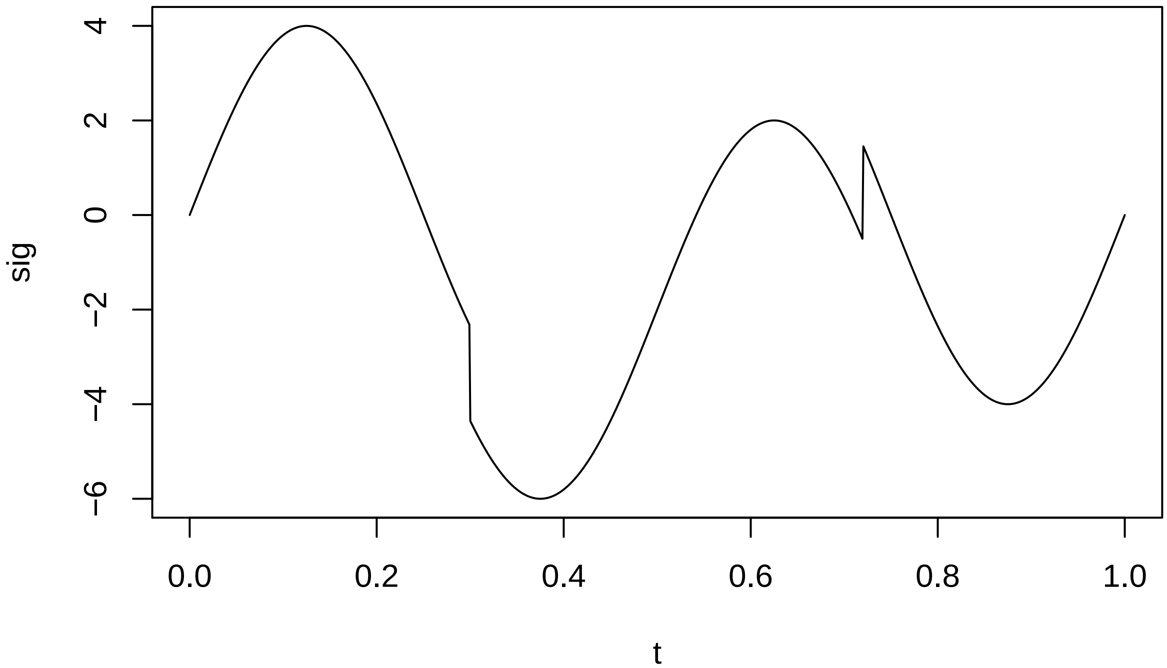
Assignment1

Tiffany Woodley

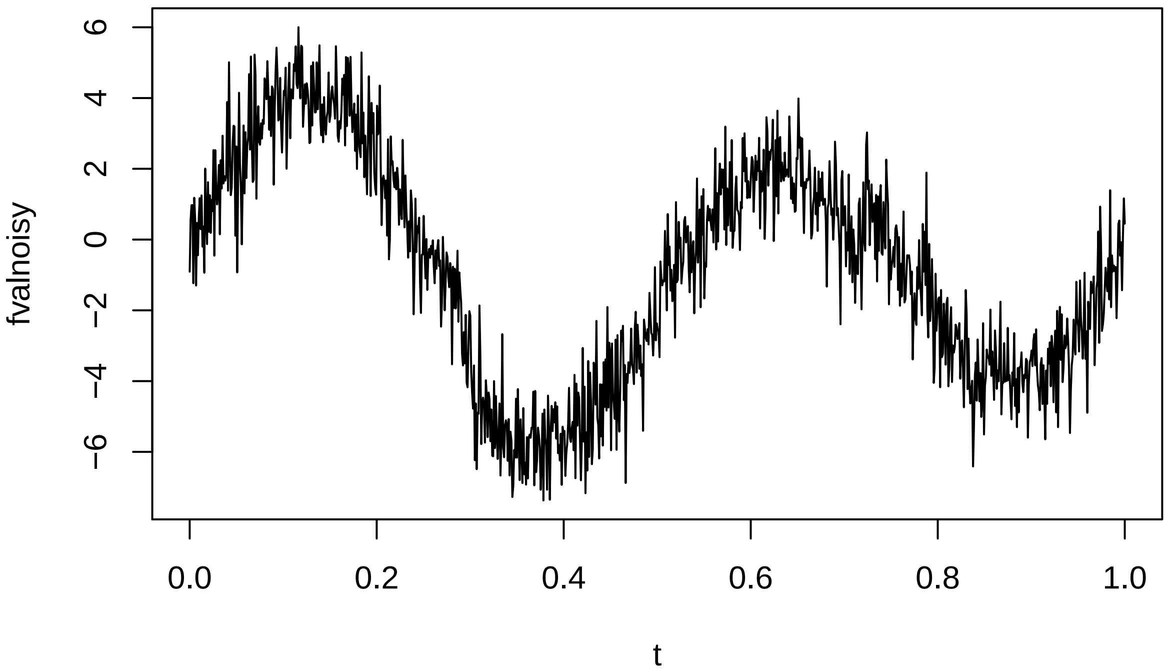
5/25/2020

Question 4: Wavelets

The data used for this question is simulated from the HeaviSine function, it consists of a smooth curve with two sharp edges.

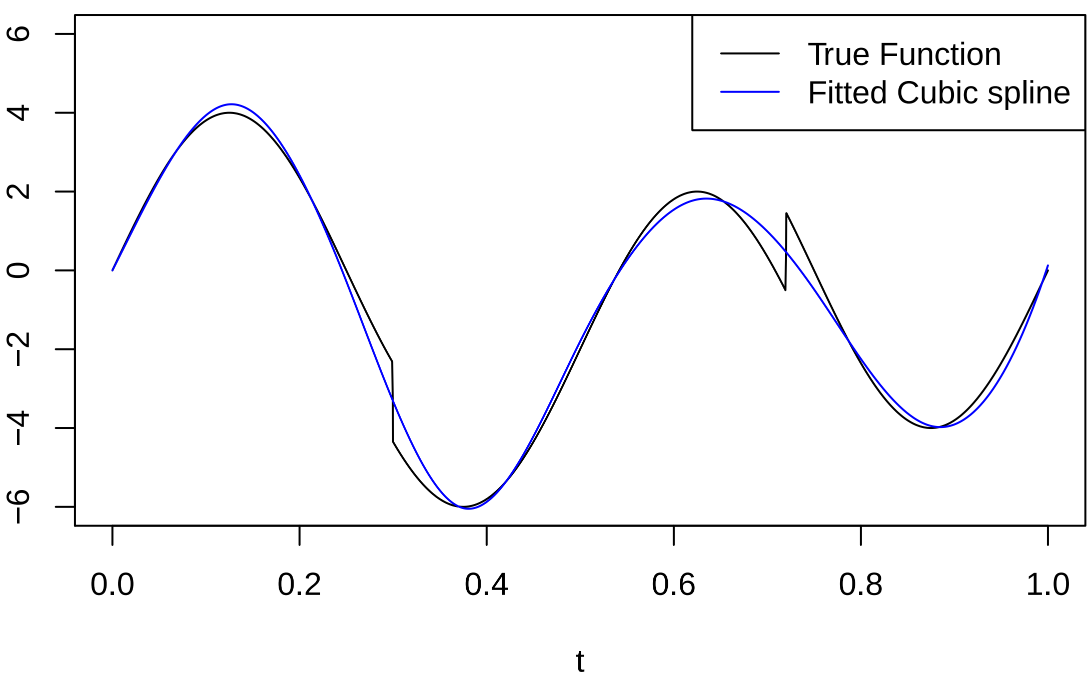


Noise was added to the function using a signal to noise ratio of 3. The addition of the noise makes the sharp edges harder to distinguish.



A cubic spline is able to capture smooth functions, a basis for a cubic spline is created using a B-spline basis with degree of freedom equal to 10, which creates evenly spaced local basis functions.

