Department of Mathematics and Mathematical Statistics

GeUmetric Deep Learning

June 11-13 2024 Umeå

Tuesday		Wednesday		Thursday	
		9:00-10:00	Bekkers	9:00-10:00	Chmiela
		10:00-10:40	Nyholm	10:00-10:40	Mereta
		COFFEE BREAK			
		11:00-11:40	Walden	11:00-11:40	Henry
		11:40-12:20	Hendi	11:40-12:20	Vermant
				12:20-13:00	Nordenfors
		LUNCH			
14:00-14:10	Intro	13:30-14:30	Kessel		
14:10-14:50	Bökman	14:30-15:10	Misof		
COFFEE BREAK					
15:10-15:50	Marchetti	15:30-16:30	Cluster Meeting		
15:50-16:30	Andersdotter	17:00-18:20	Social event		
16:30-17:10	Carlsson	19:00-	Conference dinner		

Tuesday

14.00 – 14:50 **Welcome**

Georg Bökman *On the usefulness of equivariant image representations*

In representation learning, the aim is to obtain useful latent representations of the data, such that the latent representation facilitates downstream tasks. Often, neural networks are trained to output latent representations that are invariant to perturbations of the input data. Recently, several independent works have argued that letting the latent representation change predictably when the input is perturbed is beneficial. Examples include Joint-Embedding Predictive Architectures (JEPAs) and our work on Steerers. In the Steerers work, we specifically let the latent representation be linearly equivariant to input transformations. We study keypoint description and matching under geometric transformations of the input images and demonstrate how equivariance is both a natural consequence of the data and outperforms invariant latent representations.

14:50 – 15:10 **Coffee break**

15:10 – 17:10 **Giovanni Luca Marchetti** *Harmonics of Learning*

In this talk, I will present a recent work in collaboration with C. Hillar, D. Kragic and S. Sanborn. We formally prove that, under certain conditions, if a neural network is invariant to a finite group then its weights recover the Fourier transform on that group. This provides a mathematical explanation for the emergence of Fourier features -- a ubiquitous phenomenon in both biological and artificial learning systems. The results hold even for non-commutative groups, in which case the Fourier transform encodes all the irreducible unitary group representations. Our findings have consequences for the problem of symmetry discovery. Specifically, we demonstrate that the algebraic structure of an unknown group can be recovered from the weights of a network that is at least approximately invariant within certain bounds. Overall, this work contributes to a foundation for an algebraic learning theory of invariant neural network representations.

Emma Andersdotter *Equivariant Manifold Neural ODEs and Differential Invariants*

Neural ODEs are neural network models where the network is not specified by a discrete sequence of hidden layers. Instead, the network is defined by a vector field describing how the data evolves continuously over time governed by an ordinary differential equation (ODE). These models can be generalized for data living on non-Euclidean manifolds, a concept known as manifold neural ODEs. In our paper, we develop a geometric framework for equivariant manifold neural ODEs. Our work includes a novel formulation of equivariant neural ODEs in terms of differential invariants, based on Lie theory for symmetries of differential equations. We also construct augmented manifold neural ODEs and show that they are universal approximators of equivariant diffeomorphisms on any path-connected manifold.
Link to paper being presented: arXiv/2401.14131

Oscar Carlsson *HEAL-SWIN: A Vision Transformer On The Sphere*

High-resolution wide-angle images, such as fisheye images, are increasingly important in applications like robotics and autonomous driving. Traditional neural networks struggle with these images due to projection and distortion losses when operating on their flat projections. In this presentation, I will introduce the HEAL-SWIN model, which addresses this issue by combining the SWIN transformer with the Hierarchical Equal Area iso-Latitude Pixelation (HEALPix) grid from astrophysics. This integration enables the HEAL-SWIN model to process inherently spherical data without projections, effectively eliminating distortion losses.

