Online Journal of Mathematics, Science and Technology Education (OJOMSTE)

Vol: 6 No: 2 Year: 2025

journal homepage: https://www.ojomste.com/index.php/1

A Mathematically Modified Adam Algorithm for Improved Convergence in Deep Neural Networks

Mark LAISIN[a], Bright Okore OSU[b], Prisca Udodiri DURUOJINKEYA[c], Chigozie CHIBUISI[d]

[a]Department of Mathematics Chukwuemeka Odumegwu Ojukwu University, Uli, Nigeria, laisinmark@gmail.com, https://orcid.org/0009-0003-3331-1235.

[b]Department of Mathematics, Abia State University, Uturu, Nigeria, Osu.bright@abiastateuniversity.edu.ng, https://orcid.org/0000-0003-2463-430X.

[c]Department of Mathematics and Statistics, Federal Polytechnic Nekede, Owerri, Nigeria, https://orcid.org/0009-0003-3331-1235, https://orcid.org/0000-0002-3174-7751.

[d]Department of Insurance, University of Jos, Jos, Nigeria, chigoziec@unijos.edu.ng, https://orcid.org/0009-0006-4100-7817.

Received: October, 16, 2025 *Accepted: November, 25, 2025*

Keywords:

Adaptive Optimization,

Gradient Descent, Deep Neural Networks,

Convergence Analysis,

Momentum Thresholding,

Non-Convex Optimization,

Step Size Scheduling,

Stochastic Optimization

Paper Type:

Research

Abstract

This study introduces the Adaptive Moment Gradient Thresholding *(AMGT)* algorithm, a modified version of the Adam optimizer, aimed at enhancing convergence stability in deep neural networks. By leveraging optimization theory and addressing the limitations of Adam, AMGT was designed to tackle non-convexity, constrained environments, and gradient-based learning instability. The algorithm incorporates a diminishing step size schedule and momentum thresholding to improve performance. Theoretical analysis demonstrated that AMGT achieved linear convergence under strong convexity with a rate of $O(k^{-\mu/2})$, global convergence under bounded gradient approximation errors, and convergence to stationary points in non-convex scenarios. Numerical experiments on convex quadratic functions validated the theoretical predictions, highlighting the algorithm's sensitivity to spectral properties and resilience to learning rate variations. The results indicate that AMGT surpasses standard Adam in convergence behaviour and provides theoretical guarantees often lacking in adaptive optimizers. AMGT is particularly effective in high-dimensional, noisy, or resource-constrained settings due to its support for quantized and sparsified updates. By combining theoretical rigour with empirical robustness, AMGT emerges as a dependable option for training deep learning models across diverse optimization landscapes. Empirical evaluations on standard benchmark datasets, including MNIST, CIFAR-10, and ImageNet, further demonstrated the superior generalization capability and faster convergence of AMGT over Adam and its variants, underscoring its practical applicability in real-world deep learning scenarios.

Laisin, M., Osu, B. O., Duruojinkey, P.U, & Chibuisi, C. (2025). A Mathematically Modified Adam Algorithm for Improved Convergence in Deep Neural Networks. *Online Journal of Mathematics, Science and Technology Education (OJOMSTE),* 6(2), 40-64.