論　文

特異点解消とニューラルネットワークのベイズ推定における汎化誤差

青柳　美輝[†a)]　　渡辺　澄夫[††b)]

Resolution of Singularities and the Generalization Error with Bayesian Estimation for Layered Neural Network

Miki AOYAGI[†a)] and Sumio WATANABE[††b)]

あらまし　階層型ニューラルネットワークのように，入力からも出力からも直接には働きを定められていない隠れた部分をもつ学習モデルは，パラメータが特定不能になり，フィッシャー計量が縮退するという特徴をもつ．このようなモデルを特異モデルと呼ぶ．論文 [1]〜[3] において，学習モデルのゼータ関数の極が，特異モデルのベイズ推測に関する確率的複雑さの漸近形を与えることが示されている．ここで学習モデルのゼータ関数は，カルバック情報量と事前分布より定まる 1 変数複素関数である．しかし，今まで，計算が複雑なため，数例において，その漸近形の主要項の上限は得られていたが，値は求められていなかった．この論文では，階層型ニューラルネットワークにおいての主要項を求めるための計算方法を示し，確率的複雑さの厳密な漸近展開を与える．

キーワード　確率的複雑さ，階層モデル，特異モデル，ベイズ推定，特異点解消

1. まえがき

階層型ニューラルネットワークや混合正規分布，ボルツマンマシンなどの学習モデルは，パラメータが特定不能になり，フィッシャー計量が縮退するという特徴をもつ．例えば，三層ニューラルネットワークにおいて，パラメータ w が，学習モデルよりもより少ない隠れユニット数をもつモデルを表現する時，フィッシャー情報行列 $I(w)$ が特異（$\det I(w) = 0$）になる．すなわち，小さなモデルを表現するパラメータが大きなモデルを表現するパラメータの解析的集合となる．このようなモデルを特異モデルと呼ぶ．特異モデルには，正則モデルに用いられる方法，例えば，モデル選択法である AIC [4], TIC [5], HQ [6], NIC [7], BIC [8], MDL [9] などは適用できない．したがって，様々な情報工学などの問題に適用されている特異モデル解析の研究は，早急に必要である．

近年，ベイズ推測において重要な役割を果たす確率的複雑さが，次のように特異点解消により構成された方法論と数理的に緊密な関係をもつことが明らかになった [1]〜[3]．サンプル数が n のときの学習モデルの確率的複雑さ $F(n)$ は，ある有理数 λ と自然数 ℓ が存在して $F(n) = \lambda \log n - (\ell - 1) \log\log n + O(1)$ となる．ここで $O(1)$ は n の有界な関数である．ベイズ推測における汎化誤差（真の分布からベイズ推測された分布までのカルバック情報量の平均値）$G(n)$ は，漸近展開可能であれば，確率的複雑さの差分 $F(n+1) - F(n)$ に等しいものであるが，上記の定数 λ 及び ℓ を用いて $G(n) \cong \lambda/n - (\ell - 1)/(n \log n)$ となる．この λ 及び ℓ が，学習モデルのゼータ関数より決まり，特異点解消の方法を用いて計算される量である．ここで，学習モデルのゼータ関数とは，カルバック情報量と事前分布より定まる 1 変数複素関数である．正則な統計モデルでは，$\lambda =$ パラメータ数$/2$, $\ell = 1$ である．特異モデルのベイズ推測に関する学習曲線は，たくさんの特異点が局所的に引き起こすすべての学習曲線の下側にあり，この結果，正則な統計モデルよりも常に優れた学習が実現されている．現在，λ を確率的に計算する方法が提案されたり [10]，三層ニューラルネットワークの場合 [11] や縮小ランク回帰モデル [12] の場合に，λ の上限が得られているが，実はこのような数理の確立

† 上智大学理工学部，東京都
Faculty of Science and Technology, Sophia University, 7–1 Kioi-cho, Chiyoda-ku, Tokyo, 102–8554 Japan
†† 東京工業大学精密工学研究所，横浜市
Precision and Intelligence Laboratory, Tokyo Institute of Technology, 4259 Nagatsuda, Midori-ku, Yokohama-shi, 226–8503 Japan
a) E-mail: miki-a@sophia.ac.jp
b) E-mail: swatanab@pi.titech.ac.jp

にもかかわらず，計算を行うことは非常に困難で，λ の値そのものを計算することはまだ行われていなかった．実際，ゼータ関数の極の値に関しては，数学の分野においても概均質ベクトル空間などの特別な場合にしか研究されていないのが実情である．また，特異点解消は，広中の定理により，原理的には有限の手続きにより行うことが知られており，その存在が保証されているが，具体的に求めるのは難しいとされている．更に情報学における学習モデルのカルバック情報量は中間ユニット数などのパラメータ（例えば，**4.** における式 (3) の p）を含んでいるため，確定された多項式の特異点解消よりも，高度な面を含んでいる．

この論文では，中間ユニット数のようなパラメータを含む学習モデルの確率的複雑さの主要項の係数 λ を求めるために，帰納的に特異点の解消を行う方法を提案し，三層ニューラルネットワークへの応用を通して，その有用性を明らかにする．

本論文で提案する方式の情報論的な意義としては，まず，本論文の方法により初めて，三層パーセプトロンの確率的複雑さの挙動が解明でき，複雑な学習モデルの性質が，他の単純なモデルとは，どのように異なるのかについて科学的な知識が得られたということである．次に，三層パーセプトロンのように複雑な学習モデルの選択やハイパパラメータの設計において，周辺ゆう度の値を理論計算することができなかったので，特異点をもつ学習モデルに MCMC 法を適用した場合，どの程度正しい値が得られているかを確認することができなかった．三層パーセプトロンの周辺ゆう度の理論値が得られたことにより，今後，MCMC 法の精度を解明することや，MCMC 法の精度の改良法などを考察すること，更には事後分布の近似に優れた学習アルゴリズムを作るときの基盤が作られることになる．

本論文の構成は次のとおりである．**2.** では，ベイズ学習理論の枠組みを，**3.** では特異点解消について説明し，**4.** では主目的である三層型ニューラルネットワークにおいての主要項 λ を求める．

2. ベイズ学習理論

ここでは，論文 [1]〜[3] において得られているベイズ学習理論の概略を説明する．

入出力の空間を $x \in \mathbb{R}^N$，パラメータの空間を $w \in W \subset \mathbb{R}^d$ とする．学習モデル $p(x|w)$ とその事前分布 $\psi(w)$ が与えられているものとし，真の分布は学習モデルに含まれているものとする．その真の分布を $p(x|w_0)$ と仮定する．$p(x|w_0)$ に従う独立なサンプルを $X^n = (X_1, X_2, ..., X_n)$ とすれば，事後確率 $p(w|X^n)$ を $p(w|X^n) = 1/Z_n \psi(w) \prod_{i=1}^n p(X_i|w)$，によって定める（$Z_n$ は正規化定数）．また，ベイズ推測は $p(x|X^n) = \int p(x|w)p(w|X^n)dw$, と定義される．このとき，汎化誤差 $G(n)$ は，サンプル X^n に関する平均である $E_n\{\cdot\}$ を用いて，

$$G(n) = E_n\left\{\int p(x|w_0)\log\frac{p(x|w_0)}{p(x|X^n)}dx\right\}, \quad (1)$$

で定義される．カルバック情報量を $K(w) = \int p(x|w_0)\log(p(x|w_0)/p(x|w))dx$, 経験カルバック情報量を $K_n(w) = 1/n\sum_{i=1}^n \log(p(X_i|w_0)/p(X_i|w))$, とおく． 確率的複雑さを

$$F(n) = -E_n\left\{\log\int \exp(-nK_n(w))\psi(w)dw\right\}, \quad (2)$$

とすれば，$G(n) = F(n+1) - F(n)$, が成り立つことが分かっている．

学習モデルのゼータ関数を

$$J(z) = \int K(w)^z \psi(w)dw,$$

とおけば，この関数 $J(z)$ の最も原点に近い極を $-\lambda$，その位数を ℓ とすると，

$$F(n) = \lambda\log n - (\ell-1)\log\log n + O(1),$$
$$G(n) \cong \frac{\lambda}{n} - \frac{\ell-1}{n\log n},$$

が成り立つ．ここで，$O(1)$ は n の有界な関数である．この λ 及び ℓ が $K(w) = 0$ の近傍での特異点解消によって求められる．

3. 特異点解消

この章では特異点解消定理と blow up 操作について説明し，学習理論への応用について述べる．

最初に，Atiyah [14] の論文で利用された，特異点解消の広中の定理 [13] の解析的バージョンを紹介する．

［定理］ $f(x)$ を $\mathbb{R}^n$ において，原点 0 の近傍で定義された解析関数とする．このとき，原点の近傍 V と，ある多様体 U と，U から V への固有（コンパクト集合の逆像もコンパクト）な解析関数 μ が存在して，次の (1)，(2) が成立する．

