数理解析研究所講究録 1349 巻 2004 年 177-186

ペナルティ関数を用いない信頼領域SQP法の大域的収束性について

東京理科大学理学部数理情報科学科　矢部　博 (Hiroshi Yabe)
Tokyo University of Science

数理システム　山下　浩 (Hiroshi Yamashita)
Mathematical Systems, Inc.

1　はじめに

本論文では次の制約付き最小化問題に対する数値解法について考える.

$$\begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & g(x)=0, \qquad x \geq 0, \end{array} \tag{1}$$

ただし $f:\mathbf{R}^n \to \mathbf{R}$, $g_i:\mathbf{R}^n \to \mathbf{R}, i=1,\ldots,m$ は滑らかな関数である. 上記の問題を解くための数値解法には乗数法, 逐次2次計画 (SQP) 法, 信頼領域 SQP 法, 内点法などがあるが, それらの特徴は目的関数の最小化と制約条件を満たすことを両立するためにペナルティ関数を利用することである. すなわち, ペナルティ関数をメリット関数に用いた直線探索法や信頼領域法に基づいて大域的収束性を示している. しかしながら, 実用的なペナルティパラメータをどのように選ぶか, Maratos 効果が生じないようにするためにはどうするかなどの課題がある. 他方, ペナルティ関数を用いることなく, 目的関数の最小化と実行可能性の実現を2目的計画として別々に実施する数値解法が注目されている. こうしたアプローチは Yamashita [5] の研究にみるように今までにもあったが, 当時はペナルティ関数をメリット関数に用いる解法が主流であり, 山下の研究は時期尚早の感があった. なお, 山下の研究は直線探索法に基づくものであった. ところが近年, Fletcher and Leyffer [2] が提案したフィルター法が脚光を浴びるにあたって, ペナルティ関数をメリット関数として用いない最適化法の研究が注目され始めた. Fletcher らは上記の論文において直線探索法に基づいたフィルター法のアルゴリズムを提案し, 文献 [1] において信頼領域 SQP 法に関連したフィルター法の大域的収束性を証明した. さらに彼らのアプローチに触発されて, Ulbrich, Ulbrich and Vicente [4] は内点法へのフィルター法の適用を試み, また Ulbrich and Ulbrich [3] は2目的計画法を意識した非単調アルゴリズムを提案した.

本論文では Yamashita [5] が扱ったアイデアに基づいて, ペナルティ関数を用いない新しい数値解法を提案する. この方法はフィルター法とは異なるものである. ここで提案するアルゴリズ

ムの基本的な考えは信頼領域SQP法に基づくもので，制約充足度改善アルゴリズムと目的関数最小化アルゴリズムを交互に実行するものである．そして本稿の後半で，提案する解法の大域的収束性について議論する．

なお，本論文を通じて $\|\cdot\|$ は l_2 ベクトルノルムもしくはそれから誘導される行列ノルムを表す．

2 基本的な記号

問題 (1) に対するラグランジュ関数を

$$L(w) = f(x) - y^t g(x) - z^t x \tag{2}$$

とする．ただし $w = (x, y, z)^t$ とし，$y \in \mathbf{R}^m$ と $z \in \mathbf{R}^n$ はそれぞれ等式制約，不等式制約に対応するラグランジュ乗数である．このとき Karush-Kuhn-Tucker (KKT) 条件は次式で与えられる．

$$r(w) = \begin{pmatrix} \nabla_x L(w) \\ g(x) \\ XZe \end{pmatrix} = 0, \quad x \geq 0,\ z \geq 0, \tag{3}$$

ただし

$$\begin{aligned} \nabla_x L(w) &= \nabla f(x) - A(x)^t y - z, \\ A(x) &= \begin{pmatrix} \nabla g_1(x)^t \\ \vdots \\ \nabla g_m(x)^t \end{pmatrix}, \\ X &= \mathrm{diag}\,(x_1, \cdots, x_n), \\ Z &= \mathrm{diag}\,(z_1, \cdots, z_n), \\ e &= (1, \cdots, 1)^t \in \mathbf{R}^n. \end{aligned}$$

