

非線形最適化アルゴリズムSL1QP法でのラグランジュ関数のヘッセ行列の近似について

山影, 裕紀 / YAMAKAGE, Yuki

(出版者 / Publisher)

法政大学大学院理工学・工学研究科

(雑誌名 / Journal or Publication Title)

法政大学大学院紀要. 理工学・工学研究科編

(巻 / Volume)

56

(開始ページ / Start Page)

1

(終了ページ / End Page)

4

(発行年 / Year)

2015-03-24

(URL)

<https://doi.org/10.15002/00011164>

非線形最適化アルゴリズム Sl_1QP 法での ラグランジュ関数のヘッセ行列の近似について

Hessian approximation of the Lagrange function
for nonlinear optimization algorithm Sl_1QP method.

山影 裕紀

Yuki YAMAKAGE

指導教員 堀端 康善

法政大学大学院理工学研究科電気電子工学専攻修士課程

This paper describes methods for the Sequential Quadratic Programing(SQP) method and Sl_1QP . SQP method is known as most effective method solution for solving the constrained nonlinear optimization problem.

On the other hand, Sl_1QP method was suggested by Fletcher. It is the method that adopted trust-region method in SQP method. It can be expected by Sl_1QP method that the performance of SQP improve more. In this paper we compare SQP method with Sl_1QP method which implemented BFGS method or SR1 method in numerical results.

key words: nonlinear optimization problem, SQP, Sl_1QP , Hessian approximation, BFGS, symmetric rank-one update

1 はじめに

SQP 法の中で信頼領域法を用いたものに Fletcher[1] による Sl_1QP 法が存在する。 Sl_1QP 法は信頼領域法を取り入れていることから、直線探索法を利用している SQP 法よりも安定して最適解に近づいていくことが期待される。本論文では、非線形最適化問題を解く Sl_1QP 法に BFGS 法と SR1 法の 2 つのヘッセ行列の近似アルゴリズムを実装し、数値実験により直線探索法を利用した SQP 法との比較、検証を行う。

2 Sl_1QP 法について

Sl_1QP 法は、非線形計画問題に対して有効な逐次 2 次計画法 (SQP 法) に信頼領域法を取り入れたものである。

・ Sl_1QP 法における 2 次計画部分問題

Sl_1QP 法では、以下のような 2 次計画問題を考える。

$$\begin{aligned} \text{目的関数: } q_\mu(\mathbf{p}) &= \nabla f(\mathbf{x})^T \mathbf{p}_k + \frac{1}{2} \mathbf{p}_k^T B_k \mathbf{p}_k \\ &+ \mu \sum_{i=1}^{m_e} |c_i(\mathbf{x}) + \nabla c_i(\mathbf{x}_k)^T \mathbf{p}| \\ &+ \mu \sum_{i=1}^{m_i} |\min\{0, c_i(\mathbf{x}) + \nabla c_i(\mathbf{x}_k)^T \mathbf{p}\}| \quad (1) \\ &\rightarrow \text{最小化} \end{aligned}$$

制約条件: $\|\mathbf{p}\|_\infty \leq \Delta_k$

ここで、 m_e は等式制約数、 m_i は不等式制約数である。

3 信頼領域半径の調整と方向 $\mathbf{p}_k$ の扱い

μ_k と $\mathbf{p}_k$ を求めたあと、

$$\begin{aligned} \phi_1(\mathbf{x}, \mu) &= f(\mathbf{x}) \\ &+ \mu \left[\sum_{i=1}^{m_e} |c_i(\mathbf{x})| + \sum_{i=1}^{m_i} |\min\{0, c_i(\mathbf{x})\}| \right] \quad (2) \end{aligned}$$

を利用し、

$$\rho_k = \frac{ared_k}{pred_k} = \frac{\phi_1(\mathbf{x}_k, \mu_k) - \phi_1(\mathbf{x}_k + \mathbf{p}_k, \mu_k)}{q_\mu(0) - q_\mu(\mathbf{p}_k)} \quad (3)$$

を用いて次ページのようなアルゴリズムで信頼領域半径を調整する。

信頼領域の更新アルゴリズム

```

Given  $\hat{\Delta} > 0, \Delta_0 \in (0, \hat{\Delta})$ , and  $\eta \in (0, \frac{1}{4})$  :
Evaluate  $\rho_k$ 
if  $\rho_k < \frac{1}{4}$ 
     $\Delta_{k+1} = \frac{3}{4}\Delta_k$ 
else
if  $\rho_k > \frac{3}{4}$  and  $\|\mathbf{p}_k\|_\infty = \Delta_k$ 
     $\Delta_{k+1} = \min(2\Delta_k, \hat{\Delta})$ 
else
     $\Delta_{k+1} = \Delta_k$ ;
if  $\rho_k > \eta$ 
     $\mathbf{x}_{k+1} = \mathbf{x}_k + \mathbf{p}_k$ 
else
     $\mathbf{x}_{k+1} = \mathbf{p}_k$ .

