Mathematics of Deep Learning, Summer Term 2020

Week 5

Harmonic Analysis

Philipp Harms Lars Niemann

University of Freiburg

Overview of Week 5

Acknowledgement of Sources

Sources for this lecture:

- Christensen (2016): An introduction to frames and Riesz bases
- Dahlke, De Mari, Grohs, Labatte (2015): Harmonic and Applied Analysis

Mathematics of Deep Learning, Summer Term 2020

Week 5, Video 1

Banach frames

Philipp Harms   Lars Niemann

University of Freiburg

Bases in Banach spaces

Definition (Schauder 1927)

Let X be a Banach space. A Schauder basis is a sequence $(e_k)_{k\in\mathbb{N}}$ in X with the following property: for every $f \in X$ there exists a unique scalar sequence $(c_k(f))_{k\in\mathbb{N}}$ such that

$$f = \sum_{k=1}^{\infty} c_k(f) e_k.$$

The Schauder basis is called unconditional if this sum converges unconditionally.

Remark:

- Any Banach space with a Schauder basis is necessarily separable.
- Not all separable Banach spaces have a Schauder basis (Enflo 1972).
- The coefficient functionals c_k are continuous, i.e., belong to X^*.

Translations, Modulations, and Scalings

Remark: Many useful bases are constructed by translations, modulations, and scalings of a given "mother wavelet."

Lemma

The following are unitary operators on $L^2(\mathbb{R})$, which depend strongly continuously on their parameters $a, b \in \mathbb{R}$ and $c \in \mathbb{R} \setminus \{0\}$:

- *Translation:* $T_a f(x) := f(x - a)$.
- *Modulation:* $E_b f(x) := e^{2\pi i b x} f(x)$.
- *Scaling (aka. dilation):* $D_c f(x) := c^{-1/2} f(x c^{-1})$.

Remark:

- These are actually group representations; more on this later.

Examples of Bases

Example: Fourier series

- The functions $(E_k 1)_{k\in\mathbb{Z}}$ are an orthonormal basis in $L^2([0,1])$.

Example: Gabor bases

- The functions $(E_k T_n \mathbb{1}_{[0,1]})_{k,n\in\mathbb{Z}}$ are an orthonormal basis in $L^2(\mathbb{R})$.

Example: Haar bases

- The functions $(D_{2^j} T_k \psi)_{j,k\in\mathbb{Z}}$ are an orthonormal basis of $L^2(\mathbb{R})$.
- Here ψ is the Haar wavelet

$$\psi(x) = \begin{cases} 1, & 0 \le x < \frac{1}{2}, \\ -1, & \frac{1}{2} \le x < 1, \\ 0, & \text{otherwise.} \end{cases}$$

Example: Wavelet bases

- Replace ψ by functions with better smoothness or support properties

Limitations of Bases

Requirements: continuous operations for

- Analysis: encoding f into basis coefficients (c_k)
- Synthesis: decoding f from basis coefficients (c_k)
- Reconstruction: writing $f = \sum_k c_k e_k$.

Limitations:

- It is often impossible to construct bases with special properties
- Even a slight modification of a Schauder basis might destroy the basis property

Idea: use "over-complete" bases, aka. frames

- Drop linear independence of (e_k) and uniqueness of (c_k)
- Require continuity of the analysis and synthesis operators
- Get additional benefits such as noise suppression and localization in time and frequency

Banach Frames

Definition (Gröchenig 1991)

Let X be a Banach space, and let Y be a Banach space of sequences indexed by $\mathbb{N}$. A Banach frame for X with respect to Y is given by

- Analysis: A bounded linear operator $A\colon X \to Y$, and
- Synthesis: A bounded linear operator $S\colon Y \to X$, such that
- Reconstruction: $S \circ A = \mathrm{Id}_X$.

Remark:

- The k-th frame coefficient is $c_k := \mathrm{ev}_k \circ A \in X^*$.
- If the unit vectors $(\delta_k)_{k\in\mathbb{N}}$ are a Schauder basis in Y, one obtains an atomic decomposition into frames $e_k := S\delta_k \in X$ as follows:

$$\forall f \in X : \qquad f = \sum_{k\in\mathbb{N}} c_k(f) e_k.$$

- Every separable Banach space has a Banach frame.

