

On a generalization for convolution formula of Hermite--Rodriguez neural networks

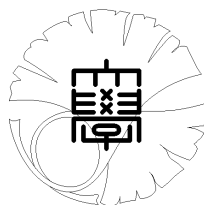
著者	レオンチエフ アレックス
雑誌名	数理科学実践研究レター
巻	LMSR 2018-24
発行年	2018-11-16
URL	http://doi.org/10.15083/00076824

数理科学実践研究レター 2018–24 November 16, 2018

**On a generalization for convolution formula
of Hermite–Rodriguez neural networks**

by

レオンチエフ アレックス



UNIVERSITY OF TOKYO
GRADUATE SCHOOL OF MATHEMATICAL SCIENCES
KOMABA, TOKYO, JAPAN

On a generalization for convolution formula of Hermite–Rodriguez neural networks

レオンチエフ アレックス¹ (東京大学 大学院数理科学研究科)

Alex Leontiev (Graduate School of Mathematical Sciences, The University of Tokyo)

概要

Orthogonal ニューラルネットワークの特別なクラスとして、Hermite–Rodriguez ニューラルネットワーク [6] がある。Hermite–Rodriguez 関数は電気刺激された器官から受信したインパルス応答信号に似ているので、そのニューラルネットワークは生物医学研究でよく使われている。

2つの Hermite–Rodriguez 関数で近似された信号の畳み込みを計算する問題がよく出てくる。Hermite–Rodriguez 関数の場合で具体的な公式が知られているため、計算でも、理論的な推計でも、簡単に得られる。本論文でその畳み込み公式の一般化を得た。その一般化はもっと広い関数クラス (Gegenbauer 関数) で記述できる。

Gegenbauer 関数に対して Orthogonal ニューラルネットワークの新しいクラスを定義することができる。その新しいニューラルネットワーククラスの性質を調べることは今後の課題である。

1 背景

1.1 Orthogonal ニューラルネットワーク

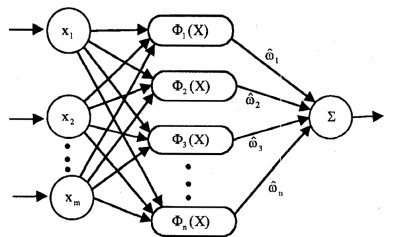


図 1: Orthogonal ニューラルネットワーク ([11] から)

考え方. 直交関数 $\{\Phi_i\}$ によって、入力信号 $y(x)$ を近似しよう :

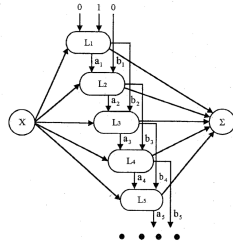
$$y(x) \approx \hat{y}(x) := \sum_{i=1}^N \hat{\omega}_i \Phi_i(x).$$

このような見方で、ニューラルネットワークを関数解析と結べる !

1.1.1 利点

- 構築が簡単 ;
- パラメータの個数が相対的に少ない ;
- 与えられた問題領域に対して、適当な直交関数システムを選べる (例、生物医工学に対して、Hermite 関数) ;
- 誤差推定を簡単にできる (⇒ 収束速度を予測できる) ;
- トレーニング戦略が多い ;
- ある直交関数システムに対して、再帰的關係を用い、計算を高速化できる (例、Lagrange, Hermite 関数/多項式) .

¹leontiev@ms.u-tokyo.ac.jp



$$iL_{i+1}(x) - (2i - 1)xL_i(x) + (i - 1)L_{i-1}(x) = 0 \quad i \geq 2 \quad (8)$$

or

$$L_{i+1}(x) = a_i x L_i(x) - b_i L_{i-1}(x) \quad (9)$$

where $L_1(x) = 1$ and $L_2(x) = x$

$$a_i = 2 - \frac{1}{i}, \quad b_i = 1 - \frac{1}{i}.$$

図 2: ([11] から)

1.2 Hermitian ニューラルネットワークの畳み込み

1.2.1 なぜ Hermitian ニューラルネットワーク?

$$h_n(t) = \frac{H_n(t)e^{-t^2/2}}{\sqrt{2^n n! \sqrt{\pi}}} \quad (\text{Hermite 関数});$$

$$H_n(t) = (-1)^n e^{t^2} \frac{d^n}{dt^n} e^{-t^2} \quad (\text{Hermite 多項式}).$$

- 低次 Hermite 関数が生物医工学で出てくる大切な信号を似ている ;
- 具体的な再帰関係が知られている : $H_{n+1}(t) = 2tH_n(t) - 2nH_{n-1}(t)$ (上記の議論を参照してください);
- ハードウェアで実装できる [1, 2, 7].

