Generalized Additive Model GAM

Wilmer Ropero Castaño
wilmer.ropero@udea.edu.co

Jairo Arturo Morán Burgos
jairo.moran@udea.edu.co

Maestría en ingeniería

Universidad de Antioquia – Medellín, 2020-1

Introduction

Founders

Robert Tibshirani and Trevor Hastie 1986

Introduction

"Our algorithms have linked funny cat videos, UFO reports and searches for tofu pizza. We're now on alert about a suspicious group of cat aliens who infiltrated our pizza industry."

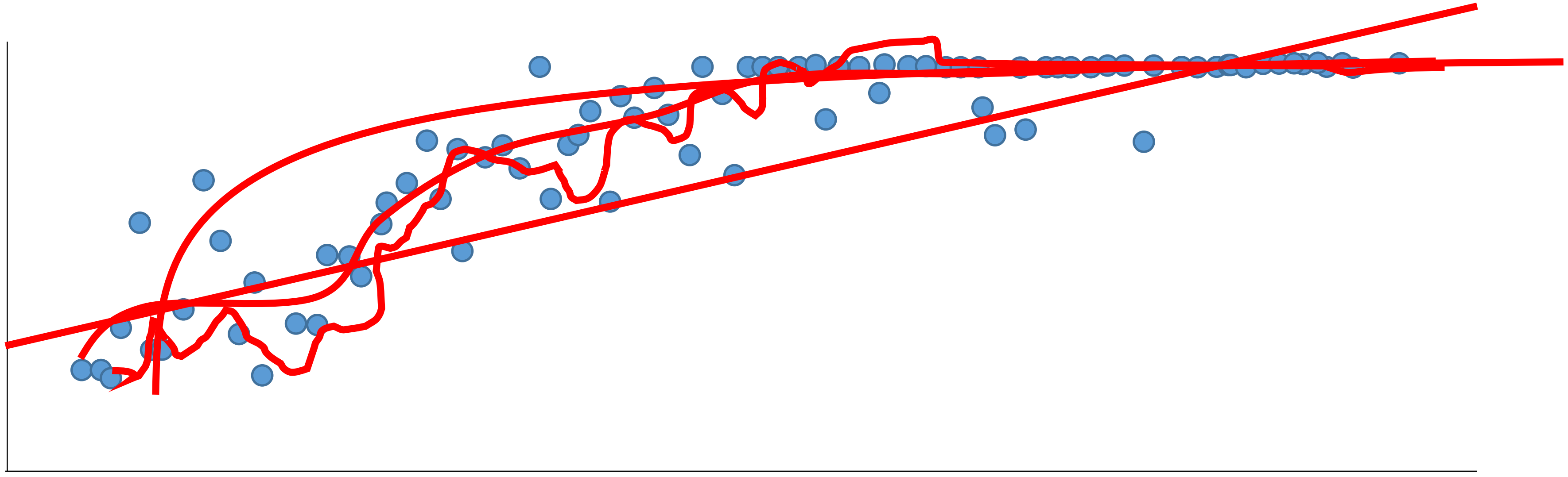
Despite its lack of popularity in the data science community, GAM is a powerful and yet simple technique.

•Easy to interpret.

•Flexible predictor functions can uncover hidden patterns in the data.

•Regularization of predictor functions helps avoid overfitting.

