線形等式つき最適化問題に対する 並列準ニュートン射影法[†]

相 吉 英太郎*・丹 内 智 得*

A Parallel Quasi-Newton Method for Linearly Constrained Optimization Problems

Eitaro AIYOSHI* and Chieri TANNAI*

A quasi-Newton method with parallel computing capabilities for unconstrained optimization problems, which is called a parallel quasi-Newton method, is converted to linearly constrained cases. A parallel quasi-Newton method is characterized by simultaneous perturbations of a trial point in plural directions, and by approximation to an inverse of the Hessian of the objective function by use of informations at the perturbed points. Then, the parallelism can be induced into the perturbation process as well as the computations of the gradients at the perturbed points. In our converted method, the perturbation process in parallel fashion is performed on the linear manifold of the equality constraints. When the objective function is quadratic, the optimum on the constraints can be obtained by use of the matrix type of the BFGS updating formula after a trial point is renewed only once.

The proposed algorithm is tested on some examples. The experiment in serial fashion on the currently used computer indicates that the algorithm effects better convergence than the gradient projection method and the quasi-Newton projection method. From these results, it is expected that the good convergence becomes pronounced on parallel computers with multi-processor.

Key Words: nonlinear optimization, linearly constrained optimization, parallel computation, parallel quasi-Newton method

1. はじめに

数値計算の計算時間を短縮する手段の一つとして、計算処理の並列化が考えられているが、そのためには並列処理装置の開発と、その並列処理に適したアルゴリズムの開発が必要不可欠である。本論文では、線形等式制約条件つき最適化問題の計算手法として、並列処理が可能となるように変形した準ニュートン射影法、いわゆる並列準ニュートン射影法を提案する。

無制約最適化問題に対する並列準ニュートン法はすでにいくつか提案され、その数値計算特性が考察されている^{1),2)}。これらの手法の特徴は、決定変数の次元の数だけ独立方向へ試行点を摂動し、その摂動点での勾配などの情報をもとに目的関数のヘッセ行列の逆行列を近似することであり、その際の試行点の複数方向への摂動作業と各摂動点での勾配計算において、並列計算処理が行えるというものである。文献1)の並列準ニュートン法では、ヘッセ行列の逆行列の近似計算に、通常の準ニュートン法のランク1公式に相当する更新公式を用いるのに対して、文献2)ではこの近似計算を一層効率化するために、複数個の摂動方向ベクトルや摂動点での勾配情報を群にまとめてそれらを行列化し、行列形式の更新公式を導入している。そして、次元数個の摂動方向ベクトルのすべてと、摂動点での勾配のすべてをそれぞれ一まとめにして行列化した場合には、ヘッセ行列の逆行列を近似する1回の更新作業が、通常の準ニュートン法の次元数回分の更新作業に相当するものとみなすことができ、実際に並列処理を行わなくても通常の準ニュートン法と比較して十分高い効率性を有し、さらに摂動作業や勾配計算および更新式の行列演算の並列化によって一層の効率化が実現できることが文献3)によって確認されている。

一方、等式制約条件つき問題に対しては、ラグランジュ関数のヘッセ行列を並列準ニュートン法の原理で

[†] 第28回計測自動制御学会学術講演会で一部発表
(1989・7)

* 慶應義塾大学理工学部 横浜市港北区日吉 3-14-1

* Faculty of Science and Technology, Keio
University, Yokohama
(Received June 11, 1991)

(Revised August 29, 1991)

近似して、補助の2次計画問題を逐次解く手法が提案された⁴⁾。この手法は、制約条件が非線形である場合に有効な手段であり、いわばラグランジュ準ニュートン法の並列化手法である。これに対して、線形状の等式制約の場合には、その線形性を生かした並列化手法のほう効率的である。そこで、本論文では、等式制約式が形成する線形多様体を張る独立な方向ベクトルを用いることによって、無制約問題に対する並列準ニュートン法の計算原理をこの多様体上に限定させたいわゆる線形多様体上の並列準ニュートン法を提案する。

文献2), 3)の無制約問題に対する準ニュートン法の並列化や、文献4)のラグランジュ準ニュートン法の並列化は、ヘッセ行列の逆行列やヘッセ行列そのものの近似計算を並列化したものにすぎない。これに対して、本論文で新たに提案する手法の並列化は、線形多様体への射影行列とヘッセ行列の逆行列との積行列の近似計算の並列化である。そこで本手法を「並列準ニュートン射影法」と称し、勾配射影法、ニュートン射影法、準ニュートン射影法との比較により、本手法の有効性を簡単な数値例で確認する。

2. 線形多様体上の並列準ニュートン法

2.1 基本手順

本論文では、線形等式制約条件つき最適化問題

$$\min_x f(\mathbf{x}) \quad (1a)$$

$$\text{subj. to } A\mathbf{x} = \mathbf{b} \quad (1b)$$

を考える。ただし、 $\mathbf{x} \in R^n$, $\mathbf{b} \in R^l$, A は $l \times n$ 行列 $[\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_l]^T$ で $\text{rank } A = l$, $f: R^n \rightarrow R^1$ は、2回連続微分可能とする。なお、 $\mathbf{a}_i \in R^n$ で、 T は転置記号を表わす。

