Note on the Applicability of the Logarithmic Number System in LLM

By Anima Anandkumar, 8/27/25

It is interesting that the new DeepSeek AI v3.1 is trained using the UE8M0 FP8 scale data format, which is nothing but the logarithmic number system (LNS), meaning it has only exponent and no mantissa. Our multiplicative weights update (Madam) for training in that format was done several years ago while at NVIDIA It yields maximum hardware efficiency with no accuracy loss (see [1]).

Logarithmic number system achieves a higher computational efficiency by transforming expensive multiplication operations in the network layers to inexpensive additions in their logarithmic representations. In addition, it attains a wide dynamic range and can provide a good approximation. Also, logarithmic number system is biologically inspired, and there is evidence that our brains use such a format for storage.

However, using standard SGD or Adam optimization for training in logarithmic format is challenging, and requires intermediate updates and optimization states to be stored in full precision (FP32). To overcome this, we proposed Multiple Weights update (Madam) that instead updates directly in the logarithmic format and leads to good training outcomes.

Our LNS-Madam when compared to training in FP32 and FP8 formats, LNS-Madam reduces the energy consumption by over 90% and 55%, respectively, while maintaining accuracy.

A diagram of a progress

AI-generated content may be incorrect.

References

[1] [LNS-Madam : Low-Precision Training in Logarithmic Number System using Multiplicative Weight pdate, Jiawei Zhao et al, 2022](https://github.com/dimitarpg13/large_language_models/blob/main/articles/LNS-Madam-Low-Precision_Training_in_Logarithmic_Number_System_using_Multiplicative_Weight_Update_Zhao_2022.pdf)