Exercise no 7

Corentin Lacroix 1812554

MTH 6312: Méthodes Statistiques D'Apprentissage

November 3, 2015

Exercise 1 Prove the $h_3(x) = (x - \xi_2)_+$ cause continuity at $x = \xi_2$ in expansion of the form $f(x) = \beta_0 + \beta_1 x + (x - \xi_1)_+ + (x - \xi_2)_+$

Solution 1 Let f be $f(x) = \beta_0 + \beta_1 x + (x - \xi_1)_+ + (x - \xi_2)_+ \forall x \in \mathcal{R}$ with $(\xi_1, \xi_2) \in \mathcal{R}^2$ and $(\beta_0, \beta_1) \in \mathcal{R}^2$

We have

$$f(x) = \begin{cases} \beta_0 + \beta_1 x \ if \ x \leq \xi_1 \\ \beta_0 + \beta_1 x + x - \xi_1 \ if \ \xi_1 < x \leq \xi_2 \\ \beta_0 + \beta_1 x + x - \xi_1 + x - \xi_2 \ if \ x > \xi_2 \end{cases}$$

Then, $\underset{\substack{x>\xi_2 \\ x\to\xi_2}}{lim} f(x) = \underset{\substack{x>\xi_2 \\ x\to\xi_2}}{lim} (\beta_0 + \beta_1 x + x - \xi_1 + x - \xi_2) = \beta_0 + \beta_1 \xi_2 + \xi_2 - \xi_1$

And $\underset{\substack{x<\xi_2 \\ x\to\xi_2}}{lim} f(x) = \underset{\substack{x<\xi_2 \\ x\to\xi_2}}{lim} (\beta_0 + \beta_1 x + x - \xi_1) = \beta_0 + \beta_1 \xi_2 + \xi_2 - \xi_1 = \underset{\substack{x>\xi_2 \\ x\to\xi_2}}{lim} f(x)$

Consequently, the f function is continous at $x = \xi_2$ thanks to the term $h_3(x) = (x - \xi_2)_+$

Exercise 2 Why don't we include terms of the form $(x - \xi_l)^i$ for $i \in \{0, 1, 2\}$ in cubic spline expansion? Prove it for the case of cubic spline with only one knot

Solution 2 Let f be $f(x) = \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3 + \beta_4 (x - \xi)_+^3 \ \forall x \in \mathcal{R}$ with $\xi \in \mathcal{R}$ and $(\beta_0, \beta_1, \beta_2, \beta_3, \beta_4) \in \mathcal{R}^5$

We have :

$$f(x) = \begin{cases} \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3 = f_1(x) \ if \ x \leq \xi \\ \\ \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3 + \beta_4 (x-\xi)^3 = f2(x) \ if \ x > \xi \end{cases}$$

With $f_1(x) = 0 \ if \ x > \xi$ and $f_2(x) = 0 \ if \ x \leq \xi$

So,

$$f'(x) = \begin{cases} f_1'(x) \ if \ x \leq \xi \\ \\ f_2'(x) \ if \ x > \xi \end{cases}$$

And we have a similar result for f".

If we derive f_1 and f_2, we get :

$$\begin{cases} f_1'(x) = \beta_1 + 2\beta_2 x + 3\beta_3 x^2 \ \forall x \leq \xi \\ \\ f_2'(x) = \beta_1 + 2\beta_2 x + 3\beta_3 x^2 + 3\beta_4 (x-\xi)^2 \ \forall x > \xi \end{cases}$$

Then, immediatly,

$$\underset{\substack{x \leq \xi \\ x \to \xi}}{lim} f_1'(x) = \underset{\substack{x > \xi \\ x \to \xi}}{lim} f_2'(x) = \beta_1 + \beta_2 \xi + 3\beta_3 \xi^2$$

Similarly,

$$\begin{cases} f_1"(x) = 2\beta_2 + 6\beta_3 x \ \forall x \leq \xi \\ \\ f_2"(x) = 2\beta_2 + 6\beta_3 x + 6\beta_4 (x-\xi) \ \forall x > \xi \end{cases}$$

And,

$$\underset{\substack{x \leq \xi \\ x \to \xi}}{lim} f_1"(x) = \underset{\substack{x > \xi \\ x \to \xi}}{lim} f_2"(x) = 2\beta_2 + 6\beta_3 \xi$$

So, in this case, f' and f" are continuous for $x = \xi$.

