Due to noise and other limitations, machine learning problems often do not have unique solutions. Therefore, to solve ML problems, we must find approximate solutions. To get the best approximate solutions for an overdetermined system, we have to minimize the squared error of a function. If we capture noise in our problem’s solution, we might overfit the data and not be able to generalize the solution. For solving underdetermined systems, we use ridge regression.