

Davies, Swann, Campey 法による関数の
極小化の一解析

京大 大型計算機センター

星野 聡

§1 序云

微係数を用いずに関数の極小点を求める Davies, Swann, Campey の方法⁽²⁾ (以下 DSC 法とよぶ) は、互いに直交する正規化された n 個のベクトル P_1, P_2, \dots, P_n を選り、直線に沿う極小化 (以下 LM と略記する) $X_{i+1} = X_i + \alpha_i P_i$ ($i = 1, \dots, n$) を行う。ここで α_i は極小化しようとする関数 $f(X_{i+1})$ が P_i 上で極小になるように決められる。

ついで、ベクトル

$$g_i = \sum_{j=1}^n \alpha_j P_j \quad (i=1, \dots, n) \quad (1)$$

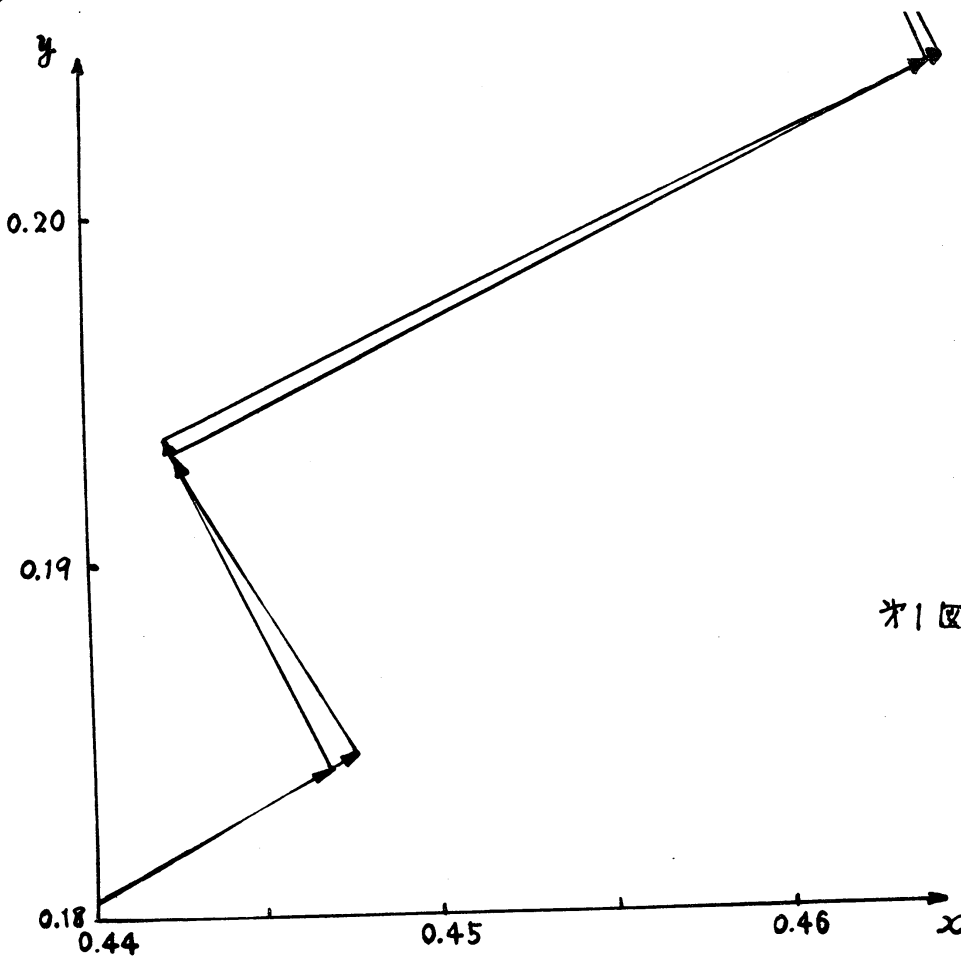
を直交正規化して、新しい P_i ($i=1, \dots, n$) とする。このプロセスを繰返してゆく。

この方法で、Rosenbrock⁽¹⁾ が与えた例題

$$f = 100 (y - x^2)^2 + (1 - x)^2 \quad (2)$$

を初期値 $x = 0.6929 \times 10^{-2}$, $y = 0.4961 \times 10^{-2}$ から出発して解くと, 丁度, 最急降下法を適用した場合のように, ジブザグ状の探索が行われ, 収束が極めて遅いことが見出される。オ1図に, その経過の一部を拡大して示す。

