

一般空間上の積分表現ニューラルネット とリッジレット変換

園田 翔

理化学研究所 革新知能統合研究センター (理研 AIP)
深層学習理論チーム 研究員

第 41 回ディープラーニングと物理学
Deep Learning and Physics (DLAP)
オンライン
2022 年 6 月 9 日

自己紹介

2018.4-present 理研 AIP 特別研究員 → 研究員

深層学習理論チーム (TL: 鈴木大慈准教授)

JST さきがけ (兼任)

2017.4-2018.3 早稲田大学 助手

2013.9-2017.3 早稲田大学 博士 (工学)

電気・情報生命工学専攻 (村田昇教授)

学振 DC2

論文「深層ニューラルネットの積分表現理論」

2012.4-2013.9 パナソニック株式会社

車載機器開発



研究分野:

機械学習, ニューラルネットの理論

実データ: 脳波解析, 工場操業データ解析, 自動運転

浅いニューラルネット

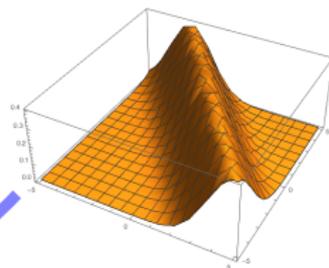
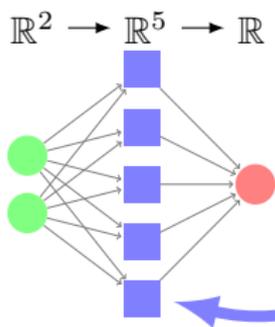
- 任意の連続関数を近似できる（普遍近似性）
- 畝（うね）状関数（ridge function）の線形結合

$$g(\mathbf{x}; \theta_p) := \sum_{i=1}^p c_i \sigma(\mathbf{a}_i \cdot \mathbf{x} - b_i), \quad \mathbf{x} \in \mathbb{R}^m,$$

where

$$\theta_p := \{(\mathbf{a}_j, b_j, c_j)\}_{j=1}^p \in (\mathbb{R}^m \times \mathbb{R} \times \mathbb{R})^p;$$

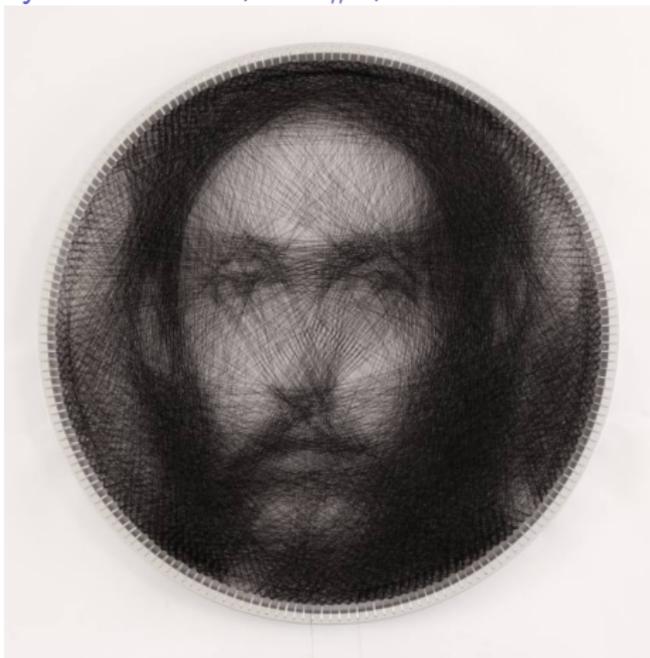
$$\sigma(z) = \exp(-z^2/2), \tanh z, \max(z, 0), \dots \quad z \in \mathbb{R}$$



eg. $\sigma(z) = \exp(-z^2/2)$

浅いニューラルネットは任意の連続関数を近似できる

String art of El Greco's Christ
by Petros Vrellis, Knit#1, 2016



$$g(\mathbf{x}; \theta_p) = \sum_{i=1}^p c_i \sigma(\mathbf{a}_i \cdot \mathbf{x} - b_i)$$

\mathbf{a} : 畝の向きと幅

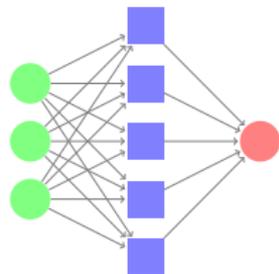
b : 位置

c : 強度

(Radon 逆変換と同じ原理)

積分表現ニューラルネット $S[\gamma]$

有限 NN (= 標準的な隠れ 1 層 NN)



- $g(\mathbf{x}; \theta_p) = \sum_{i=1}^p c_i \sigma(\mathbf{a}_i \cdot \mathbf{x} - b_i)$
- $params: \theta_p = \{(\mathbf{a}_i, b_i, c_i)\}_{i=1}^p \in \mathbb{R}^{(m+2)p}$

continuum limit

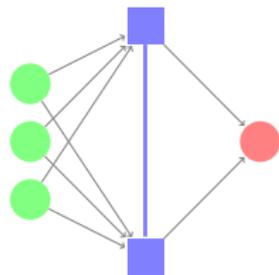


discretization



$$\gamma_p = \sum_{i=1}^p c_i \delta(\mathbf{a}_i, b_i)$$

積分表現 (無限幅, 連続) NN



- $S[\gamma](\mathbf{x}) = \int_{\mathbb{R}^m \times \mathbb{R}} \gamma(\mathbf{a}, b) \sigma(\mathbf{a} \cdot \mathbf{x} - b) d\mathbf{a} db$
- $params: \gamma \in \text{Map}(\mathbb{R}^m \times \mathbb{R} \rightarrow \mathbb{R})$