Motivation

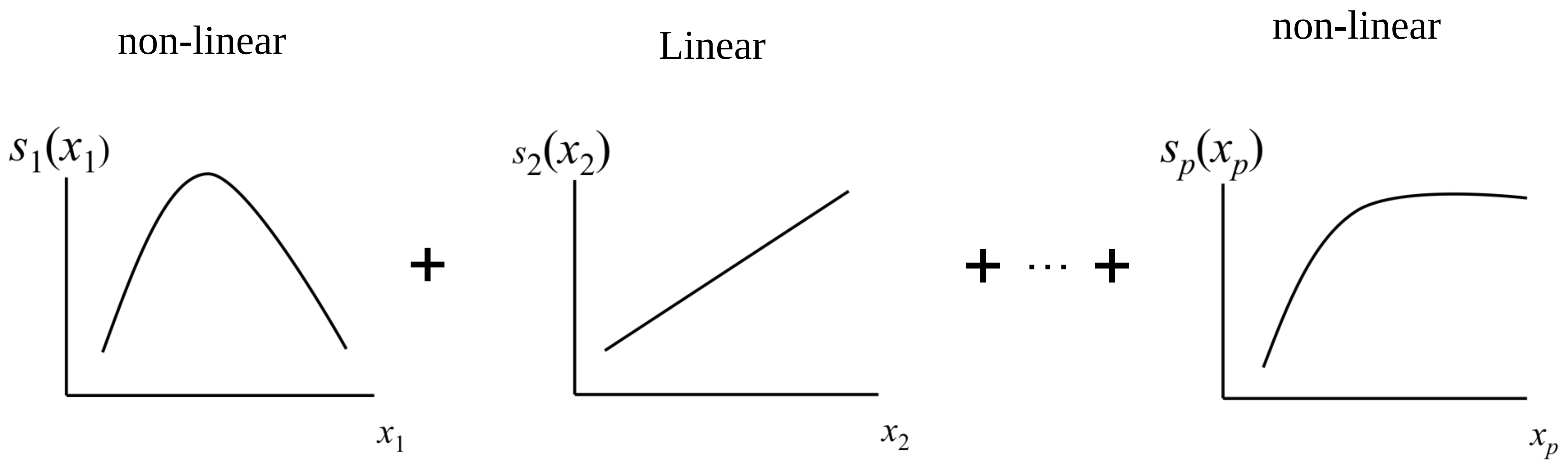


Lineal
Logarítmica
Polinómica
GAM

What is GAM?

Mathematically speaking, GAM is an **additive modeling** technique where the impact of the **predictive variables** is captured through **smooth functions** which, depending on the underlying patterns in the data, **can be nonlinear**

$g(E(Y)) =$



What is GAM?

link function dependent variable independent variables

$$g(E|y|) = \alpha + s_1(x_1) + s_2(x_2, x_3) + \cdots + s_p(x_p)$$

The feature functions $S_i()$ are built using **splines**, which allow us to **automatically model non-linear relationships** without having to manually try out many different transformations on each variable.

A brief tour or splines

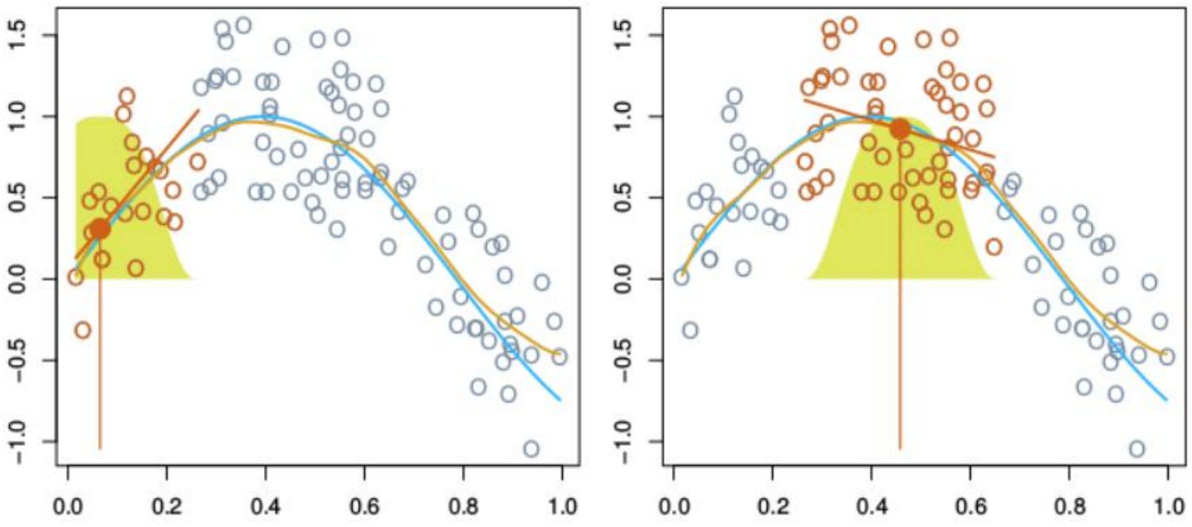
there are three classes of smoothers used for GAM

- Local regression(Loess)
- Smoothing splines
- Regression splines

A brief tour or splines

there are three classes of smoothers used for GAM

- Local regression(Loess)



