Computational Experiment: Visualizing Loss Function Convergence

Goal of the experiment: to assess how much the loss function changes when moving from one training set size to the next:

$$\Delta_k = \mathbb{E}_{p(\mathbf{w})}\big(\mathcal{L}_k(\mathbf{w}) - \mathcal{L}_{k-1}(\mathbf{w})\big)^2$$

where $\mathcal{L}$ is the loss function, $p(\mathbf{w})$ is a distribution, and k is the dataset size index.

Method: generate points according to $p(\mathbf{w})$ around the minimum point and use Monte Carlo averaging.