関連研究（積分表現）

- 有限 NN の表現力評価
 - Barron (1993) を筆頭に 90 年代の結果多数
 - **リッジレット変換**（右逆作用素 R ）もこの系譜（Murata, 1996; Candes, 1998; Rubin, 1998）
 - Bengio et al. (2006) convex neural network → 学習へ応用
- ReLU ネットの SGD 解の特徴づけ（平均場理論）
 - Nitanda-Suzuki arXiv2017; Mei-Montarari-Nguyen PNAS2018; Rotskoff-Vanden-Eijnden NeurIPS2018; Chizat-Bach NeurIPS2018; Sirignano-Spiliopoulos SIAM2020; Suzuki NeurIPS2020
- ReLU ネットのレプレゼンター定理
 - Savarese+, COLT2019; Ongie+, ICLR2020; Parhi-Nowak, arXiv2020; Unser, FCM2020.
- (一般の活性化関数) 有限 NN の RERM 解の特徴づけ
 - Sonoda-Ishikawa-Ikeda, AISTATS2020
- Bayesian NN の解析 (ridgelet prior)
 - Matsubara-Oates-Briol, JMLR2021.

本日の問題

Appendix A.3, S-Ishikawa-Ikeda (2021) arXiv:2106.04770

- 問題: 積分表現 NN が定める積分方程式

$$S[\gamma](\mathbf{x}) = \int_{\mathbb{R}^m \times \mathbb{R}} \gamma(\mathbf{a}, b) \sigma(\mathbf{a} \cdot \mathbf{x} - b) d\mathbf{a} db = f(\mathbf{x}),$$

を未知関数 γ について解く.

- Remark: 学習問題

$$\min_{\theta_p} \frac{1}{n} \sum_{i=1}^n |f(\mathbf{x}_i) - g(\mathbf{x}_i; \theta_p)|^2$$

の極限形

$$\min_{\gamma} \|f - S[\gamma]\|_{L^2(P_X)}^2$$

の強解に相当する.

$S[\gamma] = f$ の解き方. Step 1

Appendix A.3, S-Ishikawa-Ikeda (2021) arXiv:2106.04770

- Given $f \in L^1(\mathbb{R}^m) \cap L^2(\mathbb{R}^m)$ and $\sigma \in \mathcal{S}'(\mathbb{R})$,
- Step 1: 枠 $\{\sigma(\mathbf{a} \cdot \mathbf{x} - b) d\mathbf{a} db\}$ をうまく変換して Fourier 逆変換にする:

$$\begin{aligned} S[\gamma](\mathbf{x}) &:= \int_{\mathbb{R}^m \times \mathbb{R}} \gamma(\mathbf{a}, b) \sigma(\mathbf{a} \cdot \mathbf{x} - b) d\mathbf{a} db \\ &= \frac{1}{2\pi} \int_{\mathbb{R}} \left[\int_{\mathbb{R}^m} \gamma^\#(\boldsymbol{\xi}/\omega, \omega) e^{i\boldsymbol{\xi} \cdot \mathbf{x}} d\boldsymbol{\xi} \right] \sigma^\#(\omega) |\omega|^{-m} d\omega \quad (\star) \end{aligned}$$

ただし $\#$ は b に関する Fourier 変換

Step 1, cont'd (詳細)

- Step 1a: **Fourier 表示**にする:

$$S[\gamma](\mathbf{x}) = \int_{\mathbb{R}^m} \left[\int_{\mathbb{R}} \gamma(\mathbf{a}, b) \sigma(\mathbf{a} \cdot \mathbf{x} - b) db \right] d\mathbf{a} \quad (1)$$

$$= \int_{\mathbb{R}^m} [\gamma(\mathbf{a}, \cdot) *_b \sigma(\cdot)](\mathbf{a} \cdot \mathbf{x}) d\mathbf{a} \quad (2)$$

$$= \frac{1}{2\pi} \int_{\mathbb{R}} \left[\int_{\mathbb{R}^m} \gamma^\sharp(\mathbf{a}, \omega) \sigma^\sharp(\omega) e^{i\omega \mathbf{a} \cdot \mathbf{x}} d\mathbf{a} \right] d\omega. \quad (3)$$

つまり、恒等式 (\cdot^\sharp に関する Fourier 反転公式) $\phi(b) = \frac{1}{2\pi} \int_{\mathbb{R}} \phi^\sharp(\omega) e^{i\omega b} d\omega$ において $\phi = \gamma(\mathbf{a}, \cdot) *_b \sigma(\cdot)$ および $b = \mathbf{a} \cdot \mathbf{x}$ とおく。

- Step 1b: **置換積分**: $(\mathbf{a}, \omega) = (\boldsymbol{\xi}/\omega, \omega)$, $d\mathbf{a}d\omega = |\omega|^{-m} d\boldsymbol{\xi}d\omega$ により,

$$(3) = \frac{1}{2\pi} \int_{\mathbb{R}} \left[\int_{\mathbb{R}^m} \gamma^\sharp(\boldsymbol{\xi}/\omega, \omega) e^{i\boldsymbol{\xi} \cdot \mathbf{x}} d\boldsymbol{\xi} \right] \sigma^\sharp(\omega) |\omega|^{-m} d\omega. \quad (\star)$$

$S[\gamma] = f$ の解き方. Step 2

Appendix A.3, S-Ishikawa-Ikeda (2021) arXiv:2106.04770

- Recap:

$$S[\gamma](\mathbf{x}) = \frac{1}{2\pi} \int_{\mathbb{R}} \left[\int_{\mathbb{R}^m} \gamma^\sharp(\boldsymbol{\xi}/\omega, \omega) e^{i\boldsymbol{\xi} \cdot \mathbf{x}} d\boldsymbol{\xi} \right] \sigma^\sharp(\omega) |\omega|^{-m} d\omega. \quad (\star)$$

- Step 2: 変数分離形:

$$\gamma_{f,\rho}^\sharp(\boldsymbol{\xi}/\omega, \omega) := \widehat{f}(\boldsymbol{\xi}) \overline{\rho^\sharp(\omega)}. \quad (\star\star)$$

