

Índice general

1. Modelos aditivos generalizados	3
1.1. Introducción	3
1.2. Suavizado univariante	4
1.2.1. Bases de funciones	4
Bibliografía	5

Capítulo 1

Modelos aditivos generalizados

1.1. Introducción

Como bien podemos intuir por su nombre, los modelos aditivos generalizados no son más que la fusión entre los modelos lineales generalizados y los modelos aditivos, los cuales se introducen con una sección en este capítulo. Podemos ver estos dos tipos de modelos como extensiones del modelo lineal. Por un lado, como vimos en el capítulo anterior, el MLG hace uso de una función de enlace entre el predictor lineal y el valor esperado de la variable dependiente para poder expresar relaciones más complejas y relaja la hipótesis distribucional permitiendo que tal variable siga distribuciones de la familia exponencial. Por otro lado, los modelos aditivos, además de también relajar esta hipótesis de distribución, introducen las funciones de suavizado en el modelo, estas proporcionan más flexibilidad a la hora de relacionar las variables explicativas con la de respuesta.

Luego, como ya hemos mencionado y como se plantean Hastie y Tibshirani (1986,1990), el MAG reúne estas dos propuestas de forma que este tipo de modelo generaliza el modelo aditivo de la misma forma que el MLG generalizaba el modelo lineal. Sin embargo, la flexibilidad que proporciona este modelo da lugar a dos nuevos problemas teóricos: cómo estimar las funciones de suavizado y cómo de “suaves” deben ser.

Comenzaremos este capítulo viendo cómo construir los modelos aditivos generalizados, es decir, qué bases de funciones podemos elegir para obtener las funciones de suavizado y qué parámetro de suavizado se debe seleccionar o cómo se puede estimar. Luego se introduce el modelo aditivo, en el que se utilizarán los resultados vistos a lo largo del capítulo. Tras todo ello se propone la forma final del modelo aditivo generalizado.

Definición 1.1.1 (Estructura básica del modelo aditivo generalizado).

$$\mu = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_n \end{pmatrix} = \begin{pmatrix} E[Y_1] \\ \vdots \\ E[Y_n] \end{pmatrix} = E[Y]$$

$$g(\mu_i) = A_i\theta + f_1(x_{1i}) + f_2(x_{2i}) + f_3(x_{3i}, x_{4i}) + \dots, \quad \forall i = 1, \dots, n \quad (1.1)$$

Donde:

- Y_i es la variable de respuesta y sigue una distribución de la familia exponencial de media μ_i y parámetro de escalado ϕ . A partir de ahora esto lo denotaremos por: $Y_i \sim EF(\mu_i, \phi)$.
- A_i es la fila i -ésima de la matriz del modelo para aquellas componentes del modelo que son estrictamente paramétricas.
- θ es el correspondiente vector de parámetro, que antes denotábamos por β , para las variables predictoras mencionadas en el anterior punto.
- Las f_i son las funciones de suavizado para las covariables x_k .

1.2. Suavizado univariante

Dicho esto, partiremos considerando modelos que, aunque no sean adecuados para un uso práctico general, nos permitirán estudiar el marco teórico de una forma más sencilla. Es decir, en esta sección consideraremos un modelo con una sola función de suavizado, f , y una sola covariable, x , de la forma:

$$y_i = f(x_i) + \epsilon_i \quad (1.2)$$

Donde y_i es la variable de respuesta y los ϵ_i son variables aleatorias independientes e idénticamente distribuidas como $N(0, \sigma^2)$ que representan el error.

1.2.1. Bases de funciones

Nos proponemos en esta sección obtener una estimación de la función de suavizado a partir de una base de un espacio de funciones, en el que también se encontrará f (o una aproximación suya). Elegir una base equivale a tomar un conjunto de funciones $\{b_j(x)\}_{j=1}^k$ y, por tanto, podemos representar la función de suavizado como:

$$f(x) = \sum_{j=1}^k b_j(x) \beta_j \quad (1.3)$$

Para ciertos parámetros β_j a determinar.