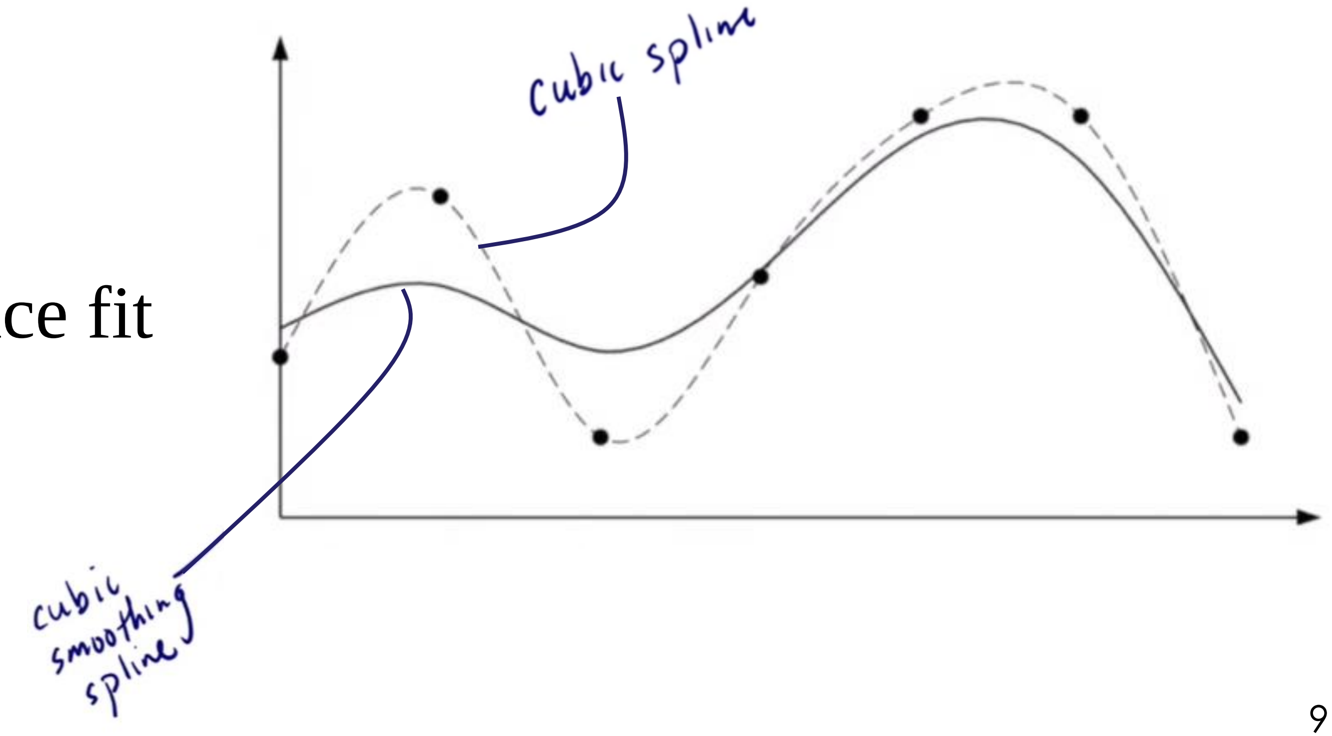
Loess produces a smoother curve than the running mean by fitting a weighted regression within each nearest-neighbor window