以下の節では，次のように目的関数の1次近似 $f_l(x;s) : \mathbf{R}^n \to \mathbf{R}$ および2次近似 $f_q(x;s) : \mathbf{R}^n \to \mathbf{R}$ を定義する．

$$\begin{aligned} f_l(x;s) &= f(x) + \nabla f(x)^t s \\ f_q(x;s) &= f(x) + \nabla f(x)^t s + \frac{1}{2} s^t H s, \end{aligned}$$

ただし，$s \in \mathbf{R}^n$ はステップであり，$H \in \mathbf{R}^{n \times n}$ は適当な対称行列である．具体的な形は後で定義する．また，各関数の差を

$$\begin{aligned} \Delta f_l(x;s) &\equiv f_l(x;s) - f(x) = \nabla f(x)^t s, \\ \Delta f_q(x;s) &\equiv f_q(x;s) - f(x) = \nabla f(x)^t s + \frac{1}{2} s^t H s, \\ \Delta f(x;s) &\equiv f(x+s) - f(x) \end{aligned}$$

と定義する．上記の式 $\Delta f_l(x;s)$，$\Delta f_q(x;s)$ は後述するアルゴリズムで解かれるLP部分問題やQP部分問題の目的関数として使われる．さらに，$\Delta f(x;s)$ および $\Delta f_q(x;s)$ の値は信頼領域半径の調整に使われる．

3　提案するアルゴリズム

ここで提案するアルゴリズムはKKT条件を満足する点を見つけるものである．この解法は制約充足度改善アルゴリズムと目的関数最小化アルゴリズムの2種類からなるもので，全過程を通じて主変数と双対変数の非負条件が満たされるように実行される．制約充足度改善アルゴリズムは与えられた許容範囲内で等式条件を近似的に満たすような点を見つける解法であり，一方，目的関数最小化アルゴリズムは上記の許容範囲内で等式条件を近似的に満たしながら目的関数値を減少させることを目指した解法である．以下では，KKT条件に対する残差として

$$\|r(w)\|_* = \max(\|\nabla_x L(w)\|,\ \|g(x)\|,\ \|XZe\|)$$

という記号を導入する．

まず主アルゴリズムについて述べる．

[提案する数値解法の主アルゴリズム]

(Step 0) 初期設定:

w_0 (ただし $x_0 \geq 0$), $\varepsilon > 0, \tau \in (0,1)$ を与える．$k=0$ とおく．

(Step 1) $\delta_k = \tau \|r(w_k)\|_*$ とおく．

(Step 2) 外部反復の点 x_k を初期点として制約充足度改善アルゴリズム（内部反復）を実行し，

$$\|g(x_{k+\frac{1}{2}})\| < \delta_k, \quad x_{k+\frac{1}{2}} \geq 0, \quad z_{k+\frac{1}{2}} \geq 0$$

を満たす点

$$w_{k+\frac{1}{2}} = (x_{k+\frac{1}{2}}, y_{k+\frac{1}{2}}, z_{k+\frac{1}{2}})^t$$

を見つける．もし $w_{k+\frac{1}{2}}$ が $\|r(w_{k+\frac{1}{2}})\|_* \leq \delta_k$ を満足するならば，$w_{k+1} = w_{k+\frac{1}{2}}$ とおいてStep 4へ行く．

(Step 3) 外部反復のSpep 2で得られた点 $x_{k+\frac{1}{2}} \geq 0$ を初期点として目的関数最小化アルゴリズム（内部反復）を実行して，

$$\|r(w_{k+1})\|_* \leq \delta_k, \quad x_{k+1} \geq 0, \quad z_{k+1} \geq 0$$

を満たす点

$$w_{k+1}=(x_{k+1},y_{k+1},z_{k+1})^t$$

を見つける.

