S 次元最急降下法における連続体の性質に関して

尾関　孝史*

On the Properties of the Continuum in the S-Dimensional Steepest Descent Method

Takashi OZEKI*

ABSTRACT

In this report, it is analyzed the convergence of the S-dimensional steepest descent method. Particularly, it is studied about the case of S=2 in detail. A probability distribution was created from the gradient, and the movement of the probability distribution was visualized on a regular tetrahedron. It was then found that the invariant set consists of four subsets and that the regular tetrahedron is divided into four regions by the two interior subsets. In the numerical calculation, it was expected that if the initial point of the even point sequence was not a point in the invariant set, it would not converge to the interior invariant set of the regular tetrahedron. Moreover, it would move between specific regions of the four domains. These proofs are the subject of future work.

キーワード：最急降下法，勾配、不変集合、閉連続体

Keywords: Steepest Descent Method, Gradient, Invariant Set, Continuum

１．まえがき

最急降下法は最適化問題を解く最も基本的な数値計算法であり、初期近似解を逐次更新して、評価関数の最小解を求める。この方法は、記録によれば、コンピュータが開発される以前の 19 世紀半ばに Cauchy により提案されている[1]。最急降下法は高次元の微分を必要としない比較的簡単な反復法であるため、近年ではディープラーニングでも利用されている。しかし、この方法は大域的収束が保証されておらず、局所解に収束してしまうことがある。また、その収束は１次収束と遅いことが知られている。特にその条件数が悪い場合は、逐次近似点の移動方向はジグザグに細かく移動し、その解への収束は非常に遅くなる[2]。この現象は、Stiefel の鳥かごと呼ばれていたが、赤池が完全な証明を与えた[3]。そして、最近、このような状態を回避するため、探索方向のステップ幅を改善する方法が多く提案されている[4]。これらの方法は、Cauchy の最急降下法のステップ幅を改良する方法であり、探索方向が１次元のクリロフ部分空間となる最急降下法である[5]。

一方、最急降下法の探索空間を２次元以上のクリロフ部分空間に拡張した S 次元最急降下法が Forsythe らによって提案されている。そして、１次元の最急降下法（Cauchy S. D.）の場合と同様に、収束方向が次第に２方向になる現象が発生することが報告されている［6］。しかし、いまだにこの予想の完全な証明はされていない。なお。Forsythe は、S 次元最急降下法で以下の成果を得ている。

1. 初期点$\boldsymbol{x}_0$が 0 でない成分が S 個以下であれば、次の近似点$\boldsymbol{x}_1$は解となり、反復は１回で終了する。
2. 初期点$\boldsymbol{x}_0$が S+1 個以上の 0 でない成分を持てば、任意の近似点$\boldsymbol{x}_k$も S+1 個以上の 0 でない成分を持つ。この場合、有限回の反復で終了しない。また、その収束速度は１次である。
3. 近似点列$\{\boldsymbol{x}_{2k}\}$での正規化された最急降下方向の点列$\{\boldsymbol{y}_{2k}\}$の集積点集合はある閉連続体Rとなる。
4. 閉連続体Rの点は、S+1 から 2S の 0 でない成分を持つ。特に、両端の固有値の成分は 0 とはならない。

*情報工学科

5.　この閉連続体Rの各点は、S次元最急降下法の２回の反復で不変である。

Cauchy S.D.を拡張したＳ次元最急降下法は、逐次更新時での探索部分空間が大きいため、収束の向上が期待できる。しかし、そのためには、Ｓ次元最急降下法の探索方向の詳細な解明が必要である。Cauchy S.D.では、閉連続体の集積点は単点であるため、その探索方向の解明は容易であった。

そこで、本報告では、Ｓ次元最急降下法の閉連続体Rの性質を解明する。

２．Ｓ次元最急降下法の性質[5,6]

n次の正定値対称行列Aに対して、その二次形式

$$f(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^t A\boldsymbol{x} \tag{1}$$

の最小値問題を考える。ここで、行列Aは、n個の異なる正の固有値$0 < \lambda_1 < \lambda_2 < \cdots < \lambda_n$を持つものとする。Ｓ次元最急降下法では、逐次近似解が、漸化式

$$\boldsymbol{x}_{k+1} = \boldsymbol{x}_k + \sum_{i=1}^{s} \gamma_i^{(k)} A^i \boldsymbol{x}_k \tag{2}$$

