Loss landscapes of neural networks through the lens of flat regions and symmetries

Fabrizio Pittorino
Bocconi University, Milan, Italy

We systematize the approach to the investigation of deep neural network landscapes by basing it on the geometry of the space of implemented functions rather than the space of parameters. In the space in which all symmetries are removed, we explore the error landscape rather than the loss. This lets us derive a meaningful notion of the flatness of minimizers and of the geodesic paths connecting them. Using different optimization algorithms that sample minimizers with different flatness we study the mode connectivity and relative distances. Testing a variety of state-of-the-art architectures and benchmark datasets, we confirm the correlation between flatness and generalization performance; we further show that in function space flatter minima are closer to each other and that the barriers along the geodesics connecting them are small. We also find that minimizers found by variants of gradient descent can be connected by zero-error paths composed of two straight lines in parameter space, i.e. polygonal chains with a single bend. We observe similar qualitative results in neural networks with binary weights and activations, providing one of the first results concerning the connectivity in this setting. Our results hinge on symmetry removal, and are in remarkable agreement with the rich phenomenology described by some recent analytical studies performed on simple shallow models: through the lens of statistical physics, we investigate the paradigmatic example of the spherical negative perceptron as the simplest model of neural network landscapes.