Examples of Banach frames

Example: Hilbert frames

- A Banach frame on a Hilbert space H with respect to ℓ^2 is a sequence $(e_k)_{k\in\mathbb{N}}$ s.t. for all $f \in H$,

$$\|f\|_H^2 \lesssim \sum_{k\in\mathbb{N}} |\langle f, e_k\rangle_H|^2 \lesssim \|f\|_H^2.$$

Example: Projections

- The projection of a Schauder basis to a subspace is a Banach frame.
- E.g., the functions $(E_k 1)_{k\in\mathbb{Z}}$ are a frame but not a basis in $L^2(I)$ for any $I \subsetneq [0,1]$.

Example: Wavelet frames

- If $\psi \in L^2(\mathbb{R}) \cap C^\infty(\mathbb{R})$ is required to have exponential decay and bounded derivatives, then $(D_{2^j} T_k \psi)_{j,k\in\mathbb{Z}}$ cannot be a basis but can be a frame.

Questions to Answer for Yourself / Discuss with Friends

- Repetition: What are Schauder bases versus frames?
- Repetition: Give some examples of frames constructed via translations, scalings, and modulations.
- Check: Is a Schauder basis a basis?
- Check: Verify the strong continuity of the translation, scaling, and modulation group actions.

Mathematics of Deep Learning, Summer Term 2020

Week 5, Video 2

Group representations

Philipp Harms   Lars Niemann

University of Freiburg

Locally compact groups

Definition (Locally compact group)

A locally compact group is a group endowed with a Hausdorff topology such that the group operations are continuous and every point has a compact neighborhood.

Theorem (Haar 1933)

Every locally compact group has a left Haar measure, i.e., a non-zero Radon measure which is invariant under left-multiplication. This measure is unique up to a constant. Similarly for right Haar measures.

Definition (Unimodular groups)

A group is unimodular if its left Haar measure is right-invariant.

Convolutions

Lemma (Young inequality)

For any $p \in [1, \infty]$, $f \in L^1(G)$, and $g \in L^p(G)$, the convolution

$$f * g(x) := \int_G f(y) g(y^{-1}x) dy = \int_G f(xy) g(y^{-1}) dy$$

is well-defined, belongs to L^p, and $\|f * g\|_{L^p(G)} \leq \|f\|_{L^1(G)} \|g\|_{L^p(G)}$.

Proof: This follows from Minkowski's integral inequality,

$$\left\| \int_G f(y) g(y^{-1} \cdot) dy \right\|_{L^p(G)} \leq \int_G |f(y)| \, \|g(y^{-1} \cdot)\|_{L^p(G)} dy,$$

and from the invariance of the L^p norm. □

Remark: The same conclusion holds for $g * f$ if G is unimodular or f has compact support.

Group Representations

Definition (Representation)

Let G be a locally compact group, and let H be a Hilbert space.

- A representation of G on H is a strongly continuous group homomorphism $\pi\colon G \to L(H)$.
- π is unitary if it takes values in $U(H)$.
- π is irreducible if $\{0\}$ and H are the only invariant closed subspaces of H, where invariance of $V \subseteq H$ means $\pi_g(V) \subseteq V$ for all $g \in G$.
- π is integrable if it is unitary, irreducible, and $\int_G |\langle \pi_g f, f\rangle_H| dg < \infty$ for some $f \in H$. Similarly for square integrability.

Remark: Unless stated otherwise, all integrals over G are with respect to the left Haar measure.

Questions to Answer for Yourself / Discuss with Friends

- Repetition: What is a square integrable representation of a locally compact group?
- Check: What condition is more stringent, integrability or square integrability? Hint: $g \mapsto \langle \pi_g f, f \rangle_H$ is continuous and bounded.
- Check: Suppose that π is reducible, can you extract a subrepresentation? Can you reduce it further down to an irreducible subrepresentation?
- Background: How are group representations related to group actions?
- Background: Look up the proof of Young's and Minkowski's inequalities!

Mathematics of Deep Learning, Summer Term 2020

Week 5, Video 3

Signal representations

Philipp Harms   Lars Niemann

University of Freiburg

Voice transform

Setting: Throughout, we fix a square-integrable representation $\pi\colon G \to U(H)$ of a locally compact group G on a Hilbert space H.

Definition (Voice transform)

For any $\psi \in H$, the voice transform (aka. representation coefficient) is the linear map

$$V_\psi\colon H \to C(G), \qquad V_\psi f(g) = \langle f, \pi_g \psi \rangle_H.$$

Remark:

- The voice transform represents signals in H as coefficients in $C(G)$.
- For any $\psi \neq 0$, injectivity of V_ψ is equivalent to irreducibility of π.

