

上下限制約をもつ非線形最小二乗問題に対する

Affine scaling を用いた ϵ -近似解法

慶知大 経営 相良信子 (Nobuko SAGARA)

京都大 工学 福島雅夫 (MASAO Fukushima)

非線形最小二乗問題の解を求める論文は決山あるが、制約条件をもつ問題を取り扱った論文はあまり多くない。そこで、
は 次の問題の数値解法を提案する。

(P ₀)	$\min_x \frac{1}{2} \ F(x)\ ^2$
	s. t. $l \leq x \leq u \quad x \in R^n$
$F \in R^m$	$f(x) = \frac{1}{2} \ F(x)\ ^2 = \frac{1}{2} \sum_{i=1}^m F_i^2(x)$
$F_i \in C^2$	l と u は x の下限と上限の n 次元ベクトル

次に展開する (P₀) の新解法の特徴は:

- ① 反復法 ② 上下限制約の特徴を考えた trust region 法。(以下 T.R. 法と書く) 通常の T.R. 法とは異なる。 ③ 内点法

本稿で提案する方法の基本的な考え

通常の T.R. 法 [2, 3, 4, 5] と比べては、現在集を含まる適当な領域 (この領域 a こと ϵ T.R. という) の中で目的関数 $f(x)$ を近似した二次関数を最小化する ϵ によって得られる集 ϵ

次の反復算とする。関数の性質や領域の大きさにより元の関数に対する近似の精度が異なるので、尤、尤法の収束を保證するように領域の大きさを調整しなければならない。本稿で提案する方法では、常に上下限制約の内部に含まれる点のみを考へる。したがって、現在の点 x を含む尤、尤として、上下限制約に含まれるような楕円を考へ、その中心次の点を定める。よって、現在の点 x から次の点 x^+ を決定する修正量 p とすると…… (わからず 次の点は $x^+ = x + p$ であるが) …… (P_0) より次の部分問題を導くことが出来る。

$$(P_1) \quad \min_{p \in \mathbb{R}^n} \frac{1}{2} \| F(x) + J(x)p \|^2$$

$$\text{s. t.} \quad \sum_{i=1}^n \left(\frac{p_i}{d_i} \right)^2 \leq \Delta^2 \quad (0 < \Delta < 1)$$

$$\text{E に対し} \quad d_i = \min [(u_i - x_i), (x_i - l_i)]$$

$$J = J(x) = \nabla F(x), \quad F = F(x) = (F_1(x), \dots, F_m(x))^T$$

(P_1) を書きかえると次の問題 (P_2) が得られる。

$$(P_2) \quad \min g^T p + \frac{1}{2} p^T B p \quad \text{s. t.} \quad \| D p \| \leq \Delta$$

$$\text{E に対し} \quad g = \nabla f(x) = J^T F \quad B = J^T J$$

$$D = \begin{bmatrix} 1/d_1 & & 0 \\ & \ddots & \\ 0 & & 1/d_n \end{bmatrix}$$

制約条件 $\| D p \| \leq \Delta$ は現在の点から上下限制約の境界への距離にもとづく metric のもとで球になっている。このことからこの方法は Affine scaling 法とわかれており、LP 問題に

に対する新しいアルゴリズムにおいてよく用いられる考え方である。^[1]

次に (P_2) で計算した p を用いて $x^+ = x + p$ の目的関数値が x の関数値より十分に減少しているかを判定することが必要。もし、この条件が満たされれば $x^+ = x + p$ として次の反復集 x^+ が決まる。そうでなければ、現在集 x を動かさずに ϵ, γ の半径を小さくして部分問題 (P_2) を再び解く。この手続きを繰返すと、必ず目的関数値を減少させるような集が得られる。この様な操作を各反復集で実行することにより、目的関数値を段階的に減少させる集列が生まれ、この集列が最適解に近づく。

○ Active set 法に基づいた (P_2) の修正

(P_0) の最適解は上下限制約の境界上にあることが一般である。上に述べた ϵ, γ 法を実行すると現在集が境界に近づくにつれて楕円の長軸/短軸 $\rightarrow \infty$ となり楕円の厚みがなくなる。従って、部分問題 (P_2) は ill-conditioned になりその解を計算する際に数値的不安定性が生ずる。これを避けるために、制約付非線形最適化手法の一つである Active set 法の考え方をとり入れた修正を行おう。まず、境界の内側には $\epsilon > 0$ の η をもたして η 変数がその領域に入ったとき、この変数はその値で固定して定数とみられる。この結果定まる部分空

