

# Estadística III

## Splines

Alejandro López Hernández

FES Acatlán - UNAM

May 22, 2020

① Introducción

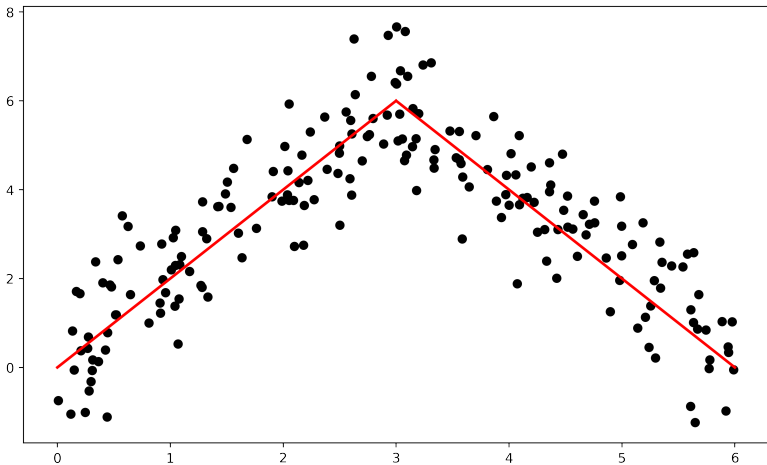
② 1-Spline

③ 2-Spline

④ p-Spline

Derivado de los métodos lineales de regresión nos enfocaremos en un método llamado **Splines Penalizados**, la idea es utilizar una base de funciones con las que podamos generar funciones continuas que suavicen nuestros datos.

La idea principal es poder ajustar datos con estructuras no lineales, si suponemos un modelo  $y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$  estamos limitados a solo ajustar datos lineales, con este modelo nuestra base de funciones es  $\{1, x\}$ , lo que se propone hacer con los 1-splines es dividir el rango la regresión en varios intervalos que de forma local si tengan un comportamiento lineal.



En este caso la regresión la obtuvimos con

$$f(x) = 0 + 2x - 4(x - 3)_+$$

donde  $u_+ = \max\{u, 0\}$ , de esta forma estamos extendiendo la base a  $\{1, x, (x - 3)_+\}$ , y nos da la facilidad de ajustarnos a estructuras mas complejas. En general, podemos hacer tantos cortes como deseamos, estos cortes son llamados **nudos**, si decidimos realizar  $K$  nudos nuestra base sera  $\{1, x, (x - \kappa_1)_+, (x - \kappa_2)_+, \dots, (x - \kappa_K)_+\}$ , por lo tanto hay que estimar  $K + 2$  parámetros  $\beta_0, \beta_1, b_1, \dots, b_K$ , y nuestra regresión es.

$$f(x) = \beta_0 + \beta_1 x + \sum_{k=1}^K b_k (x - \kappa_k)_+$$

Para expresar nuestro problema de forma lineal podemos escribir

$$\mathbf{X} = \begin{bmatrix} 1 & x_1 & (x_1 - \kappa_1)_+ & \cdots & (x_1 - \kappa_K)_+ \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & (x_n - \kappa_1)_+ & \cdots & (x_n - \kappa_K)_+ \end{bmatrix}$$

$$\boldsymbol{\beta} = [\beta_0 \quad \beta_1 \quad b_1 \quad \cdots \quad b_K]^T$$

$$\mathbf{y} = [y_1 \quad y_2 \quad y_3 \quad \cdots \quad y_n]^T$$

Por lo tanto nuestro problema se reduce a minimizar

$$\|\mathbf{y} - \mathbf{X}\boldsymbol{\beta}\|^2$$

Sin embargo, hay que restringir los coeficientes de los nudos con el objetivo de tener un control sobre la suavidad de la regresión. Esta restricción se solucionando agregando la restricción  $\beta^T D \beta \leq C$  donde:

$$D = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \cdots & 1 \end{bmatrix} = \begin{bmatrix} \mathbf{0}_{2 \times 2} & \mathbf{0}_{2 \times K} \\ \mathbf{0}_{K \times 2} & \mathbf{I}_{K \times K} \end{bmatrix} = \text{diag}(0, 0, 1, \dots, 1)$$

de esta manera nuestro problema se puede resolver utilizando un multiplicador de Lagrange:

$$\text{mínimizar } \|\mathbf{y} - \mathbf{X}\beta\|^2 + \lambda^2 \beta^T D \beta$$

para algún  $\lambda > 0$



Bueno, pues resulta que la solución es:

$$\hat{\beta}_{\lambda} = (\mathbf{X}^T \mathbf{X} + \lambda^2 \mathbf{D})^{-1} \mathbf{X}^T \mathbf{y}$$

De forma totalmente análoga que con los 1-Spline, la idea es ajustar curvas en ciertos intervalos, en este caso curvas cuadráticas. La ventaja de usar una curva cuadrática en lugar de lineales es que en los nudos tendremos funciones con derivadas (en lugar de picos). Nuestra base ahora será  $\{1, x, x^2, (x - \kappa_1)_+^2, (x - \kappa_2)_+^2, \dots, (x - \kappa_K)_+^2\}$  y nuestra función tendrá que ajustar  $K + 3$  parámetros

$$f(x) = \beta_0 + \beta_1 x + \beta_2 x^2 + \sum_{k=1}^K b_k (x - \kappa_k)_+^2$$

El problema resulta el mismo

$$\text{mínimizar } \|\mathbf{y} - \mathbf{X}\boldsymbol{\beta}\|^2 + \lambda^2 \boldsymbol{\beta}^T \mathbf{D} \boldsymbol{\beta}$$

pero con  $\mathbf{D} = \text{diag}(0, 0, 0, 1, \dots, 1)$ , y los parámetros extra a estimar. La solución resulta:

$$\hat{\boldsymbol{\beta}}_{\lambda} = (\mathbf{X}^T \mathbf{X} + \lambda^4 \mathbf{D})^{-1} \mathbf{X}^T \mathbf{y}$$

De forma mas general, se pueden definir bases que tengan derivadas de orden mas alto en los nudos. En general la base es de la forma

$$\{1, x, x^2, \dots, x^p, (x - \kappa_1)_+^p, (x - \kappa_2)_+^p, \dots, (x - \kappa_K)_+^p\}$$

por lo tanto nuestra regresión es de la forma

$$f(x) = \beta_0 + \beta_1 x + \beta_2 x^2 + \dots + \beta_p x^p + \sum_{k=1}^K b_k (x - \kappa_k)_+^p$$

El problema es totalmente análogo, solo que con más parámetros que estimar

$$\text{minimizar } \|\mathbf{y} - \mathbf{X}\boldsymbol{\beta}\|^2 + \lambda^2 \boldsymbol{\beta}^T \mathbf{D} \boldsymbol{\beta}$$

con  $\mathbf{D} = \text{diag}(\mathbf{0}_{p+1}, \mathbf{1}_K)$  La solución resulta:

$$\hat{\boldsymbol{\beta}}_\lambda = (\mathbf{X}^T \mathbf{X} + \lambda^{2p} \mathbf{D})^{-1} \mathbf{X}^T \mathbf{y}$$