Machine Learning
Exercise 2

Marc Toussaint
Machine Learning & Robotics lab, U Stuttgart
Universitätsstraße 38, 70569 Stuttgart, Germany

May 7, 2014

1 Linear Ridge Regression

On the course webpage there are two simple data sets `dataLinReg1D.txt` and `dataLinReg2D.txt`. The data files can be plotted, e.g., using gnuplot with `plot 'dataLinReg1D.txt'` or `splot 'dataLinReg2D.txt'`. Each line contains a data entry (x, y) with $x \in \mathbb{R}^d$ and $y \in \mathbb{R}$; the last entry in a line refers to y. Compute and report the optimal parameters β for a linear Ridge regression model (just linear features) for both data sets. Tips:

a) Write a routine that loads a data file and returns a matrix X containing all x_i as rows, and a vector y containing all y_i.

b) Write a routine that takes the raw X as input and returns a new $\boldsymbol{X}$ with a '1' pre-pended to each row. This routine simply computes the "linear features" including the constant 1. This routine can later be replaced by others to work with non-linear features.

c) Write a routine that returns the otpimal β from $\boldsymbol{X}$ and y.

d) Generate some test data points (along a grid) and collect them in a matrix Z. Apply routine b) to comupute features. Compute the predictions $\widehat{y} = Z\beta$ (simple matrix multiplication) on the test data and plot it.

2 Non-linear features

Test regression with quadratic features on the data sets `dataQuadReg1D.txt` and `dataQuadReg2D.txt`. Compute and report the optimal parameters β for both data sets. In principle, all you have to do is replace routine b) above (see slide 02:8).

3 Cross-validation

Implement cross-validation (slide 02:18) to evaluate the *prediction error* of the linear and polynomial regression method for `dataQuadReg2D_noisy.txt`. Report 1) the squared error when training on all data (=*training error*), and 2) the mean squared error $\widehat{\ell}$ from cross-validation.

Repeat this for different Ridge regularization parameters λ. (Ideally, generate a nice bar plot of the generalization error, including deviation, for various λ.)