を仮定すると, これは特殊解になる. ただし $\widehat{\cdot}$ は \mathbf{x} に関する Fourier 変換. $\rho \in \mathcal{S}(\mathbb{R})$ は適当な補助関数. 実際,

$$S[\gamma_{f,\rho}] = (2\pi)^{m-1} \left[\int_{\mathbb{R}} \frac{\sigma^\sharp(\omega) \overline{\rho^\sharp(\omega)}}{|\omega|^m} d\omega \right] \left[\frac{1}{(2\pi)^m} \int_{\mathbb{R}^m} \widehat{f}(\boldsymbol{\xi}) e^{i\boldsymbol{\xi} \cdot \mathbf{x}} d\boldsymbol{\xi} \right] \quad (4)$$

$$= ((\sigma, \rho)) f(\mathbf{x}). \quad (5)$$

ただし $((\sigma, \rho)) := (2\pi)^{m-1} \left[\int_{\mathbb{R}} \frac{\sigma^\sharp(\omega) \overline{\rho^\sharp(\omega)}}{|\omega|^m} d\omega \right]$ とおいた.

Step 2, cont'd (リッジレット変換 R)

- 変数分離形 $\gamma_{f,\rho}$ の正体はリッジレット変換 $R[f; \rho]$.

$$\gamma_{f,\rho}^\#(\boldsymbol{\xi}/\omega, \omega) := \widehat{f}(\boldsymbol{\xi})\overline{\rho^\#(\omega)} \implies \gamma_{f,\rho}^\#(\mathbf{a}, \omega) = \widehat{f}(\omega\mathbf{a})\overline{\rho^\#(\omega)},$$

ω に関する Fourier 逆変換:

$$\begin{aligned}\gamma_{f,\rho}(\mathbf{a}, b) &= \frac{1}{2\pi} \int_{\mathbb{R}} \widehat{f}(\omega\mathbf{a})\overline{\rho^\#(\omega)} e^{ib\omega} d\omega \\ &= \frac{1}{2\pi} \int_{\mathbb{R}^m \times \mathbb{R}} f(\mathbf{x})\overline{\rho^\#(\omega)} e^{i\omega(b-\mathbf{a}\cdot\mathbf{x})} d\mathbf{x}d\omega \\ &= \int_{\mathbb{R}^m} f(\mathbf{x})\overline{\rho(\mathbf{a}\cdot\mathbf{x}-b)} d\mathbf{x} \\ &= R[f; \rho](\mathbf{a}, b).\end{aligned}$$

- 特に, 任意の $\sigma \in S'(\mathbb{R})$ に対し一意な $\rho = \sigma_* \in L_m^2(\mathbb{R})$ が存在して,

$$R[f; \sigma_*] = S^*[f] \quad (\text{共役作用素})$$

とできる. 共役作用素 S^* は零成分を含まない特別なリッジレット変換である.

$S[\gamma] = f$ の解き方. Step 3

Theorem 10, S-Ishikawa-Ikeda (2021) arXiv:2106.04770

- $((\cdot, \cdot))$ を内積とする Hilbert 空間を $L_m^2(\mathbb{R})$ とする.
- Step 3: $\gamma \in L^2(\mathbb{R}^m \times \mathbb{R})$ の範囲において, 一般解は以下で与えられる

$$\gamma_f(\mathbf{a}, b) = R[f; \rho_0] + \sum_{ij} c_{ij} R[e_i; \rho_j] \quad (6)$$

ここで,

- $\{c_{ij}\}_{i,j \in \mathbb{N}^2}$ は任意の ℓ^2 -列
- $\{e_i\}_{i \in \mathbb{N}}$ は $L^2(\mathbb{R}^m)$ の任意の正規直交基底,
- $\{\rho_j\}_{j \in \mathbb{N} \cup \{0\}}$ は $L_m^2(\mathbb{R})$ の任意の正規直交基底であって,

$$((\sigma, \rho_j)) = \begin{cases} 1, & j = 0 \\ 0, & j \neq 0 \end{cases}$$

を満たすもの.

十分性 (解であること) は (S の有界性を認めれば) 明らか:

$$S[\gamma_f] = ((\sigma, \rho_0))f(\mathbf{x}) + \sum_{ij} c_{ij} ((\sigma, \rho_j))e_i(\mathbf{x}) = f(\mathbf{x}) + 0. \quad (7)$$

必要性 (解が必ずこの表示をもつこと) は Lemma 11.

Step 3, cont'd (必要性)

Lemma 11, S-Ishikawa-Ikeda (2021) arXiv:2106.04770

Lemma 11, modified

- $\{e_i\}_{i \in \mathbb{N}}$ と $\{\rho_j\}_{j \in \mathbb{N}}$ はそれぞれ $L^2(\mathbb{R}^m)$ と $L_m^2(\mathbb{R})$ の勝手な正規直交基底とする.

このとき, 任意の $\gamma \in L^2(\mathbb{R}^m \times \mathbb{R})$ は以下のように一意に展開できる:

$$\gamma(\mathbf{a}, b) = \sum_{ij} \langle \gamma, R[e_i; \rho_j] \rangle_{L^2(\mathbb{R}^m \times \mathbb{R})} R[e_i; \rho_j](\mathbf{a}, b) \quad (8)$$

ここまでのまとめ

Theorem 10, S-Ishikawa-Ikeda (2021) arXiv:2106.04770

$S[\gamma] = f$ を解く:

1. 枠 $\sigma(\mathbf{a} \cdot \mathbf{x} - b) d\mathbf{a} db$ を Fourier 逆変換 $e^{i\xi \cdot \mathbf{x}} d\xi \otimes \sigma^\sharp(\omega) |\omega|^{-m} d\omega$ にする
2. 変数分離形 $\gamma_{f,\rho} = R[f; \rho]$ により特殊解を得る
3. 基底展開 $\gamma_f = S^*[f] + \sum_{ij} c_{ij} R[e_i; \rho_j]$ により一般解を得る

