Experiments

Weber Jakob

October 7, 2020

1   Simulations

Following the descriptions given in the previous chapters, we are now going to test the algorithm of structured additive regression using a priori knowledge on the following problems:

- Static Function Approximation using a priori Knowledge

- Equidistant vs. Non-equidistant Knot Placement

- Different Levels and Colors of Noise λ^2

- Well-distributed Data vs. Skewed Data

These are evaluated on two functions with a priori known behavior. The first function is given by

$$f(x) = \begin{cases} 0 & \text{if } x \leq 0.8 \\ 20\sin(x-1.2) & \text{else} \end{cases}, \quad \text{for } x \in [0, 2.5]. \tag{1}$$

It is therefor a constant function till $x = 0.8$ and afterwards increasing. An example of the true function as well as noisy samples from this function is given in Fig. 1.
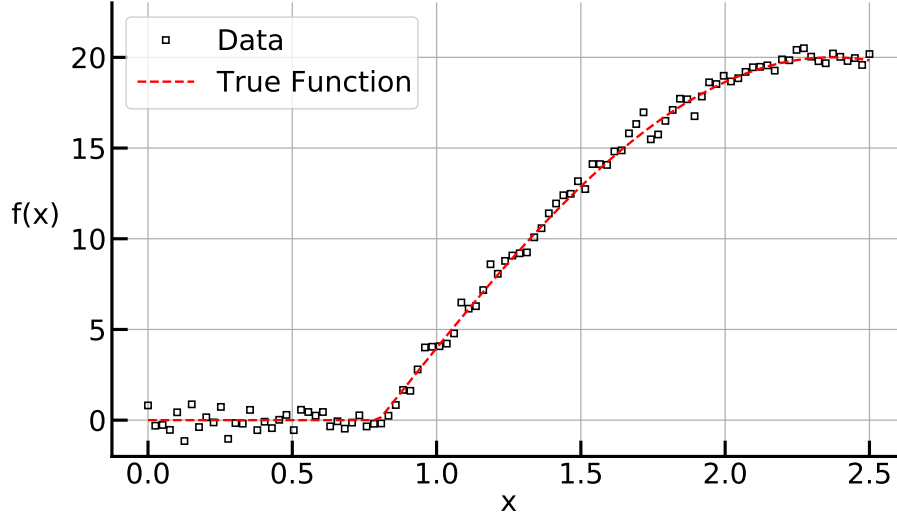
Figure 1: Test Function 1

The second function that will be investigated is given by

$$f(x) = \begin{cases} 2\exp\left(-\frac{(x-0.5)^2}{0.05}\right) + 3x & \text{, if } x \leq 0.85 \\ 2\exp\left(-\frac{(0.85-0.5)^2}{0.05}\right) + 3*0.85 & \text{, else} \end{cases}, \quad \text{for } x \in [0,1]. \quad (2)$$

The true functions depicts a unimodal behavior and is constant for $x > 0.85$. To test the algorithm, we deliberately generate "wrong" data using the following function

$$f(x) = 2\exp\left(-\frac{(x-0.5)^2}{0.05}\right) + 3x \quad \text{for } x \in [0,1]. \quad (3)$$

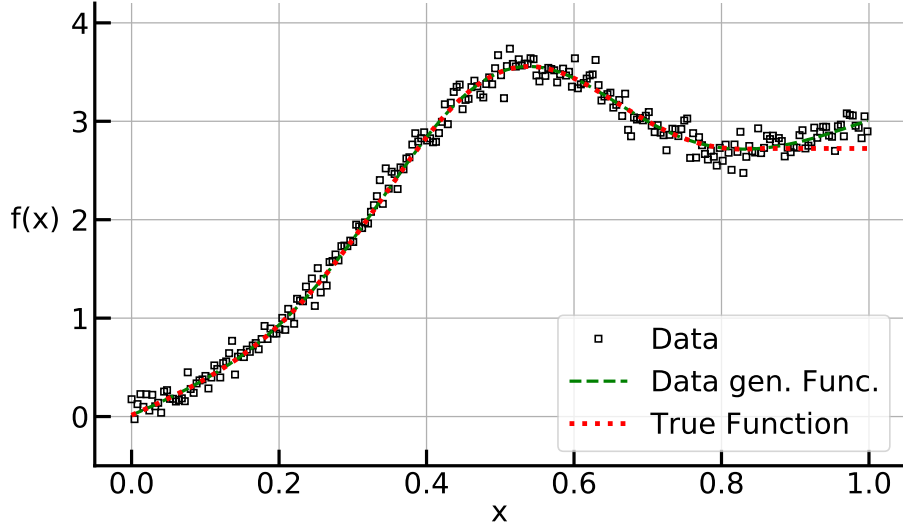An example of the true function 2 as well as noisy samples from function 3 is given in the following figure.

