



Einladung zum Oberseminar Wissenschaftliches Rechnen

Julius-Maximilians-Universität Würzburg
Lehrstuhl für Wissenschaftliches Rechnen IX

Sebastian Hofmann

Universität Würzburg, Institut für Mathematik

An optimal control based supervised learning algorithm for training Runge-Kutta structured neural networks ¹

The dynamical systems approach to machine learning is applied to a supervised learning (SL) problem for residual networks with Runge-Kutta (RK) structure. In this context, the SL problem is formulated as an optimal control problem, where the cost functional is minimized subject to a non-linear dynamical equation representing the neural network (NN) with its weights playing the role of the controls. Within this framework, the optimality of the NNs weights is characterized by the discrete Pontryagin maximum principle (PMP) that also provides the foundation of the to be presented, sequential quadratic Hamiltonian (SQH) training method for RK-NNs [1].

The SQH method represents a novel extension to the so-called successive approximation methods and has its strength in a dynamic hyperparameter-adaptation technique that guarantees the decrease of the objective functional's value in every iterations, independent of the initial hyperparameters choice. Moreover, a stopping condition serves as a control mechanism for the accuracy of the networks approximation and allows to stop training at an iteration that is optimal in the sense of the discrete PMP. Convergence and stability results for the SQH scheme are provided in the framework of training residual neural networks with RK structure. Moreover, it is shown numerically that inter alia due to the integrated, dynamic hyperparameter-adaptation technique, the efficiency and robustness of the training process of the SQH scheme is superior to a comparable algorithm, called extended method of successive approximation [2].

References

- [1] S.Hofmann and A.Borzì. *A Sequential Quadratic Hamiltonian Algorithm for Training Explicit RK Neural Networks*, Journal of Computational and Applied Mathematics 405 (2022) 113943.
- [2] Q.Li, L.Chen, C.Tai and W.E. . *Maximum Principle Based Algorithms for Deep Learning*, Journal of Machine Learning Research 18 (2018) 1-29.

Ort: Raum 30.02.003 (Mathematik West, 2.Stock)

Zeit: Freitag, 15.07.2022, 12:30 Uhr

Zu diesem Vortrag laden wir Sie herzlich ein. / You are cordially invited to this lecture.

gez. Prof. Dr. Alfio Borzì / Prof. Dr. Frank Werner