に限定して (P_2) を考えれば ill-conditioned ではなくなる。

この様にして部分問題 (P_2) は次の様に修正される。

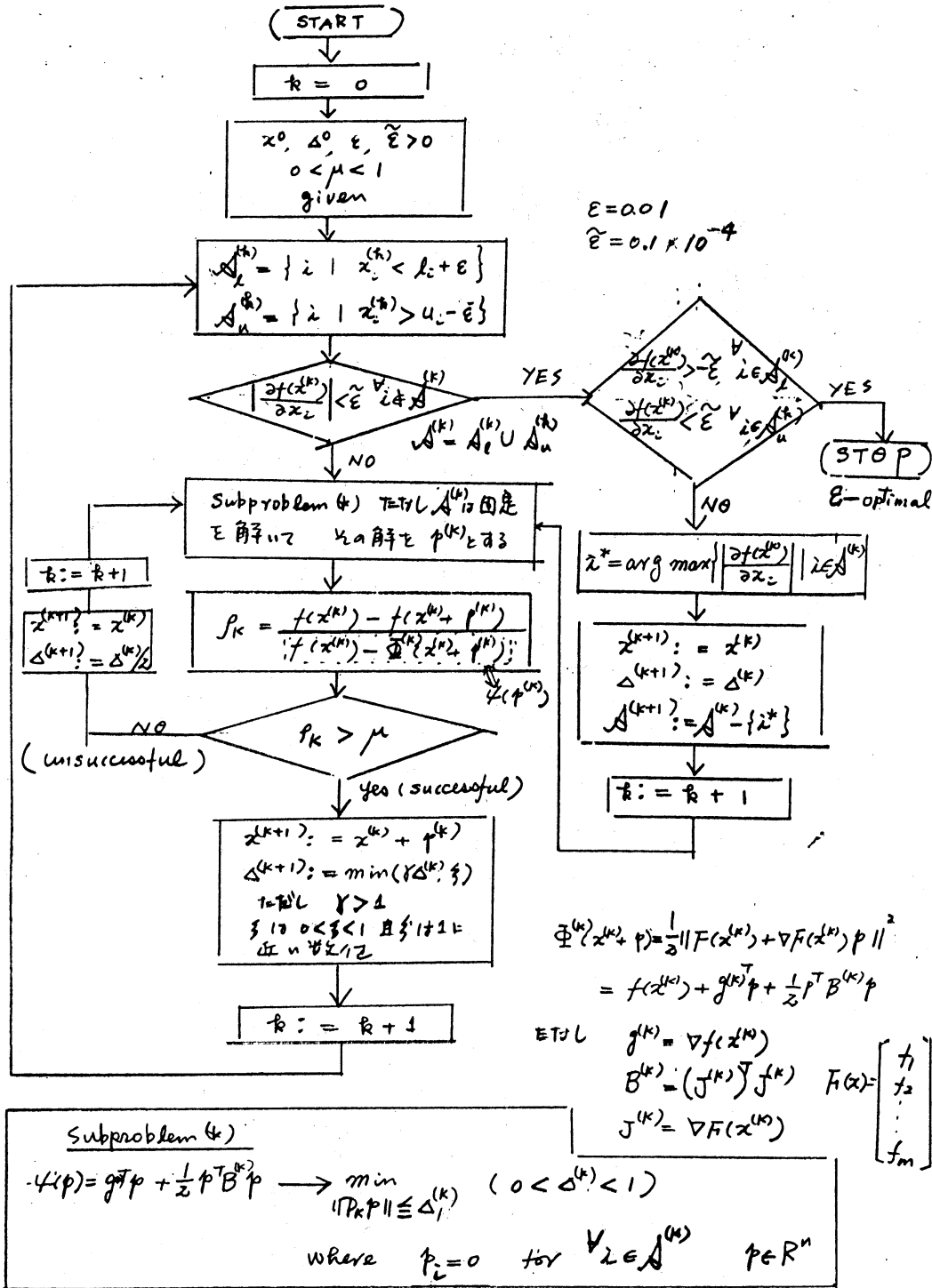
$$\begin{array}{l}
 (P_3) \quad \min_{p \in R^n} g^T p + \frac{1}{2} p^T B p \\
 \text{s. t.} \quad \|Dp\| \leq \Delta, \quad 0 < \Delta < 1 \\
 \text{ただし,} \quad p_i = 0 \quad \text{for } \forall i \in A \\
 \quad \quad \quad A = A_L \cup A_U \\
 \quad \quad \quad A_U = \{i \mid x_i > u_i - \varepsilon\}, \quad A_L = \{i \mid x_i < l_i + \varepsilon\}
 \end{array}$$

