Statistical Methods 2

Portfolio 6 – Generalized additive models

Complete the following tasks and submit your work on Blackboard by 4pm on Friday 10/03/2023

Task

Choose a dataset $\{(y_i^0, x_i^0)\}_{i=1}^n$ on which you can fit a generalized additive model (GAM), and split your dataset into a training an testing set. Then,

- Using the training set, fit a GAM model and choose the penalty parameters $\{\lambda_j\}_{j=1}^p$ using generalized cross-validation.
- Plot some of the estimated functions to illustrate the fact that, for your dataset, a GAM model is a reasonable choice (i.e. you should show that at least one estimated function is non-linear).
- Use the test set to assess the predictive performance of the fitted model.
- Using the training set, fit a generalized linear model (i.e. fit your GAM when all the functions are linear) and compute its prediction error on the test set.

For this task you can for instance fit a logistic GAM model on the `wesdr` dataset dataset[1] using the `R` package `mgcv`.

[1] available from the `R` package `gss`.