

積分表現とリッジレット変換^{*1}

理化学研究所／株式会社サイバーエージェント 園田 翔^{*2}

Sho Sonoda

RIKEN / CyberAgent Inc

概要

ニューラルネットの積分表現とその右逆作用素としてのリッジレット変換を、群表現論の立場から整理する。まず浅い全結合ネットに対する古典的リッジレット変換を復習し、Murata と Candès の仕事を Radon 変換・ウェーブレット変換と結び付けて解説する。ついで、同時群同変特微量に基づく積分表現 $M_\phi[\gamma]$ と一般化リッジレット変換 R_ψ を導入し、正則表現が既約なら

$$M_\phi \circ R_\psi = c_{\phi,\psi} \text{Id}$$

が成り立つことを説明する。この枠組みは浅い全結合ネット、深い全結合ネット、群畳み込みネットワークなどを統一的に含み、有限和離散化を通じて構成的普遍近似定理を与える。さらに、ウェーブレット変換との比較、Radon 変換との関係、Moritoh による超局所リッジレット変換に触れ、関数空間論との接点を述べる。

1 はじめに

近年の深層学習では、全結合ネット、群同変ネットワーク、注意機構など多様なモデルが用いられるが、その表現力を数学的にどう把握するかは依然として基本問題である。とりわけ、所与の関数 f をニューラルネット $\text{NN}[\gamma]$ で表す

$$\text{NN}[\gamma] = f \tag{1}$$

という「学習方程式」をどのように解くかは、本質的な問題である。

積分表現は、この問題に対する自然な線形化である。浅い全結合ネット

$$\sum_{j=1}^p c_j \sigma(a_j \cdot x - b_j)$$

^{*1} 2025 RIMS 共同研究（公開型）関数空間論とその周辺 2025 年 12 月

本研究は JSPS 科研費 24K21316, 25H01453, JST BOOST JPMJBY24E2, および JST CREST JPMJCR2015 の助成を受けたものです。

^{*2} sho.sonoda@riken.jp

において、有限個のパラメータ列 $\{(a_j, b_j, c_j)\}_{j=1}^p$ の代わりに、パラメータ空間上の分布 γ を導入して

$$S_\sigma[\gamma](x) = \int \gamma(a, b) \sigma(a \cdot x - b) da db$$

と書けば、ネットワークは γ に関して線形になる。このとき、 f から γ を構成する右逆作用素がリッジレット変換であり、

$$S_\sigma[R_\rho[f]] = f$$

のような再構成公式が成立する。この意味でリッジレット変換は、Fourier 変換やウェーブレット変換と同様に、関数を「適切な原子」の重ね合わせへ分解する解析作用素である。

Murata [4] は近似理論の立場から隠れ 1 層ネットの積分表現を導入し、Candès [1] は Radon 変換とウェーブレット変換の合成としてリッジレット変換を構成した。筆者らの最近の結果 [5] は、この見方を同時群同変 (joint-group-equivariant) という極めて一般的なクラスの特徴量へ拡張し、深い全結合ネットや群同変ネットワークを含む広いクラスに対して構成的普遍近似定理を与える。本稿では、この一般論を関数空間論・調和解析の言葉で説明し、古典的リッジレット変換、群表現論、および超局所リッジレット変換との接点を見渡す。

2 浅い全結合ネットと古典的リッジレット変換

2.1 積分表現

活性化関数 $\sigma: \mathbb{R} \rightarrow \mathbb{C}$ を固定する。入力 $x \in \mathbb{R}^m$ に対する幅 p の浅い全結合ネットは

$$S_{\sigma,p}(x; \theta) = \sum_{j=1}^p c_j \sigma(a_j \cdot x - b_j) \quad (2)$$

と書ける。ここで $a_j \in \mathbb{R}^m$ は法線方向、 $b_j \in \mathbb{R}$ は位置、 $c_j \in \mathbb{C}$ は高さを表す。 $x \mapsto \sigma(a \cdot x - b)$ は超平面 $a \cdot x = b$ に沿って一定な畝 (ridge) 関数であり、(2) は畝関数の有限和である。

