Elements of Machine Learning
Exercise Sheet 4
Winter Term 2023/2024

William LaCroix - wila00001@stud.uni-saarland.de - 7038732
Philipp Hawlitschek - phha00002@stud.unisaarland.de - 7043167

Problem 2 (T, 3 Points). **Generalized Additive Models**

In Generalized Additive Models (GAMs), we are interested in predicting our target variable $\mathbf{Y} \in \mathbb{R}$ based on the variables $X_1, ..., X_p$ as follows:

$$g(Y) = \alpha + \sum_{j=1}^{p} f_j(X_j)$$

where we assume that $\mathbb{E}(f_j(X_j)) = 0$ for all j. For the rest of this exercise, we will assume $g = id$ to be the identity function and that the dimensionality of X is $p = 2$.

1. [1pt] Without proof, will iterating this algorithm produce the same result as one of the methods you have learned about in class? Explain your reasoning.

Yes, backfitting this GAM will produce the same result as performing ridge regression, since it is essentially a sum over linear functions, where each component of the additive model is updated independently.

2. [1pt] Under which conditions will the results of the backfitting algorithm with this smoothing operator $\mathcal{S}_\lambda$ depend on the order of which $\hat{f}_j$ is updated first?

When there is strong interaction or correlation between predictors, then updating one component before another may produce different results.

3. [1pt] Write down the smoothing operator based on cubic smoothing splines.

As per lecture slides, we minimize the following: $\sum_{i=1}^{n}(y_i - f(x_i))^2 + \lambda \int f''(t)^2 \, dt$ over our function $f(Y)$ from above, based on the residuals for all splines not yet fit:

$$\mathcal{S}_\lambda(g) = \hat{f}_j(x) = argmin \sum_{i=1}^{n} \left(y_i - \alpha - \sum_{k \neq j} \hat{f}_k(x_{ki}) - f_j(x_{ji}) \right)^2 + \lambda \int f_j''(t)^2 \, dt$$