(1) $E = \mu^{-1}(f^{-1}(0))$ に対して $\mu : U \setminus E \to V \setminus f^{-1}(0)$ は同型を与える．

(2) 任意の $u \in U$ について適当な局所座標 $(u_1, \cdots, u_n)$ が存在して，$f(\mu(u)) = \pm u_1^{s_1} u_2^{s_2} \cdots u_n^{s_n}$ が成り立つ．ここで $s_1, \cdots, s_n$ は非負の整数．

この定理を，カルバック情報量 $K(w)$ に適用すれば，任意の $w \in K^{-1}(0) \cap W$ に対して，ある近傍 V_w と，ある多様体 U_w が存在して，上記の (1)，(2) を満たす，U_w から V_w への固有な解析関数 μ_w が存在する．このとき，学習モデルのゼータ関数 $J(z) = \int K(w)^z \psi(w) dw$ の V_w での局所的な積分は

$$\begin{aligned} J_w(z) &= \int_{V_w} K(w)^z \psi(w) dw \\ &= \int_{U_w} |u_1^{s_1} u_2^{s_2} \cdots u_n^{s_n}|^z \psi(\mu_w(u)) |\mu_w'(u)| du, \end{aligned}$$

となる．この積分は初等的に求めることができる．$w \in W \setminus K^{-1}(0)$ に対しては，ある近傍 V_w が存在して，$K(w') \neq 0, w' \in V_w$．このとき，$J_w(z) = \int_{V_w} K(w)^z \psi(w) dw$ は極を持たない．パラメータ空間 W がコンパクトのときは，結果として $J(z)$ の極，位数を得ることができる．

次に，blow up 操作について紹介する．

ここでは，この論文で使用する，特異点解消を求めるための主な方法である，多様体に沿った blow up 操作について説明する．

多様体 M を k 個の開集合 $U_i \cong \mathbb{R}^n$，$i = 1, 2, \cdots, k (n \geq k)$ の次のような貼合せで定義する．U_i の座標を $(\xi_{1i}, \xi_{2i}, \cdots, \xi_{ni})$ とする．U_i，U_j を $\xi_{ji} \neq 0$，$\xi_{ij} \neq 0$ において，

$$\begin{aligned} &\xi_{ij} = 1/\xi_{ji}, \qquad \xi_{jj} = \xi_{ii}\xi_{ji}, \\ &\xi_{hj} = \xi_{hi}/\xi_{ji}, 1 \leq h \leq k, h \neq i, j, \\ &\xi_{lj} = \xi_{li}, k+1 \leq l \leq n, \end{aligned}$$

で貼り合わせる．すなわち M は k 枚の $\mathbb{R}^n$ からできている．$\sim$ を貼合せから得られる同値関係とすると，$M = \coprod_{i=1}^k U_i / \sim$ となる．

$\pi : M \to \mathbb{R}^n$ を U_i 上において

$$\begin{aligned} U_i \ni (\xi_{1i}, \xi_{2i}, \cdots, \xi_{ni}) \mapsto (&\xi_{ii}\xi_{1i}, \cdots, \xi_{ii}\xi_{i-1i}, \\ &\xi_{ii}, \xi_{ii}\xi_{i+1i}, \cdots, \xi_{ii}\xi_{ki}, \xi_{k+1i}, \cdots, \xi_{ni}), \end{aligned}$$

で定義する．これは矛盾なく定義され，この写像を

$$\begin{aligned} N = \{&(x_1, \cdots, x_k, x_{k+1}, \cdots, x_n) \in \mathbb{R}^n | \\ &x_1 = \cdots = x_k = 0\}, \end{aligned}$$

に沿った blow up という．

このとき

(1) $\pi : M \to \mathbb{R}^n$ は固有写像であり，

(2) $\pi : M \setminus \pi^{-1}(N) \to \mathbb{R}^n \setminus N$ は同型写像となる．

4. 多層神経回路網の学習曲線

この章では，blow up 列を用いて三層型ニューラルネットワークの学習モデルゼータ関数の極を求める．

入力ユニット 1 個，中間ユニット p 個，出力ユニット 1 個の三層パーセプトロンを考察する．入力値を x，出力値を y，パラメータを $w = \{a_m, b_m | m = 1, \cdots, p\}$ とし，

$$f(x, w) = \sum_{m=1}^{p} a_m \tanh(b_m x),$$

と定義する．関数 f を用いて，学習モデルを，典型的ニューラルネットワークモデル

$$p(y|x, w) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}(y - f(x, w))^2\right),$$

とする．入力 x の密度関数 $q(x)$ は $[-1, 1]$ の一様分布とする．事前分布 $\psi(w)$ は C^∞ 関数でコンパクトな台 W をもち，$\psi(0) > 0$ であるとする．真の分布は $p(y|x, w_0) = \frac{1}{\sqrt{2\pi}} \exp(-\frac{1}{2} y^2)$，すなわち真の分布を与えるパラメータ集合がすべて $a_m = 0, b_m = 0$ の場合を含むものを考える．

この学習モデルのカルバック情報量は

$$K(w) = \frac{1}{2} \int_{-1}^{1} f(x, w)^2 dx,$$

となるが，テイラー展開などを用いると，次の関数

$$\int_W \left\{ \sum_{n=1}^{p} \left(\sum_{m=1}^{p} a_m b_m^{2n-1} \right)^2 \right\}^z \prod_{m=1}^{p} \mathrm{d}a_m \mathrm{d}b_m, \quad (3)$$

の原点に最も近い極が λ を与え，その位数が ℓ を与える．また，[15] では次のような評価が得られている．

$$\sum_{m=1}^{p} \frac{1}{4m - 2} \leq \lambda \leq \frac{\sqrt{p}}{2}.$$

この章では，微分形式

$$\Psi = \left\{ \sum_{n=1}^{p} \left(\sum_{m=1}^{p} a_m b_m^{2n-1} \right)^2 \right\}^z \prod_{m=1}^{p} \mathrm{d}a_m \mathrm{d}b_m,$$

の blow up を考察し，次の定理を証明する．

［主定理］ $i^2 \le p$ となる最大の i について

$$\lambda = \frac{p+i^2+i}{4i+2},\ \ell = \begin{cases} 2 & (i^2 = p) \\ 1 & (i^2 < p) \end{cases}$$

が成り立つ．

この論文で用いる記号の使用方法を述べる．例えば，

$b_1 = v_1, b_m = v_1 b_m$, で変換すれば，

という文章は，変換前と変換後の記号を区別して，

$b_1 = v_1, b_m = v_1 b'_m$, で変換すれば，

と書くべきであるが，blow up 操作を繰り返すことによる複雑さを避けるため，同じ記号を用る．

主定理の証明は，特異点解消を求める証明 1 と極の大きさの比較を行う証明 2 の 2 段階に分けて行う．

［主定理の証明 1］

部分多様体 $b_1 = 0, \cdots, b_p = 0$ で blow up する．

多様体 M を $M = \coprod_{i=1}^p U_i/\sim$，$U_1$ の座標を $(a_1, \cdots, a_p, v_1, b_2, \cdots, b_p)$ として，blow up すれば，U_1 上では $\{b_1 = v_1, b_m = v_1 b_m,\ m = 2, \cdots, p\}$ で変換すればよいので，

$$\Psi = \left\{ \sum_{n=1}^p v_1^{4n-2} \left(a_1 + \sum_{m=2}^p a_m b_m^{2n-1} \right)^2 \right\}^z v_1^{p-1}$$
$$\mathrm{d}a_1 \mathrm{d}v_1 \prod_{m=2}^p \mathrm{d}a_m \mathrm{d}b_m,$$

となる．

対称性より $\{b_i = v_i, b_m = v_i b_m,\ m = 1, 2, \cdots, p, m \neq i\}$ で変換しても同じ形の式が出て，結果として積分したときの極も同じになるので，U_1 上のみで考えればよい．

$d_1 = a_1 + \sum_{m=2}^p a_m b_m$ とおくと，

$$\Psi = \left\{ v_1^2 \left(d_1^2 + \sum_{n=2}^p v_1^{4n-4} \cdot \left(d_1 - \sum_{m=2}^p a_m b_m + \sum_{m=2}^p a_m b_m^{2n-1} \right)^2 \right) \right\}^z$$
$$\cdot v_1^{p-1} \mathrm{d}d_1 \mathrm{d}v_1 \prod_{m=2}^p \mathrm{d}a_m \mathrm{d}b_m.$$

ここで，次のような関数を定義しておく．

$$f_{n,l}(x) = \sum_{j_2+\cdots+j_l=0}^{n-l} b_2^{2j_2} \cdots b_{l-1}^{2j_{l-1}} x^{2j_l} = 1 + \cdots \ge 1.$$

