6K-5

多項式しきい値関数密度の上界の改善

早坂　智行†　　　　天野　一幸‡

†東京工業大学 大学院情報理工学研究科　　　‡群馬大学 大学院工学研究科

1　はじめに

n 変数ブール関数 $f:\{+1,-1\}^n \to \{+1,-1\}$ と n 変数多項式 $p:\mathbb{R}^n \to \mathbb{R}$ について，任意の入力 x で $\mathrm{sgn}(p(x)) = f(x)$ が成立しているとき，f は p により符号表現されている，あるいは p は f の PTF 表現であるという．ブール関数 f の多項式しきい値関数密度 (PTF 密度) とは，f を符号表現する多項式の項の個数の最小値のことである．

本研究では，n 変数ブール関数の PTF 密度の最悪時の上界として $(0.670)2^n$ を，平均時の上界として $(0.584)2^n$ を，それぞれ示した．その証明手法は，Oztop (2006) や Amano (2010) による手法をさらに押し進めたもので，線形代数による議論とコンピュータによる計算を組み合わせている．また，この手法を用いた上界改善の限界についても考察する．

2　多項式しきい値関数とその密度

本研究では，ブール関数の表現方法として，**多項式しきい値関数** による表現を扱う．多項式しきい値関数は，多項式の値の符号により表現される様なブール関数，もしくはそのような表現方法の事である．本研究では，$\{+1,-1\}$ 上のブール関数や多項式を対象として議論する．また，$x \in \{+1,-1\}$ において $x^2 = 1$ となることから，ここで扱う多項式では，変数が 2 次以上の次数で現れるような項は考慮しない．

定義 2.1. n 変数ブール関数 $f:\{+1,-1\}^n \to \{+1,-1\}$ と n 変数多項式 $p:\mathbb{R}^n \to \mathbb{R}$ について，任意の入力 $x \in \{+1,-1\}^n$ で $\mathrm{sgn}(p(x)) = f(x)$ が成立しているとき，**f は p により符号表現されている**，あるいは **p は f を表す多項式しきい値関数である**という．ここで $\mathrm{sgn}(x)$ は，$x>0, x=0, x<0$ のときそれぞれ $1, 0, -1$ を返すような関数である．

定義から分かる通り，あるブール関数 f を符号表現する多項式は一意ではなく，沢山存在する．

多項式しきい値関数は理論計算機科学の様々な分野で重要な役割を果たしている．例えば，機械学習理論や回路の複雑さや通信複雑性などの分野で研究や応用がなされている [1, 2, 3]．また，PTF は論理回路の単純化されたモデルの 1 つである threshold-of-parity 回路と同一視をする事が出来る．

ブール関数の多項式しきい値関数による表現の複雑さの尺度として「密度」というものがある．これは次のように定義される．

定義 2.2. ブール関数 f について，それを符号表現するような多項式の項の数の最小値を f の **多項式しきい値関数密度**，あるいは **PTF 密度** と呼ぶ．

前述の通り，変数が 2 次以上の次数で現れる項は考えないので，各ブール関数に対応する多項式には最大で 2^n 個の項しかない．よって，任意の n 変数ブール関数に対して，その PTF 密度の自明な上界は 2^n である．

3　過去の結果

PTF 密度の上界として，自明な上界 2^n より良いものを得られないだろうか，と考えるのは自然な事である．

Oztop [4] は線形代数の知識を使った議論により，全ての n 変数ブール関数に対する PTF 密度の上界，すなわち最悪時の上界として次を示した．

定理 3.1. ([4], Theorem 2) 任意の n 変数ブール関数 f に対し，その PTF 密度は $(0.75)2^n$ 以下である．

Amano [5] は，ここで使われた議論の自然な拡張を行う事で，ほとんど全てのブール関数に対する PTF 密度の上界，すなわち平均時の上界を解析した．その結果の紹介のため，いくつかの定義を行う．まず $\mathcal{M}^k$ を k 変数からなる単項式の集合とする ($|\mathcal{M}^k| = 2^k$ である)．次に，自由度と平均自由度なるものを定義する．

定義 3.2. k 変数ブール関数 $f:\{+1,-1\}^k \to \{+1,-1\}$ と，$\mathcal{M}^k$ の順列 π が与えられたとき，f の π に対する**自由度** $\mathbf{free}(f,\pi)$ とは，f を符号表現するような多項式が $\mathcal{M}^k - \{\pi(1),\dots,\pi(t)\}$ に含まれる項を使って書くことができるような最大の t のことである．

定義 3.3. $\mathcal{M}^k$ の順列 π に関する k 変数ブール関数の**平均自由度** $d(k,\pi)$ は，次のように定義される．

$$d(k,\pi) = \sum_{f \in \mathcal{F}_k} \frac{\mathbf{free}(f,\pi)}{2^{2^k} \cdot 2^k}$$

Improved Upper Bounds on PTF Density of Boolean Functions
†Tomoyuki HAYASAKA　‡Kazuyuki AMANO
†Department of Mathematical and Computing Sciences, Tokyo Institute of Technology
‡Depertment of Computer Science, Gunma University