Orthogonality Relations

Theorem (Duflo–Moore 1976)

There exists a unique densely defined positive self-adjoint operator $A\colon D(A) \subseteq H \to H$ *such that*

- $V_\psi(\psi) \in L^2(G)$ *if and only if* $\psi \in D(A)$, *and*
- *For all* $f_1, f_2 \in H$ *and* $\psi_1, \psi_2 \in D(A)$,

$$\langle V_{\psi_1} f_1, V_{\psi_2} f_2 \rangle_{L^2(G)} = \langle f_1, f_2 \rangle_H \langle A\psi_2, A\psi_1 \rangle_H .$$

G *is unimodular if and only if* A *is bounded, and in this case* A *is a multiple of the identity.*

Remark:

- This is wrong without the square-integrability assumption on π.
- This is difficult to show in general but easy in many specific cases.
- An immediate consequence is the existence (even density) of such ψ.
- $V_\psi\colon H \to L^2(G)$ is isometric for any $\psi \in D(A)$ with $\|A\psi\| = 1$.

Equivalence to the regular representation

Definition (Regular representation)

The left-regular representation of G is the map

$$L\colon G \to U(L^2(G)), \qquad L_g F = F(g^{-1}\cdot).$$

Lemma

π is unitarily equivalent to a sub-representation of the left-regular representation, i.e., there exists an isometry $V : H \to L^2(G)$ such that $V \circ \pi_g = L_g \circ V$ holds for all $g \in G$.

Proof: Set $V = V_\psi$ for some $\psi \in D(A)$ with $\|A\psi\| = 1$ and use that

$$V \circ \pi_{g_1}(f)(g_2) = \langle \pi_{g_1} f, \pi_{g_2}\psi \rangle_H = \langle f, \pi_{g_1^{-1} g_2}\psi \rangle_H = L_{g_1} \circ V(f)(g_2). \qquad \square$$

Analysis, Synthesis, and Reconstruction

Lemma

Let $\psi \in D(A)$ *with* $\|A\psi\| = 1$.

- *Analysis:* $V_\psi \colon H \to L^2(G)$ *is an isometry onto its range,*

$$V_\psi(H) = \{F \in L^2(G) : F = F * V_\psi \psi\}.$$

- *Synthesis: The adjoint of* V_ψ *is given by the weak integral*

$$V_\psi^* \colon L^2(G) \to H, \qquad V_\psi^*(F) = \int_G F(g)\pi_g \psi \, dg.$$

- *Reconstruction: Every* $f \in H$ *satisfies* $f = V_\psi^* V_\psi f$.

Remark:

- This can be seen as a continuous Banach frame.
- The coefficient space is the reproducing kernel Hilbert space $V_\psi(H)$.

Proof: Analysis, Synthesis, and Reconstruction

Proof:

- V_ψ is isometric thanks to the orthogonality relation and $\|A\psi\|_H = 1$.
- V_ψ^* is given by the above weak integral because

$$\langle F, V_\psi f\rangle_{L^2(G)} = \int_G F(g)\langle \pi_g\psi, f\rangle_H dg = \left\langle \int_G F(g)\pi_g\psi \; dg, f \right\rangle_H .$$

- $V_\psi V_\psi^* F = F * V_\psi\psi$ because

$$\begin{aligned} V_\psi V_\psi^* F(g) &= \langle V_\psi^* F, \pi_g\psi\rangle_H = \langle F, V_\psi(\pi_g\psi)\rangle_{L^2(G)} \\ &= \langle F, L_g V_\psi\psi\rangle_{L^2(G)} = (F * V_\psi\psi)(g). \end{aligned}$$

- As V_ψ is isometric, $V_\psi^* V_\psi = \mathrm{Id}_H$ and $V_\psi V_\psi^*$ is the orthogonal projection onto the range of V_ψ, which equals the range of $V_\psi V_\psi^*$. □

Questions to Answer for Yourself / Discuss with Friends

- Repetition: What is the voice transform, and how does it lead to signal representations?
- Check: Where is square integrability of the representation used?
- Background: There is a definition of continuous frames—can you guess what it is and/or find it in the literature?
- Transfer: What is a reproducing kernel Hilbert space, and what is the relation to the condition $F * V_\psi \psi = F$?

Mathematics of Deep Learning, Summer Term 2020

Week 5, Video 4

Regular Coorbit Spaces

Philipp Harms Lars Niemann

University of Freiburg

Orbits and Coorbits

Setting: $\pi \colon G \to U(H)$ is a square integrable representation of a locally compact group G on a Hilbert space H, and A is the Duflo–Moore operator of π.