いま、制約式(1b)の行列 A はフルランクなので、制約式を満たす領域は、 $n-l$ 次元線形多様体となり、 $n-l$ 個の独立なベクトルでこの領域全体を張ることができる。したがって、このような $n-l$ 個の方向ベクトルを用いることにより、無制約問題に対する並列準ニュートン法と同様の原理で、線形多様体上の関数 $f(\mathbf{x})$ の最小点を求めることができる。

まず、線形多様体上の並列準ニュートン法の基本原理を明らかにするために、(1a)式の目的関数 $f(\mathbf{x})$ が、2次関数

$$f(\mathbf{x}) = \frac{1}{2} \mathbf{x}^T C \mathbf{x} + \mathbf{c}^T \mathbf{x} \quad (2)$$

であるとする。ただし、 C は正定値対称な $n \times n$ 行列、 $\mathbf{c} \in R^n$ である。このとき、つぎの定理が成り立つ。

《定理1》 問題(1)において $f(\mathbf{x})$ が(2)式のよう

な2次関数であるとする。制約式(1b)を満たす任意の $\mathbf{x}^k$ に対し

$$\mathbf{s}^k = -(I - C^{-1}A^T(AC^{-1}A^T)^{-1}A)C^{-1}\nabla f(\mathbf{x}^k)^T \quad (3)$$

とすると、更新点 $\mathbf{x}^{k+1} = \mathbf{x}^k + \mathbf{s}^k$ は問題(1)の最適解である。ただし、 $\nabla f(\mathbf{x}^k)$ は $\mathbf{x}^k$ における f の勾配を表わし、横ベクトルとする。

(証明) 文献5)の練習問題9.1の間5参照。

(3)式において

$$P_e = I - C^{-1}A^T(AC^{-1}A^T)^{-1}A \quad (4)$$

とおくと、この行列 P_e はニュートン法の移動方向 $-C^{-1}\nabla f(\mathbf{x}^k)^T$ を部分空間 $\{\mathbf{s} | A\mathbf{s} = \mathbf{0}\}$ へ写像する射影行列であり、 C が単位行列の場合は正射影行列となり、(3)式の方法は勾配射影法の移動方向となる。そしてこの定理は、ニュートン法の移動方向を行列 P_e で射影した方向 $\mathbf{s}^k$ へ、任意の試行点 $\mathbf{x}^k$ を1回だけ更新しさえすれば、制約式(1b)上の2次関数(2)の最小点が得られることを保証している。もし関数 f が2次関数でない一般の非線形関数ならば、ヘッセ行列 $\nabla^2 f(\mathbf{x}^k)$ を用いた式

$$\mathbf{s}^k = -(I - [\nabla^2 f(\mathbf{x}^k)]^{-1}A^T(A[\nabla^2 f(\mathbf{x}^k)]^{-1}A^T)^{-1}A) \cdot [\nabla^2 f(\mathbf{x}^k)]^{-1}\nabla f(\mathbf{x}^k)^T \quad (5)$$

を使用して、試行点の逐次更新を行えば、試行点の更新回数に関する限り効率的であることが容易に理解される。しかしながら、この更新方向の計算には、 $[\nabla^2 f(\mathbf{x}^k)]^{-1}$ と $(A[\nabla^2 f(\mathbf{x}^k)]^{-1}A^T)^{-1}$ の2回の逆行列計算を含むため、仮にヘッセ行列 $\nabla^2 f(\mathbf{x}^k)$ が正則行列であるとしても、1回の更新作業にかかる計算量は変数 ($\mathbf{x}$ の次元数) の増加とともに著しく増大すると予想される。そこで本論文では、行列

$$(I - [\nabla^2 f(\mathbf{x}^k)]^{-1}A^T(A[\nabla^2 f(\mathbf{x}^k)]^{-1}A^T)^{-1}A) \cdot [\nabla^2 f(\mathbf{x}^k)]^{-1} \quad (6)$$

の部分に適当な行列で近似し、さらにこの近似計算を並列準ニュートン法の原理を用いて効率化する。

ところで、(6)式を適当な行列で近似するという考え方は、Goldfarb⁶⁾による準ニュートン射影法にすでに見られる。これは(6)式を行列 H^k で近似し、(5)式の代わりに

$$\mathbf{s}^k = -H^k \nabla f(\mathbf{x}^k)^T \quad (7)$$

により、現試行点 $\mathbf{x}^k$ の移動方向を計算する手法である。この場合、 $\mathbf{s}^k$ 方向へ探索して得られる更新点 $\mathbf{x}^{k+1}$ との変位 $\mathbf{p}^k = \mathbf{x}^{k+1} - \mathbf{x}^k$, および勾配差 $\mathbf{q}^k = \nabla f(\mathbf{x}^{k+1})^T - \nabla f(\mathbf{x}^k)^T$ を用い、たとえば BFGS 公式

$$H^{k+1} = \left(I - \frac{\mathbf{p}^k \mathbf{q}^{kT}}{\mathbf{q}^{kT} \mathbf{p}^k} \right) H^k \left(I - \frac{\mathbf{q}^k \mathbf{p}^{kT}}{\mathbf{q}^{kT} \mathbf{p}^k} \right) + \frac{\mathbf{p}^k \mathbf{p}^{kT}}{\mathbf{q}^{kT} \mathbf{p}^k} \quad (8)$$