(Step 4) 収束判定:

もし $\|r(w_{k+1})\|_* \le \varepsilon$ ならば終了する．さもなければ $k:=k+1$ とおいて Step 1 へ行く．□

次に制約充足度改善アルゴリズムを記述する．この部分は従来の信頼領域 SQP 法に対応するもので，各反復で QP 部分問題が 1 回だけ解かれる．このアルゴリズムを実行することによって，あらかじめ与えられた許容範囲 δ 内で等式制約条件が近似的に満たされることに注意されたい．

[制約充足度改善アルゴリズム]

(Step 0) 初期設定:

w_0 (ただし $x_0 \ge 0$), $\delta>0, \Delta_0>0, \varepsilon_0 \in (0,1)$，$\beta \in (0,1)$ を与える．$k=0$ とおく．

(Step 1) 行列 G_k を計算する．ここで G_k はヘッセ行列 $\nabla_x^2 L(w_k)$，もしくはその近似行列である．

(Step 2) QP 部分問題:

次の QP 部分問題を解いてステップ s_k および対応するラグランジュ乗数 $(y_{k+1},z_{k+1})^t$ を計算する．

$$\begin{array}{ll}\text{minimize} & \frac{1}{2}s^tG_ks+\nabla f(x_k)^ts\\ \text{subject to} & g(x_k)+A(x_k)s=0, \quad x_k+s\ge 0, \quad \|s\|_\infty \le \Delta_k.\end{array}$$

ただし，信頼領域半径 Δ_k は QP 部分問題の線形制約条件が実行可能になるように調整されるもので，必要に応じて大きくしたり小さくしたりするが限りなくゼロに近づけるようなことはしない．

(Step 3) 直線探索:

次の不等式を満たすような最小の非負整数 l_k を見つける．

$$\|g(x_k+\beta^{l_k}s_k)\| < \max\{\delta,(1-\varepsilon_0\beta^{l_k})\|g(x_k)\|\}$$

そして $\alpha_k=\beta^{l_k}$ とおく．

(Step 4) 主変数を $x_{k+1}=x_k+\alpha_ks_k$ と更新して，$w_{k+1}=(x_{k+1},y_{k+1},z_{k+1})^t$ とおく．

(Step 5) 収束判定:

もし $\|g(x_{k+1})\| < \delta$ ならば終了する.

(Step 6) $k := k+1$ とおいて Step 1 へ行く. □

アルゴリズムの Step 2 で, 信頼領域条件 $\|s\|_\infty \le \Delta_k$ は条件 $-\Delta_k e \le s \le \Delta_k e$ を意味する. ただし $e = (1, ..., 1)^t$ である. (以下で述べるアルゴリズムにおいても, 信頼領域条件は同様の意味をもつことにする)

この節を終えるにあたって, 目的関数最小化アルゴリズムを述べる. このアルゴリズムの中では, 各反復で 1 つの LP 部分問題と 2 つの QP 部分問題が解かれる. 前者はラグランジュ乗数 y, z の近似値を推定するためのものであり, 後者は主変数 x の近似解を求めるためのものである. このアルゴリズムは, あらかじめ与えられた許容範囲 δ 内で KKT 条件を近似的に満足する点を求めるまで続行される.

[目的関数最小化アルゴリズム]

(Step 0) 初期設定:

w_0 (ただし $x_0 \ge 0$), $\delta > 0, \Delta_0 > 0, \Delta_{T_0} > 0$, および $\beta \in (0,1)$ を与える. ただし $\Delta_{T_0} \le \Delta_0$ であり, x_0 は $x_0 \ge 0$ かつ $\|g(x_0)\| < \delta$ を満たす初期点である. $k = 0$ とおく.

(Step 1) 1 つの LP 部分問題を解く:

次の LP 部分問題を解いて, ステップ d_k と対応するラグランジュ乗数 $(y_{k+1}, z_{k+1})^t$ を求め, $w_k = (x_k, y_{k+1}, z_{k+1})^t$ とおく.

$(LP(x_k))$

$$\begin{array}{ll} \text{minimize} & \Delta f_l(x_k; d) = \nabla f(x_k)^t d \\ \text{subject to} & A(x_k)d = 0, \quad x_k + d \ge 0, \quad \|d\|_\infty \le 1. \end{array}$$

(Step 2) 収束判定:

もし w_k が $\|r(w_k)\| \le \delta$ を満たすならば終了する.

(Step 3) 2 つの QP 部分問題を解く:

対称行列 $H_k \in \mathbf{R}^{n\times n}$ を計算する. 次の 2 つの QP 部分問題を解いて, ステップ s_{T_k} と s_k をそれぞれ求める.

