Automated Model Building and Goodness-of-fit via Quantile Regression

Haim Bar

September 2, 2022

1 Simulated Data

1.1 Assessing Goodness of Fit

The following code shows a simulation scenario[1] in which the relationship between x and y is quadratic, and the error variance increases with x. The model is specified in line 14. We use QREM to fit the data twice – once with the correct model (line 18) and once with an incorrect model, where only a linear relationship is assumed (line 19).

```
source("initSim.R") # set up some global parameters ❺ ❶
N <- 4000
x <- runif(N,min=0,max=4*pi)
y <- sin(x) + 1.5*x + rnorm(N, 0, 0.5)
M <- data.frame(y,x)
colnames(M) <- c("y", "x1")
D <- ncol(M)-1

res <- fitQRloop(M=M, qn = qns, maxdeg = maxdeg, minDiff = minDiff)
plot(x,y, cex=0.5, pch=19, col="grey66")
```

1 is a comment

References

[1] Bar, H. Y., Booth, J. G., and Wells, M. T. (2020). A Scalable Empirical Bayes Approach to Variable Selection in Generalized Linear Models. *Journal of Computational and Graphical Statistics*, **0**(0), 1–12.

[1] In the examples in this documentation, the simulation numbers refer to Tables A3 and A4 in our paper