実対数閾値の求め方

渡辺澄夫, 東京工業大学

概 要

実対数閾値を自分で計算したい人のために基本的な性質を書きました。具体的な統計モデルの実対数閾値は不明なものも多いので、もしも解明できたら、ぜひ論文などで発表してください。

1 実対数閾値の定義

$w \in \mathbb{R}^N$ の実数値関数 $K(w) \geq 0$ と $\varphi(w) \geq 0$ について ゼータ関数 を次式で定義します（$z \in \mathbb{C}$)。

$$\zeta(z) = \int_W K(w)^z \varphi(w) dw.$$

ここで W は $\varphi(w)$ のサポート（台）です。$K(w)$ が実解析関数、$\varphi(w)$ が無限回微分可能でサポートがコンパクトであり、$K(w)\varphi(w) > 0$ となる $w \in W$ が存在すれば、ゼータ関数 は $\mathrm{Re}(z) > 0$ で複素関数として正則ですが、実は複素平面全体に有理型関数として一意に解析接続することができます（その結果得られた有理型関数も $\zeta(z)$ と書きます)。その極はすべて負の実数です。そこで、ゼータ関数の極を原点に近いほうから順に $(-\lambda_1)$, $(-\lambda_2)$, ... と名づけます。このとき λ_1 が実対数閾値です。なお、極 $(-\lambda_1)$ の位数を多重度と呼び m と書くこととします。

2 実対数閾値は応用上で何の役に立つのか

実対数閾値は、ベイズ推測における汎化誤差の平均値や対数周辺尤度の漸近挙動を与える双有理不変量です。応用上で大切な実対数閾値について成り立つ関係式をいくつかあげます。証明が必要なかたは [1] をご覧ください。

2.1 状態密度

ある $c_1 > 0$ が存在して

$$\lim_{t \to +0} \frac{1}{t^{\lambda-1}(-\log t)^{m-1}} \int \delta(t - K(w))\varphi(w) dw = c_1$$

従って

$$\lambda = 1 + \lim_{t \to +0} \frac{1}{\log t} \Big(\log \int \delta(t - K(w))\varphi(w) dw \Big)$$

2.2 分配関数

ある $c_2 > 0$ が存在して

$$\lim_{n \to \infty} \frac{n^\lambda}{(\log n)^{m-1}} \int \exp(-nK(w))\varphi(w) dw = c_2$$

従って

$$\lambda = \lim_{t \to \infty} \frac{1}{\log n} \Big(-\log \int \exp(-nK(w))\varphi(w) dw \Big)$$

2.3 体積次元

ある $c_3 > 0$ が存在して

$$\lim_{t\to+0} \frac{1}{t^\lambda(-\log t)^{m-1}} \int_{K(w)<t} \varphi(w)dw = c_3$$

従って

$$\lambda = \lim_{t\to+0} \frac{1}{\log t}\Big(\log \int_{K(w)<t} \varphi(w)dw\Big)$$

3 代数的性質

$K(w)$, $\varphi(w)$ により定まる実対数閾値を $\lambda(K,\varphi)$ と書きます。

3.1 和

$w=(w_1,w_2)$, $K(w)=K_1(w_1)+K_2(w_2)$, $\varphi(w)=\varphi_1(w_1)\varphi_2(w_2)$ とすると

$$\lambda(K,\varphi)=\lambda(K_1,\varphi_1)+\lambda(K_2,\varphi_2).$$

3.2 積

$w=(w_1,w_2)$, $K(w)=K_1(w_1)K_2(w_2)$, $\varphi(w)=\varphi_1(w_1)\varphi_2(w_2)$ のとき

$$\lambda(K,\varphi)=\min\Big\{\lambda(K_1,\varphi_1),\lambda(K_2,\varphi_2)\Big\}.$$

3.3 不等式

$K_1(w)\leq K_2(w)$ かつ $\varphi_1(w)\geq\varphi_2(w)$ であれば

$$\lambda(K_1,\varphi_1)\leq\lambda(K_2,\varphi_2).$$

3.4 イデアル不変性

$\{f_\ell(w)\}$ から生成されるイデアルと $\{g_m(w)\}$ から生成されるイデアルが等しく

$$\begin{aligned} K_1(w) &= \sum_{\ell=1}^{L} f_\ell(w)^2 \\ K_2(w) &= \sum_{m=1}^{M} g_m(w)^2 \end{aligned}$$

であるならば

$$\lambda(K_1,\varphi)=\lambda(K_2,\varphi).$$

4　幾何学的な性質

4.1　正則な場合

$K(w)=0$ を満たす w がひとつだけ w_0 であり、$\varphi(w_0)>0$ であり、$K(w)$ の w_0 でのヘッセ行列が正定値であれば、w の次元 N を用いて

$$\lambda(K,\varphi)=\frac{N}{2}.$$

4.2　部分多様体の場合

ある解析関数 $w=g(u)$ が存在して $w=g(u)$ のヤコービ行列の行列式の絶対値 $|g'(u)|$ が 0 でなく、$u=(u_1,u_2)$ とするときに

$$\begin{aligned} K(g(u)) &= \|u_1\|^2 \\ \varphi(g(u))|g'(u)| &> 0 \end{aligned}$$

とできるならば

$$\lambda(K,\varphi)=\frac{u_1 \text{ の次元}}{2}.$$

4.3　正規交差の場合

$K_0(w),\varphi_0(w)>0$ であり、

$$\begin{aligned} K(w) &= K_0(w)\prod_{i=1}^{N}(w_i)^{2k_i} \\ \varphi(w) &= \varphi_0(w)\prod_{i=1}^{N}|w_i|^{h_i} \end{aligned}$$

が成り立つとき

$$\lambda(K,\varphi)=\min_{i=1}^{N}\Big(\frac{h_i+1}{2k_i}\Big).$$

4.4　特異点解消

ある関数 $w=g(u)$ が存在して $|g'(u)|$ をその関数のヤコビアンとするとき

$$\begin{aligned} K(g(u)) &= K_0(u)\prod_{i=1}^{N}(u_i)^{2k_i} \\ \varphi(g(u))|g'(u)| &= \varphi_0(u)\prod_{i=1}^{N}|u_i|^{h_i} \end{aligned}$$

とできるならば、正規交差の場合を利用して実対数閾値を求めることができます。この場合 $w=g(u)$ は実多様体から実ユークリッド空間への関数であっても大丈夫です。このような関数を用いて変換することを特異点解消といいます。

5　注意

5.1　応用上の注意

学習理論に応用する場合には $K(w)$ として真の分布と学習モデルのカルバック・ライブラ距離を、$\varphi(w)$ として事前分布を適用します。カルバック・ライブラ距離は、パラメータ w の関数ですが積分で定義された式ですので、$K(w)$ をできるだけ扱いやすい形に等価変形することを考えるとよいことがあります。

5.2　関連事項など

実対数閾値は解析的な量のように見えるかもしれませんが、数学的に言えば関数の代数的なあるいは幾何学的な性質によって定まる値です。現代数学の様々な概念や方法と関係を有することがわかっています。

非自明な例で一般的なケースで厳密値が分かっている場合としては例えば縮小ランク回帰の場合 [2] があります。

参考文献

[1] Sumio Watanabe. Algebraic geometry and statistical learning theory, Cambridge University Press, 2009.

[2] Aoyagi, M., Watanabe, S. Stochastic complexities of reduced rank regression in Bayesian estimation Neural Networks Vol.18, No.7, pp.924-933,SEP 2005.