HW6 - Lucas Fellmeth, Sven Bergmann

2023-11-29

1

The goal of this problem is to estimate the regression function of acceleration vs time for the `mcycle` data in the package `MASS`, using cubic smoothing splines.
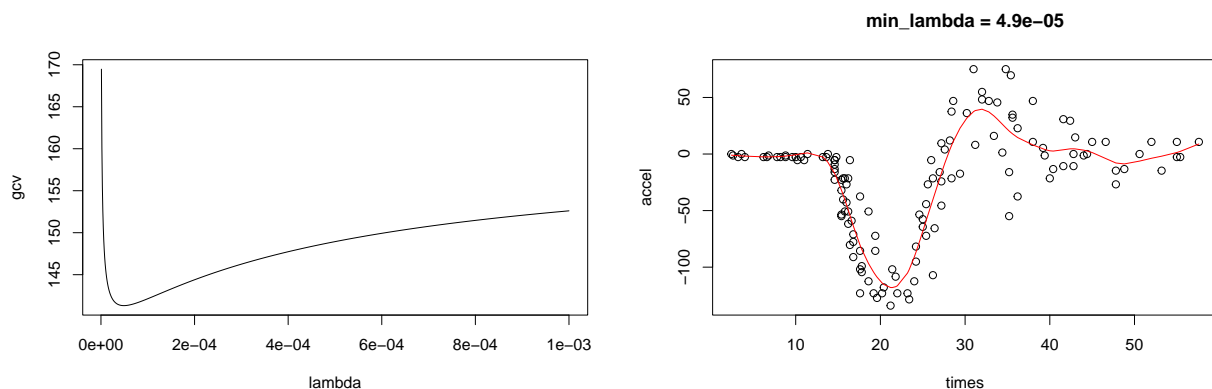
```r
library(MASS)
library(splines)
```

A

For a reasonable range of smoothing parameters λ, compute and plot the generalized cross validation measure $GCV(\lambda)$ and find the optimal smoothing parameter.

```r
with(mcycle, {
    lambda <- seq(10^(-6), 10^(-3), by = 10^(-6))

    smoothing <- function(l, x, y) {
        return(smooth.spline(x, y, lambda = l)$cv.crit)
    }

    gcv <- sapply(lambda, smoothing, x = accel, y = times)
    tmp <- data.frame(lambda = lambda, gcv = gcv)
    plot(tmp, type = "l")
    plot.new()
    plot(times, accel, main = paste("min_lambda =", tmp[which.min(tmp$gcv), ]$lambda))
    lines(smooth.spline(x = times, y = accel, lambda = tmp[which.min(tmp$gcv), ]$lambda),
        col = "red", type = "l")
})
```

B

Using the Demmler-Reinsch basis representation, how many basis functions contribute significantly to the fit, as determined, for instance, by accumulating 90% of the total degrees of freedom?

https://statisticaloddsandends.wordpress.com/2019/10/04/the-demmler-reinsch-basis-for-smoothing-splines/