が二次形式$f(\boldsymbol{x})$を最小になるようにステップ幅$\{\gamma_i\}$を定める。ここで、

$$\boldsymbol{y}_k = -\frac{A\boldsymbol{x}_k}{\|A\boldsymbol{x}_k\|} \tag{3}$$

と置くと、

$$\nabla f(\boldsymbol{x}) = A\boldsymbol{x} \tag{4}$$

であることから、$\boldsymbol{y}_k$は近似点$\boldsymbol{x}_k$における最急降下方向である。また、

$$N_i = (A^i\boldsymbol{y}, \boldsymbol{y}) \tag{5}$$

とし、$(s+1)$次モーメント行列を

$$D = \begin{pmatrix} N_0 & N_1 & \cdots & N_s \\ N_1 & \cdots & \cdots & N_{s+1} \\ \cdots & \cdots & \cdots & \cdots \\ N_s & N_{s+1} & \cdots & N_{2s} \end{pmatrix} \tag{6}$$

としたときの行列Dのi行、j列を除いたＳ次小行列式を$M_{i,j}$と置く。そして、

$$\beta_i = \frac{M_{i,s+1}}{M_{i+1,s+1}} \tag{7}$$

とする。すると、式(2)はＳ次直交多項式

$$P_s(t,\boldsymbol{y}) = t^s - \beta_1 t^{s-1} + \cdots + (-1)^s \beta_s \tag{8}$$

を用いて

$$\boldsymbol{x}_{k+1} = P_s(A,\boldsymbol{y}_k)\boldsymbol{x}_k \tag{9}$$

と表される。また、最急降下方向に対して、

$$\boldsymbol{y}_{k+1} = \frac{P_s(A,\boldsymbol{y}_k)\boldsymbol{y}_k}{\|P_s(A,\boldsymbol{y}_k)\boldsymbol{y}_k\|} \tag{10}$$

が成立する。ここで、分散を

$$V_s^{(k)} = \|P_s(A,\boldsymbol{y}_k)\boldsymbol{y}_k\|^2 \tag{11}$$

と置くと、

$$(\boldsymbol{y}_{k+1}, \boldsymbol{y}_k) = \frac{\sqrt{V_s^{(k)}}}{\sqrt{V_s^{(k+1)}}} \le 1 \tag{12}$$

であり、$V_s^{(k)}$は単調増加である。また、この分散に上界があるため、ある値V_sに収束する。そして、

$$\lim_{k\to\infty} \|\boldsymbol{y}_{k+2} - \boldsymbol{y}_k\| = 0 \tag{13}$$

が成り立つ。即ち、最急降下方向の偶数部分点列$\{\boldsymbol{y}_{2k}\}$は次第に変化が少なくなる。この現象は奇数部分点列$\{\boldsymbol{y}_{2k+1}\}$に関しても同様に示せる。更に、$S=1$即ち、Cauchy S.D.の場合は、これら２つの部分点列が収束し、近似解列$\{\boldsymbol{x}_k\}$がジグザグに解に近づいていくことが赤池によって証明されている[1]。しかしながら、$S>1$の場合の２つの部分点列の収束性は証明されていない。

偶数点列$\{\boldsymbol{y}_{2k}\}$の集積点の集合は式(13)の性質から、n次元ユークリッド空間内の単位球面S_n上の閉連続体R（continuum）となる。また、単位球面S_n上の点$\boldsymbol{y}$での変換Tを

$$T\boldsymbol{y} = \frac{P_s(A,\boldsymbol{y})\boldsymbol{y}}{\|P_s(A,\boldsymbol{y})\boldsymbol{y}\|} = \frac{P_s(A,\boldsymbol{y})}{\sqrt{V_s}}\boldsymbol{y} \tag{14}$$

とすると、$\boldsymbol{y}_{k+1} = T_k\boldsymbol{y}_k$が成立する。そして、その変換$T$の不変集合を

$$F(T) = \{\boldsymbol{r} \in S_n \mid T^2\boldsymbol{r} = \boldsymbol{r}\} \tag{15}$$