```

ここで, $\hat{\Delta}$ は信頼領域半径の最大値である.

4 ペナルティパラメータの更新アルゴリズム

ペナルティパラメータ μ の更新アルゴリズムを以下に示す. ここで, $m_k(\mathbf{p})$ は以下のように与えられる.

$$m_k(\mathbf{p}) = \sum_{i=1}^{m_e} |c_i(\mathbf{x}) + \nabla c_i(\mathbf{x}_k)^T \mathbf{p}| + \sum_{i=1}^{m_i} |\min\{0, c_i(\mathbf{x}) + \nabla c_i(\mathbf{x}_k)^T \mathbf{p}\}| \quad (4)$$

ペナルティパラメータ μ の更新アルゴリズム

```

Initail data:  $\mathbf{x}_k \mu_{k-1} > 0$  and parameters
 $\epsilon_1, \epsilon_2 \in (0, 1)$ .
Solve the subproblem (3) with
 $\mu = \mu_{k-1}$  to obtain  $\mathbf{p}(\mu_{k-1})$ ;
if  $m_k(\mathbf{p}(\mu_{k-1})) = 0$  Set  $\mu^+ \leftarrow \mu_{k-1}$ ;
else Compute  $\mathbf{p}_\infty$ ; ( $\mathbf{p}_\infty$  denotes the minimizer of
     $m_k(\mathbf{p})$  subject to the trust-region constraint)
if  $m_k(\mathbf{p}_\infty) = 0$ ;
    Find  $\mu^+ > \mu_{k-1}$  such that  $m_k(\mathbf{p}(\mu^+)) = 0$ ;
else Find  $\mu^+ \geq \mu_{k-1}$  such that
     $m_k(0) - m_k(\mathbf{p}(\mu^+)) \geq \epsilon_1[m_k(0) - m_k(\mathbf{p}_\infty)]$ ;
end(if)
end(if)
Increase  $\mu^+$  if necessary to satisfy
 $q_{\mu^+}(0) - q_{\mu^+}(\mathbf{p}(\mu^+)) \geq \epsilon_2 \mu^+[m_k(0) - m_k(\mathbf{p}(\mu^+))]$ ;
Set  $\mu_k \leftarrow \mu^+$  and  $\mathbf{p}_k \leftarrow \mathbf{p}(\mu^+)$ .

```

5 $S\ell_1$ QP 法へのSIMPLEX 法の組み込み

ペナルティパラメータの更新アルゴリズムの $\mathbf{p}_\infty$ 部分において, $\mathbf{p}_\infty$ の計算時に μ を ∞ にすることは, 目的関数を $m_k(\mathbf{p})$ で置き換えることと等価である. よって, この問題に対して線形計画問題の基本的な解法である単体法 (SIMPLEX 法) を適用する

6 SIMPLEX 法について

SIMPLEX 法は線形計画法の中では, もっとも有名なものの 1 つである. 線形計画問題には様々な形の問題があるが, 特に次のような線形等式制約及び, 非負制約のもとで線形目的関数を最小, あるいは最大にする問題を線形計画問題の標準形という.

目的関数: $\mathbf{c}^T \mathbf{x} \rightarrow \text{最小}$

制約条件: $A\mathbf{x} = \mathbf{b}$

$\mathbf{x} \geq 0$

(5)

ここで, 制約条件の中に不等式制約を含む場合, スラック変数や剰余変数を加えることで標準形に変形することが出来る. また, 非負制約 $\mathbf{x} \geq 0$ となっているが, $\mathbf{x}$ に制限がある場合もあり, 今回使用した岩波 SIMPLEX 法プログラム [2] では上限, 下限を設けている. 制約条件式の係数行列 A とそれに対応する変数 $\mathbf{x}$ を基底行列 B , 基底変数 $\mathbf{x}_B$ と非基底行列 N 非基底変数 $\mathbf{x}_N$ に区分けする. B を正則になるように選択すれば線形計画問題の解, つまり基底解が得られ, 基底解が境界制約条件を満たしていれば, それは実行可能基底解である.