特にこの関数は

$$f_{n,l}(b_m) - f_{n,l}(b_l) = (b_m^2 - b_l^2) f_{n,l+1}(b_m),$$

を満たしている．

$$c_2 = \sum_{m=i}^p a_m b_m (b_m^2 - 1),$$

及び，$i \ge 3$ に対して

$$c_i = \sum_{m=i}^p a_m b_m (b_m^2 - 1)(b_m^2 - b_2^2) \cdots (b_m^2 - b_{i-1}^2),$$

とおき，$f_{n,l}$ を用いて，Ψ を表せば，

$$\Psi = \left\{ v_1^2 \left(d_1^2 + \sum_{n=2}^p v_1^{4n-4} (d_1 + f_{n,2}(b_2) c_2 + \cdots + f_{n,i}(b_i) c_i + \cdots + f_{n,n}(b_n) c_n)^2 \right) \right\}^z$$
$$\cdot v_1^{p-1} \mathrm{d}d_1 \mathrm{d}v_1 \prod_{m=2}^p \mathrm{d}a_m \mathrm{d}b_m, \tag{4}$$

となる．

明らかに，$d_1^2 + \sum_{n=2}^p v_1^{4n-4} (d_1 + f_{n,2}(b_2) c_2 + \cdots + f_{n,i}(b_i) c_i + \cdots + f_{n,n}(b_n) c_n)^2 = 0$ は $d_1 = c_2 = \cdots = c_p = 0$ と同値である．

以下，$f_{n,i}(b_i)$ の b_i は省略する．また，$J^{(\alpha)}$ を $\mathbb{R}^\alpha$ の元とする．煩雑さを避けるため $^{(\alpha)}$ を省略し J と記すこともある．$J^{(\alpha)} > J^{(\alpha')} (\alpha \ge \alpha')$ で，$J^{(\alpha)} = (J^{(\alpha')}, *)$ であることを示し，$J^{(\alpha)} = 0^{(\alpha)}$ または $J^{(\alpha)} = 0$ で，$J^{(\alpha)} = (0, \cdots, 0)$ を表す．

一連の blow up によって次の形になることを k, K, α に関する帰納法で示す．証明に使われる過程を図 1 に示す．

［帰納的仮定］

任意の $J \in \mathbb{R}^\alpha$ に対して，

$s(J) = \#\{m; k \le m \le p, J_m^{(\alpha)} = J\}$,

$s(i, J) = \#\{m; k \le m \le i-1, J_m^{(\alpha)} = J\}$,

とする．# は集合の個数を表す．

(a) $K \ge k$,

(b)

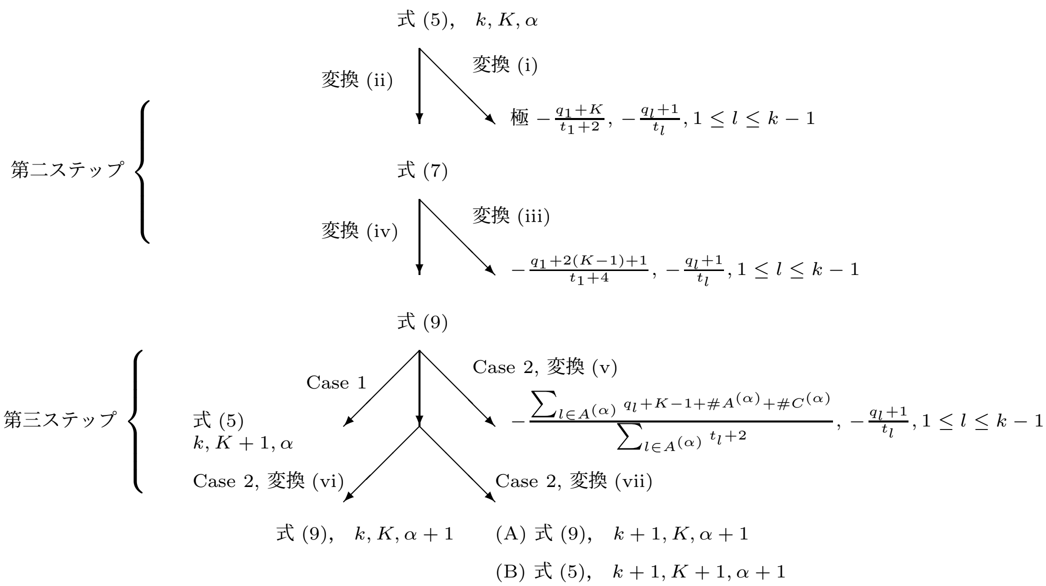

図 1 主定理の証明の流れ
Fig. 1 The flow chart of main theorem.

$$
\begin{aligned}
\Psi = \Bigg\{ & v_1^{t_1} v_2^{t_2} v_3^{t_3} \cdots v_{k-1}^{t_{k-1}} \Bigg(d_1^2 + (d_1 v_1^2 + d_2)^2 \\
& + \cdots + (d_1 v_1^{2K-4} + d_2 v_1^{2K-6} f_{K-1,2} + \cdots \\
& + d_{K-1} f_{K-1,K-1})^2 + \sum_{n=K}^{p} v_1^{4n-4K+4} \Bigg(d_1 v_1^{2K-4} \\
& + d_2 v_1^{2K-6} f_{n,2} + \cdots + d_{K-1} f_{n,K-1} \\
& + \sum_{i=K}^{p} f_{n,i} c_i \Bigg)^2 \Bigg) \Bigg\}^z \\
& \cdot \prod_{m=1}^{k-1} v_m^{q_m} \prod_{m=1}^{K-1} \mathrm{d}d_m \prod_{m=K}^{p} \mathrm{d}a_m \prod_{m=1}^{k-1} \mathrm{d}v_m \prod_{m=k}^{p} \mathrm{d}b_m . \quad (5)
\end{aligned}
$$

ここで，t_l，q_m は非負整数を表し，c_i は，ある適当な $RJ^{(\alpha)} \subset \mathbb{R}^\alpha$，非負整数 $t(i, J, l)$, 0 とならない関数 $g(i, m)$, $(K \le i \le p, 2 \le l \le k-1, i \le m \le p)$ が存在して，

$$
\begin{aligned}
c_i = & v_2^{t(i,0,2)} v_3^{t(i,0,3)} \cdots v_{k-1}^{t(i,0,k-1)} \\
& \cdot \sum_{\substack{i \le m \le p \\ J_m^{(\alpha)}=0}} g(i,m) a_m b_m \prod_{\substack{k \le i' < i \\ J_{i'}^{(\alpha)}=0}} (b_m^2 - b_{i'}^2) \\
& + \sum_{J \in RJ^{(\alpha)}} v_2^{t(i,J,2)} v_3^{t(i,J,3)} \cdots v_{k-1}^{t(i,J,k-1)} \\
& \cdot \sum_{\substack{i \le m \le p \\ J_m^{(\alpha)}=J}} g(i,m) a_m b_m \prod_{\substack{k \le i' < i \\ J_{i'}^{(\alpha)}=J}} (b_m - b_{i'}) \\
& + \sum_{J \notin RJ^{(\alpha)}, J \ne 0} v_2^{t(i,J,2)} v_3^{t(i,J,3)} \cdots v_{k-1}^{t(i,J,k-1)} \\
& \cdot \sum_{\substack{i \le m \le p \\ J_m^{(\alpha)}=J}} g(i,m) a_m \prod_{\substack{k \le i' < i \\ J_{i'}^{(\alpha)}=J}} (b_m - b_{i'}),
\end{aligned}
$$

となる．

(c) $k \le i < K$ に対して，$J_i^{(\alpha)}$ はすべて異なり，$J_i^{(\alpha)} \notin RJ^{(\alpha)}, J_i^{(\alpha)} \ne 0$ である．

(d) $J \in \mathbb{R}^\alpha$, $K \le i \le p, 2 \le l \le k-1$ に対して，$\tilde{t}(i, J, l) := t_l/2 + t(i, J, l)$ とおく．整数 $D_{J^{(\mu)},l} \ge 0$ が存在して

$$
\begin{aligned}
\tilde{t}(i, J, l) = & \sum_{J > 0^{(\mu)}} D_{0^{(\mu)},l} \quad (2s(i, 0^{(\mu)}) + 1) \\
& + \sum_{\substack{J > J^{(\mu)} \\ J^{(\mu)} \in RJ^{(\mu)}}} D_{J^{(\mu)},l} (s(i, J^{(\mu)}) + 1) \\
& + \sum_{\substack{J > J^{(\mu)} \\ J^{(\mu)} \notin RJ^{(\mu)}, J^{(\mu)} \ne 0}} D_{J^{(\mu)},l} s(i, J^{(\mu)}).
\end{aligned}
$$