Amano の結果は次のようなものである．

定理 3.4. ([5], Theorem 8) $\varepsilon > 0$ を任意の定数とする．$k \geq 1$ を整数とし，π を $\mathcal{M}^k$ の順列であるとする．このとき，ε と k に依存したある定数 $c > 0$ が存在して，n 変数ブール関数のうち $1 - 2^{-c2^n}$ の割合において，その PTF 密度は $(1 - d(k, \pi) + \varepsilon)2^n$ 以下である．

$k = 4$ のとき，$d(k, \pi) = 0.3833...$ となるような $\mathcal{M}^k$ の順列 π が存在することも Amano は示した．従って，定理 3.4 と組み合わせることで，次のような結果が得られる．

命題 3.5. ([5], Corollary 9) ほぼ全ての n 変数ブール関数 f について，その PTF 密度は $(0.617)2^n$ 以下である．

4 本研究の結果

本研究では，最悪時と平均時の PTF 密度の上界を共に改善した．どちらについても，Oztop と Amano の手法をさらに応用した議論を用いている．証明等のより詳細な議論については文献 [6] を参照されたい．

まず，平均時の PTF 密度の上界については，より良い平均自由度 $d(k, \pi)$ の値を得ることで改善した．そこでは $d(k, \pi)$ の計算を行う際に，定義から直接計算するのではなく，超平面アレンジメントの知識を用いるようなプログラムを作成した．そのプログラムを用いることで，Amano の結果にある $k = 4$ より一歩進んだ $k = 5$ のときにも $d(k, \pi)$ を現実的な時間で計算できるようになった．それにより，$d(k, \pi) = 0.416...$ となるような $\mathcal{M}^k$ の順列 π が存在することを発見した．これと定理 3.4 を合わせて，平均時の PTF 密度の上界は次のように改善される．

定理 4.1. $\varepsilon > 0$ を任意の定数とする．$k \geq 1$ を整数とし，π を $\mathcal{M}^k$ の順列であるとする．このとき，ε と k に依存したある定数 $c > 0$ が存在して，n 変数ブール関数のうち $1 - 2^{-c2^n}$ の割合において，その PTF 密度は $(0.584)2^n$ 以下である．

また，平均自由度 $d(k, \pi)$ について，具体的な k と π に対する値を求めるだけでなく，一般の場合における解析も行った．

まず，$\mathcal{M}^k$ に対する "単純" な順列として次のような順列の族 $\pi_k^{\mathcal{S}}$ を考える．

$$\pi_1^{\mathcal{S}} = \{1, x_1\}$$
$$\pi_2^{\mathcal{S}} = \{1, x_1, x_2, x_1x_2\}$$
$$\pi_3^{\mathcal{S}} = \{1, x_1, x_2, x_1x_2, x_3, x_1x_3, x_2x_3, x_1x_2x_3\}$$
$$\vdots$$

この順列 $\pi_k^{\mathcal{S}}$ について，平均自由度 $d(k, \pi_k^{\mathcal{S}})$ の極限を解析し，次の結果を得た．

定理 4.2. $\lim_{k\to\infty} d(k, \pi_k^{\mathcal{S}}) = 0$.

すなわち，この単純な順列では，平均時の PTF 密度の上界の結果を改善するには役に立たない順列だという事になる．また，任意の k と π において，平均自由度 $d(k, \pi)$ は次のような上界を持つ事も示した．

定理 4.3. 任意の $k \geq 1$ と $\mathcal{M}^k$ の順列 π において，$d(k, \pi) < 0.5$ が成立する，

これは，Amano の手法をそのまま適応しただけでは，PTF 密度の上界として $(0.5)2^n$ を証明するのが限界である，という事を示している．

一方，最悪時における PTF 密度の上界については，次のように改善した．

定理 4.4. $n \geq 3$ において，任意の n 変数ブール関数の PTF 密度は $(0.670)2^n$ 以下である．

この証明では，Oztop の議論と Amano の議論の良いところ組み合わせている．もう少し具体的に言えば，Oztop の最悪時に対する議論で用いていた "場合分け" の部分を，対応する線形計画問題に変換することで，Amano の平均時の議論を最悪時へと拡張することに成功した．そして，この線形計画問題の作成と最適値の計算をコンピュータを使った計算で行うことで上記の結果を得た．

参考文献

[1] Adam R. Klivans, Ryan O'Donnell, and Rocco A. Servedio. Learning intersections and thresholds of halfspaces. *J. Comput. Syst. Sci.*, Vol. 68, No. 4, pp. 808–840, 2004.

[2] Jehoshua Bruck and Roman Smolensky. Polynomial threshold functions, AC^0 functions, and spectral norms. *SIAM J. Comput.*, Vol. 21, pp. 33–42, February 1992.

[3] Alexander A. Sherstov. Separating AC^0 from depth-2 majority circuits. *SIAM J. Comput.*, Vol. 38, No. 6, pp. 2113–2129, 2009.

[4] Erhan Oztop. An upper bound on the minimum number of monomials required to separate dichotomies of $\{-1, 1\}^n$. *Neural Comput.*, Vol. 18, pp. 3119–3138, December 2006.

[5] Kazuyuki Amano. New upper bounds on the average PTF density of boolean functions. In *Proc. of ISAAC '10*, pp. 304–315, 2010.

[6] 早坂智行. 多項式しきい値関数密度の上界の改善. 修士論文, 東京工業大学 大学院情報理工学研究科, 2011.