The model below is fit without penalization, it can be seen that the fitted model is approximating the smooth part of the function and not over fitting to noise, and so the spline basis in this question will be fit without penalization, rather keeping the degrees of freedom lower in order to not fit to noise. The fitted model is unable to predict the sudden jumps due to it's smooth nature.



In order to capture the sharp edges, wavlet smoothing was looked into. Choosing an appropriate mother wavelet function depends on finding a function which looks most similar to the signals you are trying to pick up as it will convolve this function with the signal and pick up areas where it is highly correlated. Since the part of the signal we are trying to pick up using the wavelet functions are sudden jumps, the Haar wavelet was chosen.

The Haar mother wavelet is defined as

$$\psi(x) = \begin{cases} 1 & if\ 0 < x < \frac{1}{2} \\ -1 & if\ \frac{1}{2} < x < 1 \\ 0 & otherwise \end{cases}$$

Daughter wavelets are created through dilations and translations of the mother wavelet

$$\psi_{jk}(x) = 2^{\frac{j}{2}}\psi(2^j - k)$$

where

$$j = 0, 1, ..., J - 1,\ with\ J = log_2(N)$$

for each resolution level j

$$k = 0, ..., 2^j - 1$$

The Haar wavelet basis is created using the **GenW** function from the **wavethresh** library, using 1024 data points the resolution level J is equal to 10.

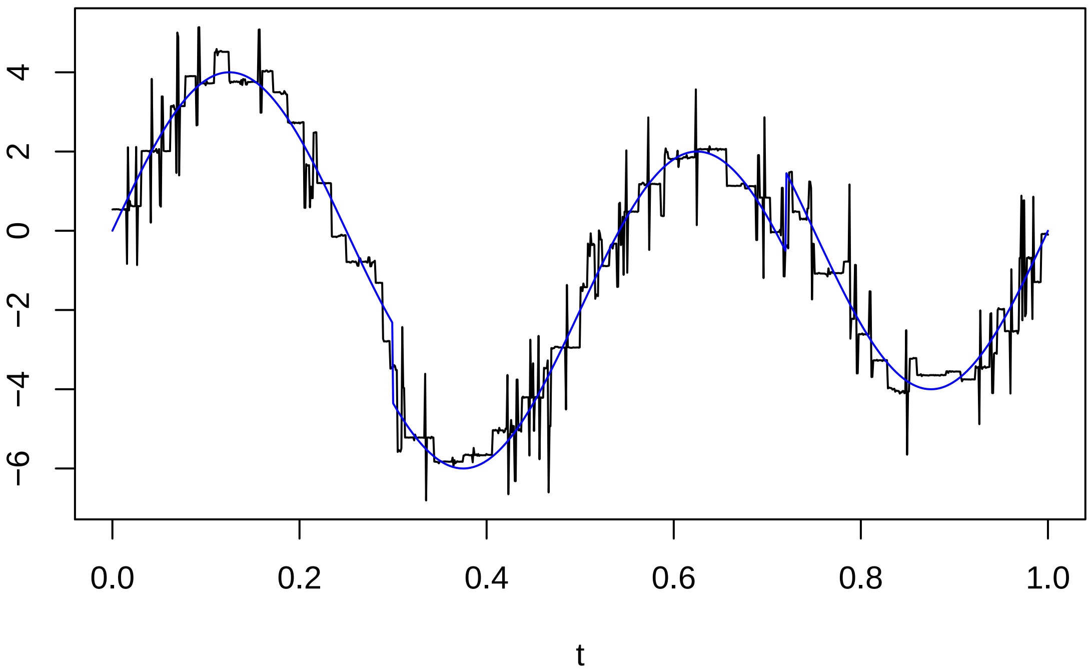
To fit a model the following loss function (Antoniadis A. 2007.) was minimised to obtain the penalised least squares wavelet estimator.

$$\ell(\theta) = ||Wy - \theta||_n^2 + 2\lambda \sum_{i<i_0} p(|\theta|)$$

An L1-penalty was added using the following equation (Antoniadis A. 2007.).

$$p_\lambda(|\theta|) = \lambda^2 - (|\theta| - \lambda)^2 I_{\{|\theta|<\lambda\}}(|\theta|)$$

The generated Haar basis was fitted to the data using the preformentioned method with a $\lambda = 2$. The fitted model captures the sharp edges but does a poor job capturing the smoothness of the function.