積分表現では、有限個のパラメータを分布 γ に置き換えて

$$S_\sigma[\gamma](x) = \int_{\mathbb{R}^m \times \mathbb{R}} \gamma(a, b) \sigma(a \cdot x - b) da db \quad (3)$$

と定める。有限和は $\gamma = \sum_{j=1}^p c_j \delta_{(a_j, b_j)}$ とおけば復元されるので、(3) は通常のニューラルネットを包含する。重要なのは、非線形だったパラメータ依存性が、 γ に関して線形になっている点である。

2.2 リッジレット変換と再構成公式

$\rho: \mathbb{R} \rightarrow \mathbb{C}$ をリッジレット関数（双対活性化関数）とし，十分良い $f: \mathbb{R}^m \rightarrow \mathbb{C}$ に対して

$$R_\rho[f](a, b) = \int_{\mathbb{R}^m} f(x) \overline{\rho(a \cdot x - b)} dx \quad (4)$$

をリッジレット変換と呼ぶ．これは f をパラメータ空間 (a, b) 上の関数へ送る変換である．

定理 2.1 (古典的再構成公式). $\sigma \in \mathcal{S}'(\mathbb{R})$, $\rho \in \mathcal{S}(\mathbb{R})$ とし，許容条件 $((\sigma, \rho)) \neq 0$ が成り立つとする．このとき任意の $f \in L^2(\mathbb{R}^m)$ に対して

$$S_\sigma[R_\rho[f]](x) = \int_{\mathbb{R}^m \times \mathbb{R}} R_\rho[f](a, b) \sigma(a \cdot x - b) da db = ((\sigma, \rho)) f(x) \quad (5)$$

が L^2 の意味で成り立つ．

ここで $((\sigma, \rho))$ は σ と ρ のみに依存する定数であり，Fourier 表現を用いて書き下せる [1, 5]．したがって $((\sigma, \rho)) = 1$ となるように規格化すれば， R_ρ は S_σ の右逆作用素である．定理 2.1 は，普遍近似定理よりも強い「構成的普遍性」を与えている．すなわち，任意の f に対して閉形式のパラメータ分布 $\gamma = R_\rho[f]$ が得られ，それを積分表現へ代入すれば f が再生される．有限和近似を施せば，通常の有限幅ネットの普遍近似定理が従う．

Murata の視点では，(3) は近似誤差評価の出発点であり [4]，Candès の視点では，(4) は高次元の線特異性を捉える解析変換であった [1]．両者は出発点こそ異なるが，同じ積分表現と再構成公式に到達している．

2.3 Radon 変換・ウェーブレット変換との関係

古典的リッジレット変換は，Radon 変換と一次元ウェーブレット変換の合成として理解できる． $u \in S^{m-1}$, $p \in \mathbb{R}$ に対して超平面積分

$$(\text{Rad} f)(u, p) = \int_{x \cdot u = p} f(x) d\mu(x)$$

を考える． $a = ru$ と極座標表示すると，(4) は

$$R_\rho[f](ru, b) = \int_{\mathbb{R}} (\text{Rad} f)(u, p) \overline{\rho(rp - b)} dp \quad (6)$$

と書ける. すなわちリッジレット変換係数は, 各方向 u に沿った Radon データ $(\text{Rad}f)(u, \cdot)$ に対する一次元ウェーブレット型変換である. この視点は, なぜリッジレット変換が「超平面状の幾何」に適合するかを説明している.

関数空間論の立場から見ると, (5) は Calderón reproducing formula の高次元かつ幾何学的な変種とみなせる. 特に ReLU のような緩増加関数を扱うには, σ を超関数として解釈し, 双対関数 ρ 側に十分な消失モーメントを課するのが自然である [1, 5].

3 同時群同変特微量と一般化積分表現

浅い全結合ネットの式 (3) は, 群作用の言葉で自然に一般化できる. 以下では G を局所コンパクト群, X と Ξ を G -空間, $\mathcal{H}$ を G のユニタリ表現とする. 簡単のため, X と Ξ には G -不変測度 $dx, d\xi$ が入っていると仮定する.

