Frame Theory for Understanding Neural Networks

Peter Balazs, Martin Ehler

Although artificial neural networks have prevailed as the workhorses of modern machine learning approaches, the mathematical theory behind them might be considered still lacking in some areas [5]. To comply with this demand, we evoke *frame theory*, an active field in functional analysis that has successfully improved the understanding of many engineering problems by connecting branches of pure and applied mathematics with the applied sciences [3]. It is a linear theory, while a layer of a neural network is a composition of an (affine) linear function and a non-linearity. To study such a layer with frame theoretic tools, we propose to

incorporate non-linearities into frame theory

for enabling a better theoretical understanding of neural networks.

Specific questions include:

- *Non-linear Frames*: Let $\sigma : \mathbb{C} \to \mathbb{C}$ be a non-linear function. For $(\psi_k)_{k\in\mathbb{N}} \subset \mathcal{H}$ in some Hilbert space, the standard frame inequalities are generalized by

$$A \|f\|_{\mathcal{H}}^2 \leq \sum_{k\in\mathbb{N}} |\sigma(\langle f, \psi_k \rangle)|^2 \leq B \|f\|_{\mathcal{H}}^2 . \quad (1)$$

  When does this hold and how do A and B depend on σ? Can we derive an analogous machinery as for standard - linear - frames?

- *Generalized Phase Retrieval*: As an important form of signal reconstruction [1], we aim to understand the relations between σ and $(\psi_k)_{k\in\mathbb{N}}$, so that the operator

$$T : f \mapsto (\sigma(\langle f, \psi_k \rangle))_{k\in\mathbb{N}} ,$$

  is injective respectively boundedly invertible.

- *Non-linear Kernels*: Let $\rho : \mathcal{H} \to \mathcal{H}'$ be a non-linear function mapping from one Hilbert space into another one - in a machine learning context usually a higher dimensional one. Define $\langle f, \psi_k \rangle_\rho := \langle \rho(f), \rho(\psi_k) \rangle_{\mathcal{H}'}$ and investigate a nonlinear frame inequality

$$A \|f\|_{\mathcal{H}}^2 \leq \sum_{k\in\mathbb{N}} \left| \langle f, \psi_k \rangle_\rho \right|^2 \leq B \|f\|_{\mathcal{H}}^2 \quad (2)$$

  and consequences thereof. Can this provide more theory for the “kernel trick” in support vector machines? Can this be compared to results as in e.g. [2, 4]?

Further questions will arise along the way. The successful applicant will be integrated into ARI and NuHAG. He/she will closely work with the two PhD advisors and a current PhD student studying a connected topic.

References

[1] C. Bachoc and M. Ehler. Signal reconstruction from the magnitude of subspace components. IEEE Trans. Inform. Theory, 61(7):1-13, 2015.

[2] P. Balazs, K. Gröchenig, and M. Speckbacher. Kernel theorems in coorbit theory. *T. Am. Math. Soc.*, 6:346–364, 2019.

[3] P. Balazs, N. Holighaus, T. Necciari, and D. Stoeva. Frame theory for signal processing in psychoacoustics. In R. Balan, J. J. Benedetto, W. Czaja, and K. Okoudjou, editors, *Excursions in Harmonic Analysis Vol. 5,*, pages 225–268. Springer, 2017.

[4] M. Ehler and M. Gräf. Reproducing kernels for the irreducible components of polynomial spaces on unions of Grassmannians. Constr. Approx., 49(1):29-58, 2018.

[5] I. Goodfellow, Y. Bengio, and A. Courville. *Deep Learning*. MIT Press, 2016. `http://www.deeplearningbook.org`.