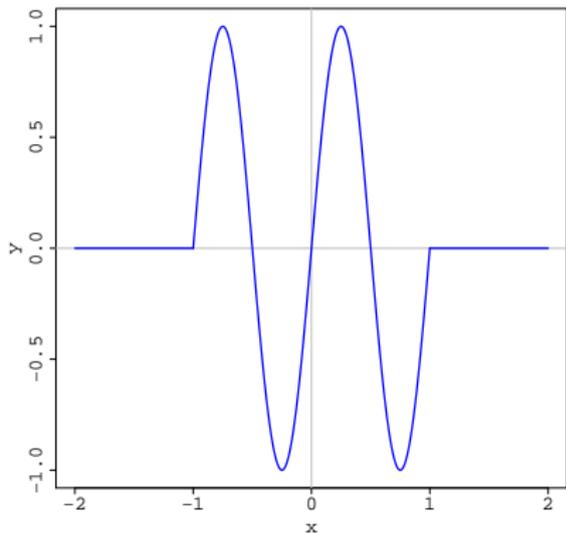
主定理 (Theorem 10, simplified)

- $\sigma \in \mathcal{S}'(\mathbb{R}), f \in L^2(\mathbb{R}^m), \gamma \in L^2(\mathbb{R}^m \times \mathbb{R})$
- $\{e_i\}_{i \in \mathbb{N}}$ an ONB of $L^2(\mathbb{R}^m)$,
- $\{\rho_i\}_{i \in \mathbb{N}}$ an ONS of $L_m^2(\mathbb{R})$ satisfying $((\sigma, \rho_i)) = 0$

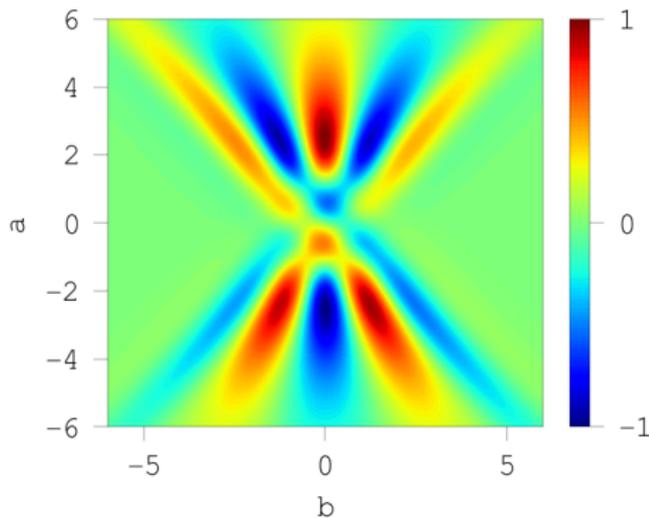
$$S[\gamma] = f \quad \Leftrightarrow \quad \exists c_{ij} \text{ s.t. } \gamma = S^*[f] + \sum_{ij} c_{ij} R[e_i; \rho_j].$$

リッジレット変換 $R[f](a, b)$ の数値計算例

- $f(x) = \sin(2\pi x) 1_{[-1,1]}(x)$
- $R[f](a, b) = \int f(x)\rho(ax - b)dx \approx \sum_i \sin(2\pi x_i)\rho(ax_i - b)\Delta x$
- $\sigma(z) = \tanh(z)$
- $\rho(z) = H[\rho_0^{(2)}](z)$ with $\rho_0(z) := e^{-z^2/2}$, Hilbert transform H



data $f(x)$



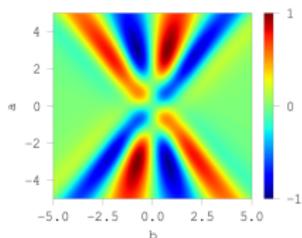
spectrum $R[f](a, b)$

再構成 $S[R[f; \rho]]$ の数値計算例

- $S[R[f; \rho]] = ((\sigma, \rho))f$ なので $((\sigma, \rho)) = 0$ の場合は縮退

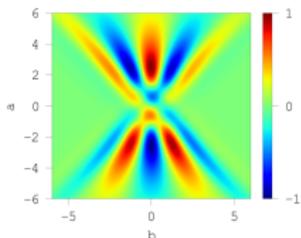
$$\rho_1 = H\rho_0^{(1)}$$

$$((\sigma, \rho_1)) = 0$$



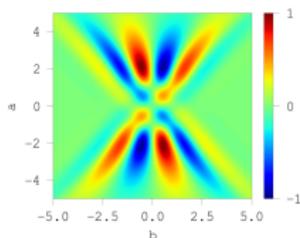
$$\rho_2 = H\rho_0^{(2)}$$

$$((\sigma, \rho_2)) \neq 0$$



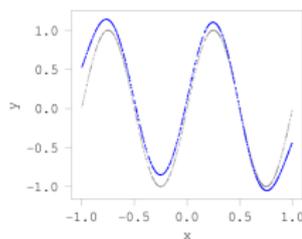
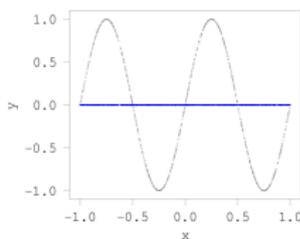
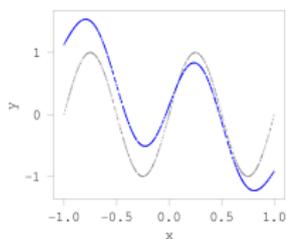
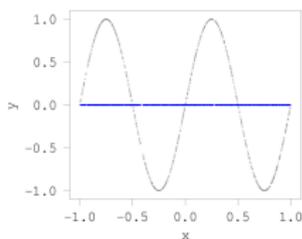
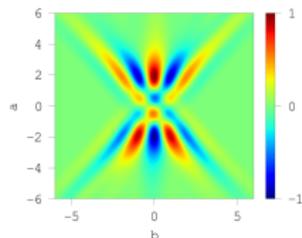
$$\rho_3 = H\rho_0^{(3)}$$

$$((\sigma, \rho_3)) = 0$$



$$\rho_4 = H\rho_0^{(4)}$$

$$((\sigma, \rho_4)) \neq 0$$



where $\rho_0(z) := e^{-z^2/2}$ (Gaussian) and H is the Hilbert transform

離散化誤差の一般論

Cor 5.4, Kurkova (2012), Complexity estimates based on integral transforms ...