Base polinómica

Si consideramos la base \mathcal{B} del espacio de polinomios de grado k , es decir, $\mathcal{B} = \{1, x_i, x_i^2, \dots, x_i^k\}$, la función de suavizado toma la forma:

$$f(x) = \beta_1 + \beta_2 x + \beta_3 x^2 + \dots + \beta_{k+1} x^k$$

Y, por tanto, el modelo 1.2 queda:

$$y_i = \beta_1 + \beta_2 x_i + \beta_3 x_i^2 + \dots + \beta_{k+1} x_i^k + \epsilon_i$$

Observación 1.2.1 (Problema de la base polinómica). Como indica Wood, 2017, p.162, por el teorema de Taylor, la base polinomial nos será útil cuando nuestro interés sea el de estudiar las propiedades de la función de suavizado en el entorno de un punto concreto, pero nos encontramos con problemas cuando queremos hacerlo en todo el dominio de f .

El principal problema se debe a que la interpolación de los datos puede resultar en una función muy oscilante o que no ajuste bien la información, dependiendo del valor de k . Esto se puede solucionar de cierta manera con el siguiente tipo de base de funciones.

Base lineal por parte

Consideremos ahora una partición de nodos $\{x_j^* : j = 1, \dots, k\}$ del rango de la variable predictora x tal que $x_j^* > x_{j+1}^*$ y la base de funciones $\mathcal{B} = \{b_j(x)\}_{j=1}^k$ donde:

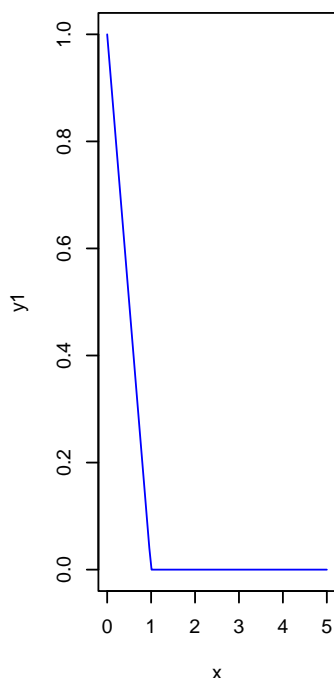
$$b_1(x) = \begin{cases} \frac{x_2^* - x}{x_2^* - x_1^*} & , \text{ si } x < x_2^* \\ 0 & \text{c.c.} \end{cases}$$

$$b_j(x) = \begin{cases} \frac{x - x_{j-1}^*}{x_j^* - x_{j-1}^*} & , \text{ si } x_{j-1}^* < x < x_j^* \\ \frac{x_{j+1}^* - x}{x_{j+1}^* - x_j^*} & , \text{ si } x_j^* < x < x_{j+1}^* \\ 0 & \text{c.c.} \end{cases}$$

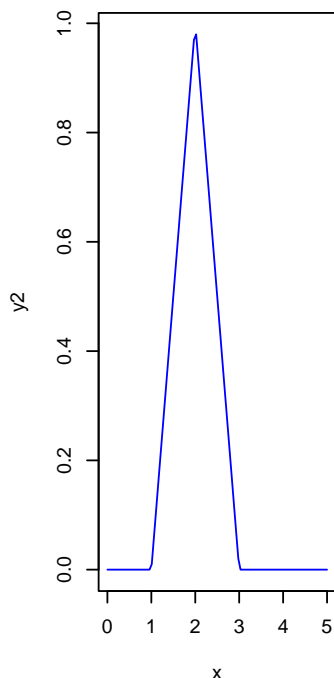
$$b_k(x) = \begin{cases} \frac{x - x_{k-1}^*}{x_k^* - x_{k-1}^*} & , \text{ si } x > x_{k-1}^* \\ 0 & \text{c.c.} \end{cases}$$

Ejemplo 1.2.1. Supongamos que el rango de x es de 0 a 5 y consideremos 6 nodos: $\{0, 1, 2, 3, 4, 5\}$, entonces podemos representar las funciones $b_0(x)$, $b_2(x)$ y $b_5(x)$ como:

Función Lineal por Partes b1()



Función Lineal por Partes b2()



Función Lineal por Partes b5()

