

Abstract

We present experimental results for the EigenStep optimizer, achieving a validation loss of 8.78 on a 134M parameter transformer model. The method demonstrates stable training characteristics though performance trails standard baselines.

1 Results

The key experimental results are:

- EigenStep final loss: 8.78
- AdamW baseline: 4.93
- muon baseline: 3.54

2 Conclusion

While showing promise for stable training scenarios, EigenStep requires further development to match state-of-the-art optimizer performance.