

BAI

智源学者成果展示——人工智能的数理基础

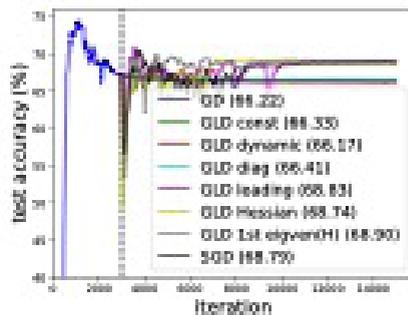
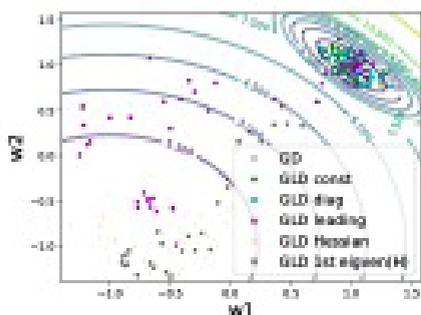
作者 朱占星（北京大学研究员、智源青年科学家）

2020年6月

深度学习训练算法的理论分析

北京大学研究员、智源青年科学家朱占星等首次发现了在深度学习中随机梯度噪声的结构对于逃离局部极小有重要作用，对于设计新的更加高效有效的训练方法有重要的指导意义。理解深度学习的训练算法的行为一直是深度学习理论的关键难题，这个发现对于理解深度学习中的随机梯度法有新的见解，部分解释了算法对于深度学习的关键作用，对于未来理解新的算法以及算法设计开启了新的可能。

Zhanxing Zhu*, Jingfeng Wu*, Bing Yu, Lei Wu and Jinwen Ma. *The Anisotropic Noise in Stochastic Gradient Descent: Its Behavior of Escaping from Minima and Regularization Effects*. 36th International Conference on Machine Learning (ICML 2019)



The SGD-type algorithms escape from sharp minima. Left: the trajectory of different dynamics in 2D toy example. Right: generalization performance w.r.t. iterations. 图片来自论文。

Beijing Academy of Artificial Intelligence



微信关注
北京智源人工智能研究院