## TESTING WITH COVARIANCE FUNCTIONS

A white paper with black text

Description automatically generated

A screenshot of a math problem

Description automatically generated

A math equations on a white background

Description automatically generated

A screenshot of a math problem

Description automatically generatedA math equations on a white background

Description automatically generated

A screenshot of a computer code

Description automatically generated

A screenshot of a computer code

Description automatically generated

A screenshot of a computer

Description automatically generated

### Results

Simulated over 5 predictors, **2 active.**

* **Always selected the correct predictors.**
* **Very high predictive capabilities**





Generally, very good results even when increasing the sig2 hyper parameter.

A graph with a line

Description automatically generated

Figure PREDICTOR 1

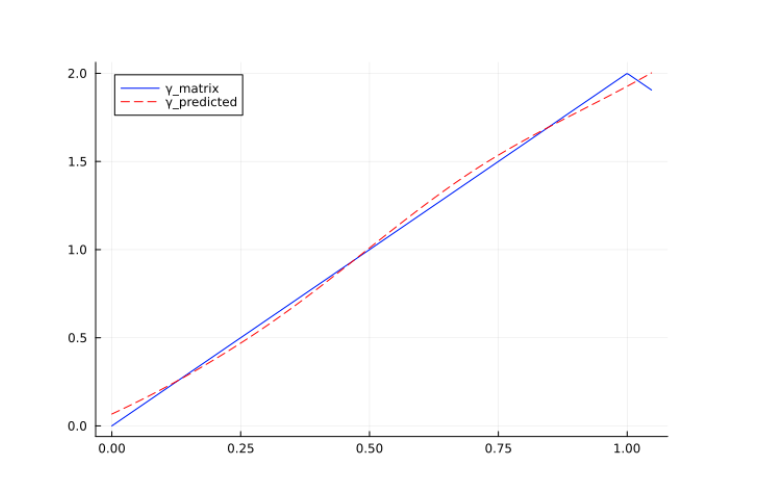


Figure 2 Predictor 2

## ORIGINAL PAPER PROBLEM



A math equations on a white background

Description automatically generated



Variability cannot be explained by the model. Even with the increase in sample size, the model overfits the data and it not able to generate the correct BETA coefficients.

The feature selection part of the model works and the corresponding predictors are selected.

## SIMPLIFIED SIMULATION

A screenshot of a white background

Description automatically generatedA screenshot of a math equation

Description automatically generated with medium confidence

In this simplified setup, the model confidence increases with the sample size. Better results were found with minor set of of bsplines (4 instead of the suggested 6 ). Still the model is not converging to the exact values of the BETA coefficients.

## SUPER SIMPLIFIED SIMULATION

A screenshot of a math test

Description automatically generated A screenshot of a math equation

Description automatically generated

In this supersimplified setup, the model confidence increases with the sample size. Better results were found with minor set of of bsplines (4 instead of the suggested 6 ).

Maybe the simulation is too complex and the author of the paper just wanted to show up the capabilities of the model to select the correct features *(without caring too much about the estimated betas, even because they never mentioned their values).*

## The importance of M-BOUNDS

Without correctly bounding the betas to specific values of M, generally the biggest absolute value in the expected beta matrix, the estimated coefficients do not look good.

## Data Hungry

I have kept testing the model, removing the added noise but keeping the same variability induced by the coefficients defined in the predictors.

I have stressed the limitations of the model and I found out that:

* The model is sensible to the time domain of the data. I might need to set different BIG M specific to each of the functional predictors.
  + I simplified the testing setup and I fixed the time domains to the same one for all the predictors
* If beta functions are “hard to learn”, It could be because the model requires a greater number of basis functions.
  + Increasing the model complexity, requires to feed more data to the model before having good results
    - Only these 2 beta functions needs to be estimated (out of 6)A white background with black text

      Description automatically generated
* All the predictors share the same time domain ( [0,1] ).

A screen shot of a computer

Description automatically generated

Figure simulation configuration

A graph of a function

Description automatically generated

Figure 4, the X predictor can be smoothed like this

A graph with red and blue lines

Description automatically generated

Figure 5 Beta 1 estimated form in RED, true form in Blue

A graph with red and blue lines

Description automatically generated

Figure 6 Figure 5 Beta 2 estimated form in RED, true form in Blue

**NOTE:** reducing the number of observations, drastically reduces the model performances.

A white background with black text

Description automatically generated

Figure 7 Increasing the number of measurements (improves the fitting of the predictors curves) and observations

A graph of a graph

Description automatically generated

A graph with red and blue lines

Description automatically generatedA screenshot of a computer

Description automatically generated A graph of a function

Description automatically generated

Figure 8 Before..

Figure 9 Curve fitting improved

But the gap was still high….Gurobi reached the time limit..

## Another test

Increased observations and measurements

* + - * Reached the time limit

A graph with a red line

Description automatically generatedA screenshot of a computer program

Description automatically generated A screenshot of a computer

Description automatically generated

A graph of a graph

Description automatically generated