

Generalization error of feature ridge regression

The circles represent the empirical generalization error, the lines the deterministic equivalent. The colors encode the size of the regularization parameter of ridge regression. The features in the first plot are untrained random features, the other plots are trained features after some number of steps of gradient descent (the facet titles have the format 'epoch'.'step'). The horizontal lines represent the loss of the neural network on the test set.

■ NN loss ■ $\lambda=0.1$ ■ $\lambda=1.0$ ■ $\lambda=10.0$ ■ $\lambda=200.0$

