TTIC 31230 Fundamentals of Deep Learning

Regularization and Generalization Problems

Problem 1. The Stationary Points for L_2 Regularization. Consider the regularized objective

$$\Phi^* = \underset{\Phi}{\mathrm{argmin}}\ E_{(x,y)\sim\mathrm{Train}}\ \left(\mathcal{L}(\Phi,x,y) + \frac{1}{2N_{\mathrm{train}}\sigma^2}||\Phi||^2\right)$$

By setting the gradient of the objective to zero, solve for Φ as a function of the average gradient g defined by

$$g = E_{\langle x,\,y\rangle\sim\mathrm{Train}}\nabla\Phi\mathcal{L}(\Phi,x,y).$$