EPH705 Homework 9. Generalized Additive Models

**Nina Dyatchenko**

April 4, 2019

1. Fit model Volume ~ Height + Girth with Gaussian link function:

library(mgcv)

## Warning: package 'mgcv' was built under R version 3.5.3

## Loading required package: nlme

## This is mgcv 1.8-28. For overview type 'help("mgcv-package")'.

log.volume <- log(trees$Volume)  
log.volume

## [1] 2.332144 2.332144 2.322388 2.797281 2.933857 2.980619 2.747271  
## [8] 2.901422 3.117950 2.990720 3.186353 3.044522 3.063391 3.058707  
## [15] 2.949688 3.100092 3.520461 3.310543 3.246491 3.214868 3.540959  
## [22] 3.456317 3.591818 3.645450 3.751854 4.014580 4.019980 4.065602  
## [29] 3.941582 3.931826 4.343805

model\_vol\_grth <- gam(log.volume ~ s(Height) + s(Girth),   
 data = trees,  
 family = gaussian(link = identity))  
model\_vol\_grth

##   
## Family: gaussian   
## Link function: identity   
##   
## Formula:  
## log.volume ~ s(Height) + s(Girth)  
##   
## Estimated degrees of freedom:  
## 1.00 2.41 total = 4.41   
##   
## GCV score: 0.00812451

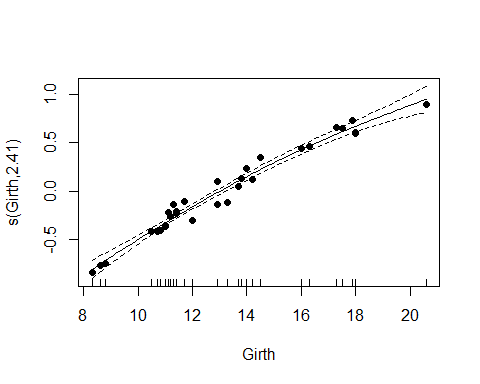
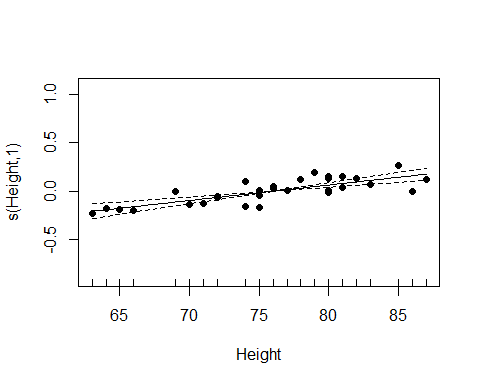
1. Fit GAM using REML with penalized cubic spline basis function:

model\_penal <- gam(log.volume ~ s(Height, bs = "cr") + s(Girth, bs = "cr"),  
 data = trees,  
 family = gaussian(link = identity),  
 method = "REML")  
model\_penal

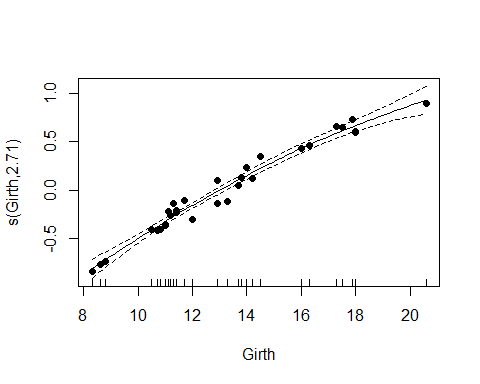
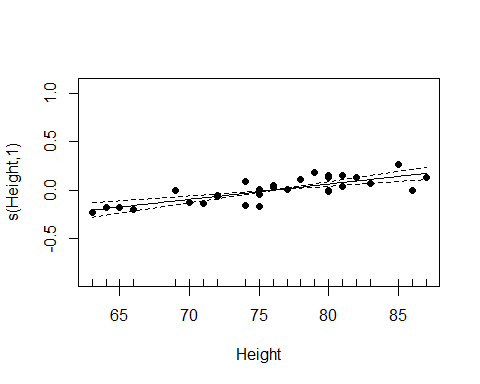
##   
## Family: gaussian   
## Link function: identity   
##   
## Formula:  
## log.volume ~ s(Height, bs = "cr") + s(Girth, bs = "cr")  
##   
## Estimated degrees of freedom:  
## 1.00 2.71 total = 4.71   
##   
## REML score: -26.04044

1. Plot fitted functions s(Height) vs. Height and s(Girth) vs. Girth:

plot(model\_vol\_grth,  
 residuals = TRUE,  
 pch = 19) ## calls plot.gam



plot(model\_penal,  
 residuals = TRUE,  
 pch = 19)



1. Predict Volume of a tree when Heoght = 70 and Girth = 10 for model obtained by REML:

vol\_predict <- data.frame(Height = 70,  
 Girth= 10)  
predict(model\_penal,  
 newdata = vol\_predict)

## 1   
## 2.675701