Remark:

- The orbit of π through $\psi \in H$ is $\{\pi_g \psi : g \in G\}$.
- V^* extends the action $\pi \colon G \times H \to H$ to

$$V^* \colon L^2(G) \times D(A) \to H, \qquad V_\psi^* F = \int_G F(g) \pi_g \psi \, dg.$$

Definition

Let X be a Banach subspace of $L^2(G)$, and let $\psi \in D(A)$.

- The orbit space associated to X and ψ is the subset $\{V_\psi^* F : F \in X\}$ of H with norm $\|f\| := \inf\{\|F\| : F \in X, V_\psi^* F = f\}$.
- The coorbit space associated to X and ψ is the set of all $f \in H$ such that $V_\psi f \in X$ with norm $\|f\| := \|V_\psi f\|_X$.

Weighted Spaces

Remark:

- The definitions of orbit and coorbit spaces work best when further structure is imposed on X.
- The main examples for X are weighted L^p spaces.

Definition

- A weight function is a continuous function $w\colon G \to \mathbb{R}_+$ which is submultiplicative and symmetric, i.e.,

$$w(gh) \leq w(g)w(h), \qquad w(g) = w(g^{-1}).$$

- The weighted space $L^p_w(G)$, $p \in [1, \infty]$, is defined as

$$L^p_w(G) \coloneqq \{F : Fw \in L^p(G)\}, \qquad \|F\|_{L^p_w(G)} \coloneqq \|Fw\|_{L^p(G)}.$$

Remark: $L^p_w(G)$ makes sense for arbitrary measurable functions w.

Properties of Weighted Spaces

Lemma

Let w be a weight function and $p \in [1, \infty]$.

1. *$L^p_w(G)$ is continuously included in $L^p(G)$.*
2. *The space $L^p_w(G)$ is L-invariant.*
3. *L acts strongly continuously on $L^p_w(G)$.*

Proof:

1. The symmetry of w implies $w(g)^2 = w(g)w(g^{-1}) \geq w(e) \geq 1$.
2. The submodularity of w implies that

$$\begin{aligned} \|L_g F\|_{L^p_w(G)} &= \|(L_g F)w\|_{L^p(G)} = \|F(L_{g^{-1}} w)\|_{L^p(G)} \\ &\leq w(g)\|Fw\|_{L^p(G)} = w(g)\|F\|_{L^p_w(G)}. \end{aligned}$$

3. It suffices to verify $\lim_{g \to e} \|L_g F - F\|_{L^2(G)} = 0$ for $F \in C_c(G)$. □

Regular Coorbit Spaces

Remark:

- The following coorbit space $H_{1,w}$ plays the role of test functions in the theory of distributions.
- More general coorbit spaces, which are not subspaces of H, are defined later on.

Definition

Let w be a weight function.

- An analyzing vector is a function $\psi \in D(A)$ with $\|A\psi\|_H = 1$ such that $V_\psi \psi \in L^1_w(G)$.
- $H_{1,w}$ is defined as the coorbit space associated to $L^1_w(G)$ and an analyzing vector ψ, i.e.,

$$H_{1,w} := \{f \in H : V_\psi f \in L^1_w(G)\}, \qquad \|f\|_{H_{1,w}} := \|V_\psi f\|_{L^1_w(G)}.$$

Correspondence Principle

Setting: We fix a weight function w and an analyzing vector ψ.

Theorem

The voice transform is an isometric isomorphism

$$V_\psi \colon H_{1,w} \to \{F \in L^1_w(G) : F = F * V_\psi \psi\}.$$

Proof:

- $X := \{F \in L^1_w(G) : F = F * V_\psi \psi\}$ is well-defined and a Banach subspace of $L^2(G)$ thanks to Young's inequality and $w \geq 1$:

$$\|F * V_\psi \psi\|_{L^2(G)} \leq \|F\|_{L^1(G)} \|V_\psi \psi\|_{L^2(G)} \leq \|F\|_{L^1_w(G)} \|V_\psi \psi\|_{L^2(G)}.$$

- The definition of the orbit and coorbit spaces is unaffected when $L^1_w(G)$ is replaced by X. □

Independence of the Analyzing Vector

Lemma

$H_{1,w}$ does not depend on the choice of analyzing vector ψ.

