UNIVERSITÀ
DEGLI STUDI
FIRENZE

Multivariate Analysis and Statistical Learning

Backfitting process in Generalized Additive Model

Massimiliano Mancini

October, 14 2022

UNIVERSITÀ DEGLI STUDI FIRENZE

Generalized Additive Model

Generalized additive models or simply GAMs extend the standard linear model by allowing non-linear functions of each of the variables, while maintaining additivity.

The GAM is in the form

$$y_i = \beta_0 + \sum_{j=1}^{p} f_j(x_{ij}) + \varepsilon_i$$

It is called an additive model because we calculate a separate f_j for each X_j , and then add together all of their contributions

UNIVERSITÀ DEGLI STUDI FIRENZE

Backfitting

Standard software such as the `gam()` function in R can be used to fit GAMs using smoothing splines, via an approach known as *backfitting*.

This method fits a model involving multiple predictors by *backfitting* repeatedly updating the fit for each predictor in turn, holding the others fixed.

To understand the backfitting process, we can simulate it on a traditional linear model.

Backfitting

Let's pretend we don't have R functions able to fit more than one parameter.

To estimate parameters we keep fix one estimated (hat) parameter at time e repeat the process until results converge.

For each parameter let's estimate the following formula

$$Y - \widehat{\beta_1} X_1 - \widehat{\beta_2} X_2 - \ldots - \widehat{\beta_{k-1}} X_{k-1} - \widehat{\beta_{k+1}} X_{k+1} - \ldots - \widehat{\beta_p} X_p = \beta_0 + \beta_k X_k + \varepsilon$$

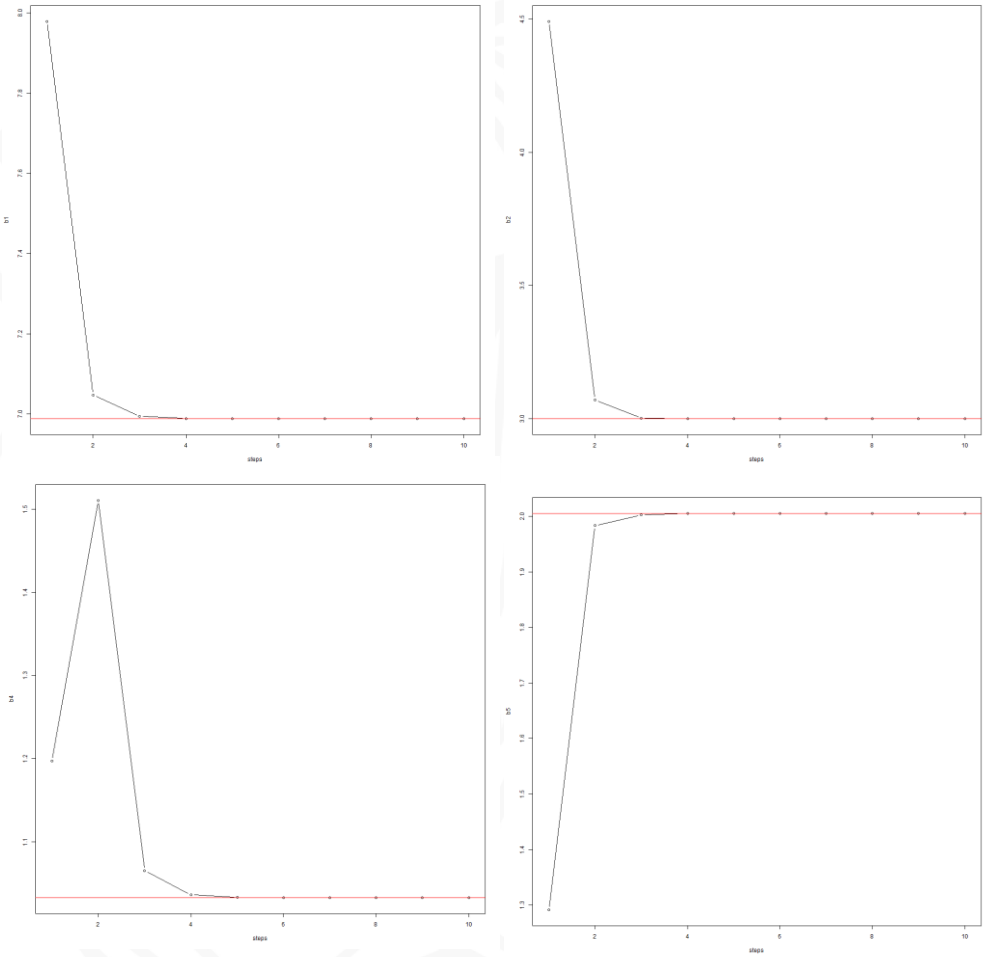
until it converges

UNIVERSITÀ DEGLI STUDI FIRENZE

Results

In few steps all five parameters are found even if we start from random values.

In the same way, parameters in GAM are found one by one, using backfitting



UNIVERSITÀ DEGLI STUDI FIRENZE

References

G. James, D. Witten, T. Hastie, R. Tibshirani

An Introduction to Statistical Learning with Applications in R

Second Edition