一般の積分表現モデル

$$S[\gamma](\mathbf{x}) := \int_{\mathcal{V}} \gamma(\mathbf{v}) \phi(\mathbf{x}, \mathbf{v}) d\nu(\mathbf{v}), \quad (9)$$

は、 p -項有限モデル $S[\gamma_p] := \sum_{i=1}^p c_i \phi(\mathbf{x}, \mathbf{v}_i)$ を用いて $O(1/\sqrt{p})$ -近似できる。

Theorem (Barron's bound)

- (\mathcal{X}, μ) the data space with σ -finite measure μ
- (\mathcal{V}, ν) the parameter space with measure ν
- $\phi : \mathcal{X} \times \mathcal{V} \rightarrow \mathbb{R}$ the feature map with $C_\phi := \sup_{\mathbf{v} \in \mathcal{V}} \|\phi(\cdot, \mathbf{v})\|_{L^2(\mu)}$
- $\Gamma_p := \{\sum_{i=1}^p c_i \delta_{\mathbf{v}_i} \mid c_i \in \mathbb{R}, \mathbf{v}_i \in \mathcal{V}\}$ p -term model

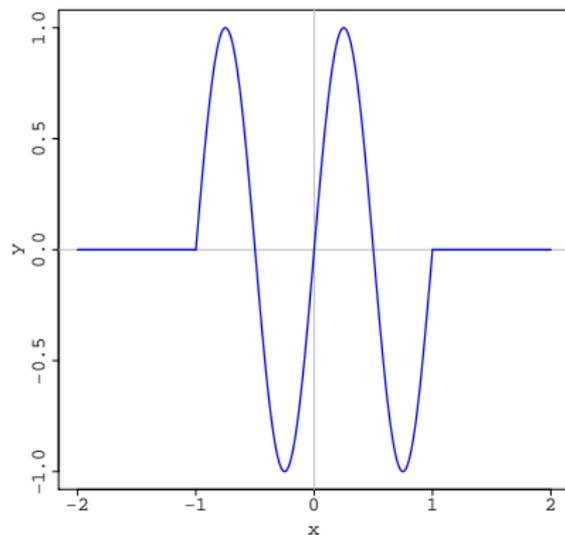
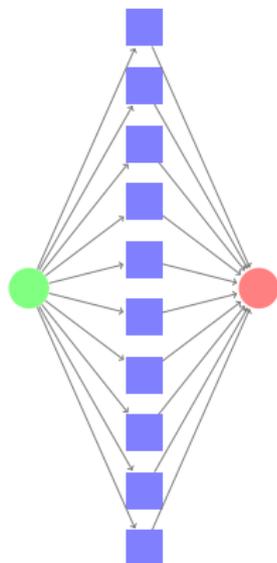
For any function $\gamma \in L^1(\nu)$ and integer p , we have

$$\inf_{\gamma_p \in \Gamma_p} \|S[\gamma] - S[\gamma_p]\|_{L^2(\mu)}^2 \leq \frac{C_\phi^2 \|\gamma\|_{L^1(\nu)}^2 - \|S[\gamma]\|_{L^2(\mu)}^2}{p}. \quad (10)$$

学習済パラメータの可視化実験

幅 $p = 10$ のニューラルネット $\sum_{j=1}^p c_j \sigma(a_j x - b_j)$ を $n = 1,000$ 体独立に学習し、学習済パラメータ (a_j, b_j, c_j) の分布を可視化する

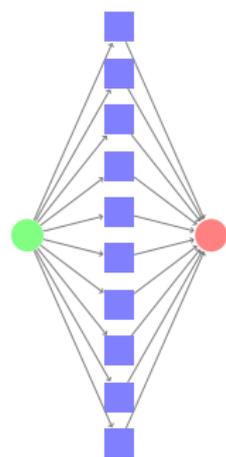
- $f(x) = \sin(2\pi x) 1_{[-1,1]}(x)$
- $\sigma(z) = \tanh(z)$
- SGD w. weight decay



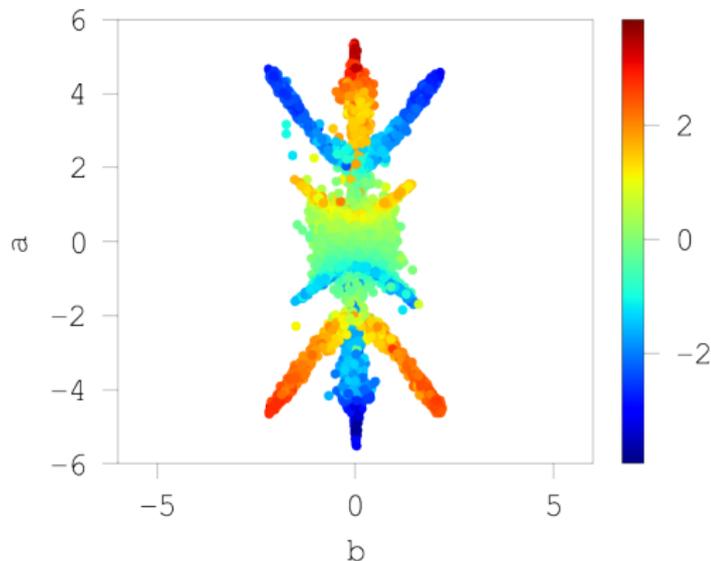
data $f(x)$

学習後パラメータの分布

幅 $p = 10$ のニューラルネット $\sum_{j=1}^p c_j \sigma(a_j x - b_j)$ を $n = 1,000$ 体学習して得られた計 $p \times n = 10,000$ 点の学習済パラメータ (a_j, b_j, c_j) の分布