Proof:

- Let ψ_1, ψ_2, ψ_3 be analyzing vectors. We will show that $V_{\psi_1} f \in L^1_w(G)$ implies $V_{\psi_3} f \in L^1_w(G)$.
- By the orthogonality relations, one has for any $g \in G$ that

$$\begin{aligned} V_{\psi_1} f * V_{\psi_2} \psi_2(g) &= \langle V_{\psi_1} f, L_g V_{\psi_2} \psi_2 \rangle_{L^2(G)} = \langle V_{\psi_1} f, V_{\psi_2}(\pi_g \psi_2) \rangle_{L^2(G)} \\ &= \langle A\psi_2, A\psi_1 \rangle_H \langle f, \pi_g \psi_2 \rangle_H = \langle A\psi_2, A\psi_1 \rangle_H V_{\psi_2} f(g), \\ V_{\psi_1} f * V_{\psi_2} \psi_2 * V_{\psi_3} \psi_3 &= \langle A\psi_2, A\psi_1 \rangle_H V_{\psi_2} f * V_{\psi_3} \psi_3 \\ &= \langle A\psi_2, A\psi_1 \rangle_H \langle A\psi_3, A\psi_2 \rangle_H V_{\psi_3} f. \end{aligned}$$

- The left-hand side belongs to $L^1_w(G)$ by Young's inequality. Assuming wlog. that ψ_2 satisfies $\langle A\psi_1, A\psi_2 \rangle_H \neq 0 \neq \langle A\psi_2, A\psi_3 \rangle_H$, one deduces that $V_{\psi_3} f$ on the right-hand side belongs to $L^1_w(G)$. □

Further Properties

Lemma

$H_{1,w}$ is π-invariant, and π acts strongly continuously on it.

Proof: Correspondence $H_{1,w} \cong X := \{F \in L^1_w(G) : F = F * V_\psi \psi\}$

- $H_{1,w}$ is π-invariant because X is L-invariant.
- π acts strongly continuously on $H_{1,w}$ because L acts strongly continuously on X. □

Lemma

$H_{1,w}$ coincides with the orbit space associated to $L^1_w(G)$ and ψ.

Proof:

- $H_{1,w}$ is an orbit space because $H_{1,w} = V_\psi^* V_\psi H_{1,w} = V_\psi^* L^1_w(G)$. □

Questions to Answer for Yourself / Discuss with Friends

- Repetition: What is a (regular) coorbit space?
- Check: Are weighted L^p spaces Banach? Do they increase or decrease in p?
- Check: If $\lim_{g \to e} \|L_g F - F\|_{L^2(G)} = 0$ holds for all F in a dense subset of $L^2(G)$, why does it then hold for all F?

Mathematics of Deep Learning, Summer Term 2020

Week 5, Video 5

Duals of Coorbit Spaces

Philipp Harms Lars Niemann

University of Freiburg

Gelfand triples

Definition

A Gelfand triple is a triple (K, H, K^*), where K is a topological vector space, which is densely and continuously included in a Hilbert space H.

Lemma

Let (K, H, K^*) be a Gelfand triple. Then the inner product $\langle \cdot, \cdot \rangle_H$ extends to a sesquilinear form on $K^* \times K$.

Proof: Let $i \colon K \to H$ be the inclusion, and let $j = \langle \cdot, \cdot \rangle_H \colon H \to H^*$. Then $i^* \colon H^* \to K^*$ is injective because i has dense range, $i^* \circ j$ includes H into K^*, and the desired extension is just the duality $K^* \times K \to \mathbb{R}$. □

Gelfand Triples of Coorbit Spaces

Setting: $\pi\colon G \to U(H)$ is a square-integrable representation with Duflo–Moore operator A, w is a weight function, and ψ is an analyzing vector.

Lemma

The spaces $(H_{1,w}, H, H^*_{1,w})$ form a Gelfand triple.

Proof:

- $H_{1,w}$ is isomorphic via the voice transform to the space $\{F \in L^1_w(G) : F = F * V_\psi\psi\}$, which is continuously included in the space $\{F \in L^2(G) : F = F * V_\psi\psi\}$, which is isomorphic via the inverse voice transform to H.
- $H_{1,w}$ contains the orbit $\{\pi_g\psi : g \in G\}$ because

$$\|\pi_g\psi\|_{H_{1,w}} = \|V_\psi(\pi_g\psi\|_{L^1_w(G)} = \|L_g V_\psi\psi\|_{L^1_w(G)} \lesssim \|V_\psi\psi\|_{L^1_w(G)} < \infty.$$

  The orbit is dense in H because π is irreducible. □

Duals of Coorbit Spaces

Remark: As $H_{1,w}$ plays the role of test functions, $H^*_{1,w}$ plays the role of distributions.

