

Exercise 6

Advanced Methods for Regression and Classification

November 29, 2018

Load the data `bone` from the package `ElemStatLearn` – see help page for details. For the following tasks, use only the data from the males. Split the data randomly into training (2/3) and test (1/3) data. The task is to use nonlinear regression smoothing with the predictor *age* and the response *spnbmd*.

1. Use the function `loess()` for local polynomial regression – see help. The tuning parameter *span* determines the fraction of observations that is considered for the local fit.
 - (a) Fit a model to the training data; select once a smaller and once a bigger value of *span*. Visualize the fitted data in the plot of the response versus the explanatory variable. You will have to sort the values first.
 - (b) Take the model(s) and predict the response for the test data. Visualize the predictions in the plot (again you will have to sort the values). Compute the MSE for the test data.
2. Compute B-splines (`bs()` from the package `splines`) based on the training data of the variable *age*, once with a “low” and once with a “higher” number of basis functions. B-splines create a specific basis expansion which yield splines, but they are numerically more stable than those basis functions mentioned in the course notes.
 - (a) Apply least-squares regression with the B-splines as predictors, for the training data, and show the fitted values as line(s) in the plot (sort the values first).
 - (b) Take the model(s) and predict the test data. Visualize the predictions in the plot (sort the values first). Compute the MSE for the test data.
3. Carry out the tasks from 2. for natural cubic splines (`ns()` from the package `splines`). These spline basis functions are similar to B-splines, but they are linear outside of the outer knots, and thus more appropriate for extrapolation.

Save your (successful) R code together with short documentations and interpretations of results in a text file (= R script file), named as *Matrikelnummer.6.R* (no word document, no plots). Submit this file to Exercise 6 of our tuwel course (deadline November 28).