(e) 整数 $g_l \ge 0$，$\eta_{k',l}^{(\xi)} \ge 0$ $(2 \le k' \le K-1, 1 \le$

$\xi \le g_l, 2 \le l \le k-1)$ が存在して，

$$\frac{t_l}{2} = \sum_{\xi=1}^{g_l}(1+\eta_{2,l}^{(\xi)}+\cdots+\eta_{K-1,l}^{(\xi)}),$$

$$0 \le \eta_{2,l}^{(\xi)} \le 2, 0 \le \eta_{2,l}^{(\xi)} + \eta_{3,l}^{(\xi)} \le 4,$$

$$\vdots$$

$$0 \le \eta_{2,l}^{(\xi)} + \eta_{3,l}^{(\xi)} + \cdots + \eta_{K-1,l}^{(\xi)} \le 2(K-2).$$

(f)　$\varphi_l^{(\xi)} := p + 2\eta_{2,l}^{(\xi)} + \cdots + (K-1)\eta_{K-1,l}^{(\xi)}$ とおく．整数 $\phi_l \ge 0 (2 \le l \le k-1)$ が存在して，$g_l \le \sum_{J_m^{(\alpha)} > J^{(\mu)}} D_{J^{(\mu)},l}$ かつ

$$q_l + 1 = \sum_{\xi=1}^{g_l} \varphi_l^{(\xi)} + \phi_l + \sum_{m=k}^{p}(-g_l + \sum_{J_m^{(\alpha)} > J^{(\mu)}} D_{J^{(\mu)},l}).$$

（帰納的仮定終）

帰納法の仮定 (d) から (f) は，極の大きさを比べるときに必要なものである．

すべてのパラメータは証明の中で随時与えられ，または更新されていく変数である．

特にすべて $J_m^{(\alpha)} = 0$ のとき，$\alpha = k-1$ で

(a′)　$k = K$.

(b′)　$c_i (k \le i \le p)$ は

$$c_i = v_2^{t(i,0,2)} v_3^{t(i,0,3)} \cdots v_{k-1}^{t(i,0,k-1)} \cdot \sum_{i \le m \le p} a_m b_m \prod_{k \le i' < i} (b_m^2 - b_{i'}^2),$$

となる．

(d′)　$D_{0^{(l-1)},l} = 1$, 他は 0.

$$\tilde{t}(i, 0^{(k-1)}, l) = D_{0^{(l-1)},l}(2(i-l)+1) = 2(i-l)+1.$$

(e′)　$\frac{t_l}{2} = 1 + \eta_{2,l}^{(1)} + \cdots + \eta_{k-1,l}^{(1)}$,

$$\eta_{2,l}^{(1)} = 0, \ldots, \eta_{l-1,l}^{(1)} = 0,$$
$$\eta_{l,l}^{(1)} = 2, \ldots, \eta_{k-2,l}^{(1)} = 2,$$
$$0 \le \eta_{k-1,l}^{(1)} \le 2, t(k, 0^{(k-1)}, l) + \eta_{k-1,l}^{(1)} = 2.$$

(f′)　$\varphi_l^{(1)} := p + 2\eta_{2,l}^{(1)} + \cdots + (k-1)\eta_{k-1,l}^{(1)}$ に対して，$q_l + 1 = \varphi_l^{(1)}$.

4.1　第一ステップ

$k = K = 2$ とする．任意の $j_2^{(1)}, \cdots, j_p^{(1)} \in \mathbb{R}$ に対して，$d_1 = 0, b_m = j_m^{(1)}, m = 2, \cdots, p$ の近傍を，$|j_m^{(1)}| \ne |j_{m'}^{(1)}|$ ならばその近傍内では $|b_m| \ne |b_{m'}|$，$|j_m^{(1)}| \ne 0, 1$ ならばその近傍内では $|b_m| \ne 0, 1$ であるように十分小さくとり，固定する．以下，変数 d_1, b_m はその近傍内を動くものとする．このとき，ある $g(i,m) \ne 0$ 存在して，

$$c_i = \sum_{\substack{i \le m \le p, \\ j_m^{(1)} = 0}} g(i,m) a_m b_m \prod_{\substack{2 \le i' < i, \\ j_{i'}^{(1)} = 0}} (b_m^2 - b_{i'}^2) + \sum_{\substack{i \le m \le p, \\ |j_m^{(1)}| = 1}} g(i,m) a_m b_m \prod_{\substack{2 \le i' < i, \\ |j_{i'}^{(1)}| = 1}} (b_m - b_{i'}) + \sum_{\substack{i \le m \le p, \\ |j_m^{(1)}| \ne 0,1}} g(i,m) a_m \prod_{\substack{2 \le i' < i, \\ |j_{i'}^{(1)}| = |j_m^{(1)}|}} (b_m - b_{i'}),$$

と表せる．

ここで，

$$\begin{cases} \alpha = 1, & J_m^{(1)} = ((j_m^{(1)})^2), \\ t_1 = 2, & q_1 = p-1, \\ RJ^{(1)} = \{J = ((j_m^{(1)})^2); |j_m^{(1)}| = 1\} \end{cases}$$

とする．

$RJ^{(1)}$ は c_i において，

$$\sum_{\substack{i \le m \le p \\ J_m^{(1)} = J}} g(i,m) a_m b_m \prod_{\substack{2 \le i' < i \\ J_{i'}^{(1)} = J}} (b_m - b_{i'}),$$

となるような J の集合である．このようにおけば，これらは帰納的仮定を満たす．(d)∼ (f) の成立に関しては，まだ，パラメータ $t(i,J,l)$，q_l, $(l \ge 2)$ が現れていないのですべて 0 としておけばよい．

4.2　第二ステップ

k, K のとき，成立していると仮定する．

(5) を $\{v_1 = d_1 = \cdots = d_{K-1} = 0\}$ で blow up する．

(5) を (i) $\{d_1 = u_1, d_l = u_1 d_l, 2 \le l \le K-1, v_1 = u_1 v_1\}$ で変換すれば，

$$-\frac{q_1 + K}{t_1 + 2},$$

$$-\frac{q_l + 1}{t_l}, 1 \le l \le k-1, \tag{6}$$

の極が求められ，この近傍に関しては特異点解消はもう必要ない．

(5) を (ii) $\{d_l = u_l d_l, 1 \le l \le K-1, v_1 = v_1\}$ で変換すると

$$\Psi = \Biggl\{ v_1^{t_1+2} v_2^{t_2} v_3^{t_3} \cdots v_{k-1}^{t_{k-1}} \Biggl(d_1^2 + (d_1 v_1^2 + d_2)^2 + \cdots + (d_1 v_1^{2K-4} + d_2 v_1^{2K-6} f_{K-1,2} + \cdots + d_{K-1} f_{K-1,K-1})^2 + \sum_{n=K}^{p} v_1^{4n-4K+2} \Biggl(d_1 v_1^{2K-3} + d_2 v_1^{2K-5} f_{n,2} + \cdots + d_{K-1} f_{n,K-1} + \sum_{i=K}^{p} f_{n,i} c_i \Biggr)^2 \Biggr) \Biggr\}^z v_1^{q_1+K-1} \prod_{m=1}^{k-1} v_m^{q_m} \cdot \prod_{m=1}^{K-1} \mathrm{d}d_m \prod_{m=K}^{p} \mathrm{d}a_m \prod_{m=1}^{k-1} \mathrm{d}v_m \prod_{m=k}^{p} \mathrm{d}b_m. \quad (7)$$

更に (7) を $\{v_1 = d_1 = \cdots = d_{K-1} = 0\}$ で blow up する．

(7) を (iii) $\{d_1 = u_1, d_l = u_l d_l, 2 \le l \le K-1, v_1 = u_1 v_1\}$ で変換すれば

$$-\frac{q_1 + 2(K-1) + 1}{t_1 + 4}, \quad -\frac{q_l + 1}{t_l}, 1 \le l \le k-1, \quad (8)$$

の極が出てくる．

(7) を (iv) $\{d_l = u_l d_l, 1 \le l \le K-1, v_1 = v_1\}$ で変換すると

$$\Psi = \Biggl\{ v_1^{t_1+4} v_2^{t_2} v_3^{t_3} \cdots v_{k-1}^{t_{k-1}} \Biggl(d_1^2 + (d_1 v_1^2 + d_2)^2 + \cdots + (d_1 v_1^{2K-4} + d_2 v_1^{2K-6} f_{K-1,2} + \cdots + d_{K-1})^2 + \sum_{n=K}^{p} v_1^{4n-4K} \Biggl(d_1 v_1^{2K-2} + d_2 v_1^{2K-4} f_{n,2} + \cdots + d_{K-1} v_1^2 f_{n,K-1} + \sum_{i=K}^{p} f_{n,i} c_i \Biggr)^2 \Biggr) \Biggr\}^z v_1^{q_1+2(K-1)} \cdot \prod_{m=1}^{k-1} v_m^{q_m} \prod_{m=1}^{K-1} \mathrm{d}d_m \prod_{m=K}^{p} \mathrm{d}a_m \prod_{m=1}^{k-1} \mathrm{d}v_m \prod_{m=k}^{p} \mathrm{d}b_m. \quad (9)$$

