたテスト問題に関する限り，関数評価回数の意味では拡張ラグランジュ関数法よりも逐次2次計画法の方が優れていた．

なお今回は問題 NLP のラグランジュ関数のヘッセ行列を近似することを考えてきたが，ヘッセ行列の逆行列を直接近似することも同様に議論できる．従来，逆行列の近似行列に関する更新公式は数値的に不安定であると言われているけれども，このことに関してはこれから検討していきたい．

参考文献

1) Brodlie, K. W., Gourlay, A. R. and Greenstadt, J.: Rank-one and Rank-two Corrections to Positive Definite Matrices Expressed in Product Form, *J. Inst. Math. Appl.*, Vol. 11, pp. 73-82 (1973).

2) Dennis, J. E., Jr. and Schnabel, R. B.: A New Derivation of Symmetric Positive Definite Secant Updates, in Mangasarian, O. L., Meyer• R. R. and Robinson, S. M. (eds.), *Nonlinear Programming 4*, pp. 167-199, Academic Press, New York (1981).

3) Dongarra, J. J., Moler, C. B., Bunch, J. R. and Stewart, G. W.: *LINPACK Users' Guide*, SIAM, Philadelphia (1979).

4) Gill, P. E., Murray, W. and Saunders, M. A.: Methods for Computing and Modifying the LDV Factors of a Matrix, *Math. Comp.*, Vol. 29, No. 132, pp. 1051-1077 (1975).

5) Goldfarb, D. and Idnani, A.: A Numerically Stable Dual Method for Solving Strictly Convex Quadratic Programs, *Math. Program.*, Vol. 27, pp. 1-33 (1983).

6) Han, S. P.: Superlinearly Convergent Variable Metric Algorithms for General Nonlinear Programming Problems, *Math. Program.*, Vol. 11, pp. 263-282 (1976).

7) Han, S. P.: Variable Metric Methods for Minimizing a Class of Nondifferentiable Functions, *Math. Program.*, Vol. 20, pp. 1-13 (1981).

8) Mayne, D. Q. and Polak, E.: A Superlinearly Convergent Algorithm for Constrained Optimization Problems, *Math. Program. Study*, Vol. 16, pp. 45-61 (1982).

9) Pierre, D. A. and Lowe, M. J.: *Mathematical Programming via Augmented Lagrangians*, pp. 232-241, Addison-Wesley Publishing Company, London (1975).

10) Powell, M. J. D.: A Fast Algorithm for Nonlinearly Constrained Optimization Calculations, in Watson, G. A. (ed.), *Numerical Analysis, Lecture Notes in Mathematics*, Vol. 630, pp. 144-157, Springer-Verlag, Berlin (1978).

11) Powell, M. J. D.: Variable Metric Methods for Constrained Optimization, in Bachem, A., Grotschel, M. and Korte, B. (eds.), *Mathematical Programming: The State of the Art*, pp. 288-311, Springer-Verlag, Berlin (1983).

12) Powell, M. J. D.: On the Quadratic Programming Algorithm of Goldfarb and Idnani, *Math. Program. Study*, Vol. 25, pp. 46-61 (1985).

13) Takahashi, S., Yamaki, N. and Yabe, H.: Application of Factorized Form to Quasi-Newton Update, *TRU Math.*, Vol. 20, pp. 265-276 (1984).

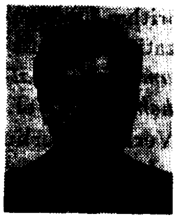
14) 高橋，本郷，宮田，八巻，矢部：非線型最適化問題解決のパッケージ(Ⅲ)，青山コンピュータ・サイエンス，Vol. 13, No. 1, pp. 53-79 (1985).

15) Tanabe, K.: Feasibility-Improving Gradient-Acute-Projection Methods: A Unified Approach to Nonlinear Programming, in Yamamoto, T. and Tanabe, K. (eds.), *The Newton Method and Related Topics, Lecture Notes in Numerical and Applied Analysis*, Vol. 3, pp. 57-76, Kinokuniya Company Ltd., Tokyo (1981).

16) Yamashita, H.: Quadratic Programming Approximation for Nonlinear Optimization, in *Proceedings of the 4th Mathematical Programming Symposium Japan*, pp. 35-58 (1983).

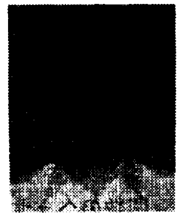
(昭和61年3月12日受付)
(昭和61年9月10日採録)

矢部　博（正会員）
昭和30年生．昭和52年東京理科大学理学部応用数学科卒業．昭和57年同大学院博士課程（数学専攻）修了．同年東京理科大学理学部応用数学科助手，現在に至る．理学博士．非線形最適化問題の数値解法に関する研究に従事．日本数学会，日本オペレーションズ・リサーチ学会各会員．



高橋　悟

昭和25年生．昭和50年東京理科大学理学部応用数学科卒業．昭和52年同大学院修士課程（数学専攻）修了．同年10月同大学院博士課程中退，同年東京理科大学理学部応用数学科助手，現在に至る．非線形最適化問題や偏微分方程式などの数値解法に関する研究に従事．日本数学会，日本オペレーションズ・リサーチ学会各会員．



八巻　直一（正会員）

昭和19年生．昭和42年東京理科大学理学部応用数学科卒業．昭和45年同大学院修士課程（数学専攻）修了．東京理科大学理学部応用数学科助手を経て，現在システム計画研究所(株)取締役，青山学院大学非常勤講師．非線形計画法の研究を続けるとともに，ソフトウェアの生産管理にも興味をもつ．著書「社内標準化便覧」（共著，日本規格協会，1985年）．日本数学会，日本オペレーションズ・リサーチ学会，日本経営工学会，日本ソフトウェア科学会各会員．