(P_3) は固定されている変数だけで取り扱うので次元が落ちる。その部分空間に対しては問題は ill-conditioned にはならず、数値的不安定性は生じない。

(P_3) の解が (P_0) に対する最適性の条件を充せば停止する。そうであれば、固定された変数を変化させることにより更に目的関数値を減少させることができる。この様な手続きにより (P_0) の最適解の ε -近似解が得られる。この計算の手順を示したのが次頁のフローチャートである。

最後に得られた解の精度は ε に依存する。したがって元の問題の最適解を得るためには修正が必要である。これには、固定されている変数の値をその境界値に直す。すなわち、 $l_i < x_i < l_i + \varepsilon$ ならば $x_i = l_i$ にする。そして、それらを固定する。この値を初期値として固定してしまった変数を除いた部分空間で λ, γ 法で解くと最適解を得ることができ

3.



• 新しい手法の計算の動きを見るために

Rosenbrock function ϵ 修正した関数 [7]

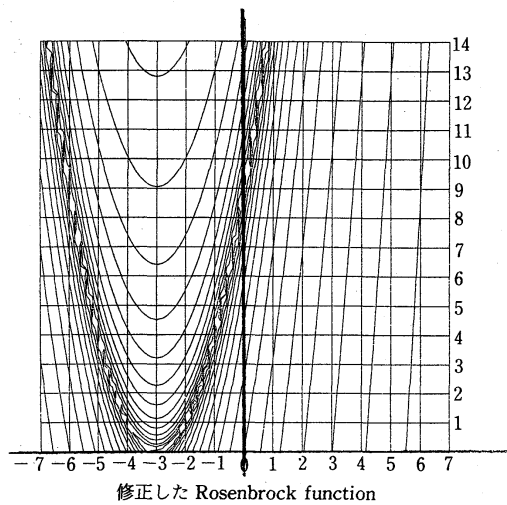
$$F_1(x) = 10 \cdot (x_2 - (x_1 + 3)^2)$$

$$F_2(x) = 2 + x_1$$

よって 次の問題を考える。

$$\min_x f(x) = \frac{1}{2} (F_1^2(x) + F_2^2(x)) \quad \text{s. t. } x \geq 0, \quad x \in \mathbb{R}^2$$