DSG法では, 探索の方向は, 互いに直交しているので, Powellの方法⁽³⁾におけるように, 探索の方向が一次独立でなくなるために生ずる不都合はないが, このようなジブザグ現象が発生する恐れがあり, これは3次元以上の場合にも起りうる。



以下においては、2次元の場合について解析を行い、ついで
数値計算の結果について説明する

§2 解析

ゆるやかに傾斜する谷をもつ関数

$$f = y^2 - \varepsilon x \quad (\varepsilon > 0) \quad (3)$$

を考える。x-y 平面上で、ジグザグ状探索の様子を図2の
ようになる。

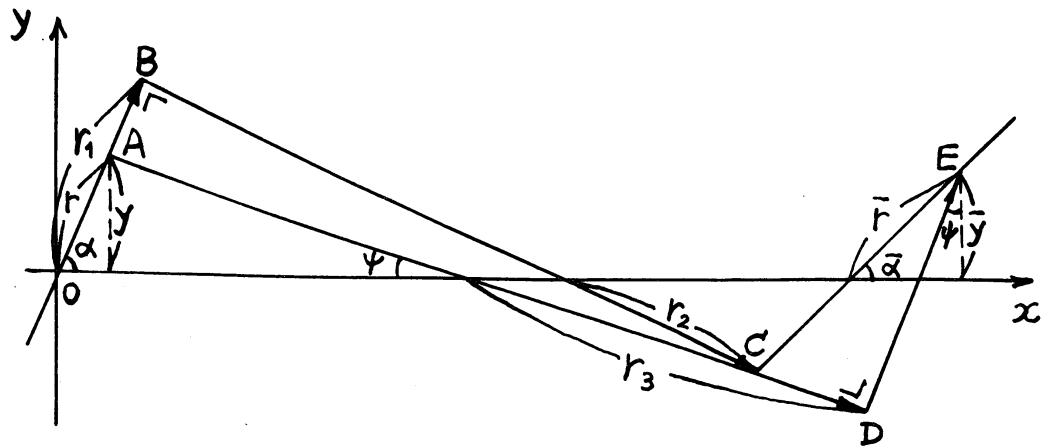


図2

すなわち、点Aより \vec{AB} , \vec{BC} , \vec{CD} , \vec{DE} とLMを行っ
て点Eに達し、これでジグザグ探索の一周期を構成する。
($\vec{AB} \perp \vec{BC}$, $\vec{CD} \perp \vec{DE}$)

x 軸と角 θ をなす方向に LM を行うとき, $x = r \cos \theta$,
 $y = r \sin \theta$ を式 (3) に用い, $\partial f / \partial r = 0$ とおくと, 極
 小点として

$$\left. \begin{aligned} r &= \frac{\varepsilon \cos \theta}{2 \sin^2 \theta} \\ x &= \frac{\varepsilon \cot^2 \theta}{2} \\ y &= \frac{\varepsilon \cot \theta}{2} \end{aligned} \right\} (4)$$

をうる。この関係とオ 2 図から、次の関係を導くことが出来る。

$$\bar{z} = \frac{(z + \tan \alpha) \tan \alpha}{1 + \tan^2 \alpha + \tan^4 \alpha - z \tan \alpha} \quad (5)$$

$$\tan \bar{\alpha} = \frac{(\tan \alpha + \bar{z}) \bar{z}^2}{1 + \bar{z}^2 + \bar{z}^4 - (\tan \alpha) \bar{z}} \quad (6)$$

たゞし

$$z = \frac{2\gamma}{\varepsilon}, \quad \bar{z} = \frac{2\bar{\gamma}}{\varepsilon} \quad (7)$$

周期的な探索をひきおこす条件として, $\bar{z} = z$, $\bar{\alpha} = \alpha$
 とおき, 式 (5) に用いると $z = \cot \alpha$ をうる。したがって,

$y = (\varepsilon/2) \cot \alpha$. これと (4) 式から, A, B 実および C, D 実が互いに一致する場合に限られることがわかる。

次に, ジグザグ探索の安定性について考えるため, z と α が $z = \cot \alpha$ をみたす値から少しずれたと考えると z は $\cot(\alpha + \lambda_n)$, α は $\alpha + \varepsilon_n$ に, また $\bar{\alpha}$ は $\alpha + \varepsilon_{n+1}$ に \bar{z} は $\cot(\alpha + \lambda_{n+1})$ になったとする。

$$\tan(\alpha + k) = \tan \alpha + k \sec^2 \alpha + O(k^2) \quad (8)$$

$$(k = \varepsilon_n, \varepsilon_{n+1}, \lambda_n, \lambda_{n+1})$$

を用い, k の二乗以上の項を無視すると

$$\lambda_{n+1} = \mu_1 \varepsilon_n + \mu_2 \lambda_n$$

$$\varepsilon_{n+1} = \nu_1 \varepsilon_n + \nu_2 \lambda_{n+1}$$

(9)