などの公式により、近似行列の更新がイテレーションの更新とともに逐次的に行われる。そして、初期点 $\mathbf{x}^1$ を制約式(1b)式を満たすように選択し、初期近似行列 H^1 を

$$H^1 = I - A^T(AA^T)^{-1}A \quad (9)$$

にとることによって、帰納的に

$$AH^k = O, \quad As^k = 0, \quad Ax^k = b \quad k=1, 2, 3, \dots$$

が成立するようにした手法である。

本論文の並列準ニュートン射影法では、以上の準ニュートン射影法のように、(6)式の近似を逐次的に行うのではなく、試行点を制約式上で複数独立方向へ摂動させることによって得られる情報により、(6)式の近似行列を生成する。

そこでまず $f(\mathbf{x})$ が2次関数の場合に、行列 P_c とヘッセ行列の逆行列 $C^{-1} (= [\nabla^2 f(\mathbf{x}^*)]^{-1})$ との積行列 $V = P_c C^{-1}$ が満たすべき条件を与えるために、互いに独立な $n-l$ 個の方向ベクトル $\mathbf{d}_1, \mathbf{d}_2, \dots, \mathbf{d}_{n-l}$ を考え、また

$$A\bar{\mathbf{x}} = b$$

なる任意の $\bar{\mathbf{x}}$ とその摂動点 $\bar{\mathbf{x}} + \mathbf{d}_i$, ($i=1, 2, \dots, n-l$) との勾配差を

$$\mathbf{y}_i = \nabla f(\bar{\mathbf{x}} + \mathbf{d}_i)^T - \nabla f(\bar{\mathbf{x}})^T, \quad i=1, 2, \dots, n-l$$

とする。さらにこれらを行列にまとめて

$$D = [\mathbf{d}_1, \mathbf{d}_2, \dots, \mathbf{d}_{n-l}] \quad D: n \times (n-l) \text{ 行列}$$

$$Y = [\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_{n-l}] \quad Y: n \times (n-l) \text{ 行列}$$

とおく。このときつぎの定理が成り立つ。

《定理2》 問題(1)において目的関数 $f(\mathbf{x})$ は(2)式のような2次関数であるとする。このとき、

$$AD = O \quad (10)$$

を満たす $n \times (n-l)$ 行列 D およびこれに対応した $n \times (n-l)$ 行列 Y , さらにある $n \times n$ 対称行列 $\tilde{V}$ に対し

$$A\tilde{V} = O \quad (11)$$

$$\tilde{V}Y = D \quad (12)$$

が成り立つならば、移動方向

$$\mathbf{s} = -\tilde{V}\nabla f(\bar{\mathbf{x}})^T \quad (13)$$

に対する更新点 $\bar{\mathbf{x}} + \mathbf{s}$ は問題(1)の最適点である。

(証明) まず、 $f(\mathbf{x})$ が(2)式のような2次関数ならば $\nabla f(\mathbf{x}) = \mathbf{x}^T C + \mathbf{c}^T$ であるから

$$\begin{aligned} \bar{\mathbf{x}} + \mathbf{s} &= \bar{\mathbf{x}} - \tilde{V}\nabla f(\bar{\mathbf{x}})^T \\ &= \bar{\mathbf{x}} - \tilde{V}(C\bar{\mathbf{x}} + \mathbf{c}) \end{aligned}$$

であり、これを制約式(1b)の左辺に代入し(11)式を用いると

$$\begin{aligned} A(\bar{\mathbf{x}} + \mathbf{s}) &= A(\bar{\mathbf{x}} - \tilde{V}(C\bar{\mathbf{x}} + \mathbf{c})) \\ &= A\bar{\mathbf{x}} - A\tilde{V}(C\bar{\mathbf{x}} + \mathbf{c}) = b \end{aligned}$$

となり、新点 $\bar{\mathbf{x}} + \mathbf{s}$ も制約式を満たす。また、 $f(\mathbf{x})$

が2次関数ならば

$$CD = Y$$

が成り立つから、両辺の左側から $\tilde{V}$ を掛け、(12)式を用いると

$$\tilde{V}CD = \tilde{V}Y = D \quad (14)$$

となる。ところで新点 $\bar{\mathbf{x}} + \mathbf{s}$ での勾配は、

$$\begin{aligned} \nabla f(\bar{\mathbf{x}} + \mathbf{s}) &= [\bar{\mathbf{x}} - \tilde{V}(C\bar{\mathbf{x}} + \mathbf{c})]^T C + \mathbf{c}^T \\ &= \bar{\mathbf{x}}^T C - \bar{\mathbf{x}}^T C \tilde{V} C - \mathbf{c}^T \tilde{V} C + \mathbf{c}^T \end{aligned}$$

となるから、この式に右側から D を掛け(14)式を用いると

$$\begin{aligned} \nabla f(\bar{\mathbf{x}} + \mathbf{s})D &= \bar{\mathbf{x}}^T CD - \bar{\mathbf{x}}^T C \tilde{V} CD \\ &\quad - \mathbf{c}^T \tilde{V} CD + \mathbf{c}^T D \\ &= \bar{\mathbf{x}}^T CD - \bar{\mathbf{x}}^T CD - \mathbf{c}^T D + \mathbf{c}^T D \\ &= 0 \end{aligned}$$

