Boletín de Estadística e Investigación Operativa
Vol. 25, No. 3, Octubre 2009, pp. 195-205

Estadística

An introduction to smoothing with penalties: P-splines

María Durbán
Departamento de Estadística
Universidad Carlos III de Madrid
✉ marialuz.durban@uc3m.es

Abstract

Many of the new areas of scientific research yield large amount of complex data (microarray, images, etc.) that can not be analysed using standard parametric models. Smoothing techniques provied a flexible framework for modelling these type of data. In this paper we focus on a recent and very atractive approach in the context of smoothing methods: penalized splines.

Keywords: Smoothing, Penalty, P-splines
AMS Subject classifications: 62G08, 62J07, 62J12.

1. Introducción

La mayoría de los lectores se habrá encontrado con el término *smoothing* en repetidas ocasiones, sea cual sea su área de investigación en Estadística. Dos son los factores que han hecho que esta técnica haya alcanzado tanta popularidad en los últimos años: 1) la creciente complejidad de los datos con los que trabajamos, y que dejan obsoletos a los modelos paramétricos tradicionales, y 2) los avances informáticos que han facilitado el ajuste de este tipos de modelos, reduciendo significativamente el coste computacional.

Supongamos que disponemos de pares de datos (x_i, y_i), un modelo suavizado (para datos normales) vendrá dado por:

$$y_i = f(x_i) + \epsilon_i\,, \quad \epsilon_i \sim N(0, \sigma^2)\,, \tag{1.1}$$

donde $f(.)$ es una función **suave** de los datos (siempre hay que partir de la premisa de que la variable respuesta es una función suave de la variable explicativa. Un modelo de suavizado no debe ser utilizado si hay clara evidencia de que esta premisa no se cumple), a la que no se le impone ninguna forma. Es evidente,

que un modelo de este tipo es una generalización de un modelo de regresión, que indudablemente, tendrá un coste computacional, pero que nos permitirá estimar la función de una forma más precisa. La estimación de la función $f(.)$ se puede realizar mediante distintos métodos divididos en dos grandes grupos: regresión tipo *kernel* y regresión con *splines*.
Los modelos tipo **kernel** se basan en la idea de que al estimar la función en un punto x_0, es deseable dar más peso a las observaciones que están próximas a ese punto. Los pesos son asignados mediante una *función kernel* e irán disminuyendo conforme nos vayamos alejando del punto x_0. Hay muchos tipos de funciones kernel (Gaussianos, tricúbicos, etc.) pero todas ellas dependen de un parámetro llamado *bandwidth* que controla los pesos en función de la distancia al punto x_0. En este artículo no se va a profundizar en este tipo de métodos, por lo que los lectores interesados en estas técnicas pueden encontrar un buen resumen en Hardle (1990), y Chu y Marron (1992) entre otros.
El segundo grupo de técnicas de suavizado están basados en **splines** que son polinomios a trozos que se unen en puntos llamados *nodos*. Hay dos grandes familias dentro de los modelos de suavizado con splines:

1. **Splines de regresión** (regression splines). En estos modelos es necesario seleccionar el número y la localización de los nodos (para controlar la suavidad de la función ajustada) e imponer restricciones para que los trozos de polinomio se unan de forma suave. Una vez hecha la elección, el modelo se ajusta por mínimos cuadrados.

2. **Splines de suavizado** (smoothing splines). Aparecen como la solución al siguiene problema de regresión no-paramétrica: encontrar la función (con dos derivadas continuas) que minimiza la *suma de cuadrados penalizada*:

$$SCP = \sum_{i=1}^{n} (y_i - f(x_i))^2 + \lambda \int_x \left(f^{''}(x)\right)^2 dx\,,$$

   donde el último término es una penalización en la segunda derivada de la curva y λ es el *parámetro de suavizado* que controla la suavidad de la misma, de modo que si $\lambda = 0$ estaremos interpolando los datos, y si $\lambda \to \infty$ tendremos un ajuste lineal (más detalles en Green y Silverman (1994)).

