FDA Homework 1

Seokjun Choi

September 30, 2019

1 Chapter 1

1.1 Problem 1

The pinch is a dataset included in the fda package. It consists of 151 measurements of pinch force for 20 replications (curves).

(a) Convert the pinch data to functional objects using 15 B-splines of order four (cubic splines) and plot the 20 smoothed curves on one graph.

To get the result we want, call the fda package and execute the code below.

```
#R code

b_spline_basis<-create.bspline.basis(c(1,151), nbasis=15)
pinch.F<-Data2fd(1:151, pinch, b_spline_basis)
plot(pinch.F)
```
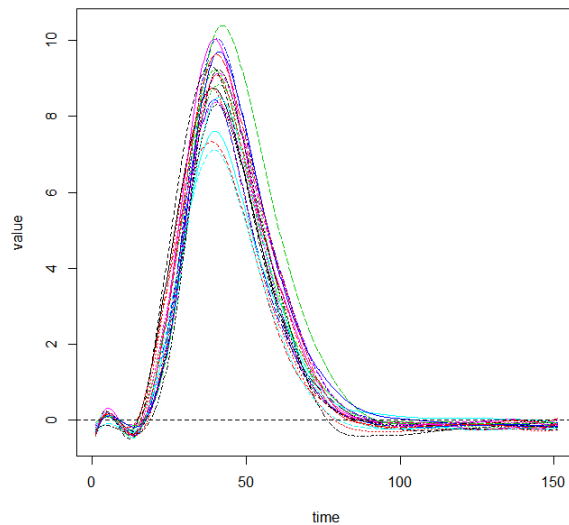
Then we get the result.

Figure 1: fitting pinch data using 15 b-splines

(b) Calculate the pointwise mean and SD and add them to the plot.

For calculation, run the code below.

```
#R code

pinch.F.mean <- mean(pinch.F)
pinch.F.mean$coefs

pinch.F.std <- std.fd(pinch.F)
pinch.F.std$coefs

par(mar=c(4,4,1,1))
plot(pinch.F, col="grey")
plot(pinch.F.mean,lwd=4,add=TRUE)
plot(pinch.F.std, lwd=4, add=TRUE, col="red")
```

For analytic expression of pointwise mean and point sd, denote b_i as i-th b-spline of order 4 on domain $[0, 120] \subset \Re$. Then

$$Mean(x) = \sum_{i=1}^{16} c_i b_i(x)$$

$$Sd(x) = \sum_{i=1}^{16} s_i b_i(x)$$

where c_i : -0.31109576, 0.73142704, -1.63349675, 2.72475346, 10.93667264, 6.03218354, 2.28791198, 0.34957598, -0.09607592, -0.16637125, -0.11240041, -0.15664938, -0.10293668, -0.16311715, -0.11897440 and s_i : 0.0846463, 0.1182031, -0.0996403, 0.9735739, 0.6517423, 1.1140145, 0.5002106, 0.3255844, 0.1166337, 0.1232983, 0.0639047, 0.0869009, 0.0824064 0.0776430, 0.07207136 respectively.
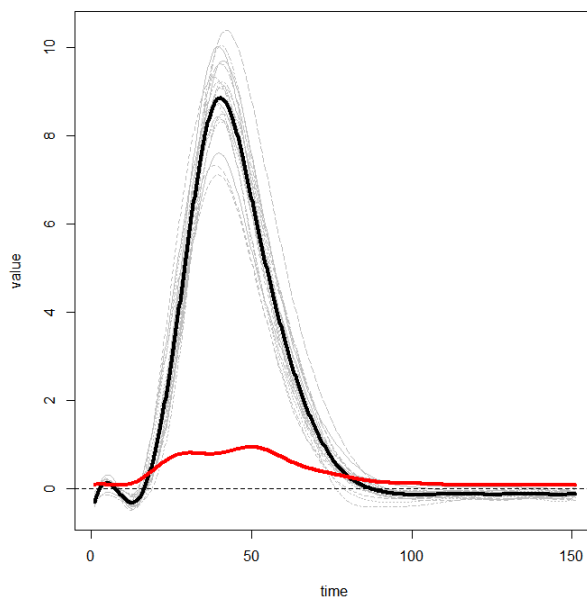
And here is a plot of the result.

Figure 2: fitting pinch data using 15 b-splines:
Mean is black bold curve, and Sd is red bold curve.

(c) Graph the perspective and contour plots of the sample covariance function $\hat{c}(t, s)$ of the pinch curves.