で定義する。即ち、不変集合を 2 回の変換Tで不変な単位球面上の点の集合とする。すると、

$$R \subset F(T) \tag{16}$$

となる。即ち、点列$\{\boldsymbol{y}_{2k}\}$のすべての集積点は不変集合に含まれる。更に Forsythe は数値実験により$S=2$の場合は$S=1$（即ち、Cauchy S.D.）の場合と同様に、この点列が収束し、この閉連続体が常に単点になることを数値実験から予想している。そこで、以下では、<u>$S=2$**の場合**</u>の不変集合の解析を行う。

３．$S=2$の場合の不変集合と特性多項式

点列$\{\boldsymbol{y}_{2k}\}$の任意の集積点を$\boldsymbol{r}=(r_i), (i=1,2,\cdots,n)$とすると、Forsythe の成果４により、S 次元最急降下法の集積点$\boldsymbol{r}$では、0 でない成分は S+1 個から２S 個の間であることが知られている。従って、特に、2 次元最急降下法では、集積点の 0 でない成分は３個または４個である。そこで、以下では、<u>**行列Aを最初から４次とし**</u>、その 4 個の固有値を$0<\lambda_1<\lambda_2<\lambda_3<\lambda_4$とする。また、座標軸を正値対称行列$A$の異なる固有ベクトル方向に定めておく。0 でない成分が３個の場合は、常に不変集合の点となるため[6]、以下では、<u>**0 でない成分が４個の場合**</u>のみを扱う。

2 次元最急降下法の 4 次の特性多項式を２つの 2 次直交多項式$P_2(t,\boldsymbol{y})$と$P_2(t,T\boldsymbol{y})$の積

$$Q_4(t)=P_2(t,\boldsymbol{y})P_2(t,T\boldsymbol{y}) \tag{17}$$

で定義する。すると、集積点rでは、特性多項式は２つの２次の直交多項式の積であることから、その４次の多項式は４つの実数α_iで因数分解できる。また特性多項式は、集積点での分散V_2を用いて、

$$\begin{aligned} Q_4(t) &= (t-\lambda_1)(t-\lambda_2)(t-\lambda_3)(t-\lambda_4)+V_2 \\ &= (t^2-\beta_1 t+\beta_2)(t^2-T\beta_1 t+T\beta_2) \\ &= (t-\alpha_1)(t-\alpha_2)(t-\alpha_3)(t-\alpha_4) \end{aligned} \tag{18}$$

と表すことができる[5]。ここで、

$$N_i=(A^i\boldsymbol{r},\boldsymbol{r})=\sum \lambda_i r_i^2 \tag{19}$$

$$\beta_1=\frac{N_3-N_1N_2}{N_2-N_1^2} \tag{20}$$

$$\beta_2=\frac{N_1N_3-N_2^2}{N_2-N_1^2} \tag{21}$$

が式（5）と（7）から得られる。また、式（17）から 0 でない成分が４個の場合は、

$$Q_2(\lambda_i)=V_2 (i=1,2,3,4) \tag{22}$$

が不変集合の点であるための必要条件であることがわかる。

４．不変集合の可視化

本節では不変集合の可視化を行う。成分が４個ある場合、3 次元空間での可視化は難しい。集積点$\boldsymbol{r}=(r_i)$が不変集合$F(T)$の点であれば、符号が異なる$(\pm r_i)$の組み合わせである16点は、特性多項式が同一のため、すべて不変集合の点となる。そこで、赤池[1]に従い、点$\boldsymbol{y}_k$の各成分の二乗を各成分に持つ点$\boldsymbol{p}=(p_i)=(y_i^2)$を導入する。$\sum p_i=1$だから、点列$\{\boldsymbol{p}_k\}$は確率変数$\lambda_i$上の変化する確率分布と見ることができる。また、この確率分布上の変換Sは

$$\boldsymbol{p}_{k+1}=S_k\boldsymbol{p}_k=(T_k)^2\boldsymbol{p}_k \tag{23}$$

と表すことができる。

もし、この点列$\{\boldsymbol{p}_{2k}\}$が収束したら、最急降下点列$\{\boldsymbol{y}_{2k}\}$の集積点は高々有限個となる。しかし、それらは非連続な孤立した点集合でから、式（13）により、点列$\{\boldsymbol{y}_{2k}\}$は Forsythe の予想通りにただ一つの集積点に収束することになる。そこで、以下では点列$\{\boldsymbol{p}_k\}$の収束性に関して解析する。