更に、Hermite 関数が良い解析性質を持つ :

- フーリエ変換で不変性を持つ :

$$\mathcal{F}(h_n) = (-i)^n \sqrt{2\pi} h_n$$

$$\left(\Rightarrow \mathcal{F}(\hat{y}) = \sum_{j=1}^N \hat{\omega}_j (-i)^j \sqrt{2\pi} h_j \right);$$

- h_0 はガウスフィルタである :

$$\Rightarrow e^{-t^2/2} * y \approx e^{-t^2/2} * \hat{y} = \sum_{i=1}^N \hat{\omega}_i (h_i * h_0)$$

1.2.2 なぜ畳み込み?

2つ Hermite ニューラルネットワークの畳み込みを計算する必要がある (例、ノイズ信号をガウス関数で復元するため)

$$y_1 * y_2 \approx \hat{y}_1 * \hat{y}_2 = \sum_{i,j=1}^N a_i^{(1)} a_j^{(2)} (h_i * h_j).$$

具体的な畳み込み公式

$$(h_i * h_j)(t) = \begin{cases} (-1)^{i+j} l_j^{i-j} \left(\frac{t^2}{2}\right), & t \geq 0 \\ l_j^{i-j} \left(\frac{t^2}{2}\right), & t < 0 \end{cases} \quad (j \leq i).$$

が知られているお陰で、

- 特殊関数/特殊多項式の計算のため、早い方法が知られている (例、再帰的關係で);
- 漸近分析と誤差推定を正確にできる.

1.3 Hermite-Rodriguez ニューラルネットワーク

1.3.1 Hermite-Rodriguez 関数

Hermite-Rodriguez 関数に対して同じ理論を応用できる:

$$hr_i(t) = \frac{h_i(t)e^{-t^2/2}}{\pi^{1/4}} = \frac{H_i(t)e^{-t^2}}{\sqrt{2^i i!} \pi}.$$

Hermite-Rodriguez 関数に対しても、具体的な畳み込み公式も知られている:

$$(hr_i * hr_j)(t) = \sqrt{\frac{(i+j)!}{2^{i+j} i! j!}} hr_{i+j}\left(\frac{t}{\sqrt{2}}\right). \quad (1.1)$$

Hermite-Rodriguez 関数は高周波信号 (すなわち、グリッチ) の解析で応用される [5].

2 主結果

命題 1 $\text{Re } \lambda > 0, \text{Re } \mu > 0, \text{Re } \nu > 0$ の時、次の積分公式が成り立つ。

$$\begin{aligned} & \int_{-1}^1 \int_{-1}^1 |s+t|^{2\nu} (1-s^2)^{\lambda-\frac{1}{2}} (1-t^2)^{\mu-\frac{1}{2}} C_\ell^\lambda(s) C_m^\mu(t) ds dt \\ &= \frac{(-\nu)_{\frac{\ell+m}{2}} (-1)^{\frac{\ell+m}{2}} \pi^{\frac{1}{2}} (2\lambda)_\ell (2\mu)_m \Gamma\left(\lambda + \frac{1}{2}\right) \Gamma\left(\mu + \frac{1}{2}\right) \Gamma\left(\nu + \frac{1}{2}\right) \Gamma(\lambda + \mu + 2\nu + 1)}{\ell! m! \Gamma\left(\lambda + \nu + \frac{\ell-m}{2} + 1\right) \Gamma\left(\mu + \nu - \frac{\ell-m}{2} + 1\right) \Gamma\left(\lambda + \mu + \nu + \frac{\ell+m}{2} + 1\right)}. \end{aligned} \quad (2.1)$$

ここで $C_\ell^\lambda(s)$ は ℓ 次 Gegenbauer 多項式である。また、 $(y)_n$ は Pochhammer 記号を表し、以下のように定義される。

$$(y)_n := \frac{\Gamma(y+n)}{\Gamma(y)} = y(y+1)\cdots(y+n-1).$$

2.1 特殊化と hierarchy

命題 1 の $\ell = m = 0$ に対する積分公式は以下の形で与えられる

$$\begin{aligned} & \int_{-1}^1 \int_{-1}^1 |s-t|^{2\nu} (1-s^2)^{\lambda-\frac{1}{2}} (1-t^2)^{\mu-\frac{1}{2}} ds dt \\ &= \frac{\pi^{\frac{1}{2}} \Gamma\left(\lambda + \frac{1}{2}\right) \Gamma\left(\mu + \frac{1}{2}\right) \Gamma\left(\nu + \frac{1}{2}\right) \Gamma(\lambda + \mu + 2\nu + 1)}{\Gamma(\lambda + \nu + 1) \Gamma(\mu + \nu + 1) \Gamma(\lambda + \mu + \nu + 1)}. \end{aligned} \quad (2.2)$$