となり、 $D = [\mathbf{d}_1, \mathbf{d}_2, \dots, \mathbf{d}_{n-l}]$ であるから、新しく得られた点での勾配 $\nabla f(\bar{\mathbf{x}} + \mathbf{s})$ が、すべての方向ベクトル $\mathbf{d}_i$, $i=1, 2, \dots, n-l$ と直交することを示している。方向ベクトル $\mathbf{d}_i$, $i=1, 2, \dots, n-l$ は部分空間 $\{\mathbf{s} | A\mathbf{s} = 0\}$ を張る独立なベクトルであるから、 $A\mathbf{s} = 0$ を満たすすべての $\mathbf{s}$ に対し

$$\nabla f(\bar{\mathbf{x}} + \mathbf{s})\mathbf{s} = 0$$

であり、 $\bar{\mathbf{x}} + \mathbf{s}$ は線形多様体上 $\{\mathbf{x} | A\mathbf{x} = b\}$ における最小点となる。■

この定理により、もし $f(\mathbf{x})$ が2次関数ならば、

$$A\mathbf{d}_i = 0, \quad i=1, 2, \dots, n-l$$

なる互いに独立な $n-l$ 個の方向ベクトル $\mathbf{d}_i$, $i=1, 2, \dots, n-l$ を用意し、これを用いて(11), (12)式を満たす $n \times n$ 対称行列 $\tilde{V}$ を適当な行列から生成すれば、射影行列 P_c を計算しなくてもよいことがわかる。このような行列 $\tilde{V}$ の具体的生成法を後述することになれば、線形多様体上の2次関数の最小点を求める手法としてつぎのような基本手順を考えることができる。

〈基本手順〉

step 1: 制約式 $A\mathbf{x} = b$ を満たす試行点 $\bar{\mathbf{x}}$ と適当な初期対称行列 $\bar{V}$ を与える。

step 2: $A\mathbf{d}_i = 0$ を満たす互いに独立な $n-l$ 個の方向ベクトル $\mathbf{d}_i$ ($i=1, 2, \dots, n-l$) を用意する。

step 3: 基点 $\bar{\mathbf{x}}$ からの $n-l$ 個の摂動点

$$\mathbf{x}_i = \bar{\mathbf{x}} + \mathbf{d}_i, \quad i=1, 2, \dots, n-l$$

と、基点 $\bar{\mathbf{x}}$ との勾配差 $\mathbf{y}_i$

$$\mathbf{y}_i = \nabla f(\mathbf{x}_i)^T - \nabla f(\bar{\mathbf{x}})^T, \quad i=1, 2, \dots, n-l$$

を計算し、行列

$$D = [\mathbf{d}_1, \mathbf{d}_2, \dots, \mathbf{d}_{n-l}]$$

$$Y = [\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_{n-l}]$$

を構成する。

step 4: 行列 D, Y を用いて $\bar{V}$ を修正し、(11), (12)

式を満たす行列 $\tilde{V}$ を生成する.

step 5: $\tilde{V}$ を用いて

$$\mathbf{s} = -\tilde{V} \nabla f(\bar{\mathbf{x}})^T$$

を計算し, $\bar{\mathbf{x}} + \mathbf{s}$ を最適解とする.

上記の基本手順中 step 3 の振動点 $\mathbf{x}_i$, $i=1, 2, \dots, n-l$ および勾配差 $\mathbf{y}_i$, $i=1, 2, \dots, n-l$ の計算がそれぞれ並列計算可能であることに注意されたい. また目的関数 f が非 2 次関数の一般の非線形関数の場合には, $\bar{\mathbf{x}} + \mathbf{s}$ は必ずしも最適点になるとは限らないため, 新点 $\bar{\mathbf{x}} + \mathbf{s}$ をあらためて試行点として上記の基本手順を繰り返すことになる. この場合, 行列 $\tilde{V}$ は (6) 式で計算される行列の近似行列とみなすことができる. 以下でこの近似行列を具体的に計算するための更新公式を導入する.

2.2 近似行列の更新式

近似行列 $\tilde{V}$ が満たすべき条件式 (12) は, 文献 2) で提案された並列準ニュートン法における並列準ニュートン条件と一致する. そこで (11), (12) 式を満たす $\tilde{V}$ の生成法として, 並列準ニュートン法で使した更新公式である行列型 BFGS 公式

$$\tilde{V} = \left(I_n - \frac{DY^T}{\varepsilon^2} \right) \bar{V} \left(I_n - \frac{YD^T}{\varepsilon^2} \right) + \frac{DD^T}{\varepsilon^2} \quad (15)$$

I_n : $n \times n$ 単位行列

を, 文献 2) の場合と同様に D と Y に対する条件

$$D^T Y = \varepsilon^2 I_{n-l} \quad (I_{n-l}: (n-l) \times (n-l) \text{ 単位行列}) \quad (16)$$

のもとで用いることが考えられる. このときつぎの定理の性質が容易に示される.

《定理 3》 問題 (1) において $f(\mathbf{x})$ が (2) 式のような 2 次関数であるとする. このとき初期行列 $\bar{V}$ に対して

$$A\bar{V} = O \quad (17)$$

が成り立ち, かつ (10) 式を満たす $n \times (n-l)$ 行列 D と対応する $n \times (n-l)$ 行列 Y に対して

$$D^T C D = \varepsilon^2 I_{n-l} \quad (I_{n-l}: (n-l) \times (n-l) \text{ の単位行列}) \quad (18)$$

が成り立つならば, 更新公式 (15) によって得られる $\tilde{V}$ に対して条件式 (11), (12) が成立する.

(証明) 比較的容易なので省略する ((12) 式の成立は文献 2) の付録参照).