Wednesday

9:00 – 10:40 **Erik Bekkers** *Neural Ideograms and Geometry-Grounded Representation Learning*

In this talk I will present the idea of Neural Ideograms, which I frame as learnable geometric representation of data. This work is inspired by the utilitarian function of pictograms and ideograms in conveying abstract concepts effectively; I present several methods for learning ideographic representations of data [1,2]. I start with the assertion that in order for a geometric representation to be meaningful, it should be grounded in the same geometry as the data. That is, the learned representations should be amenable to the same transformation laws that are applicable to the data that it represents, and thus the representation should be obtained in an equivariant manner. I'll present two methods for constructing neural ideograms: (1) via a Kendall-Shape space VAE approach [1] and (2) via our recently proposed Equivariant Neural Fields (ENFs) [2]. In particular, we show that the ENFs enable a whole new approach to geometry-grounded deep learning, which we demonstrate with the application to PDE forecasting [3]. Here we learn to associate latent geometric representations (neural ideograms) to dense fields, and model the PDE as a (learnable) ODE in latent space, in a fully equivariant manner. Finally, I present a new general purpose equivariant architecture that is both efficient and effective [4], and can be used to solve all sorts of tasks that require equivariance, such as the analyses of the above neural ideograms.

[1] Vadgama, S., Tomczak, J. M., & Bekkers, E. J. (2022, November). Kendall shape-VAE: Learning shapes in a generative framework. In _NeurIPS 2022 Workshop on Symmetry and Geometry in Neural Representations_.
[2] D. Wessels, D. M. Knigge, S. Papa, R. Valperga, S. Vadgama, S. Gavves, and E. J. Bekkers. “Grounding Continuous Representations in Geometry: Equivariant Neural Fields”. In: arXiv preprint arXiv (2024).
[3] D. M. Knigge, D. Wessels, R. Valperga, S. Papa, J.-J. Sonke, S. Gavves, and E. J. Bekkers. “Space-Time Continuous PDE Forecasting using Equivariant Neural Fields”. In: arXiv preprint arXiv (2024).
[4] E. J. Bekkers, S. Vadgama, R. Hesselink, P. A. V. der Linden, and D. W. Romero. “Fast, Expressive SE(n) Equivariant Networks through Weight-

Sharing in Position-Orientation Space". In: The Twelfth International Conference on Learning Representations. 2024.

Elias Nyholm *A Mathematician's Guide to Equivariant Transformers*

Equivariant neural networks of many kinds are nowadays well-established and extensively used in a range of tasks including molecular modelling, autonomous driving and particle simulations. However the bulk of such models, and the mathematical theories developed to describe them, are based on convolutional and graph neural networks. Similar theory for equivariant transformers and (self-)attention remains largely unexplored. In this presentation we survey current equivariant transformer efforts and point to possible ways to unify them under one mathematical framework.

10:40 – 11:00 **Coffee break**

11:00 – 12:20 **Moritz Walden** *Learning Group Invariant CY Metrics - Part I*

In this talk, we give an introduction to Calabi-Yau manifolds and discuss the concept of a Ricci-flat metric on them. Furthermore, we give an introduction to cymetric, a library that allows approximating said metric, leveraging machine learning techniques. We discuss point sampling as well as choices of different models and loss functions.

Yacoub Hendi *Learning Group Invariant CY Metrics - Part II*

In this talk, we present the results of our recent work on how to encode discrete symmetries of an NN-model that predicts the Ricci flat metric on a given Calabi-Yau manifold. The main idea is to use projections on fundamental domians of the symmetry groups to construct untrainable layers that are invariant under the action of the groups. We show that this leads to improvement in some of the losses. We also use this scheme to approximate the Ricci metric on the quotient space of a smooth Calabi-Yau by a freely acting group.

12:20 – 13:30 **Lunch**

13:30 – 15:10 **Pan Kessel** *Emergent Equivariance in Deep Ensembles*

Deep Ensembles are a powerful and widely used tool to estimate uncertainty of neural network predictions. Using Neural tangent kernel theory, we show that they enjoy emergent equivariance. Specifically, deep ensembles trained with full data augmentation are equivariant at all training times and for any input (even off-manifold).