Figure 2: Test Function 2

1.1 Static Function Approximation using a priori Knowledge

For the test using function 1, the data set contains 100 points. We use $k = 35$ as number of used splines. The smoothing parameter λ_s is optimized using cross-validation. The parameter regulating the influence of the constraint is set to the 1000-fold of the smoothing parameter λ_s. The resulting constrained fit, and for comparison, the unconstrained fit using the optimal smoothing parameter $\lambda_s = 0.1615$ are shown in the following figure.
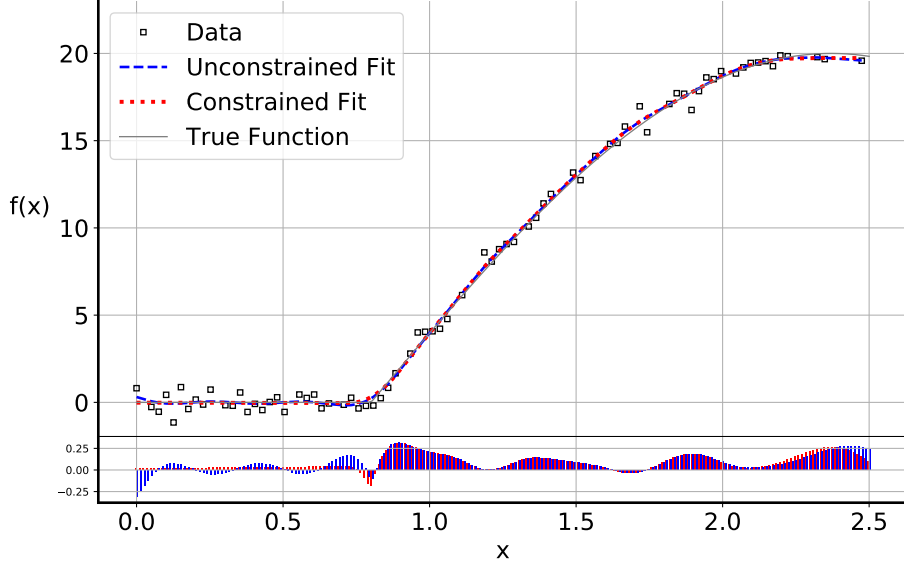
Figure 3: Constr. and Unconstr. Fit and Residuals (lower plot) for Function 1

The red line follows the a priori knowledge of increasing behavior better than the blue, unconstrained line. Especially in the constant area, e.g. $x \in [0, 0.8]$, the constraint improves the static function approximation and lowers the residual. In the area of increasing behavior, the two fits are almost identical. The mean squared error on the training data, on the test data and on the true function are given in the Table Results.