4. 3 第三ステップ

さて，c_K に注目して，

$$JB^{(\alpha)} = \{J \in \mathbb{R}^\alpha; \exists l, t(K,J,l) > 0\},$$
$$\Omega = \{A' \subset \{2 \le l \le k-1\}; \forall J \in JB^{(\alpha)}, \exists l \in A' \text{ s.t. } t(K,J,l) > 0\},$$
$$A^{(\alpha)} = \Omega\text{の元の中で個数最小のものの一つ},$$
$$C^{(\alpha)} = \{m \ge k; t(K, J_m^{(\alpha)}, l) = 0 \text{ for all } l\},$$
$$JC^{(\alpha)} = \{J \in \mathbb{R}^\alpha; t(K,J,l) = 0 \text{ for all } l\},$$

とおく．

更に，

$$T_{i,J} = \begin{cases} \sum_{l \in A^{(\alpha)}} t(i,J,l) + 2s(i,J), & \text{if } J = 0, J \in JC^{(\alpha)}, \\ \sum_{l \in A^{(\alpha)}} t(i,J,l) + s(i,J), & \text{if } J \in RJ^{(\alpha)} \cap JC^{(\alpha)}, \\ \sum_{l \in A^{(\alpha)}} t(i,J,l) + s(i,J) - 1, & \text{if } J \in JC^{(\alpha)} \setminus (RJ^{(\alpha)} \cup \{0\}), \\ \sum_{l \in A^{(\alpha)}} t(i,J,l) - 1, & \text{otherwise}, \end{cases} \quad (10)$$

$$T = \sum_{l \in A^{(\alpha)}} t_l + 2, \quad (11)$$

$$Q = \sum_{l \in A^{(\alpha)}} q_l + K - 1 + \#A^{(\alpha)} + \#C^{(\alpha)} - 1, \quad (12)$$

とおく．

Case 1

もし，ある $m \in C^{(\alpha)}$ に対して，$J_m^{(\alpha)} \notin RJ^{(\alpha)}, J_m^{(\alpha)} \neq 0$ かつ，任意の $k \le i' < K$ に対して，$J_{i'}^{(\alpha)} \neq J_m^{(\alpha)}$ となる場合．

4. 3. 1 Case 1 での変換

順番を入れ換えて，その $m \in C^{(\alpha)}$ を K であるとしてよい．このとき，$c_K = a_K g(K,K) + \cdots$ となる．したがって，a_K から d_K へ，$d_K = c_K$ で変数変換を行う．

そこで

$K \to K+1, t_1 \to t_1 + 4, q_1 \to q_1 + 2(K-1)$,

と更新させて帰納的仮定を得る．

特にすべて $J_m^{(\alpha)} = 0$ のとき，Case 1 の条件よりこの場合はない．

Case 2

Case 1 が成り立たない場合，(9) を $\{d_1 = \cdots = d_{K-1} = v_{k''} = b_m = 0, k'' \in A^{(\alpha)}, m \in C^{(\alpha)}\}$ で blow up する．

4.3.2 変 換 (v)

(9) を (v) $\{d_1 = u_1, d_l = u_l d_l, 2 \le l \le K-1, v_{k''} = u_1 v_{k''}, k'' \in A^{(\alpha)}, b_m = u_1 b_m, m \in C^{(\alpha)}\}$ で変換すると

$$-\frac{\sum_{l\in A^{(\alpha)}} q_l + K - 1 + \#A^{(\alpha)} + \#C^{(\alpha)}}{\sum_{l\in A^{(\alpha)}} t_l + 2}, \quad (13)$$

$$-\frac{q_l+1}{t_l}, 1 \le l \le k-1, \quad (14)$$

の極が出る．

特にすべて $J_m^{(\alpha)} = 0$ のとき，$\#A^{(\alpha)} \neq \phi$ ならば，$\#A^{(\alpha)} = 1$, $\#C^{(\alpha)} = 0$ であるから，

$$-\frac{q_{k''}+k}{t_{k''}+2}, -\frac{q_l+1}{t_l}, 1 \le l \le k-1,$$

の極が出る．

4.3.3 変 換 (vi)

$k' \in A^{(\alpha)}$ を一つ固定し，

(9) を (vi) $\{d_l = v_{k'} d_l, 1 \le l \le K-1, v_{k'} = v_{k'}, v_{k''} = v_{k'} v_{k''}, k'' \in A^{(\alpha)} - \{k'\}, b_m = v_{k'} b_m, m \in C^{(\alpha)}\}$ で変換する．

式 (10)，(11)，(12) を用いて，

$t_{k'} \to T, t(i,J,k') \to T_{i,J}, q_{k'} \to Q, c_i \to c_i/v_{k'}$, に更新する．

このとき，$t_{k'}/2 = (\sum_{l\in A^{(\alpha)}} t_l + 2)/2 = \sum_{l\in A^{(\alpha)}} \sum_{\xi=1}^{g_l} (1 + \eta_{2,l}^{(\xi)} + \cdots + \eta_{K-1,l}^{(\xi)}) + 1$,

$$\begin{aligned} q_{k'}+1 &= \sum_{l\in A^{(\alpha)}} (q_l+1) + K - 1 + \#C^{(\alpha)} \\ &= \sum_{l\in A^{(\alpha)}} \left\{ \sum_{\xi=1}^{g_l} \varphi_l^{(\xi)} + \sum_{m=k}^{p} \left(-g_l \right.\right. \\ &\quad \left.\left. + \sum_{J_m^{(\alpha)} > J^{(\mu)}} D_{J^{(\mu)},l} \right) + \phi_l \right\} + K - 1 \\ &\quad + \sum_{J_m^{(\alpha)} \in JC^{(\alpha)}} 1 \\ &= \sum_{l\in A^{(\alpha)}} \sum_{\xi=1}^{g_l} \varphi_l^{(\xi)} + \sum_{\substack{m=k \\ J_m^{(\alpha)} \notin JC^{(\alpha)}}}^{p} \left(-\sum_{l\in A^{(\alpha)}} g_l \right. \\ &\quad \left. + \sum_{J_m^{(\alpha)} > J^{(\mu)}} \sum_{l\in A^{(\alpha)}} D_{J^{(\mu)},l} \right) + \sum_{\substack{m=k \\ J_m^{(\alpha)} \in JC^{(\alpha)}}}^{p} \\ &\quad \cdot \left(-\sum_{l\in A^{(\alpha)}} g_l + \left(\sum_{J_m^{(\alpha)} > J^{(\mu)}} \sum_{l\in A^{(\alpha)}} D_{J^{(\mu)},l} \right) + 1 \right) \\ &\quad + \sum_{l\in A^{(\alpha)}} \phi_l + K - 1, \end{aligned}$$

となる．

(I) ある $l' \in A^{(\alpha)}$ 及び，ある ξ' が存在して $0 \le \eta_{2,l'}^{(\xi')} + \eta_{3,l'}^{(\xi')} + \cdots + \eta_{K-1,l'}^{(\xi')} < 2(K-2)$ であると仮定する．

$$g_{k'} \to \sum_{l\in A^{(\alpha)}} g_l, \quad \phi_{k'} \to \sum_{l\in A^{(\alpha)}} \phi_l,$$

$$D_{J^{(\alpha)},k'} \to \sum_{l\in A^{(\alpha)}} D_{J^{(\alpha)},l} + 1 \text{ if } J^{(\alpha)} \in JC^{(\alpha)},$$

$$D_{J^{(\mu)},k'} \to \sum_{l\in A^{(\alpha)}} D_{J^{(\mu)},l} \text{ if } J^{(\mu)} \notin JC^{(\alpha)}$$

で更新し，

$$\varphi_{k'}^{(1)}, \cdots, \varphi_{k'}^{(g_{k'})}$$

は $\varphi_l^{(\xi)}$, $l \in A^{(\alpha)}$, $\xi = 1\cdots, g_l$ を並べたものとし，

$$\eta_{\ell,k'}^{(1)}, \cdots, \eta_{\ell,k'}^{(g_{k'})}$$

は $\eta_{\ell,l}^{(\xi)}$, $l \in A^{(\alpha)}$, $\xi = 1\cdots, g_l$ を同じ順で並べたものとする．

$$t_{k'}/2 = \sum_{\xi=1}^{g_{k'}} (1 + \eta_{2,k'}^{(\xi)} + \cdots + \eta_{K-1,k'}^{(\xi)}) + 1,$$

$$\begin{aligned} q_{k'}+1 &= \sum_{\xi=1}^{g_{k'}} \varphi_{k'}^{(\xi)} + K - 1 \\ &\quad + \sum_{m=k}^{p} (-g_{k'} + \sum_{J_m^{(\alpha)} > J^{(\mu)}} D_{J^{(\mu)},k'}) + \phi_{k'}, \end{aligned}$$