この特殊化は Warnaar [10]、Varchenko、Tarasov [9] などによる Selberg [8] 型の積分の一般化の特別な場合と関係する。それらを図 3 にまとめる。青い数字は公式に含まれる連続パラメータの個数である。

2.2 命題 1 の極限值

最後に、命題 1 に関して、パラメータが無限になったときの極限值を考える。

まず、エルミート多項式を復習する。エルミート多項式 $\{H_n(x)\}_{n=1}^\infty$ は、 $L^2\left(\mathbb{R}, e^{-\frac{x^2}{2}} dx\right)$ の直交多項式であり、Gegenbauer 多項式の極限として再現できる:

$$H_n(x) = n! \lim_{\lambda \rightarrow \infty} \lambda^{-\frac{n}{2}} C_n^\lambda\left(\frac{x}{\sqrt{\lambda}}\right).$$

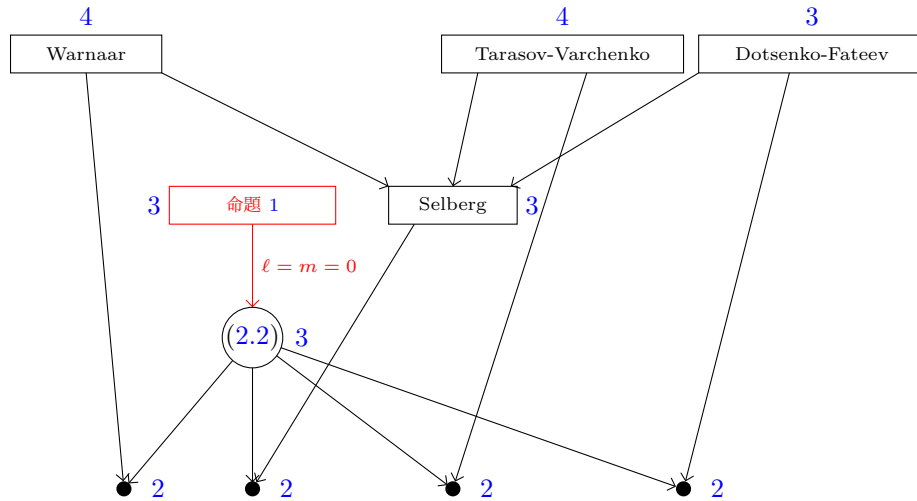


図 3: 命題 1 の $\ell = m = 0$ の特別の場合とその関連結果。ただし、欄外の青い数字は積分公式に含まれる連続パラメータの個数である。

最初の数項を挙げよう。

$$\begin{aligned}
 H_0(x) &= 1, \\
 H_1(x) &= 2x, \\
 H_2(x) &= 4x^2 - 2, \\
 H_3(x) &= 8x^3 - 12x, \\
 H_4(x) &= 16x^4 - 48x^2 + 12.
 \end{aligned}$$

命題 1 において μ/λ を一定にした上で $\lambda, \mu \rightarrow \infty$ という極限をとれば、Gegenbauer 多項式に関する積分公式からエルミート多項式に関する以下の積分公式を得る。

系 2 $\operatorname{Re} \nu > -\frac{1}{2}$ をみたす複素数 ν と実数 w に対して、 $\ell + m$ が偶数のとき、以下の積分公式が成り立つ:

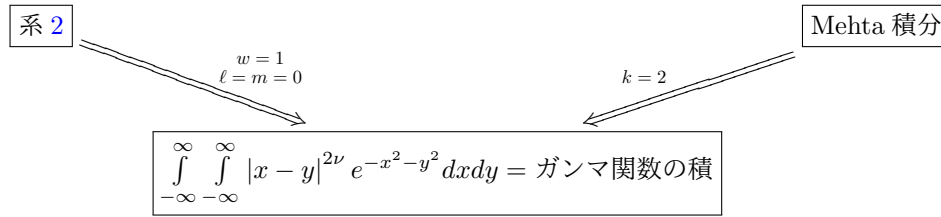
$$\begin{aligned}
 \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |x - wy|^{2\nu} e^{-x^2 - y^2} H_\ell(x) H_m(y) dx dy \\
 = (-\nu)_{\frac{\ell+m}{2}} (-1)^{\frac{\ell-m}{2}} 2^{\ell+m} \pi^{\frac{1}{2}} \Gamma\left(\frac{1}{2} + \nu\right) (w^2 + 1)^{\nu - \frac{\ell+m}{2}} w^m. \quad (2.3)
 \end{aligned}$$