Sin embargo, ambas técnicas presentan inconvenientes: en los splines de regresión, la suavidad de la función ajustada depende de la elección de los nodos, y ha de hacerse, en general, mediante complicados algoritmos que son difíciles de extender al caso multidimensional. En el caso de los splines de suavizado, los problemas son de tipo computacional, ya que este tipo de splines, utilizan tantos nodos (y por lo tanto parámetros) como observaciones.
Los **splines con penalizaciones** o **P-splines** (Eilers y Marx (1996)) combinan

lo mejor de ambos enfoques: utilizan menos parámetros que los splines de suavizado, pero la selección de los nodos no es tan determinante como en los splines de regresión. Son splines de rango bajo, el número de nodos es mucho menor que la dimensión de los datos, al contrario de lo que ocurre en el caso de los splines de suavizado. El número de nodos, en el caso de los P-splines, no supera los 40, lo que hace que sean computacionalmente eficientes, sobre todo cuando se trabaja con gran cantidad de datos. Además, la introducción de penalizaciones relaja la importancia de la elección del número y la localización de los nodos, cuestión que es de gran importancia en los splines de rango bajo sin penalizaciones (ver por ejemplo Rice y Wu (2001)). Por último, la correspondencia entre los P-splines y el BLUP (Best Linear Unbiased Predictor) en un modelo mixto permite, en algunos casos, utilizar la metodología existente en el campo de los modelos mixtos y el uso de software como `PROC MIXED` en `SAS`, y `lme()` en `S-PLUS` y `R`.

2. Bases y penalizaciones

La metodología utilizada en los P-splines se puede resumir de la siguiente forma: (a) utilizar una base para la regresión, y (b) modificar la función de verosimilitud introduciendo una penalización basada en diferencias entre coeficientes adyacentes. En el caso de datos normalmente distribuidos como en (1.1), consideramos un modelo de regresión $\boldsymbol{y} = \boldsymbol{B}\boldsymbol{a} + \boldsymbol{\epsilon}$, y una base $\boldsymbol{B} = \boldsymbol{B}(\boldsymbol{x})$, construida a partir de $\boldsymbol{x}$. Para estimar los coeficientes de regresión se minimiza la función de mínimos cuadrados penalizados:

$$S(\boldsymbol{a}; \boldsymbol{y}, \lambda) = (\boldsymbol{y} - \boldsymbol{B}\boldsymbol{a})'(\boldsymbol{y} - \boldsymbol{B}\boldsymbol{a}) + \lambda \boldsymbol{a}'\boldsymbol{P}\boldsymbol{a}\,,$$

donde $\boldsymbol{P}$ es una matriz que penaliza los coeficientes de forma suave y λ es el parámetro de suavizado. Fijado un valor de λ, los valores ajustados vienen dados por,

$$\hat{\boldsymbol{y}} = \boldsymbol{B}(\boldsymbol{B}'\boldsymbol{B} + \lambda \boldsymbol{D}'\boldsymbol{D})^{-1}\boldsymbol{B}'\boldsymbol{y} = \boldsymbol{H}\boldsymbol{y}.$$

$\boldsymbol{H}$ no es una matriz de proyección, ya que no es idempotente, pero su forma hace que el método de suavizado sea lineal. La traza de $\boldsymbol{H}$ corresponde a la dimensión del modelo (el número equivalente de parámetros que estaríamos estimando). El papel del parámetro de suavizado en los P-splines, es el mismo que tiene en cualquier otro método de suavizado: controlar la suavidad de la curva, pero aquí lo que hace es penalizar los coeficientes que están muy separados entre sí, y cuanto mayor sea λ, más se aproximarán los coeficientes a cero. Entre los criterios para elegir λ se encuentran AIC (criterio de información de Akaike, Akaike (1973)), GCV (validación cruzada generalizada, Green y Silverman (1994)) y BIC (criterio de información Bayesiano, Schwarz (1978)).
Las diferencias entre los distintos modelos de P-splines vienen dadas por la elección de la base $\boldsymbol{B}$ y de la penalización $\boldsymbol{P}$.