となる．

仮定より，ある ξ' が存在して，$0 \le \eta_{2,k'}^{(\xi')} + \eta_{3,k'}^{(\xi')} + \cdots + \eta_{K-1,k'}^{(\xi')} < 2(K-2)$ だから，

$$\varphi_{k'}^{(\xi')} \to p + 2\eta_{2,k'}^{(\xi')} + \cdots + (K-1)(\eta_{K-1,k'}^{(\xi')} + 1),$$

$$\eta_{K-1,k'}^{(\xi')} \to \eta_{K-1,k'}^{(\xi')} + 1,$$

と変換すれば，帰納的仮定 (d), (f) の形を得る．

(I 終)

(II) 任意の $l \in A^{(\alpha)}$ 及び，ξ に対して $0 \le \eta_{2,l}^{(\xi)} + \eta_{3,l}^{(\xi)} + \cdots + \eta_{K-1,l}^{(\xi)} = 2(K-2)$ ならば，(d) の仮定より，任意の $J \in \mathbb{R}^{\alpha}$ に対して，

$$\begin{aligned}\frac{t_l}{2} &= \sum_{\xi=1}^{g_l}(1+\eta_{2,l}^{(\xi)}+\cdots+\eta_{K-1,l}^{(\xi)}) \\ &= g_l(1+2(K-2)) \\ &\le \sum_{J>0^{(\mu)}} D_{0^{(\mu)},l}(2s(i,0^{(\mu)})+1) \\ &\quad + \sum_{\substack{J>J^{(\mu)} \\ J^{(\mu)}\in RJ^{(\mu)}}} D_{J^{(\mu)},l}(s(i,J^{(\mu)})+1) \\ &\quad + \sum_{\substack{J>J^{(\mu)} \\ J^{(\mu)}\notin RJ^{(\mu)},J^{(\mu)}\neq 0}} D_{J^{(\mu)},l}s(i,J^{(\mu)}).\end{aligned}$$

特に $i = K$ のときを考えると，$s(K, J^{(\mu)}) \le K-2$ より，すべて $J^{(\mu)} = 0$ でなければ

$$g_l < \sum_{J>J^{(\mu)}} D_{J^{(\mu)},l}, \tag{15}$$

である．

$$\begin{aligned}q_{k'}+1 = &\sum_{l\in A^{(\alpha)}}\sum_{\xi=1}^{g_l}\varphi_l^{(\xi)} + p + \sum_{m=k}^{p}\Bigg(-\sum_{l\in A^{(\alpha)}} g_l \\ &-1 + \sum_{J_m^{(\alpha)}>J^{(\mu)}}\sum_{l\in A^{(\alpha)}} D_{J^{(\mu)},l}\Bigg) \\ &+\sum_{l\in A^{(\alpha)}}\phi_l + p - k + 1 - p + K - 1 \\ &+\#C^{(\alpha)},\end{aligned}$$

より，

$$D_{J^{(\alpha)},k'} \to \sum_{l\in A^{(\alpha)}} D_{J^{(\alpha)},l} + 1 \text{ if } J^{(\alpha)} \in JC^{(\alpha)},$$

$$D_{J^{(\mu)},k'} \to \sum_{l\in A^{(\alpha)}} D_{J^{(\mu)},l} \text{ if } J^{(\mu)} \notin JC^{(\alpha)},$$

$$g_{k'} \to \sum_{l\in A^{(\alpha)}} g_l + 1,\ \phi_{k'} \to \sum_{l\in A^{(\alpha)}} \phi_l + K - k,$$

で更新し，$\varphi_{k'}^{(g_{k'})} = p$ と定義する．

$$\varphi_{k'}^{(1)}, \cdots, \varphi_{k'}^{(g_{k'}-1)}$$

は $\varphi_l^{(\xi)}, l \in A^{(\alpha)}, \xi = 1 \cdots, g_l$ を並べたものとし，

$$\eta_{\ell,k'}^{(1)}, \cdots, \eta_{\ell,k'}^{(g_{k'}-1)}$$

は $\eta_{\ell,l}^{(\xi)}, l \in A^{(\alpha)}, \xi = 1 \cdots, g_l$ を同じ順で並べたものとする．また，任意の ℓ について $\eta_{\ell,k'}^{(g_{k'})} = 0$ とする．

式 (15) より，$g_{k'} \le \sum_{J_m^{(\alpha)}>J^{(\mu)}} D_{J^{(\mu)},k'}$ も成立する． (II 終)

特にすべて $J_m^{(\alpha)} = 0$ のとき，

$$t_{k'} \to t_{k'} + 2, t(i,0,k') \to t(i,0,k') - 1,$$
$$q_{k'} \to q_{k'} + k - 1, c_i \to c_i/v_{k'},$$

とおく．

このとき，$t_{k'}/2 = 1 + \eta_{2,k'}^{(1)} + \cdots + \eta_{k-1,k'}^{(1)} + 1$ となる．

$$\eta_{k-1,k'}^{(1)} \to \eta_{k-1,k'}^{(1)} + 1,$$

とおけば，条件

(d′) $\tilde{t}(i, 0^{(k-1)}, k') = D_{0^{(k'-1)},k'}(2(i-k')+1)$,

(e′) $t(k, 0^{(k-1)}, k') + \eta_{k-1,k'}^{(1)} = 2$,

(f′) $q_{k'} + 1 = \varphi_{k'}^{(1)} = p + 2\eta_{2,k'}^{(1)} + \cdots + (k-1)\eta_{k-1,k'}^{(1)}$,

も成り立つ．

$JC^{(\alpha)}$ が空集合でない場合は，次の Step(∗) を行う．

Step(∗)

$J_m^{(\alpha)} \in JC^{(\alpha)}$ に対して，次に考察する点 $b_m = j_m^{(\alpha+1)}$ とその十分小さな近傍を決める．$J_m^{(\alpha)} \notin JC^{(\alpha)}$ に対しては，前と同じ点 $j_m^{(\alpha+1)} = j_m^{(\alpha)}$，すなわち，$J_m^{(\alpha)} = (J', j_m^{(\alpha)})$, $J' \in \mathbb{R}^{\alpha-1}$, $j_m^{(\alpha)} \in \mathbb{R}$，を考察するが，$J$ の長さをそろえるためにすべての $J_m^{(\alpha+1)}$ を

$$J_m^{(\alpha+1)} = \begin{cases} (J_m^{(\alpha)}, {j_m^{(\alpha+1)}}^2) & \text{if } J_m^{(\alpha)} = 0, \\ (J_m^{(\alpha)}, j_m^{(\alpha+1)}) & \text{if } J_m^{(\alpha)} \neq 0, \end{cases}$$

とおく．また，$J \in \mathbb{R}^{\alpha+1}$ に対して，$t(i,J,l) = t(i,(J',*),l)$ とおき，$g(i,m) \neq 0$ を適当に変化させる．$RJ^{(\alpha+1)}$ を c_i において，

$$\sum_{\substack{i\le m\le p \\ J_m^{(\alpha+1)}=J}} g(i,m)a_m b_m \prod_{\substack{k+1\le i'<i \\ J_{i'}^{(\alpha+1)}=J}} (b_m - b_{i'}),$$

となるような J の集合，すなわち，

$$RJ^{(\alpha+1)} = \{J \in \mathbb{R}^{\alpha+1} \mid J = (J', 0), J' \in RJ^{(\alpha)}\},$$

とおく．$\alpha \to \alpha+1$ に更新．　　　　(Step($*$) 終)

これらは帰納的仮定 (a)〜(f) を満たす．ただし，K が更新されていないので式 (5) の代わりに式 (9) を満たしている．

4.3.4 変 換 (vii)

$C^{(\alpha)} \neq \phi$ のとき，$k' \in C^{(\alpha)}$ を一つ固定して，(9) を (vii) $\{d_l = v_k d_l, 1 \le l \le K-1, b_{k'} = v_k, v_{k''} = v_k v_{k''}, k'' \in A^{(\alpha)}, b_m = v_k b_m, m \in C^{(\alpha)} - \{k'\}\}$ で変換する．

(A)　$k \le k' < K$ のとき，対称性から $k' = k$ としてよい．

$k \le i \le p$ に対して

式 (10)，(11)，(12) を用いて，

$$t_k \to T, t(i,J,k) \to T_{i,J}, q_k \to Q, c_i \to c_i/v_k,$$

とおく．

(III)　$A^{(\alpha)} \neq \phi$ のときは，(I) または (II) を k' を k に変えて，後は全く同様にしておきかえればよい．ただし，後で k が $k+1$ に更新されるので ϕ_l は上記更新を行った後，更に，$\phi_l \to (-g_l + \sum_{J_k^{(\alpha)} > J^{(\mu)}} D_{J^{(\mu)},l}) + \phi_l$ と更新する必要がある．