For the test using Function 2 the data set contains 500 points. There are two possibilities to include the a priori knowledge of a peak. First we can use the peak constraint matrix given in Chapter Penalty Matrices. On the other hand, we can use the concavity constraint, since concave functions are likely to have a peak when the data shows a peak. We examine both possibilities in Figure 4. For each fit, we used $k = 35$ as number of splines and the smoothing parameter $\lambda_s = 0.6189$ was again optimized using cross-validation. The constraint parameter λ_c was set to 618.9.
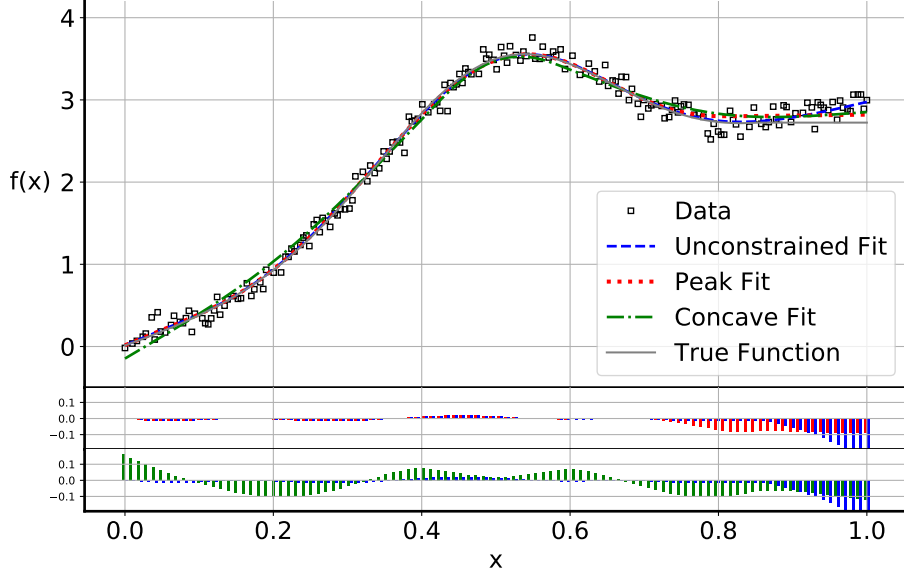
Figure 4: Constr. and Unconstr. Fit and Residuals (lower plots) for Function 2

The estimate for the concave constraint shows a the desired peak behavior, and is more smooth than the peak constraint estimate. This smoothness leads to changes in the sign of the residuals compared to the unconstrained fit. The peak constraint identifies the peak correctly and produces a sufficiently smooth fit with low residual. Especially for large $x > 0.8$, i.e. the region where the data is on purpose "wrong", the residual of the peak constraint fit is lower than the unconstrained ones. The mean squared errors are given in the results table. It is important to notice that the concavity constraint must be used carefully, since not all functions with a peak are also concave or vice versa.

1.2 Constraint vs. Noise

In this section, we investigate the effect of different noise levels as well as noise types on the fitting process. In the first part, various levels of Gaussian noise are applied to function 2. In the second part, we will use different colors of noise acting on function 1. For both examples we use $k = 35$ splines and optimize the smoothing parameter λ_s using cross-validation.

1.2.1 Noise Level

Gaussian noise is characterized by the following two parameters:

- Location: The mean value μ of the Gaussian distribution.

- Scale: The variance value σ^2 of the Gaussian distribution.

We set the location equal to zero and vary the scale value. The effect of the noise on the fitting procedure is shown in the following figure. The chosen noise levels are $\sigma^2 = \{0.01, 0.05, 0.1\}$.

5

For these values of σ^2, the correspond optimized smoothing parameters are $\lambda_{0.01} = 9.0$, $\lambda_{0.05} = \lambda_{0.1} = 68.1$.
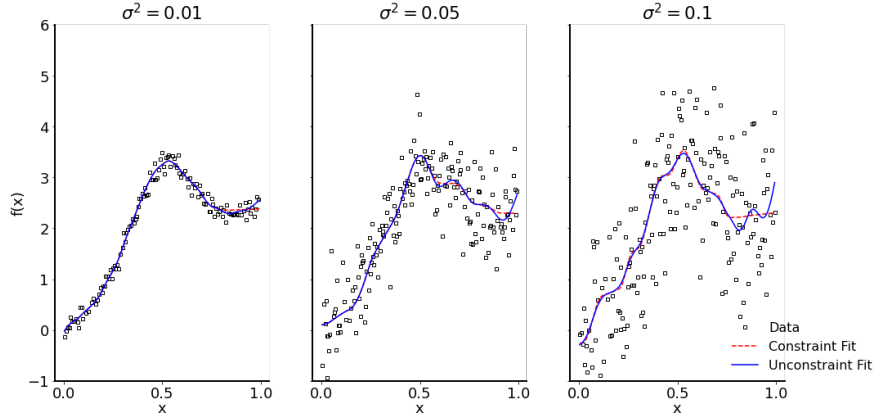

