Oberwolfach Seminar 2448b
24. - 29. November 2024

Control and Machine Learning

Borjan Geshkovski (MIT)
Domènec Ruiz-Balet (Imperial College London)
Enrique Zuazua (FAU Erlangen-Nürnberg)

Program

This seminar will address the recent developments and new horizons of the burgeoning field found at the intersection of control theory and machine learning. This perspective has shown to be fruitful in analyzing and improving a multitude of problems of interest in machine learning. These include the understanding of interpolation and generalization properties of deep neural networks, certificates for pruning of architectures via turnpike properties, normalizing flows for generative modeling, and the clustered representations of transformers–to name a few.

Simultaneously, this has challenged and still challenges the state of the art in mathematical control theory. Indeed, tackling the new questions arising in machine learning requires new control formulations and tools, both at the analysis and computational level and, oftentimes, integrating know-how from other areas such as high dimensional functional analysis and probability.

The intention of the three-part lecture series is to address the collective body of work in this field in a synergistic manner. Beginning with fundamental concepts, we gradually delve into a comprehensive analysis of the underlying models, paving the way for a deeper understanding of practical applications.

The contents of the Seminar are planned to be:

Lecture 1:
Control theoretic perspective to supervised learning and the evolutionary-steady correspondence
Enrique Zuazua

This lecture will establish the fundamental control theory perspective on machine learning, by interpreting supervised learning for ResNets as a simultaneous control problem for neural ODEs ([E17]). The latter involves the standard two-layer network ansatz of the classical Theorem of Universal Approximation, iterated over continuous time. We will also discuss second-order developments such as momentum ResNets [RBAZ22], which circumvent the issue of trajectory crossing common in classification tasks. To dissect the interpolation and approximation capabilities of these models, we will introduce the most relevant general elements from control theory, namely the Kalman rank condition and Lie brackets. We will also address their pitfalls when applied to the simultaneous control of neural ODEs, raising the need for intrinsic nonlinear strategies, which will also be discussed in the second lecture. Following the philosophy of fundamentals, the lecture will evolve on the presentation of the Pontryagin maximum principle and qualitative theory of optimal controls. To this end, the lecture will also present the turnpike property from optimal control [GZ22], which is the fundamental mechanism establishing correspondence between evolutionary and steady models, and discuss its applications.

Lecture 2:
Geometry of neural ODEs through control
Domènec Ruiz-Balet

Building on the introduced fundamentals, this lecture will further probe the working mechanisms of neural ODEs and the simultaneous control process. We will present the nonlinear control strategy conceived in [RBZ23] which ensures interpolation and approximation for neural ODEs by leveraging the geometry endowed within these models. The implications of this approach on generalization bounds and adversarial attacks will be addressed, drawing parallels with classical techniques based on Rademacher complexity estimates [EL18] and other statistical techniques. The lecture will also highlight the relevance of the choice of the activation function to achieve these goals. Recognizing the established correspondence between ODEs and their associated transport PDEs, the application of normalizing flows in generative modeling will also be explored. We will illustrate how neural ODE flows can be effectively harnessed to devise a transport map between two probability distributions [RBZ24], building analogies with other classical topics in transportation theory. The lecture will finish presenting a mean-field formulation of deep learning [EL18] and observing the vector transport formulation of classification tasks [RBZ23].

Lecture 3:
Transformers, clustering, and control of interacting particle systems
Borjan Geshkovski

This lecture will focus on transformers, which are the cornerstone of contemporary large language models. These models involve the combination of neural ODE blocks along with a self-attention mechanism. Transformers can be interpreted as interacting particle systems, for which the learning problem can once again be seen as a simultaneous control problem. We will survey the prevailing literature on the control of interacting particle systems, primarily developed for models on collective behavior such as the Cucker-Smale model. Additionally, we will delve into the idiosyncrasies of the associated mean-field transport PDE. We will then describe the geometry in which particles arrange over a long time (corresponding to many layers) for various weights of the network, leading to different clustering results [GLPR23]. The intrinsic relationship to Voronoi tessellations will also be uncovered.

Conclusions and complements

This series of lectures will be complemented by the presentation of the computational tools developed by the speakers and some relevant challenging open problems.

References

[E17] Weinan E. A proposal on machine learning via dynamical systems. *Communications in Mathematics and Statistics*, 5(1):1–11, 2017.

[EL18] Han Jiequn E, Weinan and Qianxiao Li. A mean-field optimal control formulation of deep learning. *Research in the Mathematical Sciences*, 1(6):1–41, 2018.

[GLPR23] Borjan Geshkovski, Cyril Letrouit, Yury Polyanskiy, and Philippe Rigollet. The emergence of clusters in self-attention dynamics. *arXiv preprint arXiv:2305.05465*, 2023.

[GZ22] Borjan Geshkovski and Enrique Zuazua. Turnpike in optimal control of PDEs, ResNets, and beyond. *Acta Numerica*, 31:135–263, 2022.

[RBAZ22] Domenec Ruiz-Balet, Elisa Affili, and Enrique Zuazua. Interpolation and approximation via momentum resnets and neural ODEs. *Systems & Control Letters*, 162:105182, 2022.

[RBZ23] Domènec Ruiz-Balet and Enrique Zuazua. Neural ODE Control for Classification, Approximation, and Transport. *SIAM Review*, 65(3):735–773, 2023.

[RBZ24] Domènec Ruiz-Balet and Enrique Zuazua. Control of neural transport for normalising flows. *Journal de Mathématiques Pures et Appliquées*, 181:58–90, 2024.