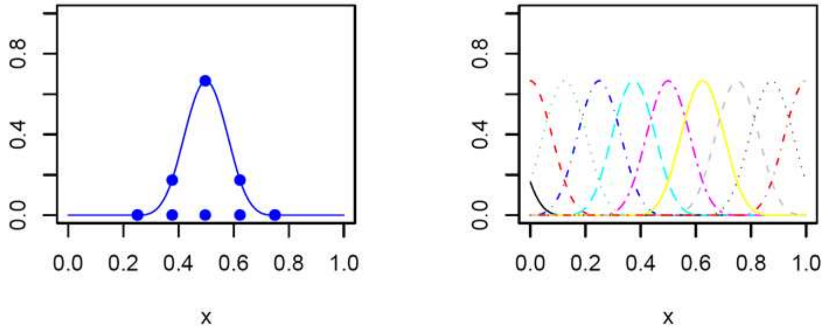

Figura 1: Base de B-splines de orden 3

2.1. Bases y nodos

Hay muchas maneras de calcular la base para la regresión: polinomios truncados (Ruppert et. al (2003)), thin plate splines (Green y Silverman (1994), Wood (2003)), etc. Nosotros recomendamos el uso de B-splines (De Boore (1977) y Dierckx (1993)), ya que son numéricamente más estables que otras bases (como es el caso de los polinomios truncados). Un ejemplo muy simple aparece en la Figura 1: un B-spline de grados 3 está formado por 4 trozos de polinomios unidos entre si que se unen en nodos. Todas las funciones de la base tienen la misma forma, pero están desplazadas horizontalmente (el desplazamiento es una función de la distancia entre los nodos). En general un B-spline de grado p:

- Consiste en $p + 1$ trozos de polinomio de orden p.
- Se unen en p nodos internos.
- En los puntos de unión las derivadas hasta el orden $p - 1$ son continuas.
- El B-spline es positivo en el dominio expandido por $p + 2$ nodos y 0 en el resto.
- Excepto en los extremos, se solapa con $2p$ trozos de polinomios de sus vecinos.
- Para cada valor de $\boldsymbol{x}$, $p + 1$ B-splines son no nulos.

La selección y localización de los nodos no está hecha de antemano (como ocurre en el caso de los smoothing splines). Es suficiente con elegir un número moderadamente grande de nodos equidistantes. Autores como Ruppert et al. (2003) aconsejan elegir los K nodos en los K-quantiles de $\boldsymbol{x}$, es decir, cada nodo t_k sería el cuantil $k/(K + 1)$ de $\boldsymbol{x}$. En cuanto al número de nodos, la mayoría de los autores utilizan como regla:

$$\text{número de nodos} = min\{40, \text{valores únicos de } \boldsymbol{x}/4\}.$$

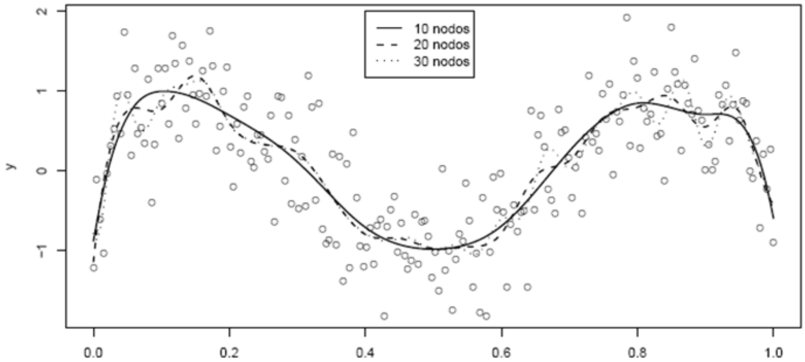

Figura 2: Curva estimada con 3, 10 y 20 nodos.

2.2. Penalizaciones

Supongamos que tenemos una base $\boldsymbol{B}$ construida con k nodos. Si utilizamos mínimos cuadrados para ajustar el modelo, la función objetivo será:

$$S(\boldsymbol{a};\boldsymbol{y}) = (\boldsymbol{y}-\boldsymbol{B}\boldsymbol{a})'(\boldsymbol{y}-\boldsymbol{B}\boldsymbol{a}) \Rightarrow \hat{\boldsymbol{a}} = (\boldsymbol{B}'\boldsymbol{B})^{-1}\boldsymbol{B}'\boldsymbol{y}\,,$$

y la curva ajustada $\hat{f}(x) = \boldsymbol{B}\hat{\boldsymbol{a}}$ dependerá del tamaño de la base.

La Figura 2 muestra el efecto que tiene el tamaño de la base en la curva ajustada. Cuanto mayor sea la base menos suave es la curva, de modo que cuando el número de nodos coincide con el número de datos obtenemos una curva que interpola los datos.

Para evitar este problema, O'Sullivan (1986) introdujo una penalización en la segunda derivada de la curva, de modo que la función objetivo pasó a ser:

$$S(\boldsymbol{a};\boldsymbol{y},\lambda) = (\boldsymbol{y}-\boldsymbol{B}\boldsymbol{a})'(\boldsymbol{y}-\boldsymbol{B}\boldsymbol{a}) + \lambda\int_{\boldsymbol{x}}\left(\boldsymbol{B}''\boldsymbol{a}\right)^2 d\boldsymbol{x}\,.$$

Este tipo de penalización es bastante común (es la que se utiliza en los splines de suavizado), sin embargo, no hay nada de particular en la segunda derivada, se puede utilizar derivadas de cualquier orden. La novedad que introducen los P-splines es que la penalización es discreta y que se penalizan los coeficientes directamente, en vez de penalizar la curva, lo que reduce la dimensionalidad del problema.