Figure 5: Constrained and Unconstraint Fit for various Noise Levels

The left plot shows the fit for $\sigma^2 = 0.01$. The noise level is quite moderate, which leads to a smooth and satisfying fit. The peak constraint is well satisfied. The constraint parameter λ_c was set to $\lambda_c = 1000\lambda_s$.

The noise level in the middle plot is already high. The unconstrained fit is therefore wiggly and shows 2 peaks. The constrained fit holds the peak constraint and generalizes better to the test set, as seen in *table resulst MSES.* The constraint parameter was again set to $\lambda_c = 1000\lambda_s$.

For the right plot and $\sigma^2 = 0.1$, both constrained and unconstrained fit are wiggly but for larger x, the constraint fit follows the constraint better. The constraint parameter was set to $\lambda_c = 1000\lambda_s$.

1.2.2 Noise Colors

We investigate the effect of different noise color on the function fitting process. The tested colors are the following power-law noise types:

- White Noise

- Pink Noise

- Brownian Noise

The unconstrained as well as the increasing constrained fit for all noise colors are shown in the following figure.
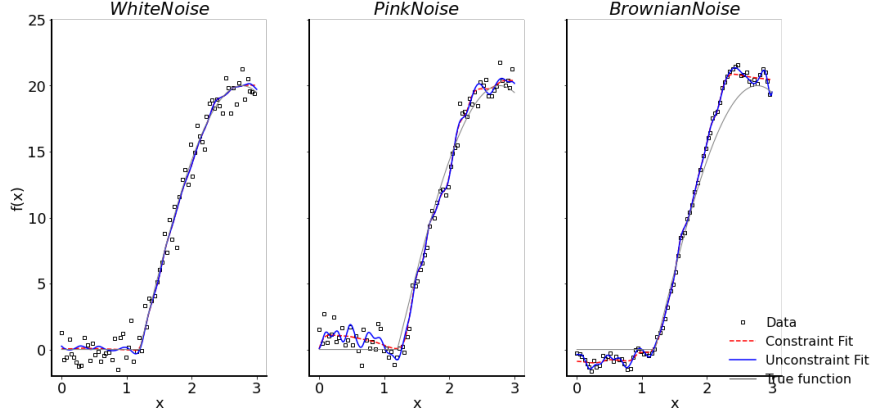
6

Figure 6: Constrained and Unconstrained Fit for various Noise Colors

For the white noise data in the left plot, the increasing constraint produces a very good fit according to the constraint. The optimized smoothing parameter for the unconstrained fit was given by $\lambda_s = 9.08$ and the constraint parameter was set to $\lambda_c = 9080$. The fit recovers the true function quite well.

For the pink noise data in the middle plot, the constraint helps regularizing the fit, but the increasing constraint is not satisfied especially for low x. The increasing part of the function, i.e. $f(x)$ for $x \geq 1.2$, was fit well. The smoothing parameter was optimized using cross-validation and set to $\lambda_s = 0.3$. The constraint parameter was set to $\lambda_c = 6000$.

For the brownian noise data in the right plot, the increasing part of the function was fit well, but for larger $x \geq 2$, the fit diverges the true function because of the strong noise component. The increasing constraint is hold quite well. The optimized smoothing parameter used was $\lambda_s = 0.0056$ and the constraint parameter was set to $\lambda_c = 6000$.

1.3 Well-distributed vs. Skewed Data and Knot Placement

- Phasendiagram: $\sigma^2 vs. \lambda_c$

- Ebner Data