A brief tour or splines

there are three classes of smoothers used for GAM

- Smoothing splines

Is a spline designed to balance fit with smoothness

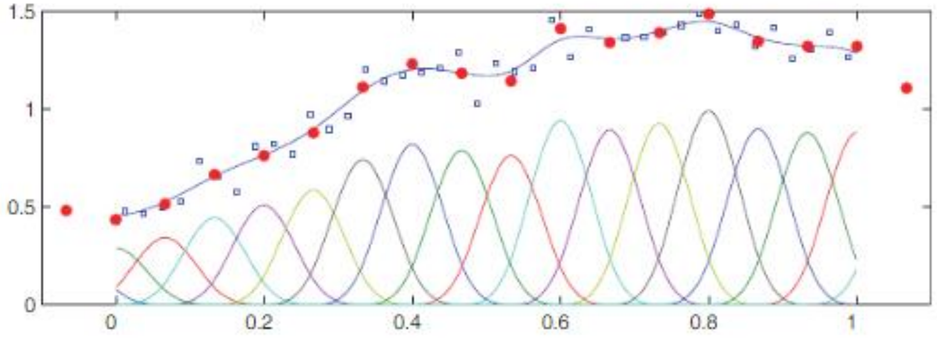


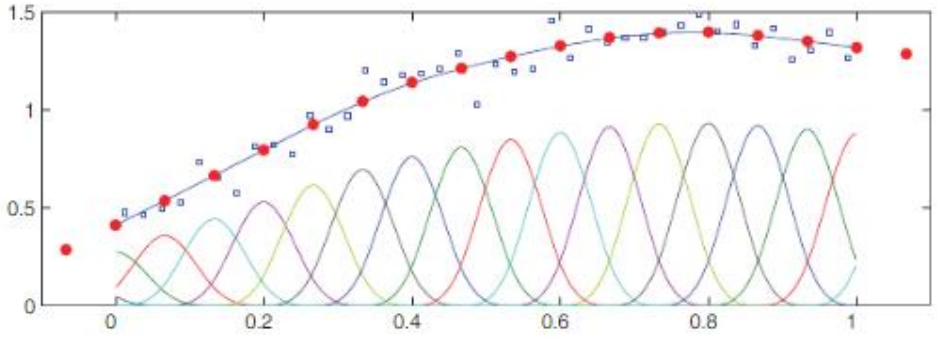
A brief tour or splines

there are three classes of smoothers used for GAM

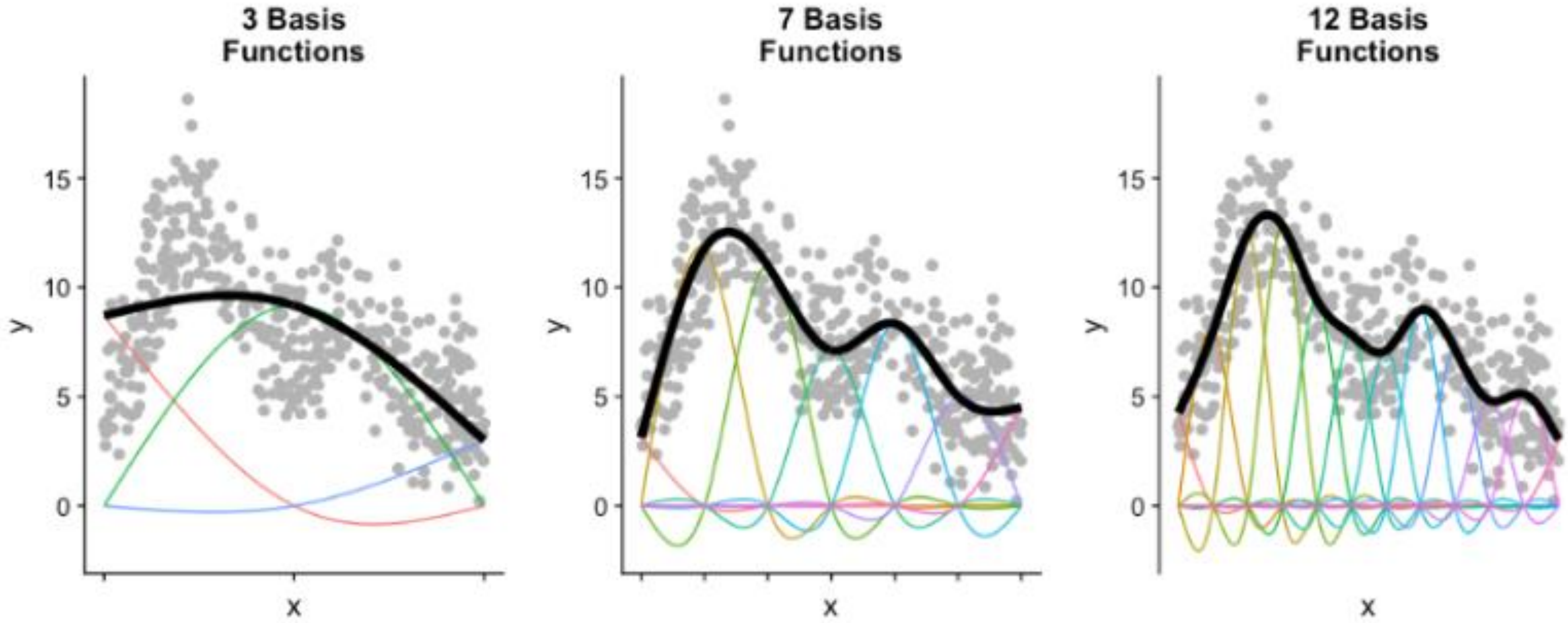
- Regression splines

they can be expressed as a linear combination, that do not depend on the dependent variable Y