また $A^{(\alpha)} = \phi$ のときは，すべての $J_m^{(\alpha)}$ は $JC^{(\alpha)}$ のもとになるから，$\#C = p-(k-1)$，$q_k+1 = p+K-k$，$t_k = 2$ に注意すれば，

$$D_{J_m^{(\alpha)},k} = 1, D_{J_m^{(\alpha)},l} = 0 (l \neq k),$$
$$\eta_{\ell,k}^{(1)} = 0, \varphi_k^{(1)} = p, \phi_k = K-k,$$

とおく．　　　　　　　　　　　　(III 終)

Step($*$) を行い，$k \to k+1$ と更新する．

これらは帰納的仮定 (a)〜(f) を満たす．ただし，K が更新されていないので式 (5) の代わりに式 (9) を満たしている．$J_m^{(\alpha)} = 0$ の場合はこのケースはない．

(B)　$K \le k' < p$ のとき，対称性から $k' = K$ としてよい．もし，$J_K^{(\alpha)} \notin RJ^{(\alpha)}, J_K^{(\alpha)} \neq 0$ ならば，$J_{i'}^{(\alpha)} = J_K^{(\alpha)}$, $i' < K$ となる $J_{i'}^{(\alpha)}$ は Case 2 の条件から必ず存在する．したがって，この場合は対称性より (A) に帰着してよい．よって，$\tilde{J} := J_K^{(\alpha)} \in RJ$ または $\tilde{J} := J_K^{(\alpha)} = 0$ としてよい．

このとき (vii) で変換すれば，$c_K = v_k(a_K g(K,K) + \cdots)$ と書け，$c_i, i \ge K+1$ には a_K は存在しないので，これを a_K から d_K へ $d_K = a_K g(K,K) + \cdots, (c_K = v_k d_K)$ で変数変換する．

ここで，a_K，b_K の記号はなくなったので，$b_k, \cdots, b_{K-1}$ の変数を $b_{k+1}, \cdots, b_K$ に変換しておく．これに付随して $J_i^{(\alpha)}$，$RJ^{(\alpha)}$ も変化させておく．

$K+1 \le i \le p$ に対して，式 (10)，(11)，(12) を用いて，

$$t_k \to T, t(i,J,k) \to T_{i,J}, q_k \to Q, c_i \to c_i/v_k,$$

と更新する．

(III) を行い，Step($*$) を行う．

$$k \to k+1, K \to K+1,$$
$$t_1 \to t_1+4, q_1 \to q_1 + 2(K-1),$$

と更新すれば，帰納法仮定 (a)〜(f) は満たされる．

特にすべて $J_m^{(\alpha)} = 0$ のとき，$A^{(k)} = \phi$ より，

$$t_k = 2, t(i, 0^{(k-1)}, k) = 2(i-k), q_k = p-1,$$
$$c_i \to c_i/v_k, D_{0^{(k-1)},k} = 1, D_{0^{(k-1)},l} = 0 (l \neq k),$$
$$\eta_{\ell,k}^{(1)} = 0, \varphi_k^{(1)} = p,$$

とおく．

$t(k, 0^{(k-1)}, l) = 0 (2 \le l \le k-1)$ と条件 (e′) より $\eta_{k-1,l}^{(1)} = 2$．更に条件 (d′)，(e′) より $t(k+1, 0^{(k-1)}, l) = \tilde{t}(k+1, 0^{(k-1)}, l) - t_l/2 = 2$.

このことを踏まえて，Step($*$) 操作を行う前は，

$$\text{(d}'\text{)}\quad \tilde{t}(i, 0^{(k-1)}, k) = D_{0^{(k-1)},k}(2(i-k)+1) = 2(i-k)+1,$$
$$\text{(e}'\text{)}\quad t(k+1, 0^{(k-1)}, l) + \eta_{k,k}^{(1)} = 2+0, \quad (2 \le l \le k),$$
$$\text{(f}'\text{)}\quad q_k + 1 = \varphi_k^{(1)} = p = p + 2\eta_{2,k}^{(1)} + \cdots + (k-1)\eta_{k-1,k}^{(1)},$$

を満たす．

Step($*$) 操作により $\tilde{t}(i, 0^{(k-1)}, k) = \tilde{t}(i, 0^{(k)}, k)$, $t(i, 0^{(k-1)}, k) = t(i, 0^{(k)}, k)$ であることから，

$$k \to k+1, K \to K+1,$$
$$t_1 \to t_1+4, q_1 \to q_1 + 2(K-1),$$

と更新すれば，帰納法仮定 (a′)〜(f′) は満たされる．

Case 2 変換 (vi)，(vii)(A) は式 (5) の代わりに式 (9) を満たした形であるが，これらの場合は有限回繰り返すことによって，このケースの条件を満たさなくなる．すなわち，(第三ステップ) を繰り返せば，有限回で式 (5) の K を増加させた帰納的仮定が得られる．

$K=p+1$ になれば終了である．

［主定理の証明 2］

最後に最大の極を求めるために次の四つの補題を準備する．

［補題 1］　すべて $J_m^{(\alpha)}=0$ のときは，v_k に関する極は

$$
\begin{aligned}
&-\frac{p}{2},\\
&-\frac{p+k}{4}, -\frac{p+2k}{6},\\
&-\frac{p+2k+k+1}{8}, -\frac{p+2k+2(k+1)}{10},\\
&\qquad\vdots\\
&-\frac{p+(i-1)(2k-2+i)+k+i-1}{4i},\\
&-\frac{p+(i-1)(2k-2+i)+2(k+i-1)}{4i+2},\\
&\qquad\vdots\\
&-\frac{(p-k-1)(k-2+p)+p-1}{4(p-k)},\\
&-\frac{p+(p-k-1)(k-2+p)+2(p-2)}{4(p-k)+2}.
\end{aligned}
$$

（証明）上記 blow up で求められる．　　（証明終）

すべて $J_m^{(\alpha)}=0$ の場合は，どの段階でも原点の近傍で blow up を行ったときに相当する．

［補題 2］　$a_m, b_m>0$, $m=1,\cdots,M$ ならば

$$\frac{\sum_{m=1}^M a_m}{\sum_{m=1}^M b_m} \geq \min\{\tfrac{a_m}{b_m} \mid m=1,\ldots,M\}.$$

［補題 3］　K を自然数とする．整数 $\eta_{k'}\geq 0$ $(2\leq k'\leq K-1)$ が $0\leq\eta_2\leq 2,\cdots,0\leq\eta_2+\eta_3+\cdots+\eta_{K-1}\leq 2(K-2)$, を満たすとき，

$$
\begin{aligned}
&t:=1+\eta_2+\cdots+\eta_{K-1},\\
&\varphi:=p+2\eta_2+\cdots+(K-1)\eta_{K-1},\\
&t=2i+m,\quad i\in\mathbb{N},\quad m=0 \text{ or } 1,
\end{aligned}
$$

とおくと，

$$
\frac{\varphi}{2t}>\frac{p+i^2+im}{4i+2m}=\frac{p+1+1+2+2+\cdots+(i-1)+(i-1)+i+im}{2t}
$$

（証明）　分子の比較により証明できる．　（証明終）

（注）　補題 3 の

$$-\frac{p+i^2+im}{4i+2m}$$

は補題 1 の v_1 $(k=1)$ に関する極に等しい．

［補題 4］　すべて $J_m^{(\alpha)}=0$ 以外の場合の極は補題 1 の極ののいずれかよりも小さい．

（証明）　主定理の証明中の記号をそのまま用いる．式 (6)，(8)，(14) の $2\leq l\leq k-1$ においては，帰納的仮定 (e) 及び (f) より，

$$
\frac{q_l+1}{t_l}\geq\frac{\sum_{\xi=1}^{g_l}\varphi_l^{(\xi)}}{2\sum_{\xi=1}^{g_l}(1+\eta_{2,l}^{(\xi)}+\cdots+\eta_{K-1,l}^{(\xi)})}.
$$

補題 2 より

$$
\frac{q_l+1}{t_l}\geq\min_{1\leq\xi\leq g_l}\left\{\frac{\varphi_l^{(\xi)}}{2(1+\eta_{2,l}^{(\xi)}+\cdots+\eta_{K-1,l}^{(\xi)})}\right\}.
$$