式（5）から、点$\boldsymbol{p}$に対して

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ \lambda_1 & \lambda_2 & \lambda_3 & \lambda_4 \\ \lambda_1^2 & \lambda_1^2 & \lambda_3^2 & \lambda_4^2 \\ \lambda_1^3 & \lambda_2^3 & \lambda_3^3 & \lambda_4^3 \end{pmatrix}\begin{pmatrix} p_1 \\ p_2 \\ p_3 \\ p_4 \end{pmatrix}=\begin{pmatrix} 1 \\ N_1 \\ N_2 \\ N_3 \end{pmatrix} \tag{24}$$

が成り立ち、左辺の行列はファンデルモンドの行列だから

$$p_i=\frac{N_3-c_2N_2+c_1N_1-c_0}{(\lambda_i-\lambda_j)(\lambda_i-\lambda_k)(\lambda_i-\lambda_l)} \tag{25}$$

を得る。ここで、

$$c_1=\lambda_j+\lambda_k+\lambda_l \tag{26}$$

$$c_1=\lambda_j\lambda_k+\lambda_k\lambda_l+\lambda_l\lambda_j \tag{27}$$

$$c_3=\lambda_j\lambda_k\lambda_l \tag{28}$$

である。また、式（20）と（21）から

$$N_3=\beta_1N_2-\beta_2N_1 \tag{29}$$

$$N_2 = \beta_1 N_1 - \beta_2 \tag{30}$$

が成り立つから、p_iはN_1, β_1, β_2で表すことができる。特に、β_1, β_2が決まれば、点$\boldsymbol{p}$は媒介変数N_1を持つ線分となる。

次に、4 次元の確率変数$\boldsymbol{p}$の可視化のための座標軸を3次元空間内に設定する。確率分布を表す正四面体を図１に示すように3次元空間に以下のように設置する。

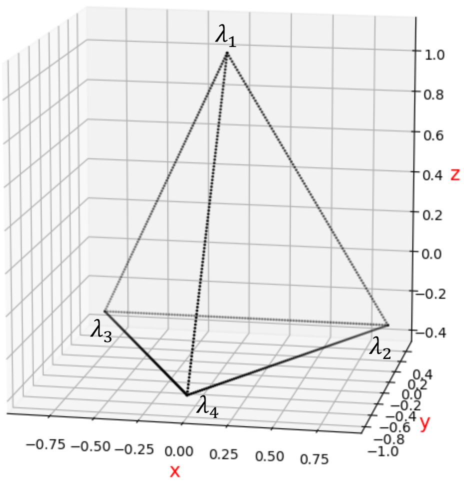

図１：確率分布の可視化

- $(1,0,0,0)$に対応する点を頂点$\left(0,0,\frac{3\sqrt{2}}{4}\right)$
- $(0,1,0,0)$に対応する点を頂点$\left(\frac{\sqrt{3}}{2},\frac{1}{2},-\frac{\sqrt{2}}{4}\right)$
- $(0,0,1,0)$に対応する点を頂点$\left(-\frac{\sqrt{3}}{2},\frac{1}{2},-\frac{\sqrt{2}}{4}\right)$
- $(0,0,0,1)$に対応する点を頂点$\left(0,-1,-\frac{\sqrt{2}}{4}\right)$

すると、一辺の長さが$\sqrt{3}$の正四面体であり、確率分布$\left(\frac{1}{4},\frac{1}{4},\frac{1}{4},\frac{1}{4}\right)$が原点$(0,0,0)$となる。各頂点ではある成分$p_i = 1$の確率分布であり、対面への垂線の足では、確率分布$p_i = 0, p_j = p_k = p_l = \frac{1}{3}$を表す。即ち、４個の各面ではある成分が０となる。