Definition

The extended voice transform is defined for any $f \in H^*_{1,w}$ and $g \in G$ as

$$V_\psi(f)(g) \coloneqq \langle f, \pi_g \psi \rangle_{H^*_{1,w} \times H_{1,w}} .$$

Remark: This extends the voice transform on H because the dual pairing between $H^*_{1,w}$ and $H_{1,w}$ extends the inner product on H.

Correspondence Principle

Remark: $L^1_w(G)^* = L^\infty_{1/w}(G)$.

Theorem (Correspondence principle)

$V_\psi \colon H^*_{1,w} \to \{F \in L^\infty_{1/w} : F = F * V_\psi \psi\}$ *is an isometric isomorphism.*

Proof: In the proof of the correspondence principle for the regular voice transform, replace the Hilbert inner product on H by the dual pairing between $H^*_{1,w}$ and $H_{1,w}$. □

Questions to Answer for Yourself / Discuss with Friends

- Repetition: How does the voice transform extend to duals of coorbit spaces?
- Check: If (K, H, K^*) is a Gelfand triple, and H is seen as a subspace of K^*, how are elements of H applied to elements of K?
- Check: Prove that the topological dual of $L^1_w(G)$ is $L^\infty_{1/w}(G)$.
- Transfer: What Gelfand triples are used to define distributions and tempered distributions?

Mathematics of Deep Learning, Summer Term 2020

Week 5, Video 6

General Coorbit Spaces

Philipp Harms Lars Niemann

University of Freiburg

Weighted Spaces

Setting: $\pi\colon G \to U(H)$ is a square-integrable representation with Duflo–Moore operator A, w is a weight function, and ψ is an analyzing vector subject to some further conditions.[1]

Definition

- A w-moderate weight is a continuous function $m\colon G \to \mathbb{R}_+$ satisfying

$$m(ghk) \leq w(g)m(h)w(k), \qquad g, h, k \in G.$$

- The weighted space $L^p_m(G)$ is defined for any $p \in [1, \infty]$ as

$$L^p_m(G) := \{F : Fm \in L^p(G)\}, \qquad \|F\|_{L^p_m(G)} := \|Fm\|_{L^p(G)}.$$

Remark:

- This extends the def. of $L^p_w(G)$ since w is a w-moderate weight.
- $\|\cdot\|_{L^p_w(G)}$ is a norm, but $\|\cdot\|_{L^p_m(G)}$ may be only a seminorm.

[1] See Theorem 3.12 in Dahlke, De Mari, Grohs, Labatte (2015).

Coorbit Spaces

Setting: We fix a w-moderate weight m.

Definition

The coorbit space $H_{p,m}$ is defined as

$$H_{p,m} := \{F \in H_{1,w}^* : V_\psi(F) \in L_m^p(G)\}.$$

Remark:

- This extends the definition of $H_{1,w}$, and $H = H_{2,1}$.
- $H_{p,m}$ is independent of the choice of analyzing vector ψ.
- $H_{p,m}$ coincides as a set with an orbit space.

Theorem (Correspondence principle)

Under an additional condition on ψ, the voice transform $V_\psi \colon H_{p,m} \to \{F \in L_m^p(G) : F = F * V_\psi \psi\}$ is an isometric isomorphism.

Structure of Coorbit Spaces

Uniqueness: $H_{p_1,m_1} = H_{p_2,m_2}$ if and only if $p_1 = p_2$ and $m_1 \lesssim m_2 \lesssim m_1$.

Duality: $H_{p,m}^* = H_{q,1/m}$ for any $p \in [1, \infty)$ and $\frac{1}{p} + \frac{1}{q} = 1$.

Embeddings: $H_{p,m}$ is increasing in p and decreasing in m.

Compact Embeddings: H_{p_1,m_1} embeds compactly in H_{p_2,m_2} if $m_1/m_2 \in L^r(G)$ for some $r \leq \frac{1}{p_2} - \frac{1}{p_1} > 0$.

Complex Interpolation: For any $\theta \in [0,1]$ and $p_1 < \infty$, $[H_{p_1,m_1}, H_{p_2,m_2}]_\theta = H_{p,m}$ with $\frac{1}{p} = \frac{1-\theta}{p_1} + \frac{\theta}{p_2}$ and $m = m_1^{1-\theta} m_2^\theta$.

Generalizations: $L_m^p(G)$ is a left- and right-invariant solid Banach function space on G, and coorbit spaces can be defined for such spaces.