We can easily do them by using for var.fd function of fda package and persp, contour function of R.

```
#R code

pinch.F.cov <- var.fd(pinch.F)
dim(pinch.F.cov$coef) #15*15
grid <- 1:15
pinch.F.cov.mat = eval.bifd(grid, grid, pinch.F.cov)
par(mfrow=c(1,2), mar=c(4,4,1,1))
persp(grid, grid, pinch.F.cov.mat, xlab="s", ylab="t", zlab="cov(s,t)")
contour(grid, grid, pinch.F.cov.mat)
```
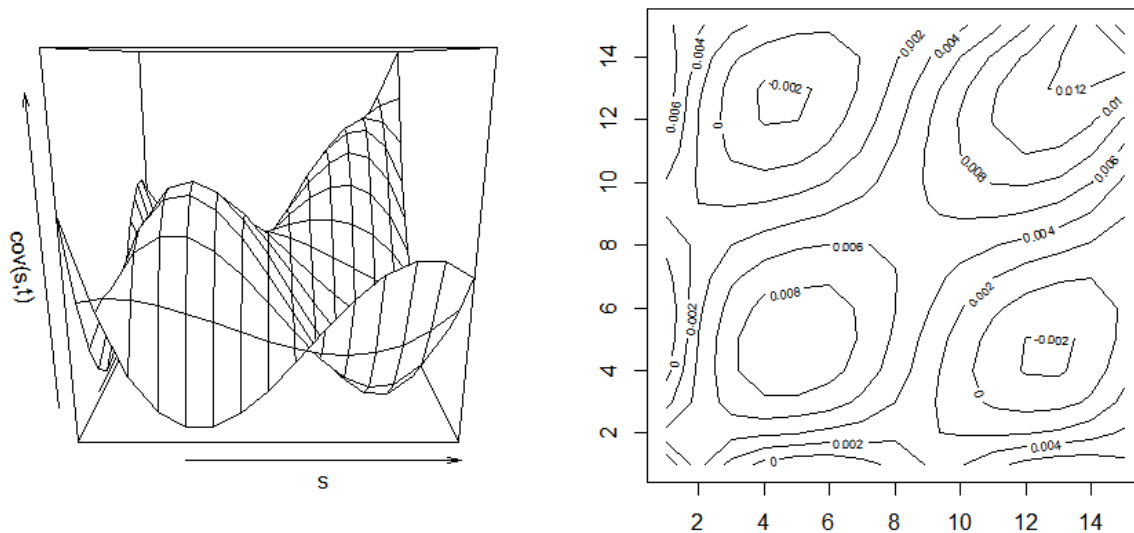
Result plots are here.

Figure 3: fitting pinch data using 15 b-splines:
covariance perspective plot and contour plot.

(d) Graph the first four EFPC's of the pinch data. How many components do you need to explain 90% of variation?

Let's do PCA with 4 components following the instruction of problem, and see the proportions of variance of each component for determining number of components for 90% explanation.

```
#R code

par(mfrow=c(1,1))
pinch.F.pca = pca.fd(pinch.F, nharm=4)
plot(pinch.F.pca$harmonics, lwd=3)
pinch.F.pca$varprop
```
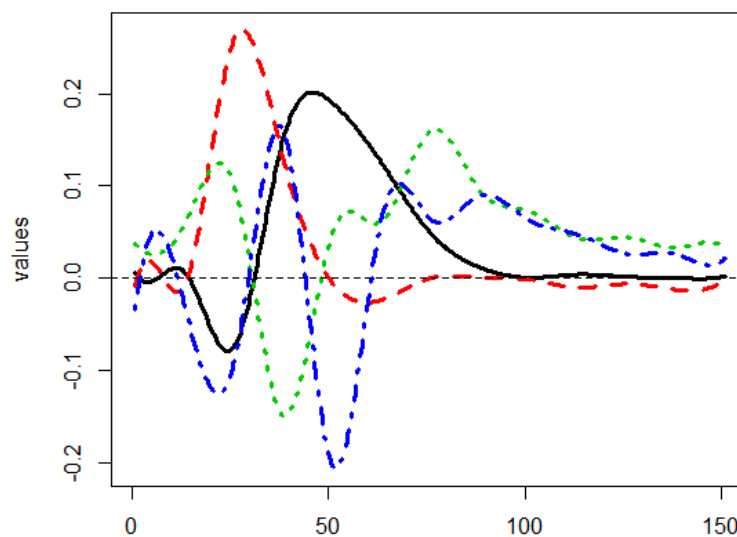
Firstly, Plot the first 4 eigenfunctions.

Figure 4: fitting pinch data using 15 b-splines and do PCA:
first 4 components.

And the ouput of last line is

```
#Console

> pinch.F.pca$varprop
[1] 0.67225632 0.24845297 0.04603548 0.01933904
```

Since the sum of first 2 proportion is 0.9207093, we need only 2 components to reach 90% level of explanation of variation .

4