可視化の例

非常に特殊な例（固有値が対称的）ではあるが、

$$\lambda_1 = 1, \lambda_2 = 2, \lambda_3 = 3, \lambda_4 = 4$$

の場合を可視化する。この場合、不変集合$F(T)$は４つの部分集合から構成される。このうち最初の２つの部分集合は、０でない成分が３個のものである。Forsythe の成果４から、λ_1とλ_4の成分が０となることはない。従って、不変集合で、０でない成分が３個の点は、正四面体の$\lambda_1\lambda_2\lambda_4$（$p_3 = r_3 = 0$）と$\lambda_1\lambda_3\lambda_4$（$p_2 = r_2 = 0$）の２個の正三角形面上の内点である。残りの２個の部分集合は、０でない成分が４個のものである。この場合は、特性方程式

$$Q_4(t) = (t-1)(t-2)(t-3)(t-4) + V_2 = 0 \tag{31}$$

が４個の実数解をもつことから、$0 < V_2 \leq 1$となる。なお、直交多項式の性質から、その４つの実数解は、図２で示すように、$1 < \alpha_1 < \alpha_2 < 2 < 3 < \alpha_3 < \alpha_4 < 4$を満たす。特に、$V_2 = 1$の場合は、２つの重根で

$$\alpha_1 = \alpha_2 = \frac{1}{2}(5 - \sqrt{5}) \tag{32}$$

$$\alpha_3 = \alpha_4 = \frac{1}{2}(5 + \sqrt{5}) \tag{33}$$

となる。

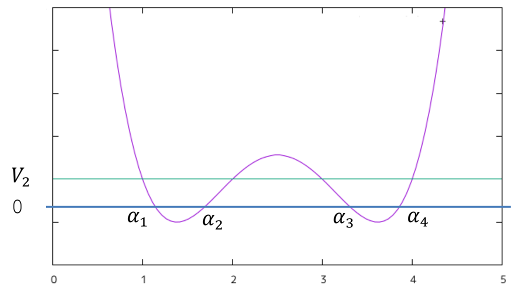

図２：特性方程式と分散V_2の関係

この結果、直交多項式$P_2(t) = t^2 - \beta_1 t + \beta_2$の$\beta_1$と$\beta_2$の決め方が以下の４種類ある。

1. $\beta_1 = \alpha_1 + \alpha_4, \beta_2 = \alpha_1 \cdot \alpha_4$
2. $\beta_1 = \alpha_2 + \alpha_3, \beta_2 = \alpha_2 \cdot \alpha_3$
3. $\beta_1 = \alpha_1 + \alpha_3, \beta_2 = \alpha_1 \cdot \alpha_3$
4. $\beta_1 = \alpha_2 + \alpha_4, \beta_2 = \alpha_2 \cdot \alpha_4$

$T\beta_1$と$T\beta_2$は残りの２つのα_iから自動的に決まる。上記の、１と２は固有値λ_iの対称性から$\beta_1 = 5$となる。そこで、以下では、１と２の場合と３と４の場合に分けて説明する。

（１）対称の場合（1.または 2.）

式（25）から、不変集合の点$\boldsymbol{p}$は、N_1とβ_2で表され

$$\begin{pmatrix} p_1 \\ p_2 \\ p_3 \\ p_4 \end{pmatrix} = \frac{1}{6}\begin{pmatrix} (6-\beta_2)(4-N_1) \\ 3(\beta_2-4)(3-N_1) \\ 3(\beta_2-4)(N_1-2) \\ (6-\beta_2)(N_1-1) \end{pmatrix} \tag{34}$$

となる。ここで、$2 \le N_1 \le 3$かつ$4 < \beta_2 < 6$である。図３はこれ不変集合を図にしたものである。正四面体を２つの領域に分割している。

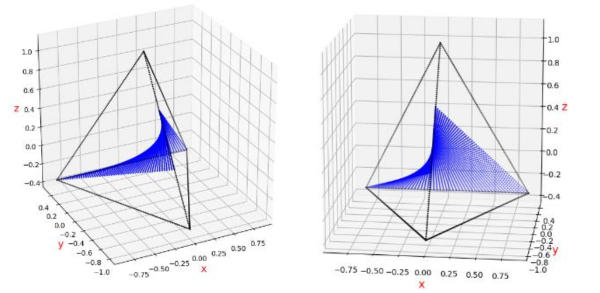

図３：$\beta_1 = 5$の場合の不変集合の部分集合

（２）非対称の場合（3.または 4.）

この場合、直交多項式$P_2(t) = t^2 - \beta_1 t + \beta_2$の頂点$\left(\frac{\beta_1}{2}, \beta_2 - \left(\frac{\beta_1}{2}\right)^2\right)$が$y = \left(x - \frac{5}{2}\right)^2 - \frac{5}{4}$上の点であるから、