系 2 で述べた積分公式の 2 通りの特殊値を考えると、それぞれ Mehta 積分 (Fact 3) の特殊値、Hermite-Rodriguez 関数の畳み込みの公式 (Fact 4) が再証明できる。

Fact 3 (Mehta 積分, [6])

$$(2\pi)^{-\frac{k}{2}} \int_{\mathbb{R}^k} \prod_{1 \leq i < j \leq k} |t_i - t_j|^{2\gamma} \exp\left(-|t|^2/2\right) dt = \prod_{j=1}^k \frac{\Gamma(1 + j\gamma)}{\Gamma(1 + \gamma)}.$$

系 2 で述べた積分公式において $w = 1, \ell = m = 0$ を代入すると、Mehta 積分 [6] の $k = 2$ の場合になる:



一方、系 2 で述べた積分公式において $s = x - y$ と変数変換し、 $w = 1$ を代入すれば、次の公式（正確には、その Mellin 変換）が得られる。

Fact 4 ([3, (18)]) *Hermite-Rodriguez* 関数 [12]

$$w_{\lambda,n}(t) := \frac{1}{\sqrt{2^n n!}} \frac{1}{\sqrt{\pi \lambda}} H_n \left(\frac{t}{\lambda} \right) e^{-\frac{t^2}{\lambda^2}}, \quad \lambda > 0$$

に関して、次の畳込みの公式が成り立つ。

$$w_{\lambda,n} * w_{\lambda,k}(t) = \sqrt{\frac{(n+k)!}{2^{n+k} n! k!}} w_{\sqrt{2}\lambda, n+k}(t). \tag{2.4}$$

以上の結果の詳しい証明とこの結果の一般化は小林俊行教授との共同研究として他の論文で出版予定である。

3 謝辞

著者は数物フロンティア・リーディング大学院と東京大学の小林俊行教授に感謝する。

参考文献

- [1] V. Chesnokov. A new adaptive architecture: Analogue synthesiser of orthogonal functions, in *Proc. Inst. Elect. Eng. Conf. Artificial Neural Networks*. pp. 714–719, 1999.
- [2] A. Cochocki and R. Unbehauen. *Neural networks for optimization and signal processing*. John Wiley & Sons, Inc., 1993.
- [3] L. R. L. Conte, R. Merletti, and G. V. Sandri. Hermite expansions of compact support waveforms: applications to myoelectric signals. *IEEE Trans. Biomed. Eng.*, **41**(12), pp. 1147–1159, Dec 1994.
- [4] V. S. Dotsenko and V. A. Fateev. Four-point correlation functions and the operator algebra in 2D conformal invariant theories with central charge $c \leq 1$. *Nuclear Phys. B*, **251**, pp. 691–734, 1985.
- [5] M. R. Mackenzie and K. Tieu. Hermite neural network correlation and application, *IEEE transactions on signal processing*, vol. **51**(12), pp. 3210–3219, IEEE, 2003.
- [6] M. L. Mehta. *Random Matrices*, Pure Appl. Math. vol. **142** (Amsterdam). Elsevier/Academic Press, Amsterdam, 2004. ISBN: 978-0-12-088409-4.
- [7] S. Osowski. Neural network for estimation of harmonic components in a power system, in *IEE Proceedings C (Generation, Transmission and Distribution)*. **139**(2), pp. 129–135, 1992.
- [8] A. Selberg. Remarks on a multiple integral. *Norsk Mat. Tidsskr.*, **26**, pp. 71–78, 1944.

- [9] V. Tarasov and A. Varchenko. Selberg-type integrals associated with \mathfrak{sl}_3 . *Lett. Math. Phys.*, **65**(3), pp. 173–185, 2003.
- [10] S. O. Warnaar. The \mathfrak{sl}_3 Selberg integral. *Adv. Math.*, **224**(2), pp. 499–524, 2010.
- [11] S.-S. Yang and C.-S. Tseng. An orthogonal neural network for function approximation. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, **26**(5), pp. 779–785, 1996.
- [12] M. A. Yusoff. Application of Hermite–Rodriguez functions to pulse shaping analog filter design. *World Acad. Sci., Eng. Technol.*, **36**, pp. 180–183, 2007.