更新式 (15) は, 準ニュートン法の BFGS 公式

$$V_{i+1} = \left(I_n - \frac{\mathbf{d}_i \mathbf{y}_i^T}{\mathbf{y}_i^T \mathbf{d}_i} \right) V_i \left(I_n - \frac{\mathbf{y}_i \mathbf{d}_i^T}{\mathbf{y}_i^T \mathbf{d}_i} \right) + \frac{\mathbf{d}_i \mathbf{d}_i^T}{\mathbf{y}_i^T \mathbf{d}_i}$$

のベクトル部分 $\mathbf{d}_i$, $\mathbf{y}_i$ を行列 D , Y で置き換えて行列化した式

$$\tilde{V} = (I_n - D(Y^T D)^{-1} Y^T) \bar{V} (I_n - Y(D^T Y)^{-1} D^T) + D(D^T Y)^{-1} D^T$$

において, (16) 式を用いて逆行列部分を簡約化した式である. なお, 関数 f が 2 次関数の場合には, (18) 式が成立するように行列 D を生成すれば, (17) 式を満たすような初期行列 $\bar{V}$ に対し, (10) 式より

$$\bar{V}[I_n - YD^T/\varepsilon^2] = O$$

が成立することが示され, 更新式 (15) の第 1 項はゼロとなる. したがって f が 2 次関数の場合には, 初期試行行列 $\bar{V}$ の選択と無関係に, 第 2 項 DD^T/ε^2 により行列 $\tilde{V}$ がきまることがわかる. 一方, f が 2 次関数でない一般の非線形関数の場合にはこの限りではなく, 更新式の第 1 項は非ゼロ行列であることを注意しておく.

ところで, (18) 式を満たす行列 D の生成は

$$\mathbf{d}_i^T C \mathbf{d}_j = \begin{cases} 0 & (i \neq j) \\ \varepsilon^2 & (i = j) \end{cases} \quad (19)$$

を満たす方向ベクトル $\mathbf{d}_i$, $i=1, 2, \dots, n-l$ を生成することと等価であるから, つぎのようにグラムシュミット法の漸化式を適用し, さらに正規化を行えばよい. すなわち, 漸化式

$$\delta_1 = \mathbf{d}_1 \quad (20a)$$

$$\delta_i = \mathbf{d}_i - \sum_{j=1}^{i-1} \gamma_{ij}^* \delta_j, \quad i=2, 3, \dots, n-l \quad (20b)$$

$$\gamma_{ij}^* = \frac{\mathbf{d}_i^T C \delta_j}{\delta_j^T C \delta_j} \quad (20c)$$

により, $\delta_1, \delta_2, \dots, \delta_{n-l}$ を生成し,

$$\mathbf{d}_i := \frac{\varepsilon \delta_i}{(\delta_i^T C \delta_i)^{1/2}}, \quad i=1, 2, \dots, n-l \quad (21)$$

とすればよい.

2.3 一般的な計算手順

2.1 節の基本手順による新試行点は, 一般の非線形目的関数に対しては最適点である保証がない. そこで, 基本手順で得られる新点を新たな基点とし, この基本手順を一つのサイクルとして繰り返す. この場合, 初期試行行列 $\bar{V}$ は計算手順開始時に設定した行列に設定しなおしてもよいが (文献 2), 3) と同様, $\bar{V}$ を更新して得られた行列 $\tilde{V}$ が, (6) 式の行列のより良い近似行列になっていることを考慮して, この新近似行列をつぎのサイクルの初期試行行列とするのが得策である. なお, 一般非線形関数では 2 次の係数行列に相当するものが定義されないため, (20c), (21) 式を C が含まない形に変形した式をこれらの代わりに使用することは, 文献 2), 3) と同様である.

以上のような, 一般非線形目的関数に対する線形多様体上の並列準ニュートン法の手法をまとめるとつぎのようになる.

〈一般手順〉

step 1: 制約式 $Ax^1=b$ を満たす初期点 x^1 と $AV^1=O$ を満たす半正定値対称な初期近似行列 V^1 , 終了判定正定数 η を与え $t=1$ とする.

step 2: $Ad_j=0$ を満たす互いに独立な $n-l$ 個の方向ベクトル d_j ($j=1, 2, \dots, n-l$) を用意する.

step 3:

$$\left| \frac{\nabla f(x^t) \cdot d_j}{\|\nabla f(x^t)\| \|d_j\|} \right| < \eta, \quad j=1, 2, \dots, n-l$$

ならば, x^t を最適解として計算を終了し, そうでなければ次ステップへいく.

step 4: 漸化式

$$\delta_1^t = d_1$$

$$\delta_i^t = d_i - \sum_{j=1}^{i-1} \gamma_{ij}^t \delta_j^t, \quad i=2, 3, \dots, n-l$$

$$\gamma_{ij}^t = \frac{d_j^T (\nabla f(x^t + \delta_i^t) - \nabla f(x^t))}{\delta_j^T (\nabla f(x^t + \delta_i^t) - \nabla f(x^t))},$$

$$i=2, 3, \dots, n-l, \quad j=1, 2, \dots, i-1$$

により, $\delta_1^t, \delta_2^t, \dots, \delta_{n-l}^t$ を生成し,

$$d_i^t := \frac{\varepsilon \delta_i^t}{(\delta_i^T (\nabla f(x^t + \delta_i^t) - \nabla f(x^t)))^{1/2}},$$

$$i=1, 2, \dots, n-l$$

とし, これらを用いて行列 $D^t = [d_1^t, d_2^t, \dots, d_{n-l}^t]$ をつくる.