したがって，補題 3 より，$-\frac{q_l+1}{t_l}$ 以上の補題 1 の（特に v_1 に関する）極が存在する．

式 (13) において

(I′) ある $l'\in A^{(\alpha)}$ 及び，ある ξ' が存在して $0\leq\eta_{2,l'}^{(\xi')}+\eta_{3,l'}^{(\xi')}+\cdots+\eta_{K-1,l'}^{(\xi')}<2(K-2)$ ならば

$$
\begin{aligned}
&\frac{\sum_{l\in A^{(\alpha)}}q_l+K-1+\#A^{(\alpha)}+\#C^{(\alpha)}}{\sum_{l\in A^{(\alpha)}}t_l+2}\\
&\geq\frac{\sum_{l\in A^{(\alpha)}}(q_l+1)+K-1}{\sum_{l\in A^{(\alpha)}}t_l+2}\\
&\geq\frac{\sum_{l\in A^{(\alpha)},l\neq l'}(q_l+1)+(q_{l'}+1)+K-1}{\sum_{l\in A^{(\alpha)},l\neq l'}t_l+t_{l'}+2}\\
&\geq\min\left\{\frac{q_l+1}{t_l},\frac{q_{l'}+1+K-1}{t_{l'}+2}\middle| l\in A^{(\alpha)},l\neq l'\right\}\\
&\geq\min\left\{\frac{\sum_\xi\varphi_l^{(\xi)}}{t_l},\frac{\sum_\xi\varphi_{l'}^{(\xi)}+K-1}{t_{l'}+2}\right.\\
&\qquad\left|\,l\in A^{(\alpha)},l\neq l'\right\}\\
&\geq\min\left\{\min_{l\in A^{(\alpha)},l\neq l',\xi}\frac{\varphi_l^{(\xi)}}{2(1+\eta_{2,l}^{(\xi)}+\cdots+\eta_{K-1,l}^{(\xi)})},\right.\\
&\qquad\min_{\xi\neq\xi'}\frac{\varphi_{l'}^{(\xi)}}{2(1+\eta_{2,l'}^{(\xi)}+\cdots+\eta_{K-1,l'}^{(\xi)})},\\
&\qquad\left.\frac{\varphi_{l'}^{(\xi')}+K-1}{2(1+\eta_{2,l'}^{(\xi')}+\cdots+(\eta_{K-1,l'}^{(\xi')}+1))}\right\}.
\end{aligned}
$$

したがって，補題3より，式(13)以上の補題1の（特に v_1 に関する）極が存在する．

(II$'$) 任意の $l \in A^{(\alpha)}$ 及び，ξ に対して $0 \leq \eta_{2,l}^{(\xi)} + \eta_{3,l}^{(\xi)} + \cdots + \eta_{K-1,l}^{(\xi)} = 2(K-2)$ ならばすべて $J^{(\mu)} = 0$ でないならば $g_l < \sum_{J > J^{(\mu)}} D_{J^{(\mu)},l}$ である．

このとき，

$$\frac{\sum_{l \in A^{(\alpha)}} q_l + K - 1 + \#A^{(\alpha)} + \#C^{(\alpha)}}{\sum_{l \in A^{(\alpha)}} t_l + 2}$$

$$\geq \frac{\sum_{l \in A^{(\alpha)}} (q_l + 1) + K - 1}{\sum_{l \in A^{(\alpha)}} t_l + 2}$$

$$\geq \min_{l \in A^{(\alpha)}} \left\{ \frac{\sum_\xi \varphi_l^{(\xi)}}{t_l}, \frac{\sum_\xi \varphi_l^{(\xi)} + p - k + 1 + K - 1}{t_l + 2} \right\}$$

$$\geq \min_{l \in A^{(\alpha)}} \left\{ \frac{\sum_\xi \varphi_l^{(\xi)}}{t_l}, \frac{p}{2} \right\}.$$

よってこの場合も，補題3より，式(13)以上の補題1の（特に v_1 に関する）極が存在する．　（証明終）

したがって，原点に近い極の候補は v_1 に関する極

$$-\frac{p}{2}, -\frac{p+1}{4}, -\frac{p+2}{6},$$

$$-\frac{p+4}{8}, -\frac{p+6}{10},$$

$$\vdots$$

$$-\frac{p+i^2}{4i}, -\frac{p+i^2+i}{4i+2},$$

$$\vdots$$

$$-\frac{p+(p-1)^2}{4(p-1)}, -\frac{p+(p-1)^2+(p-1)}{4(p-1)+2},$$

の中から選ばれる．簡単な計算によって $i^2 \leq p$ となる最大の i について

$$\frac{p+i^2+2}{4i+2}$$

が原点に最も近い極である．$i^2 = p$ となるとき，$-\frac{p+i^2}{4i} = -\frac{p+i^2+i}{4i+2}$ であることから，位数は $i^2 = p$ となるとき2，$i^2 \neq p$ となるとき1である．

よって，主定理が得られる．　（終）

5. むすび

三層ニューラルネットワークの学習モデルのゼータ関数の極を求める計算方法を示し，blow up により学習曲線の漸化式を得られることを示した．主定理の λ は，論文[15]で得られている上限 $\frac{\sqrt{p}}{2}$ と，$p = i^2$ のとき一致している．この計算方法は，他の漸近展開を求める上でも役に立ち，今後の数学的研究の基礎となる可能性がある．これは，代数幾何学の方法が，学習理論という極めて具体的な世界の複雑な問題を解決できるという意味での強さをもっていることを表している．また，数学の分野においても情報理論からでてくるゼータ関数の極の性質は，新たな問題として興味を喚起するだろうと思われる．今後は，学習モデルのゼータ関数の極の性質を代数幾何学観点から考察を行いたい．

文　献

[1] S. Watanabe, “Algebraic analysis for singular statistical estimation,” Lecture Notes on Computer Science, vol.1720, pp.39–50, 1999.

[2] S. Watanabe, “Algebraic analysis for nonidentifiable learning machines,” Neural Comput., vol.13, no.4, pp.899–933, 2001.

[3] S. Watanabe, “Algebraic geometrical methods for hierarchical learning machines,” Neural Netw., vol.14, no.8, pp.1049–1060, 2001.

[4] H. Akaike, “A new look at the statistical model identification,” IEEE Trans. Autom. Control, vol.19, no.6, pp.716–723, 1974.

[5] K. Takeuchi,“Distribution of an information statistic and the criterion for the optimal model,” Mathematical Science, no.153, pp.12–18, 1976.

[6] E.J. Hannan and B.G.Quinn, “The determination of the order of an autoregression,” J. Royal Statistical Society, Series B, vol.41, pp.190–195, 1979.

[7] N. Murata, S. Yoshizawa, and S. Amari, “Network information criterion — Determining the number of hidden units for an artificial neural network model,” IEEE Trans. Neural Netw., vol.5, no.6, pp.865–872, 1994.

[8] G. Schwarz, “Estimating the dimension of a model,” Annals of Statistics, vol.6, no.2, pp.461–464, 1978.

[9] J. Rissanen, “Universal coding, information, prediction, and estimation,” IEEE Trans. Inf. Theory, vol.30, no.4, pp.629–636, 1984.

[10] 山崎啓介，渡辺澄夫，“特異点をもつ推論モデルの学習曲線の確率的計算法,” 信学論（D-II），vol.J85-D-II, no.3, pp.363–372, March 2002.

[11] S. Watanabe, “Learning efficiency of redundant neural networks in Bayesian estimation,” IEEE Trans. Neural Netw., vol.12, no.6, pp.1475–1486, 2001.

[12] 渡辺一帆，渡辺澄夫，“縮小ランク回帰モデルのベイズ汎化誤差について，” 信学論（A），vol.J86-A, no.3, pp.278–287, March 2003.

[13] H. Hironaka, “Resolution of singularities of an algebraic variety over a field of characteristic zero,” Ann.

Math. Artif. Intell., vol. 79, pp.109–326, 1964.
[14] M.F. Atiyah, “Resolution of singularities and division of distributions,” Commun. Pure and Appl. Math., vol.13, pp.145–150, 1970.
[15] 渡辺澄夫, “ベイズ法による階層型統計モデルの推定誤差について,” 信学論（A），vol.J81-A, no.10, pp.1442–1452, Oct. 1998. (English version: Electron. Commun. Jpn., vol.83, no.6, pp.95–106, John Wiley & Sons, 2000.)

（平成 16 年 9 月 2 日受付，17 年 3 月 18 日再受付）

青柳　美輝　（正員）

1997 九大大学院数理学研究科数理学専攻博士課程了．1997 日本学術振興会特別研究員 PD．2001 上智大・理工・助手，現在に至る．博士（数理）．複素関数論，代数幾何，学習理論の研究に従事．日本数学会，応用数理学会各会員．

渡辺　澄夫　（正員）

1987 京大大学院・理学系研究科数理解析専攻博士課程単位取得退学．現在，東工大教授．学習について研究している．博士（工学）．IEEE，JNNS 各会員．