El tipo de penalización está estrechamente ligado al tipo de base que se utilice, si se utilizan polinomios truncados, la penalización es *"ridge"* independientemente del grado de los polinomios truncados, es decir:

$$S(\boldsymbol{a};\boldsymbol{y},\lambda) = (\boldsymbol{y}-\boldsymbol{B}\boldsymbol{a})'(\boldsymbol{y}-\boldsymbol{B}\boldsymbol{a}) + \lambda \boldsymbol{a}'\boldsymbol{a}\,,$$

esto equivale a poner una penalización en la derivada $p+1$ de la curva.
Por el contrario, Eilers y Marx (1996) utilizan una penalización basada en la diferencia de orden d entre los coeficientes adyacentes de la bases de B-splines, este tipo de penalización es más flexible ya que es independiente del grado del polinomio utilizado para construir los B-splines. Además, es una buena aproximación discreta a la integral de la d-ésima derivada al cuadrado:

$$S(\boldsymbol{a};\boldsymbol{y},\lambda) = (\boldsymbol{y}-\boldsymbol{B}\boldsymbol{a})'(\boldsymbol{y}-\boldsymbol{B}\boldsymbol{a}) + \lambda \boldsymbol{a}'\boldsymbol{P}_d\boldsymbol{a} \Rightarrow \hat{\boldsymbol{a}} = (\boldsymbol{B}'\boldsymbol{B} + \lambda\boldsymbol{P}_d)^{-1}\boldsymbol{B}'\boldsymbol{y}\,,$$

donde $\boldsymbol{P}_d = (\Delta^d)'\Delta^d$. En general se utiliza $d=2$, aunque se puede utilizar órdenes superiores o inferiores, dependiendo de la variabilidad de la curva y de la cantidad de ruido en los datos. Por ejemplo, una penalización de orden $d=2$ equivale a

$$(a_1 - 2a_2 + a_3)^2 + \ldots + (a_{k-2} - 2a_{k-1} + a_k)^2 = \boldsymbol{a}'\boldsymbol{D}'\boldsymbol{D}\boldsymbol{a}\,,$$

donde

$$\boldsymbol{D} = \begin{bmatrix} 1 & -2 & 1 & 0 & \ldots \\ 0 & 1 & -2 & 1 & \ldots \\ 0 & 0 & 1 & -2 & \ldots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}.$$

La Figura 3 muestra el efecto de la penalización: forzar a los coeficientes a seguir un patrón suave. Se muestra el ajuste de una curva mediante B-splines sin y con penalización, junto con las funciones que forman la base (las columnas de la matriz $\boldsymbol{B}$) multiplicadas por los coeficientes, así como los coeficientes (en un círculo). En la parte izquierda se aprecia como el patrón errático de los coeficientes da lugar a una curva poco suave, en cambio en la parte derecha, cuando se impone que se pase de un coeficiente a otro de forma suave, la curva también lo es.

Entre las propiedades de los P-splines con bases de B-splines hay que destacar que no "padecen" los efectos de frontera comunes en otros métodos de suavizado, como algunos kernels, en los que al extender la curva ajustada fuera del dominio de los datos ésta tiende hacia cero. Además se conservan los momentos, es decir, la media y la varianza de los valores ajustados será la misma que la de los datos sea cual sea el parámetro de suavizado, al contrario que los kernels que tienden a aumentar la varianza cuanto mayor es el suavizado.