定義 3.1 (同時群同変特微量). 写像 $\phi : X \times \Xi \rightarrow \mathcal{H}$ が

$$\phi(g \cdot x, g \cdot \xi) = g \cdot \phi(x, \xi), \quad g \in G, \quad (7)$$

を満たすとき, ϕ を 同時群同変 (joint-group-equivariant) という.

パラメータへの作用が自明なとき, (7) は (幾何学的深層学習における) 標準的な群同変性に帰着する. 一方, 全結合ニューロンのように入力側では群同変でない写像も, パラメータ側に双対作用を導入すると同時群同変になることが多い.

補題 3.2 (合成). $\phi_i : X_{i-1} \times \Xi_i \rightarrow X_i$ ($i = 1, \dots, n$) が同時群同変ならば,

$$\phi_{1:n}(x, \xi_1, \dots, \xi_n) = \phi_n(\cdots \phi_2(\phi_1(x, \xi_1), \xi_2) \cdots, \xi_n)$$

も $X_0 \times (\Xi_1 \times \cdots \times \Xi_n) \rightarrow X_n$ の同時群同変写像である.

この補題が, 深いネットワークを同じ枠組みで扱える理由である.

定義 3.3 (同時群同変機械と一般化リッジレット変換). $\phi, \psi : X \times \Xi \rightarrow \mathcal{H}$ を同時群同変特微量とする. パラメータ分布 $\gamma : \Xi \rightarrow \mathbb{C}$ に対し

$$M_\phi[\gamma](x) = \int_{\Xi} \gamma(\xi) \phi(x, \xi) d\xi \quad (8)$$

を 同時群同変機械と呼ぶ. また $f : X \rightarrow \mathcal{H}$ に対し

$$R_\psi[f](\xi) = \int_X \langle f(x), \psi(x, \xi) \rangle_{\mathcal{H}} dx \quad (9)$$

を一般化リッジレット変換と呼ぶ.

(8) の積分は Bochner 積分と考える. Ξ 上の係数関数 γ が特徴量族 $\{\phi(\cdot, \xi)\}_{\xi \in \Xi}$ の混合比を指定しており, 無限幅ネットワークはこの形に入る.

4 再構成公式と Schur の補題

$L^2(X; \mathcal{H})$ 上の正則表現を

$$(\pi_g[f])(x) = g \cdot f(g^{-1} \cdot x) \quad (10)$$

で定める. また Ξ 上の係数関数に対する双対表現を

$$(\widehat{\pi}_g[\gamma])(\xi) = \gamma(g^{-1} \cdot \xi) \quad (11)$$

で定める.

命題 4.1. ϕ, ψ を同時群同変特徴量とする. このとき

$$R_\psi \circ \pi_g = \widehat{\pi}_g \circ R_\psi, \quad M_\phi \circ \widehat{\pi}_g = \pi_g \circ M_\phi$$

が成り立つ.

Proof. R_ψ について計算すると

$$R_\psi[\pi_g[f]](\xi) = \int_X \langle g \cdot f(g^{-1} \cdot x), \psi(x, \xi) \rangle_{\mathcal{H}} dx = \int_X \langle f(x), \psi(x, g^{-1} \cdot \xi) \rangle_{\mathcal{H}} dx = (\widehat{\pi}_g[R_\psi[f]])(\xi).$$

第二式では dx の不変性と ψ の同時群同変性を用いた. M_ϕ についても

$$M_\phi[\widehat{\pi}_g[\gamma]](x) = \int_{\Xi} \gamma(g^{-1} \cdot \xi) \phi(x, \xi) d\xi = \int_{\Xi} \gamma(\xi) g \cdot \phi(g^{-1} \cdot x, \xi) d\xi = (\pi_g[M_\phi[\gamma]])(x)$$

となる. □

定理 4.2 (再構成公式). π が $L^2(X; \mathcal{H})$ 上の既約ユニタリ表現であり, M_ϕ と R_ψ が有界作用素だとする. このときある定数 $c_{\phi, \psi} \in \mathbb{C}$ が存在して, 任意の $f \in L^2(X; \mathcal{H})$ に対し

$$M_\phi \circ R_\psi[f] = c_{\phi, \psi} f \quad (12)$$

が成り立つ. 特に $c_{\phi, \psi} \neq 0$ なら,

$$\gamma_f := c_{\phi, \psi}^{-1} R_\psi[f]$$

は学習方程式 (1) の解である.