Questions to Answer for Yourself / Discuss with Friends

- Repetition: How are (general) coorbit spaces $H_{p,m}$ defined?
- Check: $H_{p,m} \subseteq H^*_{1,w}$ implies $L^p_m(G) \subseteq L^1_w(G)^*$—how can this be seen directly? Hint: show that $m(e) = m(gg^{-1}) \lesssim m(g)w(g^{-1})$.
- Background: Read up on duality, embedding, and interpolation properties of L^p spaces.

Mathematics of Deep Learning, Summer Term 2020

Week 5, Video 7

Discretization

Philipp Harms   Lars Niemann

University of Freiburg

Towards Banach Frames on Coorbit Spaces

Setting: $\pi \colon G \to U(H)$ is a square-integrable representation with Duflo–Moore operator A, w is a weight function, m is a w-moderate weight, $p \in [1, \infty]$, and ψ is an analyzing vector subject to some further conditions.[2]

Strategy:

- Define a Banach frame for $\{F \in L^p_m(G) : F = F * V_\psi \psi\}$ via left-translations of the kernel $V_\psi \psi$, i.e., by writing

$$F = \sum_k c_k(F) L_{g_k} V_\psi \psi$$

  for a well-chosen sequence of $g_k \in G$.
- Get a Banach frame for $H_{p,m}$ via the correspondence principle.

---

[2]See Theorem 3.19 in Dahlke, De Mari, Grohs, Labatte (2015).

Density and Separation

Remark: Intuitively, translations of a kernel by (g_k) are a frame if (g_k) spreads out over all of G and does not accumulate anywhere.

Definition

A sequence $(g_k)_{k\in\mathbb{N}}$ in G is called

- U-dense if U is a compact neighborhood of $e \in G$ and $\bigcup_k L_{g_k}U = G$.
- separated if there exists a compact neighborhood U of $e \in G$ such that $L_{g_k}U \cap L_{g_l}U = \emptyset$ for $k \neq l$.
- relatively separated if it is a finite union of separated sequences.

Banach Frames on Weighted Spaces

Definition

The weighted sequence space ℓ^p_m is defined as

$$\ell^p_m := \{\lambda\colon \lambda m \in \ell^p\}, \qquad \|\lambda\|_{\ell^p_m} := \|\lambda m\|_{\ell^p}.$$

Theorem

If U is a sufficiently small neighborhood of $e \in G$ and (g_k) is a U-dense and relatively separated sequence in G, then $(L_{g_k} V_\psi \psi)_{k\in\mathbb{N}}$ is a Banach frame for $X := \{F \in L^p_m(G) : F = F * V_\psi \psi\}$ with respect to ℓ^p_m.

Remark: the frame coefficients are specified in the proof.

Proof: Banach Frames on Weighted Spaces

Proof for $p = 1$ and $m = w$:

- Let (Ψ_k) be a partition of unity subordinated to $(L_{g_k}U)$.
- We define some preliminary analysis and synthesis operators:

$$X \ni F \mapsto (\langle \Psi_k, F\rangle_{L^2(G)})_{k\in\mathbb{N}} \in \ell^1_w, \qquad \ell^1_w \ni \lambda \mapsto \sum_k \lambda_k L_{g_k} V_\psi \psi \in X.$$

- These operators are well-defined and continuous: letting $C := \sup_{g\in U} w(z)$, one has

$$\begin{aligned}
\|(\langle \Psi_k, F\rangle_{L^2(G)})_{k\in\mathbb{N}}\|_{\ell^1_w} &= \sum_k |\langle \Psi_k, F\rangle_{L^2(G)}| w(g_k) \\
&\leq C \sum_k \langle \Psi_k, |F| w\rangle_{L^2(G)} = C\|F\|_{L^1_w(G)}, \\
\Big\| \sum_k \lambda_k L_{g_k} V_\psi \psi \Big\|_{L^1_w(G)} &\leq \sum_k |\lambda_k| \|L_{g_k} V_\psi \psi\|_{L^1_w(G)} \\
&\leq \sum_k |\lambda_k| w(g_k) \|V_\psi \psi\|_{L^1_w(G)} = \|\lambda\|_{\ell^1_w} \|V_\psi \psi\|_{L^1_w(G)}.
\end{aligned}$$

Proof: Banach Frames on Weighted Spaces (cont.)