×1000

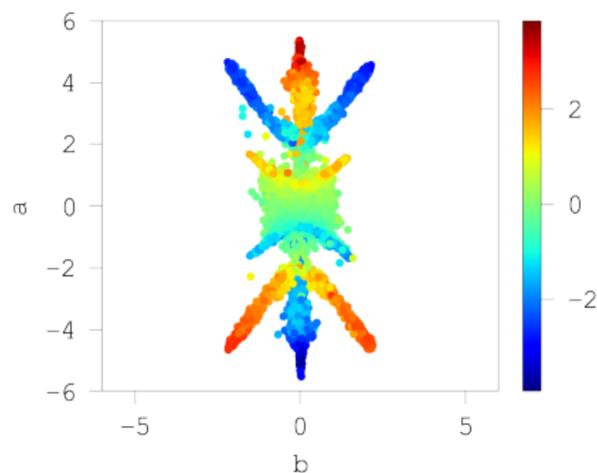


$\arg \min \hat{L}_n(\theta_p)$

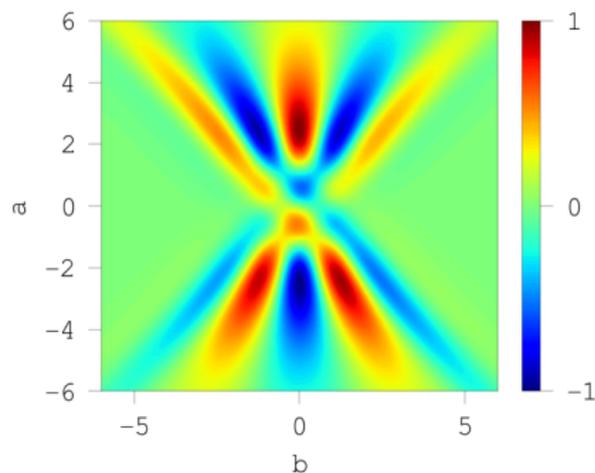
学習後パラメータの分布

Theorems 3.3 and 3.4, S-Ishikawa-Ikeda, Ridge Regression with Over-Parametrized Two-Layer Networks Converge to Ridgelet Spectrum, AISTATS2021

- リッジレット変換像 $R[f; \sigma_*]$ に弱収束する
- (OP 極限の) 学習済パラメータはリッジレット変換像を調べれば良い



scatter plot $\arg \min \widehat{L}_n(\theta_p)$



spectrum $R[f](a, b)$

畳み込みネットの場合¹

¹S-Ishikawa-Ikeda, Universality of group convolutional neural networks based on ridgelet analysis on groups, [arXiv:2205.14819](https://arxiv.org/abs/2205.14819)

“一般化” 群畳み込み

S-I-I, arXiv:2205.14819

- 群 G , Hilbert 空間 \mathcal{H} , G の表現 $T : G \rightarrow GL(\mathcal{H})$ を任意にとる
- (G, T) から定まる群畳み込み (= 双対表現 T^* の行列要素) :

$$(a *_T x)(g) := \langle T_{g^{-1}}[x], a \rangle_{\mathcal{H}} = \langle x, T_g^*[a] \rangle_{\mathcal{H}}, \quad a, x \in \mathcal{H}, g \in G.$$

- 例 ($\tilde{a}(g) := \overline{a(g^{-1})}$ として)

① (狭義) 群畳み込み: $G = G, \mathcal{H} = L^2(G, \mu), T_g^*[a](h) := a(g^{-1}h)$

$$(a *_T x)(g) = \langle x, T_g^*[a] \rangle = \int_G x(h) \tilde{a}(h^{-1}g) d\mu(h)$$

② チャネル付き画像の巡回畳み込み:

$G = \mathbb{Z}_n^2, \mathcal{H} = L^2(\mathbb{Z}_n^2)^m = \mathbb{R}^{n \times n \times m}, T_{p,q}[x](i,j) := x(i-p, j-q)$

$$(a * x)(p, q) = \langle T_{(-p, -q)}[x], a \rangle = \frac{1}{n^2 m} \sum_k \sum_{i,j} x_{(i+p, j+q)}^k a_{(i,j)}^k$$

群同変

S-I-I, arXiv:2205.14819

- 関数 $f: \mathcal{H} \rightarrow \mathcal{H}'$ が G -同変とは, 任意の $g \in G, x \in \mathcal{H}$ に対して

$$f(g \cdot x) = g \cdot [f(x)]$$

となること.

- 本研究では, $f: \mathcal{H} \rightarrow C(G)$ に対する T と正則表現の作用に限る:

$$f(T_g[x]) = f(x)(g^{-1} \bullet)$$

Hilbert 空間 \mathcal{H} 上の G -CNN

S-I-I, arXiv:2205.14819

- Hilbert 空間 \mathcal{H} を入力信号 x とフィルタ a の空間とする.
 - 論文では無限次元も可. 以下では簡単のため有限次元.
 - \mathcal{H} 上の Lebesgue 測度を λ とする
- 群畳み込みネット (G -CNN)

$$S[\gamma](x)(g) := \int_{\mathcal{H} \times \mathbb{R}} \gamma(a, b) \sigma((a * x)(g) - b) d\lambda(a) db, \quad x \in \mathcal{H}, g \in G$$

- リッジレット変換: 関数 $f: \mathcal{H} \rightarrow C(G)$ に対し, $f_e = f(\bullet)(e)$ と書いて,

$$R[f; \rho](a, b) := \int_{\mathcal{H}} f_e(x) \overline{\rho(\langle x, a \rangle_{\mathcal{H}} - b)} d\lambda(x), \quad (a, b) \in \mathcal{H} \times \mathbb{R}$$

- 再構成公式: f は群同変かつ $f_e \in L^1(\mathcal{H}) \cap L^2(\mathcal{H})$ のとき,

$$S[R[f]](x)(g) = ((\sigma, \rho))f(x)(g), \quad x \in \mathcal{H}, g \in G.$$

証明 (G -CNN 再構成公式) 1/2

S-I-I, arXiv:2205.14819

- 方程式 $S[\gamma] = f$ を解く.