Proof. 命題 4.1 より $T := M_\phi \circ R_\psi$ は π_g と可換である．実際，任意の $g \in G$ に対し

$$T \circ \pi_g = M_\phi \circ \widehat{\pi}_g \circ R_\psi = \pi_g \circ T.$$

したがって Schur の補題により T は恒等作用素のスカラー倍でなければならない．すなわち $T = c_{\phi, \psi} \text{Id}$ である． $\square$

系 4.3 (構成的普遍近似)．定理 4.2 の仮定のもとで $c_{\phi, \psi} \neq 0$ とする．このとき特徴量族 $\{\phi(\cdot, \xi) \mid \xi \in \Xi\}$ の線形包は $L^2(X; \mathcal{H})$ で稠密である．さらに，各 $f \in L^2(X; \mathcal{H})$ に対し有限個の点 $\xi_1^{(N)}, \dots, \xi_N^{(N)} \in \Xi$ と係数 $c_1^{(N)}, \dots, c_N^{(N)} \in \mathbb{C}$ を選んで

$$\sum_{j=1}^N c_j^{(N)} \phi(\cdot, \xi_j^{(N)}) \longrightarrow f \quad (N \rightarrow \infty)$$

を L^2 で成立させることができる．

Proof. $\gamma_f = c_{\phi, \psi}^{-1} R_\psi[f]$ とおけば，定理 4.2 により $M_\phi[\gamma_f] = f$ である． γ_f を単関数列で近似し，(8) の有界性を用いれば主張が従う． $\square$

定理 4.2 の重要な点は，証明が極めて短いことである．実際に使っているのは，(i) 測度の群不変性，(ii) 特徴量の同時群同変性，(iii) Schur の補題，の三つだけである．深いネットワークを含む広いクラスがこの最小限の原理で同時に扱えるのが本結果の強みである．

注意 4.4 (既約性の向こう側)．既約性は十分条件であって必要条件ではない．実際，命題 4.1 から常に

$$T_{\phi, \psi} := M_\phi \circ R_\psi \in \text{End}_G(L^2(X; \mathcal{H}))$$

である．したがって π が既約でない場合でも， $T_{\phi, \psi}$ は表現の可換環に属する．例えば，形式的に

$$L^2(X; \mathcal{H}) \simeq \int_{\widehat{G}}^{\oplus} H_\lambda \otimes \mathbb{C}^{m(\lambda)} d\mu(\lambda)$$

と分解できるなら， $T_{\phi, \psi}$ は各既約成分上で

$$T_{\phi, \psi} \simeq \int_{\widehat{G}}^{\oplus} I_{H_\lambda} \otimes A(\lambda) d\mu(\lambda)$$

という形の作用素として振る舞うはずである．既約性は $m(\lambda) = 1$ かつ $A(\lambda)$ が定数の場合に対応する．この観点は，multiplicity-free な場合の“作用素値再構成公式”や，非既約系に対する周波数別のリッジレット解析を示唆している．

5 具体例

5.1 浅い全結合ネット

$G = \text{Aff}(m) = GL(m) \ltimes \mathbb{R}^m$ をアフィン群とし, $g = (L, t)$ の $X = \mathbb{R}^m$ への作用を $g \cdot x = Lx + t$ で定める. パラメータ空間 $\Xi = \mathbb{R}^m \times \mathbb{R}$ には

$$g \cdot (a, b) = (L^{-\top} a, b + t^\top L^{-\top} a) \quad (13)$$

を作用させる. すると

$$\phi(x, (a, b)) = \sigma(a \cdot x - b), \quad \psi(x, (a, b)) = \rho(a \cdot x - b)$$

