1X-7

制約条件付き非線形最小2乗問題への逐次2次計画法の適用

八巻 直一　　矢部 博*　　高橋 悟*
(システム計画研究所、　*東京理科大学 理学部)

1. はじめに

非線形最小2乗問題は、一般の非線形関数の最小化アルゴリズムを用いることが可能で、実際その様に扱われる例も多い。しかしながら、目的関数が関数の2乗和である構造を、もっと有効に活かす試みが考えられ、幾つかの有力な手法が提案されている。一方、非線形最適化アルゴリズムとしてSQP法 (Successive Quadratic Programming method:逐次2次計画法)[4]が非常に有望視されている。本稿では制約条件付き最小2乗問題に対して、問題の構造をSQP法に組み入れることを試みる。

制約条件付き最小2乗問題は、次の様に表わされる。

『 C^2級関数 f_j :$R^n \to R^1, j=1,\cdots,m$ $(m \geqq n)$ 及び g_i :$R^n \to R^1, j=1,\cdots,r$が与えられているとき制約条件 $g(x) \leqq 0$ の下で目的関数 $F(x)= 1/2\ f(x)^T f(x)$ を $x \in R^n$ について最小にせよ。ただし $x^T = (x_1,\cdots,x_n)$, $f(x)^T = (f_1(x),\cdots,f_m(x))$, $g(x)^T = (g_1(x),\cdots,g_r(x))$ 』

上の問題において、等号制約 $h(x) = 0$ は $h(x) \leqq 0$ かつ $-h(x) \leqq 0$ と同値であるので、省いても一般性は失われない。

2. SQP法(逐次2次計画法)

一般の無制約最小化問題『F(x)を最小にせよ』に準ニュートン法を適用すれば、点列$\{x_k\}$は次のように定義される。

$$x_{k+1} = x_k + \alpha_k d_k, \quad B_k d_k = -\nabla F(x_k) \qquad \cdots(1)$$

ここに、α_k は刻み幅、B_k は$\nabla^2 F(x_k)$の近似行列であり、セカント条件

$$B_{k+1}(x_{k+1} - x_k) = \nabla F(x_{k+1}) - \nabla F(x_k) \qquad \cdots(2)$$

を満たすように決定される。(1)の d_k は $P(d)=\nabla F(x_k)^T d + 1/2\ d^T B_k d$ の停留点である。B_k が正定値であれば $P(d)$ の最小点として d_k が得られる。

他方、一般の制約条件付き最小化問題『$g_i(x) \leqq 0$ $(i=1,\cdots,r)$の下でF(x)を最小にせよ』に対するSQP法は、d_k を求めるのに、P(d)を目的関数とし g(x)の x_k における線形近似を制約条件としたQP問題を解く方法である。ただし、この場合の B_k はラグランジュ関数

$$L(x,\lambda) = F(x) + \lambda^T g(x) \qquad \cdots\cdots(3)$$

のxに関するヘッセ行列$\nabla_{xx}L$ の近似行列である。

逐次2次計画法のアルゴリズムは以下のとおりである。

ステップ 0. 初期点 x_1, n×n 正定値対称行列 B_1を与える。$k=1,2,\cdots$ に対して以下を繰り返す。

ステップ 1. x_k, B_k に対する次の2次計画問題『線形制約条件 $g_i(x_k) + \nabla g_i(x_k)^T d \leqq 0$ $(i=1,\cdots,r)$の下で、2次関数 $1/2\ d^T B_k d + \nabla F(x_k)^T d$ を最小にせよ。』 ……(4) 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。

ステップ 2. もしx_k, λ_{k+1}がもとの問題の Kuhn-Tucker 条件を満たしていれば停止し、さもなければステップ 3 へ行く。

ステップ 3. α_k を定めて、$x_{k+1} = x_k + \alpha_k d_k$ とする。

ステップ 4. 行列 B_k を更新し B_{k+1}を作成する。

3. 問題の構造を活かした公式の導出

無制約最小2乗問題の解を求める数値解法では、ニュートン法を基礎にした解法が代表的である。Gauss-Newton法は残差が非常に小さい場合には良い性能をもつが、残差が比較的大きなときには性能が落ちる。そこで Dennis, Gay and Welsch[1] は、いずれの場合にも良い性能を示すような工夫をしている。本稿では、Dennisらに従って制約条件付き最小2乗問題に適したSQP法を構成する。なお以下では添え字kを省略し、k+1を"+"で表わす。

Dennisら(DGWと書く)は、$\nabla^2 F(x) = J_f(x)^T J_f(x) + \sum f_i(x) \nabla^2 f_i(x)$の2項目を行列Aで近似し $B = J_f(x)^T J_f(x) + A$ として準ニュートン法に適用している。ただし J_f はfのヤコビ行列である。

我々は制約条件付き最小2乗問題においてDGWのアイデアを採用して、ラグランジュ関数のxについてのヘッセ行列

Application of SQP method to nonlinear least squares problems with constraints

Naokazu YAMAKI, Hiroshi YABE*, Satoru TAKAHASHI*
Research Institute of Systems Planning Ltd., * Science University of Tokyo

$$\nabla_{xx} L(x,\lambda) = J_f(x)^T J_f(x) + \sum f_i(x)\nabla^2 f_i(x) + \sum \lambda_i \nabla^2 g_i(x) \quad \cdots\cdots(5)$$

