

## Exercise

### Generalized Additive Models

Load the data `Auto` from the package `ISLR`. The data contain different characteristics, information about the origin, and the name of the car. The goal is to predict the variable `mpg`, so the miles per gallon, using the remaining variables. Select those variables which could be useful for the prediction. Randomly select a training set of about 2/3 of the observations, by using the random seed “123”. Build the following models always based on the training set, and evaluate them for the test set data consisting of the remaining observations.

*Generalized Additive Models (GAMs):* function `gam()` from the `library(mgcv)`

- (a) Apply regression with GAMs, and let the function select the optimal tuning parameters (degrees of freedom). You might have to be careful which of the input variables are appropriate for smooth functions in the model.
- (b) Which variables are significant in the model? How complex are the smooth functions?
- (c) Plot the explanatory variables against their smoothed values as they are used in the model. You can simply use:  

```
plot(gam.object, page=1, shade=TRUE, shade.col="yellow")
```

  
How can you interpret this plot?
- (d) Compute the RMSE for the test set.
- (e) Try to tune the model to get a value of the RMSE below 2.9. Some ideas can be found in the help of `step.gam`.