Práctica 1

1. Considere el conjunto de datos `pinch` de la librería `fda`. Los datos corresponden a 20 mediciones de la fuerza de pellizco tomadas cada 2 mili–segundos. Los datos originales corresponden a 300 tiempos. El sub-conjunto de datos `pinch` corresponde a haber seleccionado las observaciones de modo que el máximo de las curvas ocurra a los 0.076 segundos.

   (a) Grafique las curvas junto son el estimador $\widehat{\mu}$ de su media.

   (b) Suavice los datos usando 15 $B-$splines cúbicos. Grafique las curvas suavizadas junto con el promedio de los datos suavizados.

      i. Grafique las superficies $\widehat{\gamma}(t,s)$ para los datos orginales que llamaremos $\widehat{\gamma}$ y suavizados $\widehat{\gamma}_{\text{BS}}$.

      ii. Para $j=1,\dots,4$, grafique la $j-$ésima autofunción de $\widehat{\Gamma}$ y $\widehat{\Gamma}_{\text{BS}}$ en el mismo grafico, asegurándose que tienen la misma orientación, o sea, signo$(\widehat{\phi}_j, \widehat{\phi}_{j,\text{BS}})=1$. En cada caso, cuantas de ellas explican el 90% de la variabilidad total.

   (c) Suavice los datos usando un suavizado basado en núcleos usando el núcleo normal y distintos valores de ventana $h=0.1, 0.2$ y 0.5. Grafique las curvas suavizadas junto con el promedio de los datos suavizados. Que observa?

      i. Grafique las superficies $\widehat{\gamma}(t,s)$ para los datos orginales que llamaremos $\widehat{\gamma}$ y suavizados $\widehat{\gamma}_h$ cuando $h=0.1$.

      ii. Para $j=1,\dots,4$, grafique la $j-$ésima autofunción de $\widehat{\Gamma}$ y $\widehat{\Gamma}_h$ cuando $h=0.1$ en el mismo grafico, asegurándose que tienen la misma orientación, o sea, signo$(\widehat{\phi}_j, \widehat{\phi}_{j,h})=1$. En cada caso, cuantas de ellas explican el 90% de la variabilidad total.

2. Considere el conjunto de datos `CanadianWeather` de la librería `fda` y los promedios mensuales de las 35 estaciones, `CanadianWeather$monthlyTemp`.

   (a) Grafique las curvas junto son el estimador $\widehat{\mu}$ de su media.

   (b) Grafique la superficie $\widehat{\gamma}(t,s)$.

   (c) Realice el boxplot funcional de los datos e identifique los datos atípicos

   (d) Realice los boxplots puntuales correspondientes a cada mes. Qué diferencias observa. Comente.

   (e) Use el boxplot funcional para mostrar 3 regiones que correspondan niveles 30, 60 y 90%, usando un color más oscuro para la región central 30% y colores más claros para 60 y 90%.

3. El núcleo de covarianza de Mattern está definido por

$$\gamma_\nu(s,t) \quad = \quad C\left(\frac{\sqrt{2\nu}\,|s-t|}{\rho}\right) \qquad C(u)=\frac{\sigma^2\, 2^{1-\nu}}{\Gamma(\nu)}\, u^\nu K_\nu\left(u\right)\,, \qquad 0<t,s<1 \tag{1}$$

para $\nu > 0$. σ^2 es el parámetro de varianza, ν es el de suavidad y ρ el de rango. $K_\nu(\cdot)$ es la función de Bessel de segundo tipo y pueden obtenerse en `R` mediante la función `besselK`. Las trauyectorias de un proceso Mattern son k veces diferenciables para cualquier $\nu > k$ con probabilidad 1.

(a) Fije como semilla 1234, simule y grafique 100 trayectorias i.i.d.de un proceso con media $\mu(t) = 0$ y con operador de covarianza Γ_ν dado por el núcleo de covarianza (1) tomando $\sigma^2 = \rho = 1$, para tres valores de ν, $\nu = 0.5, 2$ y 4. Use una grilla en $[0, 1]$ de 50 puntos equiespaciados.

No use un paquete de `R` pero hagalo usando la funcion `besselK` para obtener una matriz de covarianza en la grilla de tiempos.

(b) Use el método que prefiera para crear un grafico de la superficie del núcleo de covarianza real y estimado para cada ν, comente las similaridades y diferencias que observa.

(c) Para $j = 1, \ldots, 4$ y cada valor de ν, grafique la $j-$ésima autofunción de $\widehat{\Gamma}_\nu$ y Γ_ν en el mismo grafico, asegurándose que tienen la misma orientación, o sea, signo$(\widehat{\phi}_j, \phi_j) = 1$. Comente las similaridades y diferencias que observa.

Plotee la variabilidad explicada para cada ν, comente las similaridades y diferencias que observa.

(d) Cuando $\nu = 2$, plotee las derivadas de orden 1 calculadas usando $B-$splines cúbicos con 50 elementos utilizando la función `deriv.fd` del paquete `fda` y las derivadas numéricas tradicionales.