- The reconstruction operator (i.e., analysis followed by synthesis),

$$R\colon X \to X, \qquad RF := \sum_{k\in\mathbb{N}} \langle F, \Psi_k\rangle L_{g_k} V_\psi \psi,$$

tends to Id_X as U tends to $\{e\}$ because for any $F \in X$,

$$\begin{aligned}
&\Big\| F * V_\psi\psi - \sum_k \langle \Psi_k, F\rangle_{L^2(G)} L_{g_k} V_\psi\psi \Big\|_{L^1_w(G)} \\
&\quad = \Big\| \sum_k \int_G F(g)\Psi_k(g)(L_g - L_{g_k}) V_\psi\psi dg \Big\|_{L^1_w(G)} \\
&\quad \le \sum_k \langle \Psi_k, |F|\rangle_{L^2(G)} \sup_{g\in L_{g_k}U} \|(L_g - L_{g_k})V_\psi\psi\|_{L^1_w(G)} \\
&\quad \le \sum_k \langle \Psi_k, |F|\rangle_{L^2(G)} w(g_k) \sup_{u\in U} \|(L_u - \mathrm{Id})V_\psi\psi\|_{L^1_w(G)} \\
&\quad \le C\|F\|_{L^1_w(G)} \sup_{u\in U} \|(L_u - \mathrm{Id})V_\psi\psi\|_{L^1_w(G)} \to 0.
\end{aligned}$$

Proof: Banach Frames on Weighted Spaces (cont.)

- R is invertible for sufficiently small U because Id_X is invertible and invertible operators are open.
- Any $F \in X$ can be written as

$$F = RR^{-1}F = \sum_{k\in\mathbb{N}} \langle \Psi_k, R^{-1}F\rangle_{L^2(G)} L_{g_k} V_\psi \psi.$$

- Thus, the desired Banach frame for X with respect to ℓ^1_w is

$$e_k := L_{g_k} V_\psi \psi \in X, \quad c_k := \langle \Psi_k, R^{-1}(\cdot)\rangle_{L^2(G)} \in X^*, \quad k \in \mathbb{N}. \qquad \square$$

Banach Frames for Coorbit Spaces

Corollary

If U is a sufficiently small neighborhood of $e \in G$ and (g_k) is a U-dense and relatively separated sequence in G, then $(\pi_{g_k}\psi)_{k\in\mathbb{N}}$ is a Banach frame for $H_{p,m}$ with respect to ℓ^p_m.

Proof: Apply the isomorphism $V_\psi^{-1}\colon X \to H_{p,m}$. □

Harmonic Analysis and Neural Networks

- Let G be a sub-group of the affine group $GL(\mathbb{R}^d) \ltimes \mathbb{R}^d$, and define

$$\pi\colon G \to U(L^2(\mathbb{R}^d)), \qquad \pi_{(A,b)}(f)(x) = \det(A)^{-1/2} f(A^{-1}(x-b)).$$

- Then coorbit theory provides continuous and discrete representations

$$\begin{aligned} f(x) &= \int_G F(A,b)\det(A)^{-1/2}\psi(A^{-1}(x-b))dAdb \\ &= \sum_k c_k \det(A_k)^{-1/2}\psi(A_k^{-1}(x-b_k)), \end{aligned}$$

where ψ is a suitable analyzing vector, with an equivalence of norms

$$\|F\|_{L^p_m(G)} \simeq \|c_k\|_{\ell^p_m} \simeq \|f\|_{H_{p,m}}.$$

- These representations can be interpreted as infinite-width multi-layer perceptrons with activation function ψ.

Questions to Answer for Yourself / Discuss with Friends

- Repetition: How are Banach frames of weighted spaces and coorbit spaces constructed?
- Background: Refresh your memory of the definition and construction of partitions of unity.
- Check: Why is the set of invertible operators open in the set of bounded linear operators?
- Discussion: How could coorbit theory be used to derive approximation bounds of neural networks?

Mathematics of Deep Learning, Summer Term 2020

Week 5, Video 8

Wrapup

Philipp Harms Lars Niemann

University of Freiburg

Outlook on this week's discussion and reading session

- Reading:
  - Feichtinger Groechenig (1988): A unified approach to atomic decompositions
  - Dahlke, De Mari, Grohs, Labatte (2015): Harmonic and Applied Analysis
- Numerical Example:
  - Some wavelet transforms in image analysis.

Summary by learning goals

Having heard this lecture, you can now. . .

- Describe bases and frames in Hilbert and Banach spaces.
- Build signal representations from group representations.
- Interpret such representations as multi-layer perceptrons.