To capture both the smooth function and the sudden spikes, both the previous basis functions can be combined in order to derive benefit from both.

To fit a combination of the two techniques (splines and wavelet shrinkage) the proposed method in *Wavelet regression with automatic boundary adjustment* (Oh, Hee-Seok & Lee, Thomas. (2005)) was used. Fitting each part of the model to the residuals the left behind iteratively until convergence. The steps taken as outlined in the paper are as follows:

To find $\hat{f}_H(x)$ which is a combination of a local polynomial function $\hat{f}(x)_{LP}$ and wavelet function $\hat{f}(x)_W$

$$\hat{f}_H(x) = \hat{f}(x)_{LP} + \hat{f}(x)_W$$

1. Obtain an initial estimate for $\hat{f}(x)_{LP}$ by fitting a model to the local basis functions f^0 and setting $\hat{f}^0_{LP} = f^0$
2. For $j = 1, ...$ iterate the following steps
   (a) Apply wavelet shrinkage to $y_i - \hat{f}^{j-1}_{LP}$ and obtain $\hat{f}^j_W$
   (b) Estimate $\hat{f}^j_{LP}$ by fitting local polynomial regression to $y_i - \hat{f}^j_W$
3. Stop if $\hat{f}_H = \hat{f}^j_{LP} + \hat{f}^j_W$ converges.

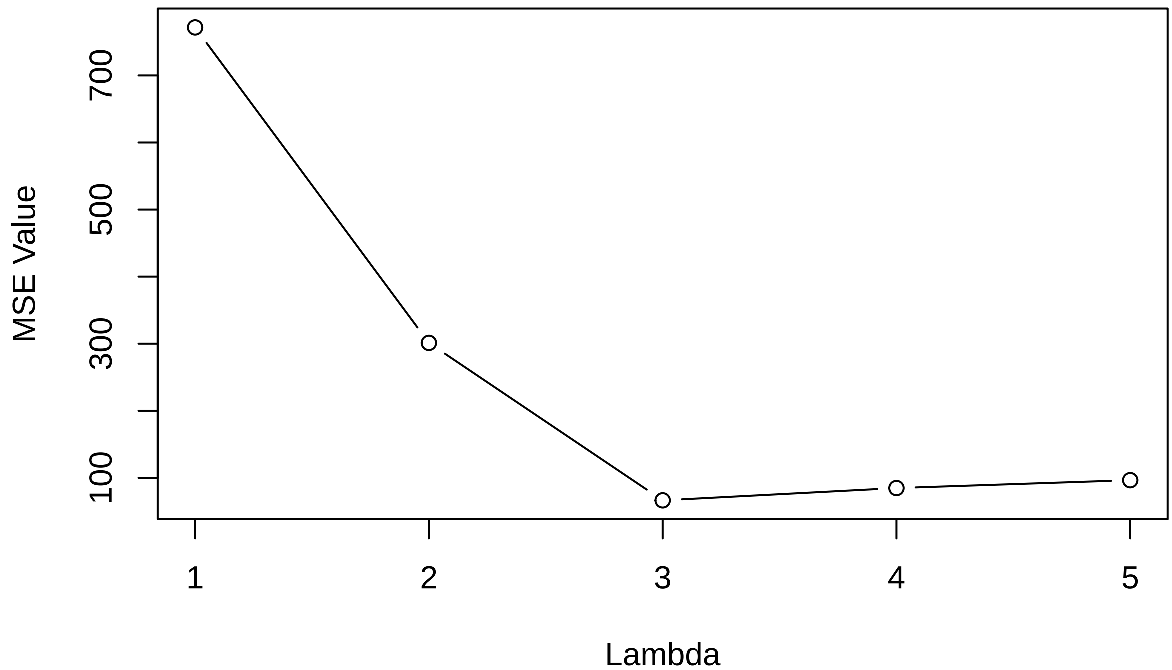
Instead of using a local polynomial function we use spline basis functions and conduct the same fitting methodology with $\hat{f}_{LP} = \hat{f}_{SPLINE}$

Choosing a threshold value was done by minimising the mean integrated squared error between the predicted function g_λ and the true function g(Antoniadis A, Bigot J, Sapatinas T. 2001.). Since we know the true underlying function, this method can be used.