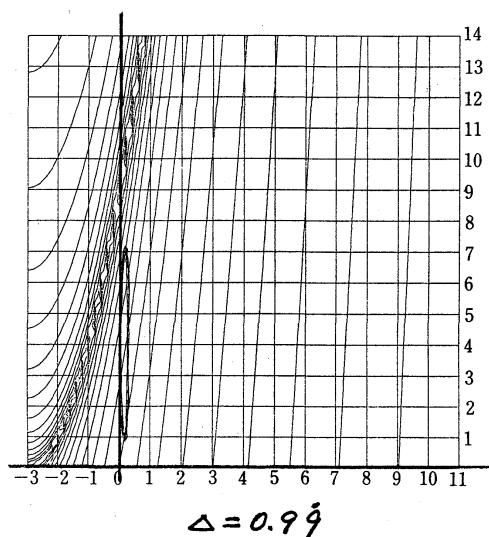
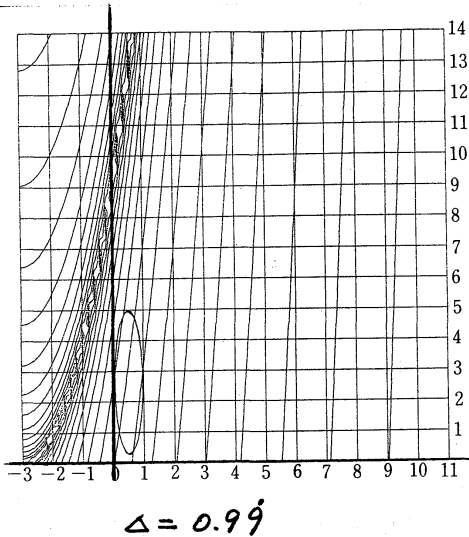
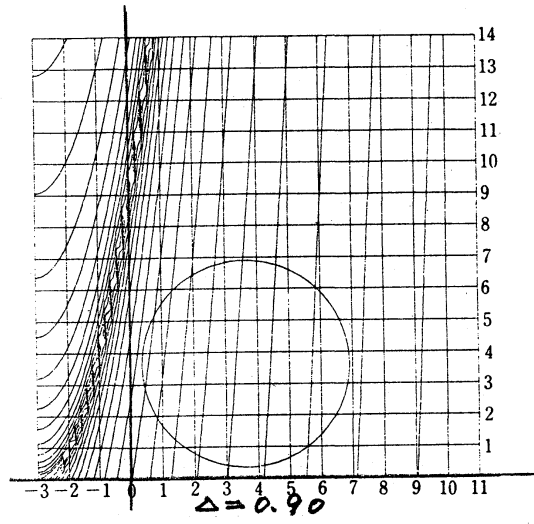
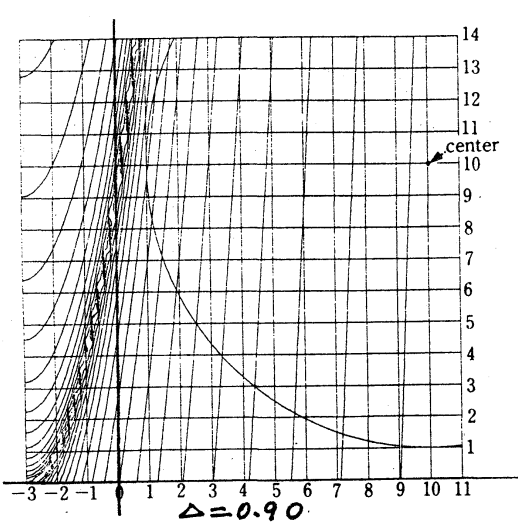
この真の解は $x^* = (0, 9)^T$



$x^0 = (10, 10)^T$ を初期値としたときの新しい手法の反復算の動きは下記の通り： $\epsilon = 0.01$

Iter.	x_1	x_2	$\ F\ $
0	10.	10.	1590.
1	3.63	3.63	404.
2	0.51	2.63	97.
3	0.14	4.03	58.
4	0.11	7.57	21.
5	0.0136	9.072	2.0
6	0.1-02	9.0(2)	2.0(2)

これらの最初のいくつかの反復算における x, γ の円を描いたのが次の図である。



○新しいアルゴリズムの収束について

このアルゴリズムは有限回の反復で ϵ -最適解に到達することが証明できる。

Lemma 1

$$-\psi(p) \geq \frac{1}{2} \min_i \{ |\hat{g}_i| \Delta, |\hat{g}_i| / \|\hat{J}_i\|^2 \}$$

$$T=0 \quad \hat{g}_i = x_i \cdot \frac{\partial f}{\partial x_i}, \quad \hat{J}_i = x_i \cdot \frac{\partial F}{\partial x_i}$$

∴ $\exists \delta$ は $|\frac{\partial f(x)}{\partial x_i}| \geq \tilde{\epsilon}$ 且 $\exists \delta$ δ 充たす特定存在も

Lemma 2 $\Delta \rightarrow 0$ は起らない

証明 $\Delta \rightarrow 0$ を仮定して $\rho \rightarrow 1$ を証明し、 Δ の更新の規則と反する矛盾を導く。この証明の途中で lemma 1 を用いる。