$$\beta_2 = \frac{1}{2}(\beta_1^2 - 5\beta_1 + 10) \tag{35}$$

を満たす。従って、式（25）から、不変集合の点$\boldsymbol{p}$は、N_1とβ_1で表され

$$\begin{pmatrix} p_1 \\ p_2 \\ p_3 \\ p_4 \end{pmatrix} = \frac{1}{12}\begin{pmatrix} (\beta_1 - 1 - N_1)(6 - \beta_1)(7 - \beta_1) \\ 3(\beta_1 - 2 - N_1)(\beta_1 - 4)(7 - \beta_1) \\ -3(\beta_1 - 3 - N_1)(\beta_1 - 3)(6 - \beta_1) \\ -(\beta_1 - 4 - N_1)(\beta_1 - 3)(\beta_1 - 4) \end{pmatrix} \tag{36}$$

となる。ここで、$1 \le N_1 \le 4$かつ$4 < \beta_2 < 6$である。図４はこれ不変集合を図にしたものである。この場合も正四面体を２つの領域に分割している。

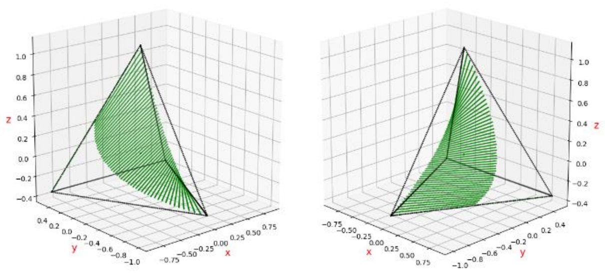

図３：$\beta_1 = 5$の場合の不変集合の部分集合

（３）共通の場合

$V_2 = 1$の場合、特性方程式は重根となり、$\alpha_1 = \alpha_2$かつ$\alpha_3 = \alpha_4$が成立する。そして、$\beta_1 = \beta_2 = 5$である。この場合、不変集合の部分集合は（１）と（２）の共通部分となる。従って、式（25）から、不変集合の点$\boldsymbol{p}$は、N_1のみで表され

$$\begin{pmatrix} p_1 \\ p_2 \\ p_3 \\ p_4 \end{pmatrix} = \frac{1}{6}\begin{pmatrix} 4 - N_1 \\ 3(3 - N_1) \\ 3(N_1 - 2) \\ (N_1 - 1) \end{pmatrix} \tag{37}$$

となる。ここで、$2 \le N_1 \le 3$である。図５の右図はこの不変集合を図にしたものである。正四面体の中心である原点$(0,0,0)$を通る線分である。

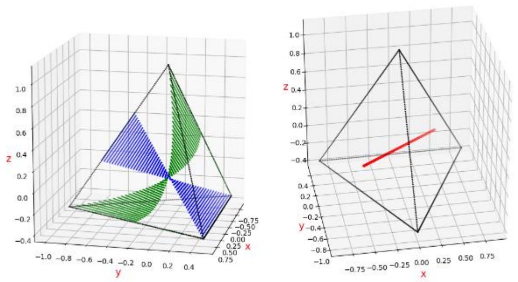

図５：確率分布$\boldsymbol{p}$の変換Sによる不動点集合

この線分上の確率分布$\boldsymbol{p}$は、変換Sの不動点である。即ち、

$$Sp_i = (\lambda_i^2 - \beta_1\lambda_i + \beta_2)^2 p_i = p_i \tag{38}$$

が成り立つ。

固有値が対称的という特殊な例ではあるが、確率分布が正四面体で表され、不変集合によって、４個の領域に分割されることがわかった。

図６は、正四面体を$p_2 = 0.3$で切断した断面で、p_2及びp_3成分が２回のS変換の後にどう変化するかを表した図である。青い領域はp_2成分が減少し、緑の領域はp_2成分が増加する。p_3成分に関してはその逆となる。赤い領域は、２ 回の変換Tでr_2成分の符号が反転する領域であり、式（13）から次第に現れなくなる領域である。青と緑の境界は、２個の不変集合と一致している。