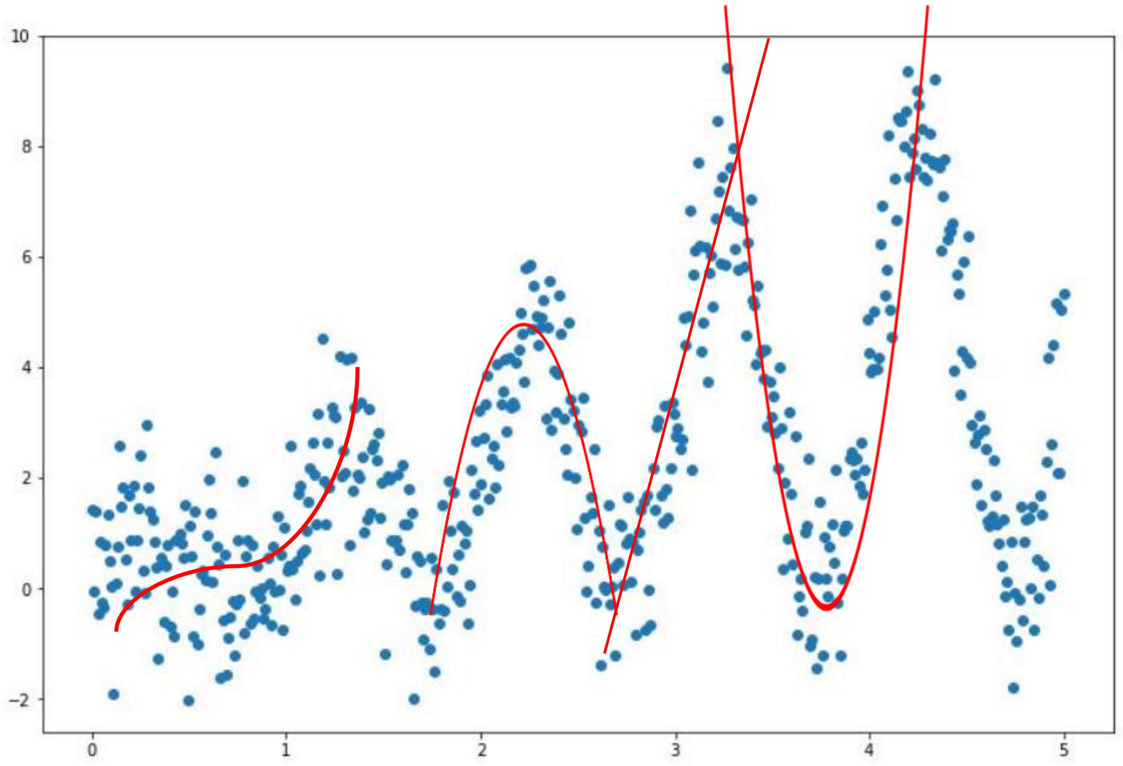




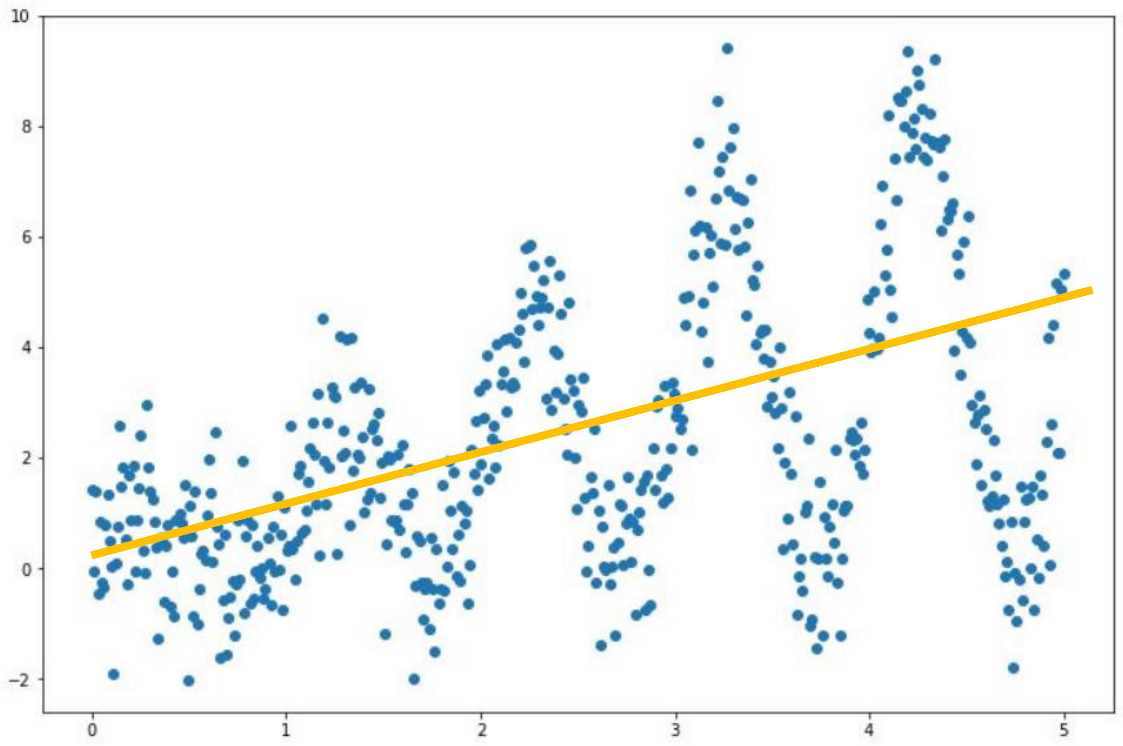
Building interpretable models

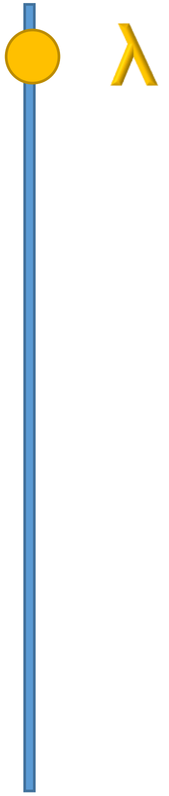


Building interpretable models

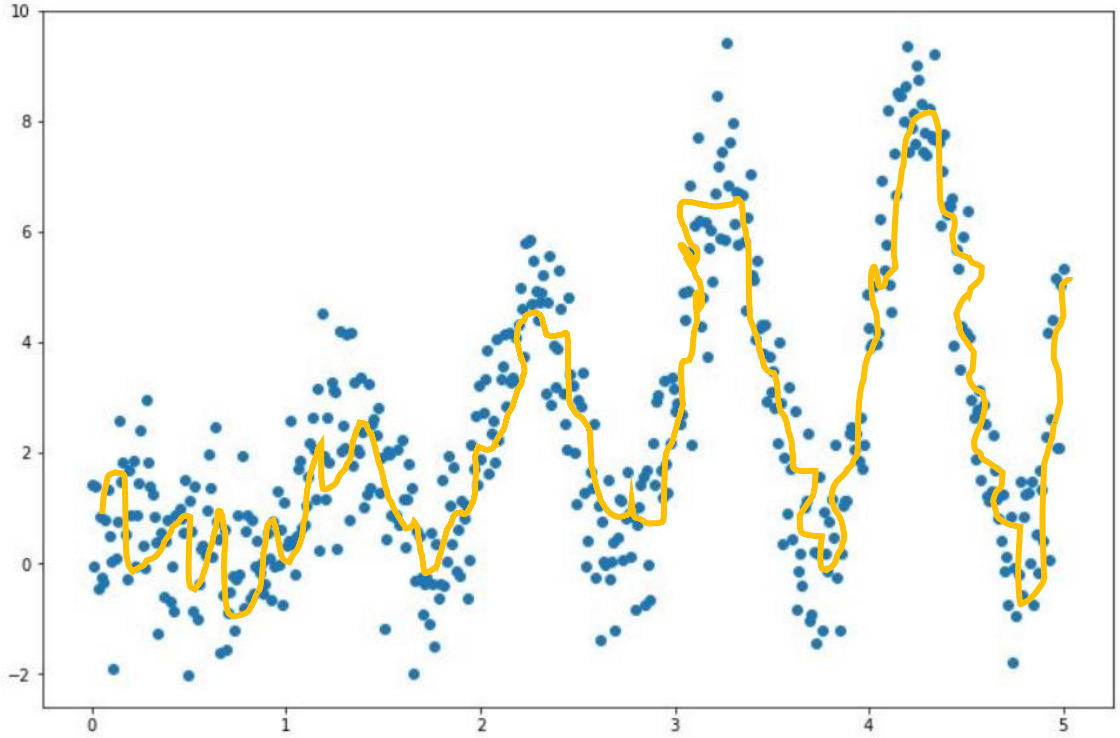


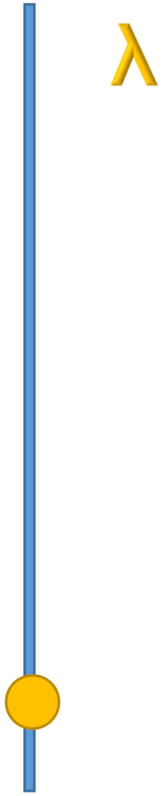
Building interpretable models



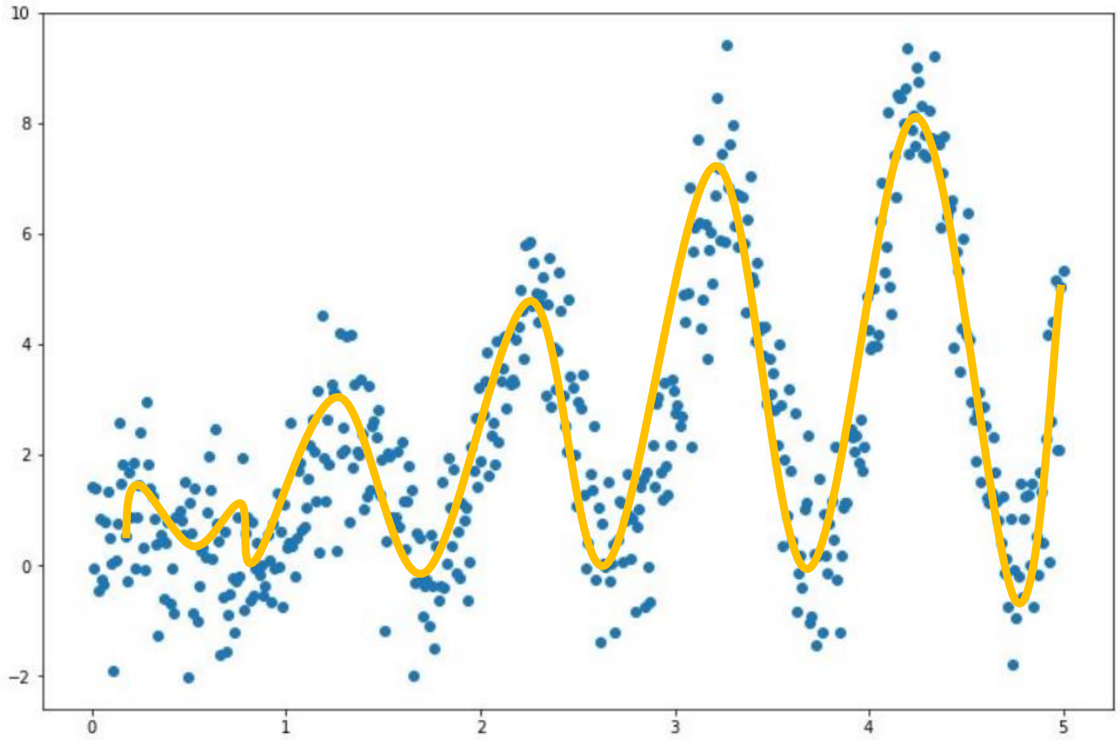


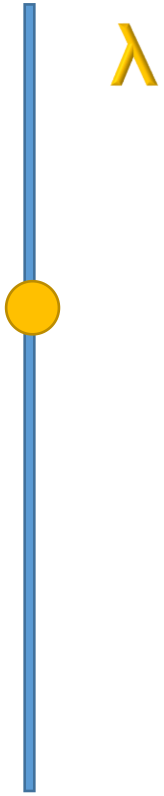