Def. $\rho_k > \mu$ (プロチャート参照) を充たるとき successful といひ、その他るとき, unsuccessful といふ。

Lemma 3 各反復集で unsuccessful は有限回しか起らない。

証明 Lemma 2 より自明。

Lemma 4

$$f(x) - f(x+p) \geq \frac{1}{2} \mu \min_i \left\{ |\hat{g}_i| \left(\Delta, |\hat{g}_i| / \|\hat{J}_i\|^2 \right) \right\}$$

$$\text{EFG} \quad \rho_i = 0 \quad \text{for } \forall i \in \mathcal{A}$$

Theorem 有限回の反復で ε -最適解に到達する。

証明 Lemma 2, 3, 4 より導かれる。

・ 数値実験

新しい手法と従来からの手法である Active set 法^[5] (Gauss-Newton) の二つの方法を用いて数値実験を行った。その結果、変数の個数が大きくなるに従って新しい手法が有効となることが分った。これらの計算は京都大学大型計算センターで行った。

Wrong Wood function^[6] $m = 3(n-2)$

$$f_{6i-5}(x) = 10(x_{2i} - x_{2i-1}^2)$$

$$f_{6i-4}(x) = 1 - x_{2i-1}$$

$$f_{6i-3}(x) = 3\sqrt{10}(x_{2i+2} - x_{2i+1}^2)$$

$$f_{6i-2}(x) = 1 - x_{2i+1}$$

$$f_{6i-1}(x) = \sqrt{10}(x_{2i} + x_{2i+2} - 2)$$

$$f_{6i}(x) = \sqrt{10}(x_{2i} - x_{2i+2})$$

$$\min_x f(x) = \frac{1}{2} \sum_{i=1}^m f_i(x) \quad \text{s. to } x \geq l$$

下限 l は適当なランダムな正の数値を選び、初期値は $l + \varepsilon$ より充分に大きいような値をとった。また、計算で用いた ε と $\tilde{\varepsilon}$ の値は $\varepsilon = 0.001$, $\tilde{\varepsilon} = 0.1 \cdot 10^{-3}$ である。この数値実験の結果の CPU time と反復回数 k の表の通り:

実験の結果の CPU time と反復回数 k の表の通り:

n	m	New Method	Active set method
20	54	0.479 (sec)	0.462 (sec)
50	144	14.674	17.757
70	204	46.727	69.731
80	234	78.598	119.137
90	264	118.143	212.276
100	294	161.191	324.306
110	324	223.162	476.646

n	New Method	Active set Method
20	23	30
50	50	72
70	59	102
80	67	117
90	71	132
100	71	147
110	74	162

n	<u>CPU time for New Method</u> CPU time for Active Set Method	<u># of iter. for New Method</u> # of iter. for Active Set Method
20	1.036	.766
50	.826	.694
70	.670	.578
80	.659	.572
90	.556	.527
100	.497	.482
110	.468	.456

ε -最適解を補正して最適解を得るために要する計算時間は、 n が50以上のときは ε -最適解の計算時間の6%から7%位であることが数値実験より分った。従って、 n が大きくなるとは新しい手法が優れていることが分る。

この研究の一部は愛知大学研究助成 C-1 によってなされた。

- 参考文献 [1] E. R. Barnes: A variation on Karmarkar's algorithm for solving linear programming problems, Math Prog. 36 (1986) 174-182 [2] J. J. Moré and D. C. Sorensen: Newton's Method, ANL-82-8 Feb (1982) [3] J. J. Moré and D. C. Sorensen: Computing a trust region step, SIAM J Sci Stat. Comput. Vol. 4 553-572 (1983) [4] J. J. Moré: The Levenberg-Marquardt Algorithm: Implementation and theory, Proceeding of the Dundee Conference on Numerical Analysis, G. A. Watson ed., Springer-Verlag (1978) 105-116 [5] R. Fletcher: Practical Methods of Optimization Vol. 1. Vol. 2 John Wiley & Sons (1981) [6] Ph. Toint: On large scale nonlinear least squares calculations, SIAM J. Sci. Stat. Comput. Vol. 8 No. 3 (1987) [7] 相良: Trust Region法を用いた非負制約をもつ非線形最小二乗問題 愛知大学法経論集第116号 昭和63年3月