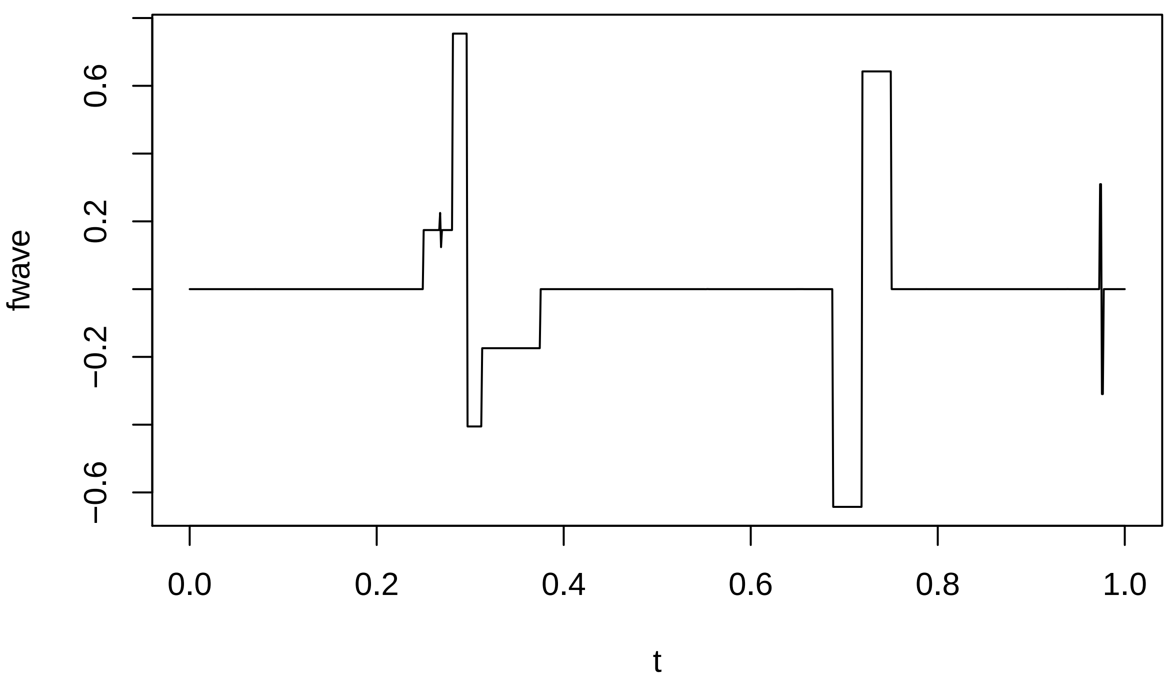
$$M(\lambda) = E\int (g_\lambda(x) - g(x))^2 dx$$

The minimum of the previous equation will occur at minimum squared pointwise sum of errors between the true values and approximated values.