If we had added to f a term $\beta(x-\xi)^i_+,\ i \in \{0,1,2\}$:

- for $i = 0$, f wouldn't be continous in $x = \xi$ because $\underset{\substack{x \leq \xi \\ x \to \xi}}{lim} f_1(x)$ and $\underset{\substack{x > \xi \\ x \to \xi}}{lim} f_1(x)$ would be separated by the constant β.

- for $i = 1$, f' wouldn't be continous in $x = \xi$ because $\underset{\substack{x \leq \xi \\ x \to \xi}}{lim} f_1'(x)$ and $\underset{\substack{x > \xi \\ x \to \xi}}{lim} f_1'(x)$ would be separated by the constant β.

- for $i = 2$, f wouldn't be continous in $x = \xi$ because $\underset{\substack{x \leq \xi \\ x \to \xi}}{lim} f_1"(x)$ and $\underset{\substack{x > \xi \\ x \to \xi}}{lim} f_1"(x)$ would be separated by the constant β.

Exercise 3 Generate the following graph using smoothing splines with $\lambda = 0.00022$ and $df = 12$.

Solution 3 R-code for generating the corresponding plot :

```
> plot(subset(bone, gender == 'female')$age,
+ subset(bone, gender == 'female')$spnbmd,
+ col='red', pch=3, xlab='Age', ylab='Relative change in spinal BMD')
> points(subset(bone, gender == 'male')$age,
 subset(bone, gender == 'male')$spnbmd,
  col='blue', pch=4)

> femaleSmooth = smooth.spline(subset(bone, gender == 'female')$age,
```

```
+ subset(bone, gender == 'female')$spnbmd, df=12)

> maleSmooth = smooth.spline(subset(bone, gender == 'male')$age,

+ subset(bone, gender == 'male')$spnbmd, df=12)


> lines(femaleSmooth, col='red', lwd=3)

> lines(maleSmooth, col='blue', lwd=3)

> legend(20, 0.125, c("Female", "Male"), lty=c(1,1),

+ lwd=c(2.5,2.5), col=c("red", "blue"))
```

The corresponding figure is in annex.

Exercise 4 Show that the matrix $\boldsymbol{S}_\lambda = \boldsymbol{N}(\boldsymbol{N}^\top \boldsymbol{N} + \lambda \boldsymbol{\Omega}_N)^{-1}\boldsymbol{N}^\top$ is symmetric, idempotent and its rank is equal to N (number of observations)

Solution 4 First of all, let's compute $\boldsymbol{S}_\lambda^\top$:

$$\begin{aligned}
\boldsymbol{S}_\lambda^\top &= N((\boldsymbol{N}^\top \boldsymbol{N} + \lambda \boldsymbol{\Omega}_N)^{-1})^\top) \\
&= \boldsymbol{N}((\boldsymbol{N}^\top \boldsymbol{N} + \lambda \boldsymbol{\Omega}_N)^\top)^{-1}) \\
&= \boldsymbol{N}(\boldsymbol{N}^\top \boldsymbol{N} + \lambda \boldsymbol{\Omega}_N^\top)^{-1}) \\
&= \boldsymbol{N}(\boldsymbol{N}^\top \boldsymbol{N} + \lambda \boldsymbol{\Omega}_N)^{-1}\boldsymbol{N}^\top \textit{ because } \boldsymbol{\Omega}_N \textit{ is symmetric}
\end{aligned} \tag{1}$$

We can conclude that :

$$\boxed{\boldsymbol{S}_\lambda \textit{ is symmetric}}$$

Secondly, it is clear that $\boldsymbol{S}_\lambda$ is inversible is equivalent to $\boldsymbol{N}$ inversible and $\boldsymbol{N}^\top \boldsymbol{N} + \lambda \boldsymbol{\Omega}_N)^{-1}$. The question assumes implicitely that $\boldsymbol{N}^\top \boldsymbol{N} + \lambda \boldsymbol{\Omega}_N)$ (and I would be very embarassed to prove it). Given that, we just have to prove that $\boldsymbol{N}$ is inversible to prove that $\boldsymbol{S}_\lambda$ is inversible.