step 5: 基点 x^t からの $n-l$ 個の摂動点

$$x_i^t = x^t + d_i^t, \quad i=1, 2, \dots, n-l$$

と, 基点 x^t との勾配差

$$y_i^t = \nabla f(x_i^t) - \nabla f(x^t), \quad i=1, 2, \dots, n-l$$

を計算し, 行列

$$Y^t = [y_1^t, y_2^t, \dots, y_{n-l}^t]$$

を構成する.

step 6: 更新公式

$$V^{t+1} = \left(I_n - \frac{D^t Y^{tT}}{\varepsilon^2} \right) V^t \left(I_n - \frac{Y^t D^{tT}}{\varepsilon^2} \right) + \frac{D^t D^{tT}}{\varepsilon^2}$$

により, D^t, Y^t を用いて V^t を修正し, V^{t+1} を生成する.

step 7: V^{t+1} を用いて, 移動ベクトル

$$s^t = -V^{t+1} \nabla f(x^t)^T$$

を計算する.

step 8: 直線探索問題

$$\min_{\alpha} \{f(x^t + \alpha s^t) | \alpha > 0\}$$

を解いて, その解を α^t とし $x^{t+1} = x^t + \alpha^t s^t$ を求め, $t := t+1$ とし step 3 へ戻る.

ところで, この計算手順を実行するには, 制約式を満たす初期点と $AV^1=O$ を満たす対称な初期行列 V^1 や制約式上を張る $n-l$ 個の独立な方向ベクトルを用意する必要があった. しかし, これらの作業は制約式

の係数行列 A によっては容易ではない. このような場合には, 元の問題を

$$\min_x f(x) + \sum_{i=1}^l w_i z_i^2 \quad (22a)$$

$$\text{subj. to } Ax + I_l z = b \quad (22b)$$

のように変換する. ただし, $z \in R^l$ は人工変数, $w_i, i=1, 2, \dots, l$ は十分大きな正の係数, I_l は $l \times l$ 単位行列である. この問題を解いて得られた解において, 人工変数 z の値が 0 になっていれば, このときの x の解が元の問題の解となる. 一方, 0 にならない成分が z にあれば, 制約式 (1b) を満たす許容解が存在しない

この変換問題では, 任意に与えた x に対して (22b) 式を満たす z が従属的に求まる. こうして, (22b) 式を満たす $n+1$ 個の点 $(x_i^t, z_i^t), i=0, 1, \dots, n$ を選び, そのうちの一つを基点として, 方向ベクトルを

$$d_i^t = (x_i^t, z_i^t) - (x_0^t, z_0^t), \quad i=1, 2, \dots, n$$

のようにとれば, (x, z) 空間で制約式 (22b) が形成する線形多様体を張る互いに独立な n 個の方向ベクトルが得られる. さらにこれらの方向ベクトルで $(n+l) \times n$ 行列 $D (= [d_1, d_2, \dots, d_n])$ をつくり, 初期近似行列として $(n+l) \times (n+l)$ の行列

$$V^1 = DD^T$$

をとると, この初期行列 V^1 は, $[A, I_l]V^1=O$ を満たすことは明らかである. このような作業を踏まえて, 変換問題 (22) に対して, (x, z) 空間上で前述の計算手順を実行すればよい.

3. 数値計算例

本論文で提案した線形多様体上の並列準ニュートン法の有効性や計算効率をいくつかの例で確認した結果を示す. まずテスト用につぎの三つの例題を考えた.

例題 1 (2 次目的関数の例)

$$\min_x f(x) = \sum_{i=1}^{n/2} \sum_{j=0}^{n/2-i} (x_{i+2j} + x_{i+2j+1})^2$$

$$\text{subj. to } Ax = b$$

$$\text{where } b^T = (n+2 \ n+2 \ \dots \ n+2)$$

$$n; \text{ multiples of } 4$$

最適解は $x^{opt} = (2 \ 2 \ \dots \ 2)$

例題 2

$$\min_x f(x) = \sum_{i=1}^{n/2} \sum_{j=0}^{n/2-i} (x_{i+2j} + x_{i+2j+1})^4$$

$$\text{subj. to } Ax = b$$

$$\text{where } b^T = (n+2 \ n+2 \ \dots \ n+2)$$

$$n; \text{ multiples of } 4$$

最適解は $x^{opt} = (2 \ 2 \ \dots \ 2)$

例題 3

$$\begin{aligned} \min_{\mathbf{x}} f(\mathbf{x}) &= \sum_{i=1}^{n/2} \sum_{j=2i-1}^{n/2+i-1} \left(x_j - \frac{1}{2} x_{j+1}^2 \right)^2 \\ \text{subj to } \mathbf{Ax} &= \mathbf{b} \\ \text{where } \mathbf{b}^T &= \left(-\frac{n}{2}-1 \quad -\frac{n}{2}-1 \quad \cdots \quad -\frac{n}{2}-1 \right) \end{aligned}$$

n ; multiples of 4

最適解は $\mathbf{x}^0 = (-1 \ -1 \ \cdots \ -1)$

いずれの例題においても、制約式の係数行列 A は $\frac{n}{2} \times n$ 行列、すなわち $l = \frac{n}{2}$ とし、行列 A を