線形計画問題において, 最適解が存在するなら, 実行可能基底解のなかに存在する. この実行可能基底解から最適解を見つけるのが SIMPLEX 法である.

7 ラグランジュ関数のヘッセ行列について

次のような非線形最適化問題を取り扱う.

目的関数: $f(\mathbf{x}) \rightarrow \text{最小化}$

制約条件: $c_i(\mathbf{x}) = 0 \quad (i = 1, \dots, m_e)$ (6)

$c_i(\mathbf{x}) \geq 0 \quad (i = m_e + 1, \dots, m)$

今, 簡単のために, 次の等式制約のみを含む問題を考える.

目的関数: $f(\mathbf{x}) \rightarrow \text{最小化}$

制約条件: $\mathbf{c}(\mathbf{x}) = 0$

(7)

ただし、 $\mathbf{c}(\mathbf{x}) = (c_1(\mathbf{x}), c_2(\mathbf{x}), \dots, c_m(\mathbf{x}))^T$ を表すものとする。

問題 (7) に対するラグランジュ関数 (Lagrangian) を

$$L(\mathbf{x}, \mathbf{u}) = f(\mathbf{x}) - \sum_{i=1}^m u_i c_i(\mathbf{x}) \quad (8)$$

で定義し、この関数の $\mathbf{x}$ に関する勾配とヘッセ行列をそれぞれ

$$\nabla_{\mathbf{x}} L(\mathbf{x}, \mathbf{u}) = \nabla f(\mathbf{x}) - \sum_{i=1}^m u_i \nabla c_i(\mathbf{x}) \quad (9)$$

$$\nabla_{\mathbf{x}}^2 L(\mathbf{x}, \mathbf{u}) = \nabla^2 f(\mathbf{x}) - \sum_{i=1}^m u_i \nabla^2 c_i(\mathbf{x}) \quad (10)$$

と書く。また、ベクトル関数 $\mathbf{c}$ の点 $\mathbf{x}$ におけるヤコビ行列 (Jacobian matrix)、すなわち偏微分係数 $\partial c_j(\mathbf{x}) / \partial x_i$ を第 (i, j) 成分とするような $n \times m$ 行列を $\nabla \mathbf{c}(\mathbf{x})$ で表す。式 (10) のヘッセ行列について補足しておく。 n 変数の実数値関数 $f(\mathbf{x})$ について、 $f(\mathbf{x})$ が 2 回連続微分可能な関数であるならば、行列関数

$$\nabla^2 f(\mathbf{x}) = \begin{pmatrix} \frac{\partial^2 f(\mathbf{x})}{\partial x_1^2} & \cdots & \frac{\partial^2 f(\mathbf{x})}{\partial x_1 \partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f(\mathbf{x})}{\partial x_n \partial x_1} & \cdots & \frac{\partial^2 f(\mathbf{x})}{\partial x_n^2} \end{pmatrix} \quad (11)$$

を $f(\mathbf{x})$ のヘッセ行列 (Hessian Matrix) という。このヘッセ行列は SQP 法のアルゴリズムを構築する上で重要なパラメータとなる。

8 準ニュートン法によるヘッセ行列の近似

前節の式 (11) の $\nabla^2 f(\mathbf{x})$ は非線形最適化においては計算できない場合がある。そこでヘッセ行列を近似して計算する準ニュートン法を使用する。ヘッセ行列を近似したものが式 (1) における B_k である。本論文においては準ニュートン法として BFGS 法と SR1 法を使用した。以下に BFGS 法と SR1 法について示す。

両方の近似式で使われるベクトル $\mathbf{s}_k$ 、 $\mathbf{y}_k$ を以下の様に定義する。

$$\mathbf{s}_k = \mathbf{x}_{k+1} - \mathbf{x}_k \quad (12)$$

$$\mathbf{y}_k = \nabla_{\mathbf{x}} L(\mathbf{x}_{k+1}, \mathbf{u}_{k+1}) - \nabla_{\mathbf{x}} L(\mathbf{x}_k, \mathbf{u}_{k+1}) \quad (13)$$

