A SIMPLE YET EFFECTIVE TENSOR-BASED ODE MODEL FOR DEEP LEARNING

D. Palitta, J. M. Tabeart, and N. Vannieuwenhoven
Dipartimento di Matematica, Università di Bologna
P.zza di Porta S. Donato 5, Bologna, Italy
davide.palitta@unibo.it

In the past few years, applied mathematicians started looking at the forward propagation step of deep learning techniques in terms of discretization methods, e.g., forward Euler, applied to an unknown, underlying differential operator. Each layer of the network is seen as a time step of the discretization method [1, 2]. This point of view paved the way for so-called neural ordinary differential equations (ODE) [3]. In the latter framework, the deep learning process is modeled by an ODE: Inputs are translated into initial values whereas outputs are viewed as the ODE solution evaluated at the final time step. Information propagates along the ODE flow in place of the net so that the extremely problem-dependent design of the latter is no longer needed. The training phase is now employed to learn the parameters definining the neural ODE. In this talk we present a novel tensor-based neural ODE, namely an ODE defined by tensors, to model a deep learning process. Preliminary results on classification problems show the potential of such new tool.

References

[1] L. Ruthotto, E. Haber, *Deep Neural Networks Motivated by Partial Differential Equations*, J Math Imaging Vis, 62 (2020), pp. 352–364.

[2] E. Haber, L. Ruthotto, *Stable architectures for deep neural networks*, Inverse Problems, 34 (2017).

[3] R. T. Q. Chen, Y. Rubanova, J. Bettencourt, and D. K. Duvenaud, *Neural Ordinary Differential Equations*, Advances in Neural Information Processing Systems, 31 (2018).