Step 1a. Fourier 表示にする:

$$\begin{aligned} S[\gamma](x)(g) &= \int_{\mathcal{H} \times \mathbb{R}} \gamma(a, b) \sigma(\langle T_{g^{-1}}[x], a \rangle_{\mathcal{H}} - b) d\lambda(a) db \\ &= \frac{1}{2\pi} \int_{\mathcal{H} \times \mathbb{R}} \gamma^{\sharp}(a, \omega) \sigma^{\sharp}(\omega) e^{i\omega \langle T_{g^{-1}}[x], a \rangle_{\mathcal{H}}} d\lambda(a) d\omega. \end{aligned}$$

Step 1b. $(a, \omega) = (\xi/\omega, \omega)$, $dad\omega = |\omega|^{-m} d\xi d\omega$ と変数変換して以下のようになる:

$$S[\gamma](x)(g) = \frac{1}{2\pi} \int_{\mathcal{H} \times \mathbb{R}} \gamma^{\sharp}(\xi/\omega, \omega) \sigma^{\sharp}(\omega) e^{i\langle T_{g^{-1}}[x], \xi \rangle_{\mathcal{H}}} |\omega|^{-m} d\lambda(\xi) d\omega.$$

証明 (G-CNN 再構成公式) 1/2

S-I-I, arXiv:2205.14819

Step 2. 未知関数 γ として以下のような変数分離形を仮定する:

$$\gamma_{f,\rho}^\sharp(\xi/\omega, \omega) := \widehat{f}_e(\xi) \overline{\rho^\sharp(\omega)}.$$

作り方から, これは特殊解である:

$$\begin{aligned} S[\gamma_{f,\rho}^\sharp](x)(g) &= \frac{1}{2\pi} \int_{\mathcal{H}} \widehat{f}_e(\xi) e^{i\langle T_{g^{-1}}[x], \xi \rangle_{\mathcal{H}}} d\lambda(\xi) \int_{\mathbb{R}} \sigma^\sharp(\omega) \overline{\rho^\sharp(\omega)} |\omega|^{-m} d\omega \\ &= ((\sigma, \rho)) f(T_{g^{-1}}[x])(e) \\ &= ((\sigma, \rho)) f(x)(g). \end{aligned}$$

結局, $\gamma_{f,\rho} = R[f; \rho]$ である:

$$\begin{aligned} \gamma_{f,\rho}(a, b) &= \frac{1}{2\pi} \int_{\mathbb{R}} \widehat{f}_e(\omega a) \overline{\rho^\sharp(\omega)} e^{i\omega b} d\omega \\ &= \frac{1}{2\pi} \int_{\mathcal{H} \times \mathbb{R}} f_e(x) \overline{\rho^\sharp(\omega)} e^{i\omega b - i\omega \langle x, a \rangle_{\mathcal{H}}} d\lambda(x) d\omega \\ &= \int_{\mathcal{H}} f(x)(e) \overline{\rho(\langle x, a \rangle_{\mathcal{H}} - b)} d\lambda(x) =: R[f; \rho](a, b). \end{aligned}$$

対称空間上の全結合ネットの場合²

²S-Ishikawa-Ikeda, Fully-Connected Network on Noncompact Symmetric Space and Ridgelet Transform based on Helgason-Fourier Analysis, to appear in [ICML2022](#), [arXiv:2203.01631](#)

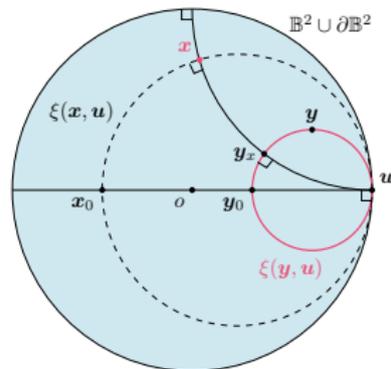
非コンパクト対称空間 $X = G/K$

S-I-I, arXiv:2203.01631

- 非コンパクト対称空間 $X = G/K$
 - 連結実単純 Lie 群 G
 - G の岩澤分解 $G = KAN$ に対し, $X := G/K = \{gK \mid g \in G\}$ を G が推移的に作用する多様体 (等質空間) とみなしたもの
 - $r := \text{rank } X := \dim \mathfrak{a}$ とおく
- α -値距離 $\langle x, u \rangle$:
 - 点 x を通り u を法ベクトルとするホロ球 $\xi(x, u)$ と原点 o との α -値距離
 - Euclid 全結合層 $\sigma(\mathfrak{a} \cdot x - b)$ のスカラー積 $\mathfrak{a} \cdot x$ を代替させる

• 具体例

- 双曲空間 \mathbb{H}^m : 言語データ, 木構造データ
- 正定値対称行列の空間 P_m : 共分散行列



Helgason-Fourier 変換

Helgason, GGA (1984, Introduction); GASS (2008, Chapter III)

- 非コンパクト対称空間 $X = G/K$ 上の Laplace-Beltrami 作用素 Δ_X の固有関数 $e^{(-i\lambda+\varrho)\langle x,u \rangle}$ による分解

$$\widehat{f}(\lambda, u) := \int_X f(x) e^{(-i\lambda+\varrho)\langle x,u \rangle} dx, \quad (\lambda, u) \in \mathfrak{a}^* \times \partial X$$

- 双曲空間 $X = \mathbb{H}^m$ のとき $\mathfrak{a}^* = \mathbb{R}$.
- Note: 通常の Euclid-Fourier 変換は Laplacian Δ の固有関数 $e^{-i\mathbf{x}\cdot\lambda\mathbf{u}} = e^{-i\lambda d(0,\xi(\mathbf{x},\mathbf{u}))}$ (= 平面波) による波面分解

- 反転公式

$$f(x) = |W|^{-1} \int_{\mathfrak{a}^* \times \partial X} \widehat{f}(\lambda, u) e^{(i\lambda+\varrho)\langle x,u \rangle} |c(\lambda)|^{-2} d\lambda du$$

ただし $f \in C_c^\infty(X)$ に対し各点収束, $f \in L^2(X)$ に対し L^2 収束. $|W|$ は定数. c は Harish-Chandra の c 関数.