It seems to be also a complicated task... The space spanned by the N cubic spline spline basis function is a N linear dimension space. By adding more constraints on these N cubic spline functions. By forcing the global resulting function to have fixed values at the knots, that is to say for a given real set $(\alpha_1, \alpha_2, \ldots, \alpha_n) \in \mathcal{R}^N)$ forcing f to respect the constraints $\forall k \in \{1, 2, \cdots, N\},\ f(x_k) = \sum_{i=1}^{N} \theta_i N_i(x_k) = \alpha_k$, we remove N degrees of freedom, that gives 0 degrees of freedom.

Thus, the linear application defined for a given set of $\mathcal{R}^N$ $(x_1, x_2, \ldots, x_n)$, $\forall f$ in the natural cubic spline space by $\phi(f) = (f(x_1), f(x_2), \ldots, f(x_n))$ is a bijection between the cubic spline space and $\mathcal{R}^N$.

Consequently, the columns vector of $\boldsymbol{N}$ are the images by the application ϕ of the function basis $(N_1, N_2, \cdots, N_N)$. The columns of the matrix form a basis of $\mathcal{R}^N$ are independent and :

$$\boxed{\boldsymbol{N} \textit{ is inversible and so is } \boldsymbol{S}_\lambda}$$

(I admit that the justification for ϕ being injective is not satisfactory..)

Exercise 5 Prove that $\boldsymbol{S}_\lambda$ can be written in form of $(\boldsymbol{I} + \lambda \boldsymbol{K})^{-1}$ where $\boldsymbol{K}$ does not depend on λ.

Solution 5 Knowing that if $\boldsymbol{A}$, $\boldsymbol{B}$ and $\boldsymbol{C}$ are inversible matrices, then $(\boldsymbol{ABC}) = \boldsymbol{C}^{-1}\boldsymbol{B}^{-1}\boldsymbol{A}^{-1}$.

We have :

$$\boxed{\begin{aligned} \boldsymbol{S}_\lambda &= ((\boldsymbol{N}^\top)^{-1}((\boldsymbol{N}^\top\boldsymbol{N} + \lambda\boldsymbol{\Omega}_N)(\boldsymbol{N}^{-1}))^{-1} \\ &= (\boldsymbol{I} + \lambda\boldsymbol{N}^{-\top}\boldsymbol{\Omega}_N\boldsymbol{N}^{-1})^{-1} \end{aligned}} \tag{2}$$

Exercise 6 Produce the following graphs using smoothing splines with the mentioned degrees of freedom.

Solution 6 R-code to generate the 2 plots :

```
> plot(airData$daggett.pressure.gradient, airData$ozone.level,
+ xlab='Daggett Pressure Gradient', ylab='Ozone Level', pch=19, col='gray')

> smoothFunction1 = smooth.spline(airData$daggett.pressure.gradient,
+ airData$ozone.level, df=5)
> smoothFunction2 = smooth.spline(airData$daggett.pressure.gradient,
+ airData$ozone.level, df=11)
> smoothFunction3 = smooth.spline(airData$daggett.pressure.gradient,
+ airData$ozone.level, df=17)

> lines(smoothFunction1, col='blue', lwd=3)
> lines(smoothFunction2, col='red', lwd=3)
```

```
> lines(smoothFunction3, col='yellow', lwd=3)

> legend(60, 28.5, c("df = 5", "df = 11", "df = 17"), lty=c(1,1,1),

+ lwd=c(2.5,2.5,2.5), col=c("blue", "red", "yellow"))
```