$(QP_T(x_k))$

$$\begin{array}{ll} \text{minimize} & \Delta f_q(x_k; s_T) = \dfrac{1}{2} s_T^t H_k s_T + \nabla f(x_k)^t s_T \\ \text{subject to} & A(x_k) s_T = 0, \quad x_k + s_T \geq 0, \quad \|s_T\|_\infty \leq \Delta_{T_k} \end{array}$$

$(QP(x_k))$

$$\begin{array}{ll} \text{minimize} & \Delta f_q(x_k; s) = \dfrac{1}{2} s^t H_k s + \nabla f(x_k)^t s \\ \text{subject to} & g(x_k) + A(x_k) s = 0, \quad x_k + s \geq 0, \quad \|s\|_\infty \leq \Delta_k, \end{array}$$

ただし，信頼領域半径 Δ_k は $\Delta_{T_k} \leq \Delta_k$ を満たし，かつ，部分問題 $(QP(x_k))$ の制約条件が実行可能になるように選ばれる．

(Step 4) $\bar{s}_k = \left(\min\left\{\dfrac{\|s_{T_k}\|_\infty}{\|s_k\|_\infty},\ 1\right\}\right) s_k$ とおいて，次の不等式が成り立つような最小の非負整数 l_k を求める．

$$\Delta f_q(x_k; (1-\beta^{l_k}) s_{T_k} + \beta^{l_k} \bar{s}_k) \leq \frac{1}{2} \Delta f_q(x_k; s_{T_k}) \tag{4}$$

$\rho_k = \beta^{l_k}$ として，$s_{\rho_k} = (1-\rho_k) s_{T_k} + \rho_k \bar{s}_k$ とおく．

(Step 5) 信頼領域半径の調整:

もし $\|g(x_k + s_{\rho_k})\| \geq \delta$ ならば，$\Delta_{T_{k+1}} = \dfrac{1}{2} \Delta_{T_k}$ とおく．さもなければ，次のようにおく．

$$\begin{cases} もし \Delta f(x_k; s_{\rho_k}) > \frac{1}{4} \Delta f_q(x_k; s_{\rho_k}) ならば, \Delta_{T_{k+1}} = \frac{1}{2} \Delta_{T_k} とおく \\ もし \Delta f(x_k; s_{\rho_k}) \leq \frac{3}{4} \Delta f_q(x_k; s_{\rho_k}) ならば, \Delta_{T_{k+1}} = 2\Delta_{T_k} とおく \\ さもなければ, \Delta_{T_{k+1}} = \Delta_{T_k} とおく. \end{cases}$$

(Step 6) もし $\Delta f(x_k; s_{\rho_k}) \leq 0$ かつ $\|g(x_k + s_{\rho_k})\| < \delta$ ならば, $x_{k+1} = x_k + s_{\rho_k}$ とおく．さもなければ $x_{k+1} = x_k$ とおく．

(Step 7) $k := k+1$ とおいて Step 1 へ行く． □

Step 3 において与えられた行列 H_k にはいくつかの形が考えられ，例えば $\nabla^2 f(x_k)$, $\nabla_x^2 L(w_k)$ あるいはその近似行列などがあげられる．ゼロベクトルが部分問題 $(LP(x_k))$, $(QP_T(x_k))$ の実行可能解になることから

$$\Delta f_l(x_k; d_k) \leq \Delta f_l(x_k; 0) = 0$$

かつ

$$\Delta f_q(x_k; s_{T_k}) \leq \Delta f_q(x_k; 0) = 0$$

が成り立つ．このアルゴリズムの基本的な考えは，部分問題 $(QP_T(x_k))$ を解いて目的関数の降下方向を生成し，一方，部分問題 $(QP(x_k))$ を解いて等式制約条件 $g(x)=0$ に対するニュートン方向を生成するものである．そしてこれら 2 つの探索方向を Step 4 で組み合わせるのである．これはちょうど，無制約最小化問題において最急降下方向とニュートン方向を組み合わせる考え方に対応していると解釈できる．なお

$$\|\bar{s}_k\| = \left(\min\left\{\frac{\|s_{T_k}\|_\infty}{\|s_k\|_\infty},\ 1\right\}\right)\|s_k\| \le \|s_{T_k}\|_\infty \le \Delta_{T_k}$$

