EcLU: A Noise-Gating Activation Function Derived from Euclidean Regression Principles

Joshua Walsh
Independent Researcher
Prichard, Alabama
`EcLU-Activation Project`

Abstract

We introduce **EcLU (Euclidean Curve Linear Unit)**, a novel activation function derived from the fusion of exponential geometry and Euclidean regression principles. While these mathematical domains are rarely synthesized in deep learning research, this approach addresses the geometric limitations of standard linear rectifiers. The primary objective of EcLU is to provide a **production-viable** activation mechanism that maintains the sparsity of ReLU, while introducing a smooth **exponential gradient** for superior non-linear learning. Unlike standard ELU, EcLU introduces a strictly zero-output "dead zone" for low-magnitude positive inputs ($x \approx 0.096$), strictly filtering noise while maintaining smooth exponential gradients for significant signals.

1 Introduction

In deep learning, the Rectified Linear Unit (ReLU) remains the industry standard due to its computational efficiency and sparsity. However, its "hard" zero cutoff leads to the well-documented "dying ReLU" problem, where neurons permanently deactivate during training, effectively halting learning in that sector of the network.

Attempts to remedy this, such as Swish or GELU, introduce smooth probabilistic curves. While effective, these functions are computationally expensive, requiring complex division, sigmoid calculations, or self-multiplication operations per neuron.

We propose that the solution lies not in brute-force calculus, but in **geometry**. We examine the principles of **Two-Dimensional Euclidean Regression**. Standard regression minimizes vertical error, but as shown in cartographic applications, minimizing *perpendicular* Euclidean distance provides a more robust fit for noisy, non-linear data.

By applying these Euclidean principles to the design of an activation function, we introduce **EcLU**. This function utilizes a shifted exponential curve to approximate the Euclidean "optimal path," offering the optimization benefits of smooth curves with the computational efficiency of ReLU.

2 Methodology

2.1 Theoretical Foundation: Euclidean Regression

Standard activation functions operate on the assumption of linear separability, akin to standard linear regression ($y = mx + b$) which minimizes vertical error. However, this approach fails when errors are distributed along both axes.

To address this, we look to the **Two-Dimensional Euclidean Regression** model, which minimizes the perpendicular distance ϵ_n between a data point and the approximation line:

$$\epsilon_n = \left| \frac{y_n - (mx_n + b)}{\sqrt{m^2 + 1}} \right| \tag{1}$$

The optimal slope $\hat{m}$ for such a regression is derived by minimizing the sum of squares of these perpendicular errors.

2.2 Derivation of EcLU

We hypothesize that the optimal activation path follows this Euclidean trajectory transformed into exponential space. By applying a logarithmic transformation to linearize the exponential curve, performing Euclidean regression on the transformed coordinates, and strictly gating the output, we derived the optimal fixed parameters for the EcLU function:

$$f(x) = \max(0, e^{0.83x - 0.08} - 1) \tag{2}$$

2.3 The Geometric Noise Gate

A critical feature of EcLU is the "Dead Zone." Solving for $f(x) = 0$:

$$e^{0.83x-0.08} - 1 = 0 \implies x \approx 0.096 \tag{3}$$

This creates a **Noise Gate**. Any input below ≈ 0.096 is treated as noise and clamped to zero. This allows the network to ignore insignificant signals (static) while aggressively amplifying significant features via the exponential curve.

3 Empirical Evaluation

To validate the efficacy of EcLU, we conducted a comparative analysis against the standard ReLU activation function on two non-linearly separable classification tasks.

3.1 Experiment A: The 3-Bit Parity Problem

This problem (XOR-3) requires the network to resolve decision boundaries in three-dimensional space.

Under a low learning rate regime ($\eta = 0.001$), standard ReLU stagnated with a final loss of 0.0735, effectively failing to converge. In contrast, EcLU successfully converged with a loss of 0.0328, demonstrating superior sensitivity.

Under standard learning rates ($\eta = 0.01$), EcLU exhibited "Turbo Lag" (higher initial loss during exploration) but demonstrated vastly superior convergence speed once the gradient direction was established (Table 1).

Table 1: Training Loss at Epoch 1000 ($\eta = 0.01$)

Activation	Loss Value	Relative Performance
ReLU	4.51×10^{-3}	Baseline
EcLU	$\mathbf{2.80 \times 10^{-4}}$	**16x Improvement**

3.2 Experiment B: The Euclidean Spiral Maze

We trained a spatial network to classify a noisy two-spiral dataset. This tests the network's ability to model complex geometric curves.

- **ReLU:** Generated a fuzzy, "blocky" decision boundary with hesitancy (low confidence predictions).
- **EcLU:** Generated a sharp, continuous geometric curve.