は同時群不変である. 実際,

$$(L^{-\top} a) \cdot (Lx + t) - (b + t^\top L^{-\top} a) = a \cdot x - b$$

より $\phi(g \cdot x, g \cdot (a, b)) = \phi(x, (a, b))$ が従う. 正則表現

$$(\pi_g f)(x) = |\det L|^{-1/2} f(L^{-1}(x - t)) \quad (14)$$

は $L^2(\mathbb{R}^m)$ 上既約であり [2], 定理 4.2 から古典的再構成公式 (5) が復元される.

注意 5.1. 厳密には $\Xi = \mathbb{R}^m \times \mathbb{R}$ は (13) のもとで一つの軌道ではない. たとえば $a = 0$ は別軌道をなす. しかしこの例外集合はリッジレット変換の本質部分ではなく, 軌道分解あるいは零測度集合の除去で処理できる. 一般の G -空間に対しては, 準不変 (quasi-invariant) 測度と直積分分解を使うのが自然である.

5.2 深い全結合ネット

深さ n の全結合ネットを

$$\phi_i(z, \xi_i) = C_i \sigma_i(A_i z - b_i), \quad i = 1, \dots, n$$

の合成として書く. ここで $\xi_i = (A_i, b_i, C_i)$ は第 i 層のパラメータである. 入力側のアフィン変換と出力側の直交変換を組み合わせた群を適切に選ぶと, 各層 ϕ_i は同時群同変となり, 補題の合成則より

$$\phi_{1:n}(x, \xi_1, \dots, \xi_n) = \phi_n(\cdots \phi_2(\phi_1(x, \xi_1), \xi_2) \cdots, \xi_n)$$

も同時群同変になる．したがって

$$D_\phi[\gamma](x) = \int_{\Xi_1 \times \dots \times \Xi_n} \gamma(\xi_1, \dots, \xi_n) \phi_{1:n}(x, \xi_1, \dots, \xi_n) d\xi_1 \dots d\xi_n \quad (15)$$

という深い積分表現に対して一般化リッジレット変換が存在し、 L^2 -再構成公式が得られる [5].

ここで強調したいのは、(15) では 全層のパラメータをまとめて一つの大きなパラメータ空間に積分しているという点である．したがって得られる構成的普遍近似定理は、深さ n の有限ネットに対する存在定理ではなく、深さ n の特徴量を持つ無限幅モデルに対する厳密表現定理である．有限幅近似はこの後段で行う．

5.3 二次形式ネットワーク

一般論の利点は、既存のネットワークだけでなく、新しい関数族の普遍性も同じ方法で示せることである．その一例として

$$Q_\sigma[\gamma](x) = \int_{\text{Sym}(m) \times \mathbb{R}^m \times \mathbb{R}} \gamma(A, b, c) \sigma(x^\top Ax + x^\top b + c) dA db dc \quad (16)$$

を考える．ここで $\text{Sym}(m)$ は実対称行列全体である． $g = (L, t) \in \text{Aff}(m)$ に対し

$$g \cdot (A, b, c) = (L^{-\top} AL^{-1}, L^{-\top} b - 2L^{-\top} AL^{-1}t, c + t^\top L^{-\top} AL^{-1}t - t^\top L^{-\top} b) \quad (17)$$

と作用させると、

$$\phi(x, (A, b, c)) = \sigma(x^\top Ax + x^\top b + c)$$

は同時群不変になる．したがって (16) も定理 4.2 の適用対象であり、 $L^2(\mathbb{R}^m)$ 上の構成的普遍近似を持つ．

5.4 群畳み込みネットワーク

同時群同変の枠組みは、いわゆる群畳み込みネットワークも自然に含む．ある群 Γ が特徴空間に作用しているとし、 T_g をその作用による平行移動とする．全結合層

$$z \mapsto C\sigma(Az - b)$$

を、

$$\phi^\tau(x, \xi)(g) = T_g[C\sigma(AT_{g^{-1}}x - b)] \quad (18)$$

と“群畳み込み化”すると、標準的な Γ -convolution, DeepSets 型モデル, さらには $E(n)$ -equivariant 層まで統一的に書ける [5]. 各層が同時群同変ならば, その合成もやはり同時群同変であり, 一般化リッジレット変換による再構成がそのまま適用できる.