なので，Step 4 において

$$\|s_{\rho_k}\| \le (1-\rho_k)\|s_{T_k}\| + \rho_k\|\bar{s}_k\| \le \Delta_{T_k}$$

が成り立つ．また，目的関数最小化アルゴリズムで生成される点列 $\{w_k\}$ が $x_k \ge 0,\ z_k \ge 0$ および $\|g(x_k)\| < \delta$ を満足することに注意されたい．．

4 大域的収束性

本節では，前節で提案した数値解法の大域的収束性に関する命題をいくつか紹介する．詳しい証明は Yamashita and Yabe [6] を参照されたい．大域的収束性を示すために以下の条件を仮定する．

仮定 G

(G1) 関数 f および $g_i, i=1,...,m$ は 2 回連続的微分可能である．

(G2) 任意の点 x_0 に対して，集合 $\{x\in\mathbf{R}^n\,|f(x)\le f(x_0)\}\cap\{x\in\mathbf{R}^n|\ \|g(x)\|\le\delta_0\}\cap\{x\in\mathbf{R}^n|\ x\ge 0\}$ はコンパクト集合である．ただし，集合 $\{x\in\mathbf{R}^n\,|f(x)\le f(x_0)\}$ は $x_0\in\mathbf{R}^n$ における目的関数の準位集合を表す．．

(G3) 行列 G_k，H_k は一様有界である．

(G4) 次式を満たす正の定数 $\bar{\Delta}$ が存在する．

$$0 < \Delta_{T_k} \le \Delta_k \le \bar{\Delta} \quad \text{for all } k$$

4.1 制約充足度改善アルゴリズムの大域的収束性

次の定理は制約充足度改善アルゴリズムの大域的収束性を示している．

Theorem 1 仮定 G が成り立つとする．$\{x_k\}$ を制約充足度改善アルゴリズムで生成される点列とする．このとき *Step 3* の直線探索手順は有限回の反復で終了し，さらに点列 $\{x_k\}$ の任意の集積点 x_∞ $(\in\mathbf{R}^n)$ は $g(x_\infty)=0$ を満足する．

4.2 目的関数最小化アルゴリズムの大域的収束性

本節では，目的関数最小化アルゴリズムの大域的収束性を示す．そのためにいくつかの補助定理を紹介する．以下では，$\Delta f_l(x_k; d_k)$ および $\Delta f_q(x_k; s_{T_k})$ という量が重要な役割を演ずる．まず，これらの量に関する性質を述べる．

Lemma 1 仮定 G が成り立つとする．d_k, y_{k+1}, z_{k+1} をそれぞれ部分問題 $LP(x_k)$ の解および対応するラグランジュ乗数とする．このとき以下のことが成り立つ．

(i) もしある k に対して $\Delta f_l(x_k; d_k) = 0$ ならば，与えられた $\delta > 0$ に対して点 $w_k = (x_k, y_{k+1}, z_{k+1})^t$ は

$$\|g(x_k)\| < \delta, \quad \nabla_x L(w_k) = 0, \quad X_k Z_{k+1} e = 0, \quad x_k \geq 0, \quad z_{k+1} \geq 0$$

を満たす．

(ii) もし部分列 $K \subset \{0, 1, 2, \ldots\}$ が存在して

$$\lim_{k\to\infty,\ k\in K} \Delta f_l(x_k; d_k) = 0$$

が成り立つならば，点列 $\{(x_k, y_{k+1}, z_{k+1})\}$ の任意の集積点 $w_\infty = (x_\infty, y_\infty, z_\infty)^t$ は

$$\|g(x_\infty)\| < \delta, \quad \nabla_x L(w_\infty) = 0, \quad X_\infty Z_\infty e = 0, \quad x_\infty \geq 0, \quad z_\infty \geq 0$$

を満足する．

(iii) ある k に対して $\Delta f_q(x_k; s_{T_k}) = 0$ となるならば，$\Delta f_l(x_k; d_k) = 0$ が成り立つ．

(iv) もし部分列 $K \subset \{0, 1, 2, \ldots\}$ が存在して

$$\lim_{k\to\infty,\ k\in K} \Delta f_q(x_k; s_{T_k}) = 0, \qquad \liminf_{k\to\infty,\ k\in K} \Delta_{T_k} > 0$$

