Expansive Networks: Linear Parameter Injections At Train-Time

Dmitri Tkatch

June 28, 2025

Abstract

The hypothesis is that, while mathematically equivalent and contrary to modern wisdom, two consecutive matrices of sizes 100x10000 and 10000x100 with no non-linearity inbetween them actually learn more easily than one 100x100 matrix. We call the act of replacing the 100x100 matrix by the former two matrices "expansion". We propose Expansive Networks – a family of neural networks that are trained with expansions which are then collapsed at test-time via matrix multiplication. Only one layer is expanded at a time for computational memory savings.