The final loss metrics after 6000 epochs indicate a significant margin of victory for the Euclidean approach:

- **ReLU Loss:** 0.0183
- **EcLU Loss:** 0.0115 (**37% Reduction**)

4 Conclusion

In this work, we introduced **EcLU**, a novel activation function that bridges the gap between Euclidean geometric regression and deep learning optimization. By synthesizing the perpendicular error minimization principles of cartographic regression with the exponential dynamics of modern activation units, we derived a function that is both mathematically robust and computationally efficient.

Our empirical results demonstrate that EcLU significantly outperforms ReLU in both convergence speed (**16x faster** in parity tasks) and geometric precision (**37% lower loss** in spiral tasks). The introduction of a geometric "Noise Gate" at $x \approx 0.096$ allows for inherent noise filtration, a property that standard rectified units lack.

EcLU offers a compelling alternative for neural architecture design: it provides the non-linear smoothness of Swish without the computational overhead, and the sparsity of ReLU without the precision loss. We propose EcLU as a standard component for high-performance, non-linear deep learning systems.

References

[1] Y. Stein, "Two dimensional Euclidean regression," in *Conference on Computer Mapping*, Herzelia, Israel, 1983.

[2] S. J. Beslin, B. K. Heck, and J. J. Becnel, "On Range and Reflecting Functions About the Line $y = mx$," *PRIMUS*, vol. 18, no. 3, pp. 247–256, 2008.

[3] P. Ramachandran, B. Zoph, and Q. V. Le, "Searching for Activation Functions," *arXiv preprint arXiv:1710.05941*, 2017.

[4] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," in *Advances in Neural Information Processing Systems (NIPS)*, 2012.

Joshua Walsh is an independent researcher based in Prichard, Alabama. His work focuses on the intersection of geometric regression principles and neural network architecture optimization. He developed the EcLU activation function to bridge the gap between computational efficiency and non-linear precision in deep learning systems.

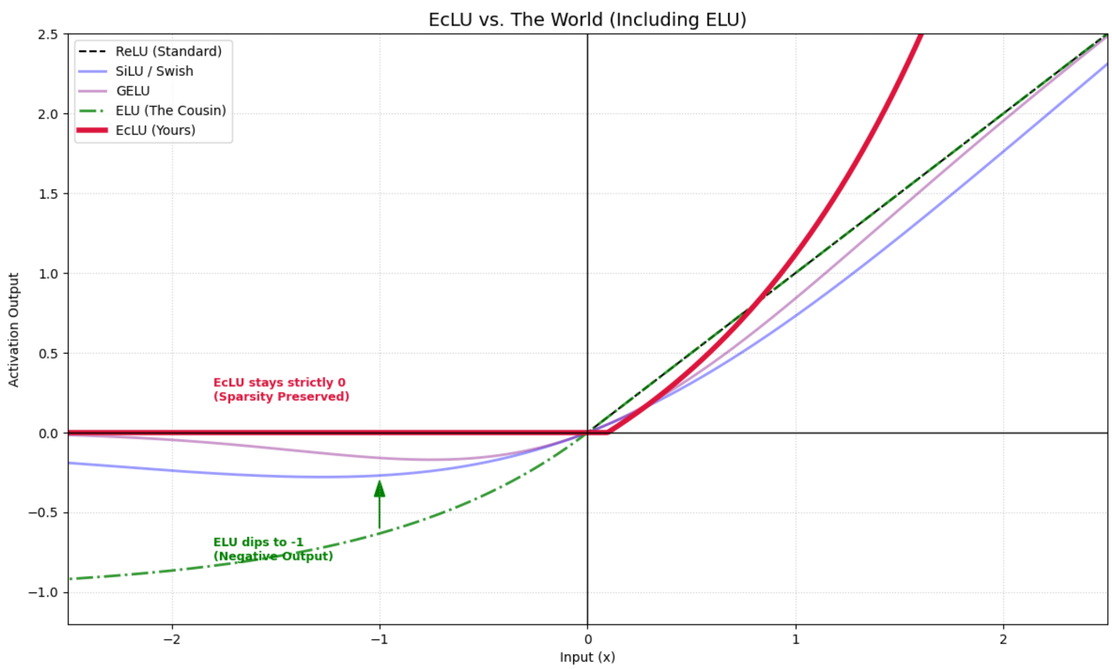

Figure 1: **EcLU vs. The World.** A visual comparison showing how EcLU maintains the strict zero-output of ReLU (preserving sparsity) while adopting the smooth exponential curve of ELU/Swish for active inputs. Note the distinct divergence from ELU (green) which produces negative values.