の２項目を行列Ａ、３項目を行列Ｃで近似することを試みる。すなわち

$$B = J_f(x)^T J_f(x) + A + C \quad \cdots\cdots(6)$$

とする。

Ａについてのセカント条件および更新公式として、ＤＧＷと同様に

$$A_+ s = q \ , \ q = [J_f(x_+) - J_f(x)]^T f(x_+) \quad \cdots\cdots(7)$$

$$A_+ = \beta A - \frac{s^T(q-\beta As)}{(s^T y)^2} yy^T + \frac{(q-\beta As)y^T + y(q-\beta As)^T}{s^T y} \quad \cdots\cdots(8)$$

$$\beta = \min(|s^T q/s^T As|, 1) \ ,$$

$$s = x_+ - x, \quad y = \nabla F(x_+) - \nabla F(x)$$

を用いる。ここで q は$f_i(x)$が２次関数を仮定したとき $q = \sum f_i(x_+)\nabla^2 f_i(x_+) s$ となる性質を持つ。またβは残差の大小や f_i の非線形性に応じて、行列Ａをスケーリングするパラメータである。

他方 $\sum \lambda_{+i}(\nabla g_i(x_+) - \nabla g_i(x)) \simeq \sum \lambda_{+i} \nabla^2 g_i(x_+)s$ であるから、C_+は

$$C_+ s = z, \quad z = [J_g(x_+) - J_g(x)]^T \lambda_+ \quad \cdots\cdots(9)$$

を満たすように定める。

以上の議論より、ＳＱＰ法におけるＱＰ部分問題(4) は次のように書き直される。

『 d について、２次目的関数
$1/2\, d^T [J_f(x)^T J_f(x) + A + C]d + f(x)^T J_f(x)d$
を線形制約条件 $g_i(x) + \nabla g_i(x)^T d \leqq 0 \ (i=1,\cdots,r)$
の下で最小にせよ。』

Han[2]は、行列Ｂが正定値ならばＳＱＰ法が凸非線形最適化問題に対して大域的収束することを示している。従ってヘッセ行列の構造を取り入れた場合にも正定値性の保存を考慮したＣの更新公式を構築すべきである。

４．数値実験および考察

数値実験では、従来のＳＱＰ法における行列Ｂの更新公式と、本稿で導いた $J^T J + A + C$ とおく方法との比較を試みた。Ｂの更新では Powell[4]の修正ＢＦＧＳ公式が用いられることが多いが、今回の実験でもこれに従った。

Ａの更新公式としてＤＧＷ公式を用いるので、本ＳＱＰ法では制約が無ければそのままＤＧＷ法と同等になることを注意しておく。Ｃの更新公式については、今回の実験ではＤＦＰおよびＢＦＧＳ公式にPowellの修正を加えた公式を用いた。それらはそれぞれ

$$C_+ = C + \left(1 + \frac{s^T Cs}{s^T \eta}\right)\frac{\eta\eta^T}{s^T \eta} - \frac{Cs\eta^T + \eta s^T C}{s^T \eta} \quad \cdots(10)$$

$$C_+ = C - \frac{Css^T C}{s^T Cs} + \frac{\eta\eta^T}{s^T \eta} \quad \cdots\cdots(11)$$

で与えられる。ただし $\eta = \theta z + (1-\theta)Cs$ であり、パラメータθは $s^T z \geqq 0.2 s^T Cs$ ならば$\theta=1$, さもなければ$\theta = 0.8 s^T Cs/s^T(Cs - z)$と選ばれる。なおＣの更新公式の構築についてはラグランジュ乗数によるスケーリングや、制約の線形性を生かすなどの工夫の余地がある。

使用したテスト問題はBetts問題(n=5,m=5,r=6)、Wong問題(n=7,m=9,r=4) ([3]の問題79、問題100) である。実験結果は表１に示す通りである。ここでＡ、Ｂの初期行列は、$B_1 = I$, $A_1 = 0$ とし、Ｃの初期行列は表中の１と３では $C_1 = 0$ 、２と４では $C_1 = I$ を用いた。実験結果を見る限りでは、概ね本稿で導いたＳＱＰ法が最小２乗問題にはよくなじみ、従来のＳＱＰ法に較べてよい成績をおさめていることがわかる。

表１

	Betts SQP	1	2	3	4	Wong SQP	1	2	3	4
Iteration Count	13	10	10	8	10	15	10	9	14	11
Evaluation Count	13	10	10	8	10	33	20	13	26	15
QP Iteration Count	40	31	30	24	30	32	20	18	28	22

1 DFP C=0　2 DFP C=I　3 BFGS C=0　4 BFGS C=I

参考文献

[1]Dennis,Jr.,J.E., Gay,D.M. and Welsch,R.E., ACM Trans. Math. Software, 7, 348-368 (1981).

[2]Han,S.P., Math. Program., 20, 1-13 (1981).

[3]Hock,W. and Schittkowski,K., Lecture Notes in Economics and Mathematical Systems 187, Springer-Verlag, Berlin (1981).

[4]Powell,M.J.D., Lecture Notes in Mathematics 630, 144-157, Springer-Verlag, Berlin (1978).