BFGS 公式

$$B_{k+1} = B_k + \frac{\mathbf{y}_k (\mathbf{y}_k)^T}{(\mathbf{y}_k)^T \mathbf{s}_k} - \frac{B_k \mathbf{s}_k (\mathbf{s}_k)^T B_k}{(\mathbf{s}_k)^T B_k \mathbf{s}_k} \quad (14)$$

SR1 公式

$$B_{k+1} = B_k + \frac{(\mathbf{y}_k - B_k \mathbf{s}_k)(\mathbf{y}_k - B_k \mathbf{s}_k)^T}{(\mathbf{s}_k - B_k \mathbf{s}_k)^T \mathbf{s}_k} \quad (15)$$

9 数値実験

SQP 法と Sl_1 QP 法を用いて非線形最適化問題を解き最適解が求まるまでの反復回数を比較する。比較に利用する SQP 法は岩波 SQP プログラム [2] を使用する。

1. 実験 1 以下のような非線形最適化問題を考える

$$\text{目的関数: } -x_1 + 10x_1^2 + 10x_2^2$$

$$\text{制約条件: } x_1^2 + x_2^2 - 1 = 0$$

$$x_1 - 0.5 \geq 0$$

$$x_2 + 0.5 \geq 0$$

(16)

Sl_1 QP 法の初期値 $\Delta_0 = 0.5$, $\mu_0 = 100$ とする。

以下に計算結果を示す。

表 1: Sl_1 QP 法 (BFGS)、SQP 法、NLPQLP 法の大反復回数 (iter)、2 次計画問題の計算回数 (QPs)、線形計画問題の計算回数 (LPs) (- は実行不可能)

初期点		Sl_1 QP(BFGS)			SQP		NLPQLP	
x_1	x_2	iter	QPs	LPs	iter	QPs	iter	QPs
1.0	0.5	4	4	0	4	6	6	6
0.5	0.5	7	7	0	4	7	10	10
-1.5	-1.5	10	12	4	9	15	-	-
3.0	3.0	8	8	2	85	167	9	9
10	10	15	15	9	55	106	10	10

表 2: Sl_1 QP 法 (SR1)、SQP 法、NLPQLP 法の大反復回数 (iter)、2 次計画問題の計算回数 (QPs)、線形計画問題の計算回数 (LPs) (- は実行不可能、* は解に未到達)

初期点		Sl_1 QP(SR1)			SQP		NLPQLP	
x_1	x_2	iter	QPs	LPs	iter	QPs	iter	QPs
1.0	0.5	4	4	0	4	6	6	6
0.5	0.5	*	*	*	4	7	10	10
-1.5	-1.5	*	*	*	9	15	-	-
3.0	3.0	8	9	2	85	167	9	9
10	10	15	15	9	55	106	10	10

表 3: $S\ell_1$ QP 法 (SR1、BFGS)、SQP 法、NLPQLP 法の目的関数の評価回数 ($\#f$)、目的関数の勾配の評価回数 ($\#g$) (-は実行不可能、*は解に未到達)

初期点		$S\ell_1$ QP 法 (SR1)		$S\ell_1$ QP 法 (BFGS)		SQP		NLPQLP	
x_1	x_2	$\#f$	$\#g$	$\#f$	$\#g$	$\#f$	$\#g$	$\#f$	$\#g$
1.0	0.5	4	4	4	4	8	8	16	6
0.5	0.5	*	*	7	6	9	8	13	10
-1.5	-1.5	*	*	10	9	18	15	-	-
3.0	3.0	8	8	8	8	762	150	27	9
10	10	15	15	15	15	337	110	11	10

10 おわりに

数値実験の結果から、計算時間のかかる 2 次計画問題を解く回数が、初期点が最適点から近いときは同程度であるが、離れるほど $S\ell_1$ QP 法のほうが単調直線探索法の SPQ 法、非単調直線探索法の NLPQLP 法に比べ、より少ない回数で収束した。更に 2 種類の直線探索法の SQP 法では実行不可能な初期点や最適点に到達しない初期点でも安定して収束した。これらより良い結果が得られたといえる。今後の課題としてさらに複雑な非線形最適化問題に対する比較や、それに伴った適切な $S\ell_1$ QP 法の初期パラメータの選択があげられる。

参考文献

- [1] R.Fletcher: Practical Methods of Optimization, Second Edition, John Wiley and Sons, 1988
- [2] 茨木俊秀.: FORTRAN77 最適化プログラミング, 岩波書店, 東京, 1989