

# Advanced Machine Learning Project Update

Si Kai Lee  
sl3950

In the process of writing the proposal, I skimmed through Kawaguchi 2016<sup>1</sup>. Since then, I had read the paper twice, once for the Columbia Advanced Machine Learning Seminar and once to start the project. However, I could not understand it as it was really dense and not very well written. Since then, I have spoken to Professor Hsu about switching to another framework to show that the local minimums of  $\frac{1}{2}||Y - W_3W_2W_1X||_F^2$  are actually global minimums.

I was suggested to look at Ge et al.'s Matrix Completion has No Spurious Local Minimum<sup>2</sup> and I have been trying to work through the proofs of the paper ever since. I am pretty sure that Ge et al. have the signs wrong for equation 3.4 but it does not affect the main gist of the paper. Currently, I am struggling to understand the proof of claim 2f. I was also pointed to Sebastian Bubeck's blogpost on the geometry of linearised neural networks and from there I have started reading Hardt and Ma's Identity Matters in Deep Learning<sup>3</sup> to try to gain an understanding of their framework and Kawaguchi's framework. I am also quite positive that the fifth line of the proof for Hardt and Ma in Bubeck's post has a sign error too.

---

<sup>1</sup>Kawaguchi, K., 2016. Deep Learning without Poor Local Minima. arXiv preprint arXiv:1605.07110.

<sup>2</sup>Ge, R., Lee, J.D. and Ma, T., 2016. Matrix Completion has No Spurious Local Minimum. arXiv preprint arXiv:1605.07272.

<sup>3</sup>Hardt, M. and Ma, T., 2016. Identity Matters in Deep Learning. arXiv preprint arXiv:1611.04231