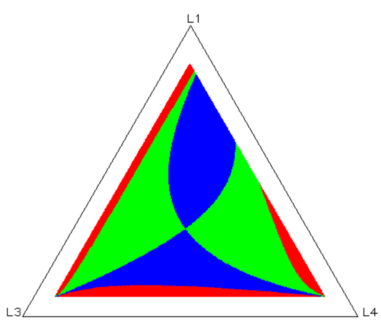

図６：$p_2 = 0.3$でのp_2及びp_3成分の増減

図７は、不動点$(0.25, 0.25, 0.25, 0.25)$の近くに初期点$\boldsymbol{p}_0 = (0.24, 0.26, 0.24, 0.26)$を定めて、100 万回の反復時の移動の様子を調べたものである。青い点が偶数点列であり、赤い点が奇数点列である。茶色の線分は図５の不動点集合である。正四面体内の不動点には収束せずに、正四面体の面$\lambda_1\lambda_3\lambda_4$（$p_2 = 0$）の方向に不動点集合にほぼ平行に移動している。

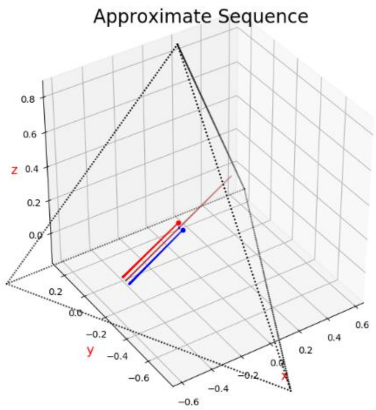

図７：偶数点列と奇数点列の移動の様子

多くの数値計算の結果によると、不変集合に近い点ほど、１回の変換での移動距離が短くなる。また、確率変数の偶数点列$\{\boldsymbol{p}_{2k}\}$の初期点が不変集合内の点でなければ、正四面体の内部の不変集合には収束せず、４つの領域の特定の領域間を移動する。しかも、その特定の領域は、$p_2 \to 0$または$p_3 \to 0$のどちらかであり、その成分が単調減少する領域に限られる。このことから、点列$\{\boldsymbol{p}_{2k}\}$が正四面体の２つの面$\lambda_1\lambda_2\lambda_4$（$p_3 = 0$）または$\lambda_1\lambda_3\lambda_4$（$p_2 = 0$）の２つのどちらかに収束することが予想される。この予想が正しければ、Forsythe の予想も証明されたことになる。

５．まとめ

本報告では、S 次元最急降下法の収束の様子を解析した。特に、$S = 2$の場合を詳しく調べた。勾配から確率分布を作成し、正四面体に確率分布の移動の様子を可視化した。そして、固有値が対称の場合は、不変集合が４個の部分集合からなり、内部の２個の部分集合によって、正四面体が４個の領域に分割されることがわかった。数値計算では、確率変数の偶数点列$\{\boldsymbol{p}_{2k}\}$の初期点が不変集合内の点でなければ、正四面体の内部の不変集合には収束せず、４個の領域の特定の領域間を移動した。今後の課題として、これらの証明が残されている。

参考文献

[1] A. Cauchy: Méthode générales pour la résolution des systems de’équations simultanées, Comp. Rend. Acad. Sci. Paris, Vol 25, 536-538, (1847).

[2] 今野浩，山下浩：非線形計画法，(1978).

[3] H. Akaike: On a Successive Transformation of Probability Distribution and Its Application to the Analysis of the Optimum Gradient Method, Ann. Inst. Stat. Math., vol.~11, pp.~1--16, (1959).

[4] J. Barzilai, J. M. Borwein, Two-point step size gradient methods, IMA J.Numer. Anal., Vol 8, pp. 141-148, (1088).

[5] 尾関孝史：S 次元最急降下法における探索方向の漸近的な振る舞いに関して，福山大学工学部紀要，Vol. 42，pp.794-138，(2019).

[6] G. E. Forsythe: On the Asymptotic Directions of the s-Dimensional Optimum Gradient Method, Numer. Math., vol. 11, pp. 57-76, (1968).