ここで

$$\mu_1 = 1 - \cot^2 \alpha$$

$$\mu_2 = \cot^2 \alpha$$

$$\nu_1 = \tan^2 \alpha$$

$$\nu_2 = 1 - \tan^2 \alpha$$

(10)

(9) 式から

$$\lambda_n = \lambda_0 + (1 - \cot^2 \alpha)(\varepsilon_0 - \lambda_0)n$$

$$\varepsilon_n = \varepsilon_0 + (1 - \cot^2 \alpha)(\varepsilon_0 - \lambda_0)n$$

(11)

をうる。 λ_n, ε_n は n に比例して変化し、その比例係数は $\lambda, \lambda_0, \varepsilon_0$ に依存する。 $\cot^2 \alpha \approx 1$ または $\varepsilon_0 \approx \lambda_0$ のときその値は小さいから、 λ_n, ε_n はゆるやかに変化する。これが、ジグザグ現象である。 $\cot^2 \alpha = 1, \varepsilon_0 = \lambda_0$ が安定を失いかどうかは、 $O(\tau^2)$ の項を調べる必要があるが、数値実験によると、安定ではないようである。たとえば $(z, \tan \alpha) = (1.0489177, 1.0437009)$ から出発させ、(5), (6)式を用いて計算してゆくと 100 回の繰返しを行つて、 $(1.1655843, 0.86463618)$ になるだけである。

§ 3 計算例

式(1)について実験した結果を表1表に示す。

No. of searches	$x=0.6929 \times 10^{-2}$ $y=0.4961 \times 10^{-2}$		$x=-1.2$ $y=1.0$	
	x	y	x	y
10	0.4099	0.1480	0.0624	0.0066
20	0.4613	0.2111	0.5109	0.2600
30	0.4938	0.2361	0.8852	0.7834
40	0.5069	0.2557	1.0017	1.0036
50	0.5265	0.2709	1.0000	1.0000

表1表

DSC法による極小化の過程，式(1)の場合

このように、ジブザグ現象は初期値により大きく変化する。
 表には参考のため $x = -1.2$, $y = 1.0$ から出発した場合
 も記入してある。極小化の探索の道程がより長いにもかかわ
 らず、速やかに収束している。次に三次元の例として

$$f = 100(z^2 + (y - z^2)^2) + (1 - x)^2 + (z - x + y)^2 \quad (12)$$

を二つの初期値から出発させた経過をオ2表, オ3表に示す。

No. of searches	x	y	z	d	f
0	-0.0300	-0.0700	0.0000		1.5651
1	0.0336	-0.0700	0.0000	0.0636	1.4505
2	0.0336	0.0014	0.0000	0.0714	0.9349
*3	0.0336	0.0014	0.0003	0.0003	0.9349
*4	0.0457	0.0150	0.0003	0.0181	0.9283
5	0.0625	0.0000	0.0003	0.0225	0.8841
6	0.0625	0.0000	0.0005	0.0002	0.8841
*13	0.3476	0.1047	0.0023	0.0806	0.5100
14	0.3424	0.1177	0.0023	0.0140	0.4823
15	0.3424	0.1177	0.0021	0.0001	0.4823
*16	0.4372	0.1761	0.0025	0.1112	0.4069
17	0.4304	0.1871	0.0023	0.0130	0.3833
18	0.4304	0.1871	0.0023	0.0000	0.3833
*43	0.9990	0.9979	-0.0002	0.0474	0.0000
44	0.9990	0.9979	-0.0002	0.0000	0.0000
45	0.9989	0.9980	0.0000	0.0002	0.0000
*46	0.9999	0.9999	-0.0000	0.0021	0.0000
47	0.9999	0.9999	-0.0000	0.0000	0.0000