6 ウェーブレット変換・超局所リッジレット変換との比較

6.1 一般化ウェーブレット変換との比較

局所コンパクト群 G の既約ユニタリ表現 $\pi : G \rightarrow U(H)$ と許容ベクトル $\eta \in H$ に対して, 一般化ウェーブレット変換は

$$W_\eta f(g) = \langle f, \pi(g)\eta \rangle_H \quad (19)$$

で定義され, Schur の補題から再構成公式

$$\int_G W_\eta f(g) \pi(g)\eta dg = c_\eta f$$

が従う. 証明の骨格は定理 4.2 と同じである.

しかし両者には大きな違いもある. ウェーブレット変換では係数の添字空間が群そのもの G であるのに対し, リッジレット変換では添字空間は一般には G -空間 Ξ である. また, ウェーブレット変換は表現空間 H の中で完結するが, 一般化リッジレット変換では入力空間 X とパラメータ空間 Ξ が別々に現れる. この違いは, ニューラルネットの「入力」と「パラメータ」を同時に扱うために本質的である.

一方, $\Xi = G$ として, ある基底特徴量 $\phi_0 \in L^2(X; \mathcal{H})$ から

$$\phi(x, g) = \pi_g \phi_0(x)$$

を作ると, 一般化リッジレット変換は形式的には一般化ウェーブレット変換と同じ構造になる. この意味でリッジレット変換は, 群そのものの調和解析というより, 群作用に対する調和解析とみなすのが適切である.

6.2 超局所リッジレット変換

古典的リッジレット変換の幾何学をさらに精密化したのが, Moritoh による超局所リッジレット変換である [3]. その出発点は, Candès–Donoho 型リッジレット変換が Radon 変換

のスライスに対する一次元ウェーブレット変換になっているという事実である。Moritoh は周波数局在パラメータ $\xi \in \mathbb{R}^n \setminus \{0\}$ を導入し、

$$\mathcal{R}_\varphi f(\omega, p; \xi) = \int_{\mathbb{R}^n} f(x) \overline{\varphi_\xi(\omega, \langle x, \omega \rangle - p)} dx \quad (20)$$

を定義した。ここで $\omega \in S^{n-1}$ は方向、 $p \in \mathbb{R}$ は超平面位置であり、 φ_ξ はリッジレット関数の超局所化である。

この変換に対して反転公式

$$f(x) = C^{-1} \int_{\mathbb{R}^n} \int_{S^{n-1}} \int_{\mathbb{R}} \mathcal{R}_\varphi f(\omega, p; \xi) \varphi_\xi(\omega, \langle x, \omega \rangle - p) dp d\omega \frac{d\xi}{|\xi|^n} \quad (21)$$

が成り立ち、さらに方向付き特異性の検出や two-microlocal Besov 空間との関係が得られる [3]。ここで重要なのは、(20) が単なる再構成だけでなく、波面集合に近い局所情報を保持している点である。

筆者らの一般化リッジレット変換は主として (1) の解を構成する立場に立っており、超局所的な情報を陽に追跡していない。しかし関数空間論の視点からは、深い特徴量 $\phi_{1:n}$ に対する超局所リッジレットを構成し、係数空間で特異性・方向性・層ごとの周波数分解を記述することは極めて魅力的な課題である。

6.3 係数空間の幾何と特異性

リッジレット係数空間そのものの幾何も重要である。古典的リッジレット変換では、係数は概ね「方向 u 」「位置 b 」「スケール r 」に対応し、超平面状の特異性を持つ関数ほど、係数がその近傍に集中する。たとえば f が $x \cdot u_0 = b_0$ に沿って折れ曲がりやジャンプを持つとき、 $R_\rho[f](ru, b)$ は $u \approx u_0$, $b \approx rb_0$ 付近で大きくなり、方向の合わない畝に対しては急速に小さくなる。ウェーブレット変換が点特異性や局所スケールを捉えるのに対し、リッジレット変換は余次元 1 の幾何に適合しているのである。