- Plancherel: $\|f\|_{L^2(X)} = \|\widehat{f}\|_{L^2(|W|^{-1}|c(\lambda)|^{-2}d\lambda du)}$

非コンパクト対称空間 $X = G/K$ 上の FNN (X -FNN)

S-I-I, arXiv:2203.01631

- $X = G/K$ 上の全結合ネット (X -FNN)

$$S[\gamma](x) := \int_{\mathfrak{a}^* \times \partial X \times \mathbb{R}} \gamma(a, u, b) \sigma(a\langle x, u \rangle - b) e^{\rho\langle x, u \rangle} da du db$$

- リッジレット変換

$$R[f; \rho](a, u, b) := \int_X c[f](x) \overline{\rho(a\langle x, u \rangle - b)} e^{\rho\langle x, u \rangle} dx.$$

ただし $c[f]$ は Harish-Chandra c -関数を用いて定義される作用素 (後述)

- 再構成公式

$$S[R[f; \rho]](x) = ((\sigma, \rho))f(x), \quad x \in X.$$

証明 (X-FNN 再構成公式) 1/2

S-I-I, arXiv:2203.01631

- $S[\gamma] = f$ を解く.

Step 1a. $b \in \mathbb{R}$ を利用して Fourier 表示にする:

$$\begin{aligned} S[\gamma](x) &:= \int_{\mathbf{a}^* \times \partial X \times \mathbb{R}} \gamma(a, u, b) \sigma(a \langle x, u \rangle - b) e^{\varrho \langle x, u \rangle} da du db \\ &= \frac{1}{2\pi} \int_{\mathbf{a}^* \times \partial X \times \mathbb{R}} \gamma^\#(a, u, \omega) \sigma^\#(\omega) e^{(i\omega a + \varrho) \langle x, u \rangle} da du d\omega. \end{aligned}$$

Step 1b. $(a, \omega) = (\xi/\omega, \omega)$, $da d\omega = |\omega|^{-r} d\xi d\omega$ と変数変換して, 以下を得る:

$$S[\gamma](x) = \frac{1}{2\pi} \int_{\mathbf{a}^* \times \partial X \times \mathbb{R}} \gamma^\#(\xi/\omega, u, \omega) \sigma^\#(\omega) e^{(i\xi + \varrho) \langle x, u \rangle} |\omega|^{-r} d\xi du d\omega.$$

証明 (X-FNN 再構成公式) 2/2

S-I-I, arXiv:2203.01631

Step 2. 未知関数 γ として次のよう変数分離形を仮定する:

$$\gamma_{f,\rho}^\#(\xi/\omega, u, \omega) = \widehat{f}(\xi, u) \overline{\rho^\#(\omega)} |\mathbf{c}(\xi)|^{-2}.$$

ただし $\widehat{\cdot}$ は Helgason-Fourier 変換を表す. $|\mathbf{c}(\xi)|^{-2}$ は Plancherel 測度 (\mathbf{c} は Harish-Chandra \mathbf{c} -関数). 作り方から以下が成り立つ:

$$S[\gamma_{f,\rho}] = ((\sigma, \rho))f.$$

最後に, リッジレット変換は以下のように計算される:

$$\begin{aligned} R[f; \rho](a, u, b) &= \frac{1}{2\pi} \int_{\mathbb{R}} \widehat{f}(\omega a, u) |\mathbf{c}(\omega a)|^{-2} \overline{\rho^\#(\omega)} e^{i\omega b} d\omega \\ &= \frac{1}{2\pi} \int_{\mathbb{R} \times X} \mathbf{c}[f](x) \overline{\rho^\#(\omega)} e^{(-i\omega a + \varrho)\langle x, u \rangle + i\omega b} dx d\omega \\ &= \int_X \mathbf{c}[f](x) \overline{\rho(a\langle x, u \rangle - b)} e^{\varrho\langle x, u \rangle} dx. \end{aligned}$$

ただし $\mathbf{c}[f]$ は $\widehat{\mathbf{c}[f]}(\xi, u) = \widehat{f}(\xi, u) |\mathbf{c}(\xi)|^{-2}$ となるように定義された Helgason-Fourier multiplier である.

まとめ

- 3つの異なる NN に対して統一的な方法でリッジレット変換と再構成公式を導出した
 - Euclid 全結合層 $\sigma(\mathbf{a} \cdot \mathbf{x} - b)$
 - 群畳み込み層 $\sigma((\mathbf{a} * \mathbf{x})(g) - b)$
 - 対称空間上の全結合層 $\sigma(\langle \mathbf{a}, \mathbf{x}, \mathbf{u} \rangle - b)e^{\langle \mathbf{x}, \mathbf{u} \rangle}$
- リッジレット変換 (= 右逆作用素, 分解作用素) があれば,
 - 普遍近似定理が構成的に証明できる
 - 離散化を通じて有限 NN が構成できる
 - 学習済パラメータの分布が調べられる
 - 零空間の構造が分かる
 - (帯域制限した場合の近似下限なども調べられる)
- Take home message:
一般空間上の Fourier 反転公式があれば NN の再構成公式が作れる