Phillipp Misof *Neural Tangent Kernel for Equivariant Neural Networks*

The neural tangent kernel (NTK) is a quantity closely related to the training dynamics of neural networks (NNs). It becomes particularly interesting in the infinite width limit of NNs, where this kernel becomes deterministic and time-independent, allowing for an analytical solution of the gradient descent dynamics under the mean squared error loss, resulting in a Gaussian process behaviour. In this talk, we will first introduce the NTK and its properties, and then discuss how it can be extended to NNs equivariant with respect to the regular representation. In analogy to the forward equation of the NTK of conventional NNs, we will present a recursive relation connecting the NTK to the corresponding kernel of the previous layer in an equivariant NN. As a concrete example, we provide explicit expressions for the symmetry group of 90° rotations and translations in the plane, as well as Fourier-space expressions for SO(3) acting on spherical signals. We support our theoretical findings with numerical experiments.

15:10 – 15:30	**Coffee break**	
15:30 – 16:30	**Cluster meeting**	
17:00 - 18:20	**Social event**	Galaxen, Universum
19:00 -	**Conference dinner**	Båten

Thursday

9:00 – 10:40	**Stefan Chmiela**	*The Unreasonable Effectiveness of Symmetric Priors in Quantum Chemistry*

The convolution operation is a cornerstone of many machine learning (ML) model architectures. Graph neural networks generalize the convolutional paradigm to unstructured domains and beyond the translation group, most notably the special orthogonal group SO(3). This methodological advance was originally driven by the demands of physical simulations, but now finds wide applicability. Specifically, the reduced-order representation of highly non-linear interactions in quantum-systems necessitates a separation into rotationally invariant and equivariant operations to accurately capture different properties. This talk will evaluate the current capabilities of GNNs, show how they are applied to accurate simulations of quantum systems over long timescales, and explore their limitations. Additionally, it will highlight how this application is poised to drive further advancements in GNN methodologies, pushing the frontiers of this rapidly evolving field.

Stefano Mereta — *Tropical differential algebra and its fundamental theorem*

After introducing some basics of tropical geometry and discussing briefly its fundamental theorem, we will introduce tropical differential algebra and discuss a fundamental theorem in this context. We extend results from Aroca et al., Falkensteiner et al. and from Fink and Toghani, holding only in the case of trivial valuation as introduced by Grigoriev, to (partial) differential equations with power series coefficients over any valued field. To do so, a crucial ingredient is the framework for tropical differential equations introduced by Giansiracusa and Mereta. As a corollary of the fundamental theorem, the radius of convergence of solutions of a differential equation over a nontrivially valued field can be computed tropically. This talk is based on results appearing in arXiv:2303.12124, joint with F. Gallinaro.

10:40 – 11:00	**Coffee break**	

11:00 – 13:00

Nathan Henry — *The Geometry of Linear Attention*

Single-headed self-attention is an essential component of many modern machine learning models. However, while its effectiveness is evident empirically, little is understood theoretically. In this paper, we use ideas from algebraic geometry to describe the hypothesis space of the linear attention mechanism. We find the dimension of this space, find its singularities, and show that it is Euclidean closed.

Joannes Vermant — *Structural rigidity and flexibility using group theory*

In kinematics, one wants to consider how articulated structures can move. One model, used often in rigidity theory, is that of a bar-joint framework, which is a graph which is embedded in a Euclidean space. We have given a description of such structure using group theory, which gives rise to a graph of groups. Motions of such a structure can also be described using group theory and find necessary conditions for a realisation to be rigid, meaning the structure does not admit any motions. Other problems, arising for example in combinatorial geometry, can also be modelled in the exact same way by varying the group.

Oskar Nordenfors — *Restricted optimization vs. data augmentation for equivariant neural networks*

In this talk I will present joint work with Axel Flinth and Fredrik Ohlsson, comparing training dynamics of two strategies for achieving equivariance in neural networks. Restricting layers to be equivariant vs. training on augmented data. Under a condition on the geometry of the admissible, equivariant layers, along with some natural assumptions on the data, network, and loss, the two strategies have the same set of equivariant stationary points for the gradient flow. Under an additional condition on the parametrization of the linear layers, the set of equivariant layers is invariant under the gradient flow of the augmentation strategy.