この観点から見ると、リッジレット変換は単なる再構成のための補助道具ではなく、パラメータ分布 $\gamma_f = R_\psi[f]$ を通じて“どの畝がどの程度使われるか”を可視化する解析でもある。浅いネットでは、係数空間で見た分布が関数の幾何や特異性を反映する。深いネットに対しても、層ごとの特徴量がどのような幾何を係数空間へ押し出すのかを理解できれば、学習で得られるパラメータ分布の解釈、さらには最適化や正則化の数理に新しい手掛かりを与えると期待される。

7 今後の課題

最後に、関数空間論の観点から重要と思われる課題を三点挙げる。

■(1) **非既約・非推移的状况** 現実のパラメータ空間 Ξ はしばしば複数の軌道に分かれ、正則表現 π も既約とは限らない。この場合、定理 4.2 はそのままではスカラー型再構成を与えないが、注意 4.4 のように可換環 End_G を通じて作用素値の再構成へ進む余地がある。関数空間論で馴染み深い直積分分解や軌道分解が、基本道具になるだろう。

■(2) **層ごとの深さの扱い** 現在の深いリッジレット変換理論は、全層パラメータをまとめて積分する (15) を扱う。これに対して、本来の深いネットワークは層ごとに独立な分布 $\gamma_1, \dots, \gamma_n$ を持つ

$$M_{\phi_n}[\gamma_n] \circ \dots \circ M_{\phi_1}[\gamma_1]$$

という形であり、こちらを直接解く理論は未完成である。どの深さがどの表現成分を担うのか特定する問題は、深層構造に由来する新しい未解決問題である。

■(3) **離散化とフレーム理論** 構成的普遍近似定理は連続表現の exact formula を与えるが、有限幅近似には離散化が必要である。ここでは coorbit 理論や可算フレーム、非直交展開の安定性が本質的になる。とくにリッジレット係数のノルムがどの関数空間を特徴付けるか、離散フレームがどの程度まで深層リッジレット変換に移植できるかが興味深い。

8 おわりに

本稿では、積分表現とリッジレット変換を、浅い全結合ネットから同時群同変機械へと一般化する流れで説明した。核心は、ネットワークをパラメータ分布に関して線形化し、その右逆を Schur の補題によって構成するという一点にある。この見方に立つと、ニューラルネットの普遍近似定理は、群作用に適合した原子分解・再構成公式として読み直すことができる。

リッジレットは Radon 変換・ウェーブレット変換・特異性解析と深く結びついた対象である。Murata の積分表現、Candès のリッジレット変換、Moritoh の超局所リッジレット変換、そして近年の深層リッジレット変換理論は、互いに別々の文脈から出発しながら、いずれも「適切な変換で係数化し、再構成する」という共通の構造を持つ。今後、この共通構

造を関数空間論の観点からさらに磨くことで、深層学習の数理に新しい見通しが与えられることを期待している。

参考文献

- [1] E. J. Candès, *Ridgelets: Theory and Applications*, PhD thesis, Stanford University, 1998.
- [2] 小林俊行, 大島利雄, リー群と表現論, 岩波書店, 2005.
- [3] S. Moritoh, Detection of singularities in wavelet and ridgelet analyses, *RIMS Kôkyûroku Bessatsu* **B57** (2016), 1–13.
- [4] N. Murata, An integral representation of functions using three-layered networks and their approximation bounds, *Neural Networks* **9** (1996), 947–956.
- [5] S. Sonoda, Y. Hashimoto, I. Ishikawa, and M. Ikeda, Deep ridgelet transform and unified universality theorem for deep and shallow joint-group-equivariant machines, in *Proc. 42nd International Conference on Machine Learning*, PMLR 267 (2025), 56480–56498.