オ2表

DSC法による極小化の過程, 式(12)の場合

初期値 $x = -0.03$, $y = -0.07$, $z = 0.0$

No. of searches	x	y	z	d	f
0	0.0090	-0.0600	0.0000		1.3478
1	0.0121	-0.0600	0.0000	0.0031	1.3428
2	0.0121	0.0002	0.0000	0.0602	0.9759
3	0.0121	0.0002	0.0001	0.0001	0.9759
*4	0.0121	0.0007	0.0001	0.0005	0.9759
5	0.1290	-0.0053	0.0001	0.1170	0.8250
6	0.1290	-0.0053	0.0012	0.0011	0.8249
*13	0.1703	0.0146	0.0016	0.0018	0.7328
14	0.1723	0.0319	0.0015	0.0173	0.7050
15	0.1723	0.0319	0.0012	0.0002	0.7050
*16	0.1725	0.0329	0.0012	0.0010	0.7049
17	0.1942	0.0281	0.0016	0.0222	0.6858
18	0.1942	0.0281	0.0019	0.0002	0.6858
*124	0.9878	0.9754	-0.0008	0.0129	0.0004
125	0.9876	0.9756	0.0000	0.0009	0.0002
126	0.9877	0.9756	0.0000	0.0000	0.0002
*127	0.9993	0.9986	-0.0003	0.0257	0.0000
128	0.9993	0.9986	-0.0000	0.0003	0.0000

表 3

DSC法による極小化の過程, 式(12)の場合

初期値 $x=0.009, y=-0.06, z=0.00$

表 2 は $(x, y, z) = (-0.03, -0.07, 0.0)$, 表 3 は $(x, y, z) = (0.009, -0.06, 0)$ から出発した場合である。*印をつけた探索は, 全変位 θ_1 の方向に対する θ_2 を示す。d は, 各探索による変位である。このように, 探索の経過は, 初期値によって変化し, 表 3 に示した場合に

は、ジグザグ現象によって収束が遅い。

ジグザグ現象を防ぐ一つの方法は、DSC法による n 回の直交する方向への探索の後で、前回の全変位方向への探索を追加することである。(3)式の場合には、この結果、次の探索方向は、谷の底の方向 (x 軸) と一致することがわかる。一般の場合には、これでジグザグ現象を完全に防ぐことは保証できない。実験によれば(2)式についてはオ4表に、(12)式についてはオ5表に示すように、収束するまでの繰返し回数を少なくすることが出来る。

No. of searches	$x=0.6929 \times 10^{-2}$ $y=0.4961 \times 10^{-2}$		$x=-1.2$ $y=1.0$	
	x	y	x	y
10	0.1557	0.0306	-0.0394	0.0060
20	0.8198	0.6646	0.7095	0.4942
30	1.0000	1.0000	1.0000	1.0000

オ4表

DSC法に附加的な探索を追加した場合の

極小化の過程，式(1)の場合

No. of searches	x	y	z	d	f
0	0.0090	-0.0600	0.0000		1.3478
1	0.0121	-0.0600	0.0000	0.0031	1.3428
2	0.0121	0.0002	0.0000	0.0602	0.9759
3	0.0121	0.0002	0.0001	0.0001	0.9759
4	0.0121	0.0007	0.0001	0.0005	0.9759
*5	0.1290	-0.0053	0.0001	0.1170	0.8250
6	0.1290	-0.0053	0.0012	0.0011	0.8249
7	0.1303	0.0185	0.0013	0.0239	0.7689
8	0.2291	0.0333	0.0023	0.0999	0.6688
*9	0.2263	0.0519	0.0032	0.0188	0.6289
10	0.2263	0.0520	0.0016	0.0163	0.6286
11	0.2650	0.0578	0.0020	0.0391	0.5979
12	0.4556	0.1879	0.0005	0.2307	0.4065
*25	0.9652	0.9349	-0.0029	0.0037	0.0042
26	0.9654	0.9349	0.0006	0.0036	0.0028
27	0.9946	0.9888	-0.0005	0.0613	0.0001
28	0.9995	0.9993	-0.0000	0.0115	0.0000

表 5

DSC法に附加的な探索を追加した場合の極小化の過程, 式(12)の場合
初期値 $x=0.009$, $y=-0.06$, $z=0.00$

参考文献

- (1) Rosenbrock, H.H. (1960). An automatic method for finding the greatest or least value of a function, Computer Journal, Vol.3, p.175-184.
- (2) Swann, W.H. (1964). Report on the development of a new direct search method of optimisation, I.C.I.Ltd., Central Instrument Laboratory Research Note 64/3.
- (3) Powell, M.J.D. (1964). An efficient method of finding the minimum of a function of several variables without calculating derivatives, Computer Journal, Vol.7, p.155-162.