$$\begin{array}{c} \left(\frac{n}{2} + 1 \right) \text{ 列} \\ \left\{ \begin{array}{cccccccccccccccc} 1 & 1 & \cdots & \cdots & 1 & 1 & 0 & 0 & \cdots & \cdots & 0 & 0 \\ 0 & 1 & 1 & \cdots & \cdots & 1 & 1 & 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & 1 & \cdots & \cdots & 1 & 1 & 1 & 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & 1 & 1 & 1 & 1 & \cdots & \cdots & 1 & 1 & 0 & 0 \\ 0 & 0 & \cdots & 0 & 0 & 1 & 1 & 1 & 1 & \cdots & \cdots & 1 & 1 & 0 \\ 0 & 0 & \cdots & 0 & 0 & 1 & 1 & 1 & 1 & \cdots & \cdots & 1 & 1 \end{array} \right\} \\ \left(\frac{n}{2} - 1 \right) \text{ 列} \\ n \text{ 列} \end{array}$$

と設定した。また制約式 $\mathbf{Ax} = \mathbf{b}$ が形成する線形多様体を張る互いに独立な $\frac{n}{2}$ 個の方向ベクトル $\mathbf{d}_1, \mathbf{d}_2, \dots, \mathbf{d}_{n/2}$ からなる $n \times \frac{n}{2}$ 行列 $D (= [\mathbf{d}_1, \mathbf{d}_2, \dots, \mathbf{d}_{n/2}])$ として

$$\begin{array}{c} \left(\frac{n}{2} \right) \text{ 列} \\ \left\{ \begin{array}{cccccccc} 1 & 1 & 1 & 1 & \cdots & 1 & 1 \\ -1 & 0 & 0 & \cdots & \cdots & 0 \\ 0 & -1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & \cdots & 0 & -1 & 0 \\ 0 & 0 & \cdots & \cdots & 0 & 0 & -1 \end{array} \right\} \left\{ \begin{array}{l} \left(\frac{n}{2} \times \frac{n}{2} \right) \text{ 単位} \\ \text{行列} \times (-1) \end{array} \right. \\ \leftarrow \text{第} \left(\frac{n}{2} + 1 \right) \text{ 行} \\ \left\{ \begin{array}{cccccccc} 1 & 1 & 1 & 1 & \cdots & 1 & 1 \\ -1 & 0 & 0 & \cdots & \cdots & 0 \\ 0 & -1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & \cdots & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & \cdots & \cdots & 0 & -1 & 0 & 0 \end{array} \right\} \left\{ \begin{array}{l} \vdots \\ \left(\frac{n}{2} \times \frac{n}{2} \right) \text{ 単位行列} \times (-1) \text{ の第} \left(\frac{n}{2} - 2 \right) \text{ 行まで} \end{array} \right. \end{array}$$

を選んだ。また、摂動ベクトルの大きさは $\varepsilon = 0.01$ ，終了判定パラメータは $\eta = 0.001$ とした。なお初期試行点に関しては、つぎのような点を選定した。ただし、 $\mathbf{b}$ は制約式右辺のベクトル $\mathbf{b}$ の成分値である。

初期試行点 ①

$$x_i = \begin{cases} i & , i = 1, 2, \dots, \frac{n}{2} \\ b - \sum_{j=1}^{n/2} x_j, & i = \frac{n}{2} + 1 \\ x_{i-n/2-1} & , i = \frac{n}{2} + 2, \frac{n}{2} + 3, \dots, n \end{cases}$$

初期試行点 ②

$$x_i = \begin{cases} (-1)^i \times (10i), & i = 1, 2, \dots, \frac{n}{2} \\ b - \sum_{j=1}^{n/2} x_j, & i = \frac{n}{2} + 1 \\ x_{i-n/2-1} & , i = \frac{n}{2} + 2, \frac{n}{2} + 3, \dots, n \end{cases}$$

計算手順中、並列処理が実行可能な部分も、単一プロセッサによる逐次処理した場合に、通常の勾配射影法および準ニュートン射影法との比較を経過時間（計算開始から計算終了までの時間）とイテレーション回数で比較したのが Table 1~3 である。計算は EPSON PC-286 VG の倍精度型計算による。なお、勾配射影法 (GP と略記) は (3) 式において、行列 C のかわりに単位行列を用いて試行点を更新する手法であり、またニュートン射影法 (NP と略記) は (5) 式をそのまま試行点の摂動方向とする手法で、また準ニュートン射影法 (QN と略記) は (7)~(9) 式を用いる手法である。また本手法は PQN と略記した。

例題 1 の目的関数は 2 次ではあるが、このヘッセ行列 C は正則ではなく、(3) 式による方向ベクトル $\mathbf{s}^k$ の計算は不能である。ただし、等式制約 $\mathbf{Ax} = \mathbf{b}$ の上では大域的な最適点をもつ。例題 2 の目的関数はこの

Table 1 Computing time (iterations) of Example 1

Initial point	Computational method	Problem size: n				
		4	8	12	16	20
Case ①	PQN	1 s (1)	4 s (1)	8 s (1)	13 s (1)	20 s (1)
	QN	1 s (2)	5 s (2)	26 s (5)	66 s (7)	137 s (9)
	NP	The matrix C in eq. (3) is singular				
	GP	5 s (4)	7 s (2)	188 s (29)	547 s (52)	1270 s (81)
Case ②	PQN	1 s (1)	5 s (1)	9 s (1)	15 s (1)	23 s (1)
	QN	2 s (2)	10 s (4)	31 s (6)	74 s (8)	150 s (10)
	NP	The matrix C in eq. (3) is singular				
	GP	6 s (5)	76 s (23)	422 s (66)	1152 s (110)	2655 s (170)

