Course Overview

Numerical Methods for Deep Learning

Part 1: Shallow Models

1. Semisupervised Learning
   - Numerics: Sparse linear algebra, eigenvalue problems
2. Linear Least-Squares
   - Numerics: QR, SVD, steepest descent, iterative solvers
3. Linear Classification
   - Numerics: Convex optimization, Newton
4. Single Layer Neural Networks
   - Reading/Presentation: [1]
   - Numerics: Stochastic optimization, Gauss-Newton
5. Convolutional Neural Nets
   - Reading/Presentation: [2, 3]
   - Numerics: Structured matrix computation, PDE-based regularization

Part 2: Deep Models

6. Multilayer Perceptron
   - Reading/Presentation: [4]
   - Numerics: Backpropagation
7. Residual Neural Networks
   - Reading/Presentation: [5, 6]
8. Continuous Models
   - Reading/Presentation: [7, 8]
   - Numerics: ODE theory, time integrators
9. Optimal Control
   - Reading/Presentation: [9]
   - Numerics: adjoint method, discretize-optimize
10. Continuous Normalizing Flows
    - Reading/Presentation: [10, 11]
    - Numerics: optimal transport, conservation laws

References

[1] Guang-Bin Huang, Qin-Yu Zhu, and Chee-Kheong Siew. Extreme learning machine: Theory and applications. *Neurocomputing*, 70(1-3):489–501, December 2006.

[2] Y LeCun, B E Boser, and J S Denker. Handwritten digit recognition with a back-propagation network. In *Advances in neural information processing systems*, pages 396–404, 1990.

[3] X Glorot and Yoshua Bengio. Understanding the difficulty of training deep feedforward neural networks. *jmlr.org*.

[4] D.E. Rumelhart, Geoffrey Hinton, and J. Williams, R. Learning representations by back-propagating errors. *Nature*, 323(6088):533–538, 1986.

[5] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Identity mappings in deep residual networks. In *European Conference on Computer Vision*, pages 630–645. Springer, 2016.

[6] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 770–778, 2016.

[7] Weinan E. A Proposal on Machine Learning via Dynamical Systems. *Communications in Mathematics and Statistics*, 5(1):1–11, March 2017.

[8] Tian Qi Chen, Yulia Rubanova, Jesse Bettencourt, and David Duvenaud. Neural Ordinary Differential Equations. In *NeurIPS*, June 2018.

References (cont.)

[9] Amir Gholami, Kurt Keutzer, and George Biros. ANODE: Unconditionally Accurate Memory-Efficient Gradients for Neural ODEs. *arXiv.org*, February 2019.

[10] Linfeng Zhang, Weinan E, and Lei Wang. Monge-Ampère Flow for Generative Modeling. *arXiv.org*, September 2018.

[11] Will Grathwohl, Ricky T Q Chen, Jesse Bettencourt, Ilya Sutskever, and David Duvenaud. FFJORD: Free-form Continuous Dynamics for Scalable Reversible Generative Models. October 2018.