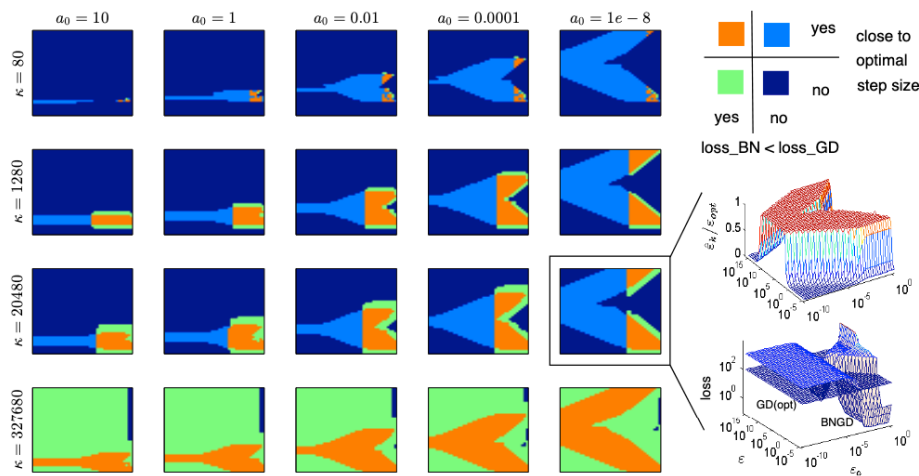


Research Highlight: The Effect of Batch Normalization on Gradient Descent

Work of Professor SHEN Zuowei, Dr LI Qianxiao and Dr CAI Yongqiang

Batch normalization [1] is a commonly applied technique to stabilize and accelerate training deep neural networks via gradient descent. However, how it actually works was a mystery from a mathematical point of view.

In this work, we investigate the quantitative effect of applying batch normalization to simplified machine learning problems. In this case, we can prove convergence and robustness properties of batch normalization, especially for high dimensional problems. We also investigate how such properties are related to hyper-parameters in training, such as learning rates. This paves way for a more thorough understanding of normalization techniques in training deep neural network models.



References:

[1] Ioffe, S. and Szegedy, C. Batch normalization: Accelerating deep network training by reducing internal covariate shift. In International Conference on Machine Learning, pp.448–456, 2015

[2] Cai, Yongqiang, Qianxiao Li, and Zuowei Shen. "A Quantitative Analysis of the Effect of Batch Normalization on Gradient Descent." International Conference on Machine Learning. 2019.