Building interpretable models





Building interpretable models





Estimating GAMs

GAMs consist of **multiple smoothing functions**. Thus, when estimating GAMs, the goal is to **simultaneously estimate all smoothers**, while **factoring in the covariance** between the smoothers

Local scoring algorithm

- Any type of smoother
- Computationally more expensive

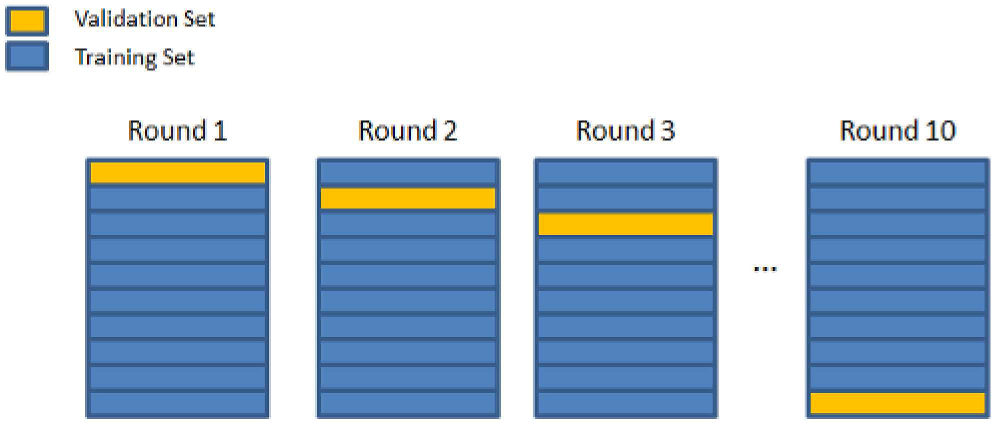
GAM as a large GLM

- Penalized Re-weighted Iterative Least Squares (PIRLS)
- Automated selection of smoothing parameters

Choosing the Smoothing Parameters

The choice of smoothing parameters, i.e., the parameters that control the smoothness of the predictive functions, is key for the aesthetics and fit of the model

Generalized cross validation criteria (GCV)



