Generalization error of feature ridge regression

The circles represent the empirical generalization error, the lines the deterministic equivalent (solid with empirical and dashed with linearized covariance). The colors encode the size of the regularization parameter used for ridge regression. The features in the first plot are untrained random features, the other plots are trained features after some number of epochs of gradient descent. The horizontal lines represent the loss of the neural network on the test set.

◯ Feature Regression   ＋ Linear Regression

■ NN loss   ■ λ=0.1   ■ λ=1.0   ■ λ=10.0   ■ λ=100.0

↑ Generalization error

Epoch

0   0.07   0.88   15   Linear Regression

Samples →