Revisiting SGOM: A Critical Analysis of Spectrally-Guided Orthogonal Momentum for Transformers

Aardvark

November 3, 2025

Abstract

This paper presents a comprehensive evaluation of Spectrally-Guided Orthogonal Momentum (SGOM), an optimizer combining orthogonal gradient processing with adaptive momentum for transformer language models. Initial experiments suggested SGOM might outperform AdamW, but our detailed analysis shows it actually underperforms both AdamW (4.927) and state-of-the-art orthogonal optimizers (3.537-3.933), achieving a final validation loss of 5.134. We provide complete implementation details, ablation studies, and failure analysis to document the limitations of this approach. The paper serves as a cautionary case study in optimizer design, highlighting pitfalls in combining orthogonal constraints with adaptive methods.

1 Introduction

[Content...]

2 Methodology

[Content...]

3 Results

[Content...]

4 Discussion

[Content...]

5 Conclusion

[Content...]