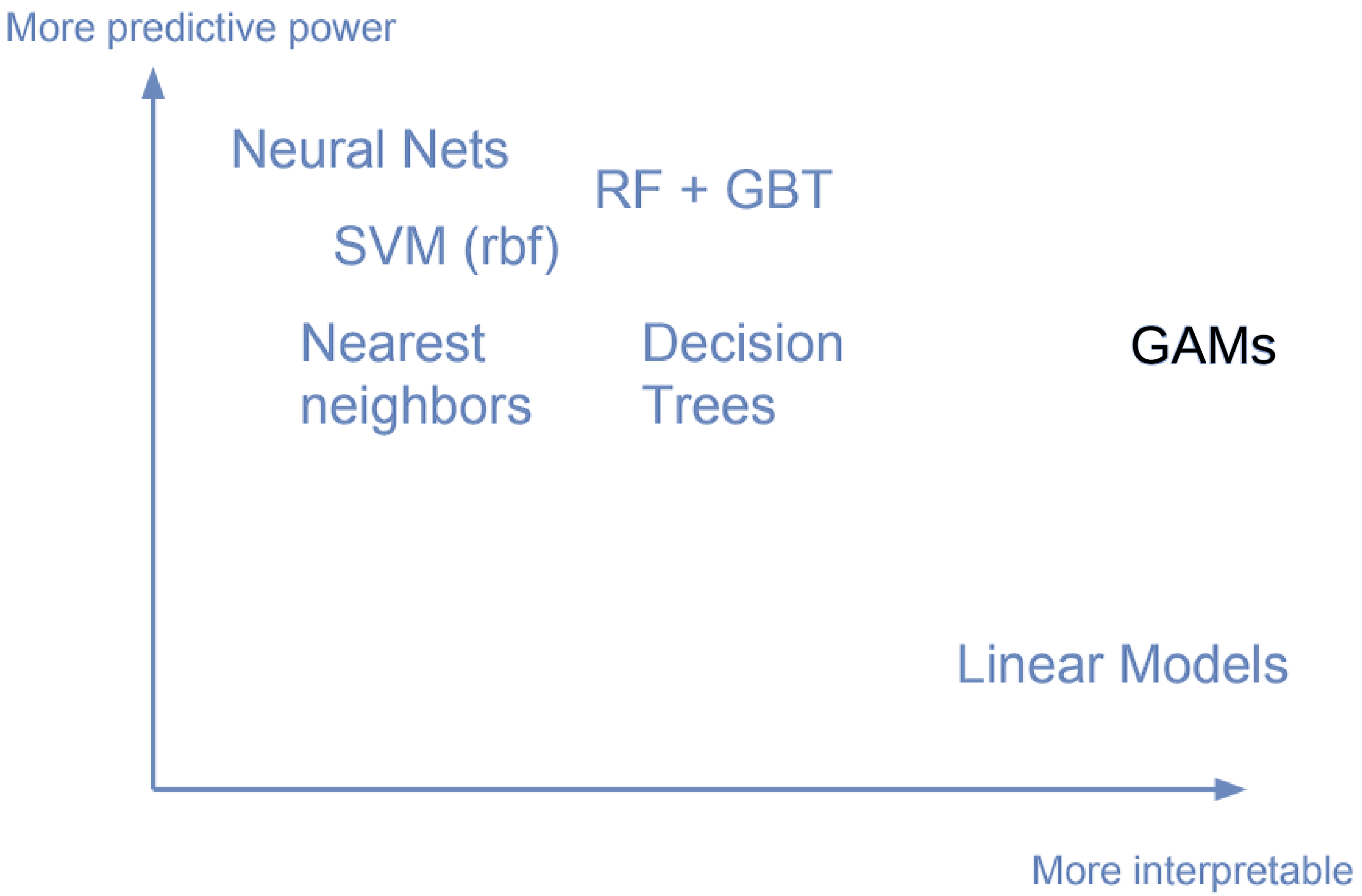
Mixed model approach via restricted maximum likelihood (<u>REML</u>).

1. Given a trial vector λ, estimate β using PIRLS.
2. Update λ by maximizing the restricted log likelihood.
3. Repeat steps 1 and 2 until convergence.

Why Use GAM?

interpretability, flexibility/automation, and regularization.

Why Use GAM?


More predictive power
Neural Nets
RF + GBT
SVM (rbf)
Nearest neighbors
Decision Trees
GAMs
Linear Models
More interpretable

Some comparations

Model	Validation AUROC	Estimation Time	Scoring Time
Random forest	0.809	6.39	39.38
GAM, lambda=0.6	0.807	3.47	0.52
GAM, estimate lambdas	0.815	42.72	0.29
GAM, estimate lambdas, extra shrinkage	0.814	169.73	0.33
SVM	0.755	13.41	1.12
Linear logit	0.800	0.1	0.006
KNN with K=100	0.800	NA	3.34

Conclusions

- Intuitive regularization via smoothing
- Automatically model non-linearities
- Constraints (convexity, monotonicity, periodicity, ...)
- Controlled extrapolation

Referencias

- Green, P. J., & Silverman, B. W. (1993). Nonparametric regression and generalized linear models: a roughness penalty approach. Chapman and Hall/CRC.
- Hastie, T. and Tibshirani, R. (1990). Generalized Additive Models, volume 43. CRC Press, 1990.
- Wood, S. N. (2017). Generalized additive models: an introduction with R. Chapman and Hall/CRC.
- Hastie, T., and Tibshirani, R. (1987). Generalized additive models: some applications. Journal of the American Statistical Association, 82(398), 371-386.
- https://cran.r-project.org/package=gam
- https://cran.r-project.org/package=mgcv
- https://www.statsmodels.org/stable/gam.html
- https://pygam.readthedocs.io/
- https://github.com/dswah/PyData-Berlin-2018-pyGAM/