Table 2 Computing time (iterations) of Example 2

Initial point	Computational method	Problem size: n				
		4	8	12	16	20
Case ①	PQN	3 s (2)	15 s (4)	39 s (5)	67 s (5)	122 s (6)
	QN	2 s (6)	13 s (6)	71 s (15)	186 s (22)	302 s (22)
	NP	The matrix C in eq. (3) is singular				
	GP	3 s (3)	31 s (10)	229 s (38)	700 s (71)	1449 s (113)
Case ②	PQN	5 s (4)	21 s (5)	43 s (5)	60 s (4)	92 s (4)
	QN	3 s (4)	27 s (13)	122 s (27)	289 s (35)	756 s (51)
	NP	The matrix C in eq. (3) is singular				
	GP	6 s (5)	139 s (46)	700 s (118)	1949 s (200)	4081 s (281)

Table 3 Computing time (iterations) of Example 3

Initial point	Computational method	Problem size: n				
		4	8	12	16	20
Case ①	PQN	2 s (4)	7 s (6)	15 s (8)	23 s (8)	38 s (9)
	QN	1 s (4)	10 s (9)	31 s (13)	56 s (12)	146 s (18)
	NP	2 s (3)	× ×	32 s (9)	60 s (9)	164 s (14)
	GP	3 s (8)	16 s (18)	30 s (24)	48 s (28)	64 s (30)
Case ②	PQN	2 s (4)	6 s (5)	9 s (5)	14 s (5)	21 s (5)
	QN	1 s (4)	10 s (10)	35 s (15)	92 s (20)	188 s (24)
	NP	3 s (4)	8 s (5)	18 s (5)	34 s (5)	59 s (5)
	GP	7 s (16)	26 s (32)	30 s (24)	86 s (52)	189 s (93)

The symbol × indicates that computation of the inverse matrix in equation (5) becomes impracticable because of the singularity

例題1の関数を4次化したものである。また2通りの初期試行点のうち、①は最適点の比較的近くに、また②は最適点より離れた遠方に選んである。本手法を準ニュートン射影法と比較してみると、問題の規模が20変数の場合、計算時間で約2.5倍から9倍の効率性を示し、対勾配射影法では、例題1、2で約12倍から約115倍の効率性を、例題3で約1.7倍から約9倍の効率性を示している。どの例題でも最適点より離れた地点に初期試行点をとるほど、他のいずれの手法と

比較しても計算の効率化が顕著になることを示している。さらに、目的関数の2次の係数行列 C が正則な例題3においては、ニュートン射影法と比較してみても、20変数の場合には、計算時間の点で約2.8倍から約4倍程度本手法のほうが優れている。

4. おわりに

制約条件のない最適化問題に対する並列準ニュートン法の自然な拡張として、線形等式制約条件つき最適化問題に対する並列計算用の最適化手法を提案し、その計算の効率性をいくつかの例題に対して確認した。その結果は、たとえ並列計算機による並列処理を実行しなくても、勾配射影法との比較では無論のこと、問題の規模が比較的大きければ準ニュートン射影法と比較しても、かなりの計算時間の短縮化が確認された。このことから、複数のプロセッサを内蔵した並列計算機の使用により、step 4の漸化式やstep 5の振動作業、step 6の行列計算などを実際に並列計算することにより、一層の計算効率の向上が期待される。

ところで、線形等式制約条件つき最適化問題は、線形多様体上の最適化問題であり、いわば通常の直線探索問題を多次元化した問題でもある。このようなタイプの問題は、線形ノルム空間における射影問題そのものや、内部経路法^{7),8)}などの一つの手法のアルゴリズムを構成する部分的問題などに見られ、これらを手間をかけずに解くためにも、本手法の適用が期待される。なお、不等式制約条件つき問題への拡張が、今後に残された課題であることを付記しておく。

参考文献

- 1) T. A. Straeter: A Parallel Variable Metric Optimization Algorithm, NASA Technical Note, D-7329 (1973)
- 2) 山田, 相吉: 群共役性を用いた並列準ニュートン法, 計測自動制御学会論文集, **25**-4, 482/489 (1989)
- 3) 相吉, 杉内: 非線形最適化の並列計算法, 計測と制御, **29**-12, 1070/1076 (1990)
- 4) 相吉, 山田: 等式制約条件つき最適化問題に対する並列準ニュートン法, 計測自動制御学会論文集, **25**-12, 1347/1354 (1989)
- 5) 相吉, 志水: 数値計画法演習, 朝倉書店 (1985)
- 6) D. Goldfarb: Extension of Davidon's Variable Metric Method to Maximization under Linear Inequality and Equality Constraints, SIAM J. Applied Mathematics, **17**-4, 739/764 (1969)
- 7) P. E. Gill, W. Murray, M. Saunders, J. Tomlin and M. H. Wright: On Projected Newton Barrier Method for Linear Programming and an Equivalence to Karmarkar's Projective Method, Mathematical Programming, **36**-2, 183/209 (1986)
- 8) 今野, 山下: 線形計画法, 第14章, 日科技連 (1987)