ならば，$\displaystyle\lim_{k\to\infty,\ k\in K} \Delta f_l(x_k; d_k) = 0$ が成り立つ．

次の補助定理は $\Delta f_q(x_k; s_{T_k})$ の評価に関するものである．

Lemma 2 $x_k \geq 0$ かつ $\|g(x_k)\| < \delta$ を満たす x_k が与えられているとする．また，Δ_{T_k} は十分に小さくとられているとする．もし $\Delta f_l(x_k; d_k) < 0$ ならば，

$$|\Delta f_q(x_k; s_{T_k})| \geq c_1 \|s_{T_k}\|_\infty$$

となるような正の定数 c_1 が存在する．

次の補助定理はアルゴリズムの Step 4 と Step 5 が実現可能であることを示している．

Lemma 3 仮定 G が成り立つとき，アルゴリズムの *Step 4* では次式を満足する ρ_k がつねに見つかる.

$$\Delta f_q(x_k; s_{\rho_k}) \le \frac{1}{2}\Delta f_q(x_k; s_{T_k}).$$

さらに，十分に小さい Δ_{T_k} に対して

$$\|g(x_k + s_{\rho_k})\| < \delta$$

が成り立つ.

次の補助定理は後述の定理 2 の証明で使われる.

Lemma 4 仮定 G が成り立つとき，もし $\liminf_{k\to\infty} |\Delta f_l(x_k; d_k)| = c_2 > 0$ ならば

$$\liminf_{k\to\infty} \rho_k > 0$$

となる. さらに反復回数 k に無関係な正の数 $\bar{\Delta}_T$ が存在して，任意の $\Delta_{T_k} \in (0, \bar{\Delta}_T)$ に対してステップ s_{ρ_k} は

$$\|g(x_k + s_{\rho_k})\| < \delta$$

を満足する.

次の定理は大域的収束性を示している.

Theorem 2 仮定 G が成り立つとき，目的関数最小化アルゴリズムは有限回の反復回数で終了するか，もしくは与えられた $\delta > 0$ に対して

$$\|g(x_\infty)\| \le \delta, \quad \nabla_x L(w_\infty) = 0, \quad X_\infty Z_\infty e = 0, \quad x_\infty \ge 0, \quad z_\infty \ge 0 \tag{5}$$

を満足する集積点 $w_\infty = (x_\infty, y_\infty, x_\infty)^t$ が存在する.

4.3 主アルゴリズムの大域的収束性

上述した制約充足度改善アルゴリズムの大域的収束性と目的関数最小化アルゴリズムの大域的収束性を組合わせれば，次のように主アルゴリズムの大域的収束性が示せる.

Theorem 3 仮定 G が成り立つとき，主アルゴリズムは有限回の反復回数で終了するか，もしくは生成される点列 $\{w_k\}$ の任意の集積点は KKT 条件を満足する.

References

[1] R. Fletcher, N. I. M. Gould, S. Leyffer, Ph. L. Toint and A. Wächter, Global convergence of a trust-region SQP-filter algorithm for general nonlinear programming, *SIAM Journal on Optimization*, 13 (2002), pp. 635-659.

[2] R. Fletcher and S. Leyffer, Nonlinear programming without a penalty function, *Mathematical Programming*, 91 (2002), pp. 239-269.

[3] M. Ulbrich and S. Ulbrich, Non-monotone trust region methods for nonlinear equality constrained optimization without a penalty function, *Mathematical Programming*, 95 (2003), pp. 103-135.

[4] M. Ulbrich, S. Ulbrich and L.N. Vicente, A globally convergent primal-dual interior point filter method for nonconvex nonlinear programming, Technical Report TR00-12, Department of Computational and Applied Mathematics, Rice University, Houston, Texas, USA, 2000.

[5] H. Yamashita, *A globally convergent quasi-Newton method for equality constrained optimization that does not use a penalty function*, Technical Report, Mathematical Systems Inc., Tokyo, Japan, June 1979 (revised September 1982).

[6] H. Yamashita and H. Yabe, A globally and superlinearly convergent trust-region SQP method without penalty functions for nonlinearly constrained optimization, Technical Report, November 2003.