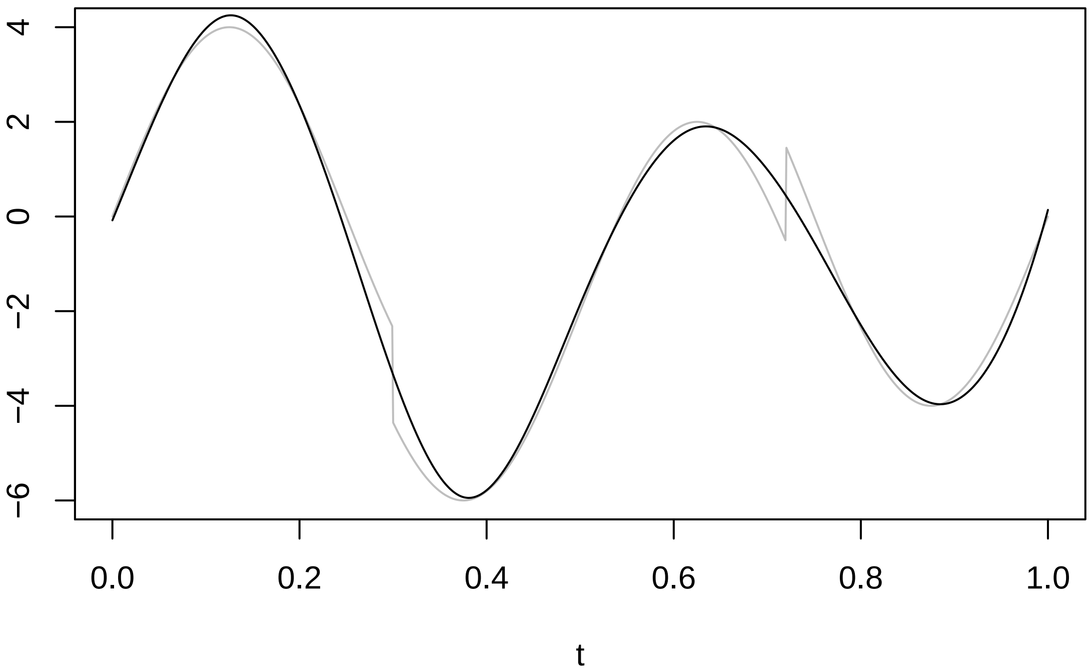
The MSE value vs λ can be seen below, with the best fit occuring at $\lambda = 3$



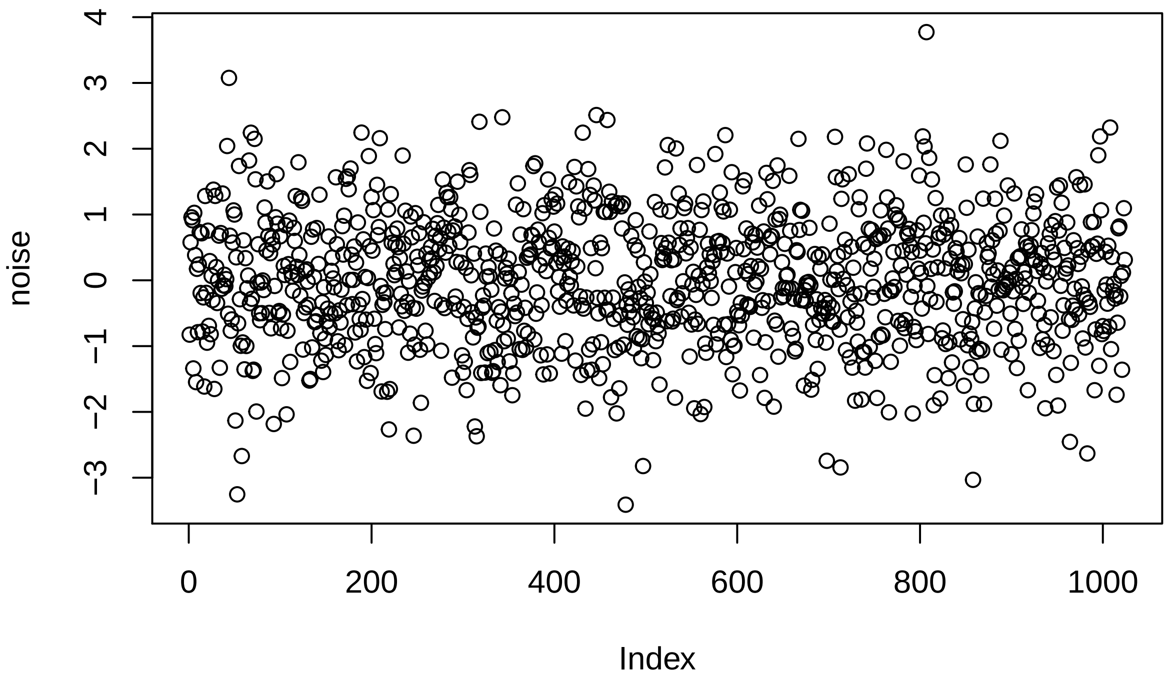
The final model was fit using $\lambda = 3$. The $\hat{f}_W(x)$ can be seen below, it picks up the two sharp edges at t = 0.3 and t = 0.72 with some fluctuations around them. The rest of the function is zero, allowing $\hat{f}_{SPLINE}(x)$ to approximate the smooth function.



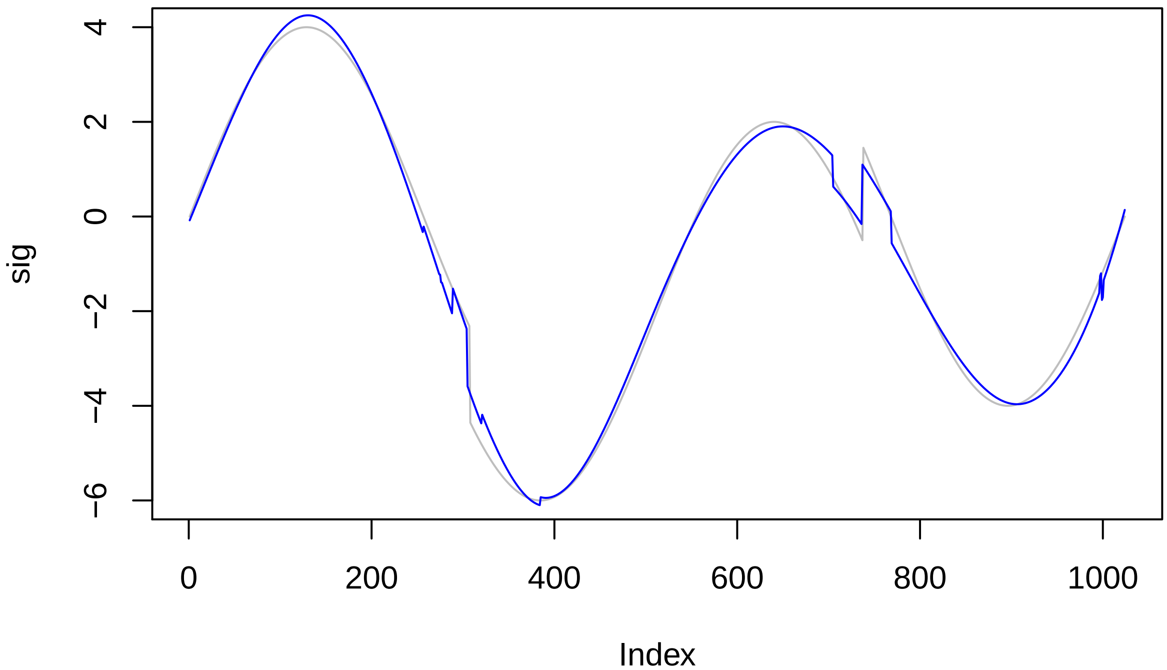
$\hat{f}_{SPLINE}(x)$ is a good approximation of the smooth part of the true function.



The residuals do not seem to contain any patterns, this is a good indication that the $\hat{f}_H(x)$ was able to capture the true function.



The fitted $\hat(f)_H(x)$ can be seen below. The function is able to capture the smooth and sharp edges.



References

1. Antoniadis A, Bigot J, Sapatinas T. 2001. Wavelet Estimators in Nonparametric Regression: A Comparative Simulation Study. J Stat Soft 6(6).

2. Antoniadis A. 2007. Wavelet methods in statistics: some recent developments and their applications. Statist Surv. 1(0):16???55.

3. Oh, Hee-Seok & Lee, Thomas. (2005). Hybrid local polynomial wavelet shrinkage: Wavelet regression with automatic boundary adjustment. Computational Statistics & Data Analysis. 48. 809-819. 10.1016/j.csda.2004.04.002.

4. Erni, Birgit, 2020, *Chapter 5: Wavelet Smoothing*, lecture notes, Advanced Analytics STA5057W, University of Cape Town, May 2020