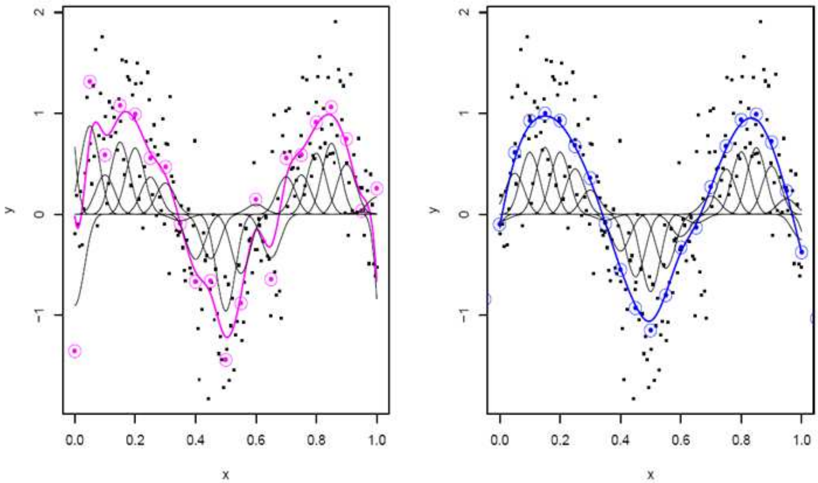

Figura 3: Curva estimada con 20 nodos en la base, sin penalizar los coeficientes (izquierda) y penalizando los coeficientes (derecha).

3. P-splines como modelos mixtos

La gran revolución de los P-splines ocurrida en los últimos años es debida a la posibilidad de escribir un modelo de suavizado donde se utilizan P-splines como un modelo mixto (o modelo con efectos aleatorios), la ventaja que tiene el utilizar este enfoque es que nos permite: por un lado utilizar toda la metodología desarrollada para los modelos mixtos, y por otro, utilizar el software para modelos mixtos que está disponible en la mayoría de los programas estadísticos.

La conexión entre regresión no paramétrica y los modelos mixtos se remonta a principio de los 90 (Speed (1991)). Más tarde, otros autores como Verbyla et al. (1999), desarrollan en más profundidad el tema del suavizado con modelos mixtos (en el contexto de los splines cúbicos) y más tarde, Wand (2003), en el contexto de los P-splines con polinomios truncados como bases. Sin embargo, no se había intentado buscar una representación de los P-splines con B-splines como bases como modelos mixtos, lo cual, entre otras cosas daría lugar a bases más estables (Currie y Durbán (2002)).

El interés por la representación de los P-splines como modelos mixtos surgió de las dificultades que aparecen en el campo de los modelos aditivos de suavizado, ya que hay problemas de identificabilidad del modelo. El uso de los P-splines hace que no sea necesario un método iterativo (como el *backfitting algorithm*) para la estimación de las curvas, pero era entonces necesario modificar las bases de forma que una curva se pueda descomponer como suma de un componente polinómico (del mismo orden que la penalización) y otro no polinómico, es decir, dado el modelo:

$$\begin{aligned} \boldsymbol{y} &= \boldsymbol{Ba} + \boldsymbol{\epsilon} \quad \boldsymbol{\epsilon} \sim N(0, \sigma^2 \boldsymbol{I}), \quad \text{se reformula como} \\ \boldsymbol{y} &= \boldsymbol{X\beta} + \boldsymbol{Zu} + \boldsymbol{\epsilon} \quad \boldsymbol{\epsilon} \sim N(0, \sigma^2 \boldsymbol{I}). \end{aligned}$$

Según la base que haya utilizado para los P-splines, $\boldsymbol{X}$ y $\boldsymbol{Z}$ tendrán una forma diferente. En el caso de los B-splines:

$$\boldsymbol{X} = [\mathbf{1} : \boldsymbol{x}] \quad \text{y} \quad \mathbf{Z} = \boldsymbol{BU}\boldsymbol{\Sigma}^{-\mathbf{1/2}},$$

donde $\boldsymbol{U}$ y Σ son matrices que forman parte de la descomposición en valores singulares de la matriz de penalización $\boldsymbol{D}'\boldsymbol{D}$.

Una vez descrita esta nueva base, es inmediato establecer la conexión con un modelo mixto, donde:

$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{Zu} + \boldsymbol{\epsilon}, \quad \boldsymbol{u} \sim N(0, \sigma_u^2 \boldsymbol{I}_{c-2}), \quad \boldsymbol{\epsilon} \sim N(0, \sigma^2 \boldsymbol{I}),$$

siendo c el número de columnas de la base original $\boldsymbol{B}$, y el parámetro de suavizado es $\lambda = \sigma^2/\sigma_u^2$, lo cual hace que sea posible estimar el parámetro de suavizado junto con los otros parámetros del modelo. En el contexto de los modelos mixtos, el método estándar para la estimación de los parámetros de la varianza es el método de máxima verosimilitud restringida (REML) cuya expresión viene dada por:

$$\begin{aligned} \ell_R(\sigma_u^2, \sigma_\varepsilon^2) &= -\frac{1}{2}\log|\boldsymbol{V}| - \frac{1}{2}\log|\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X}| \\ &\quad -\frac{1}{2}\boldsymbol{y}'(\boldsymbol{V}^{-1} - \boldsymbol{V}^{-1}\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{V}^{-1})\boldsymbol{y}, \end{aligned}$$

donde $\boldsymbol{V} = \sigma_u^2 \boldsymbol{ZZ}' + \sigma^2 \boldsymbol{I}$. El vector de parámetros $\boldsymbol{\beta}$ y el vector de coeficientes aleatorios $\boldsymbol{u}$ son estimados como:

$$\begin{aligned} \hat{\boldsymbol{\beta}} &= (\boldsymbol{X}'\hat{\boldsymbol{V}}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}'\hat{\boldsymbol{V}}^{-1}\boldsymbol{y}, \\ \hat{\boldsymbol{u}} &= \hat{\sigma}_u^2 \boldsymbol{Z}'\hat{\boldsymbol{V}}^{-1}(\boldsymbol{y} - \boldsymbol{X}\hat{\boldsymbol{\beta}}). \end{aligned}$$

4. Aplicaciones y extensiones

El hecho de que un Pspline se pueda escribir como un modelo mixto nos proporciona ventajas en dos ámbitos distintos: por un lado hace que se pueda flexibilizar la hipótesis de linealidad en infinidad de modelos, y por otro, es posible incluir estructuras complejas en los modelos usuales de suavizado.

Por ejemplo, en el caso de **datos correlados**, si intentamos ajustar una curva sin tener en cuenta la correlación que hay en los datos, veremos que los métodos de selección del parámetro de suavizado van a elegir un valor del parámetro menor al que corresponde y por lo tanto la curva no va a ser suave. Si quisiéramos estimar tanto el parámetro de suavizado como los parámetros que determinan la estructura de correlación, necesitaríamos hacerlo de forma iterativa, lo que haría que el resultado final fuera muy sensible a la elección de los parámetros iniciales. Sin embargo, con los P-splines como modelos mixtos es inmediato el introducir una estructura de correlación y estimarla simultáneamente a la curva suave (Currie y Durbán (2002)).

Otro ámbito en el que los P-splines están tomando un papel relevante es en el del análisis de **datos longitudinales**, tan frecuentes en aplicaciones médicas y biológicas. En general, este tipo de datos corresponden a medidas tomadas a varios individuos en distintos instantes de tiempo, y se caracterizan por la dependencia que hay entre las medidas repetidas hechas a un mismo individuo. Los modelos mixtos más comunes para datos longitudinales representan a cada individuo como la suma de la media de la población (que varía con el tiempo) y que se modela como un efecto fijo, y un polinomio de grado bajo (generalmente una línea) en la que los coeficientes son aleatorios y sirven para modelar la variabilidad individual. Sin embargo, estos modelos pueden no ser apropiados en algunas situaciones, por ejemplo, cuando las trayectorias individuales sean una función no-lineal del tiempo. Los P-splines permiten ajustar modelos más flexibles en los que las diferencias específicas individuales son una función suave del tiempo (Durbán et al. (2005)).

Las nuevas áreas de investigación científica que se han abierto en los últimos años, han dado lugar a datos con una compleja estructura multidimensional, en particular datos de tipo espacial y espacio temporal (en áreas como, medio ambiente, meteorología, demografía, etc.). Los P-splines tienen una extensión natural al **caso multidimensional** mediante el producto tensorial o de kronecker de las bases unidimensionales, y una penalización no-isotrópica que permite tener una cantidad de suavizado distinta en cada una de las dimensiones (Currie et al. (2006)). Estos modelos también tienen una representación como modelos mixtos (Lee y Durbán (2009)).

Por último, todas las ideas expuestas en las secciones anteriores se pueden aplicar al caso de **modelos lineales generalizados con penalizaciones**. Estos modelos fueron introducidos por Marx y Eilers (1998) en el caso unidimensional, y por Currie et al. (2004) en el multidimensional. Si los modelos con penalizaciones admiten una representación como modelos mixtos, la extensión al caso generalizado son los modelos lineales mixtos generalizados (ver por ejemplo Lee y Durbán (2009)).

Estas son sólo unas cuantas de las muchas aplicaciones que tienen los modelos de suavizado con penalizaciones. Es un área de investigación que está en pleno apogeo en la actualidad y que está traspasando rápidamente el ámbito de la Estadística para abrirse camino en otras áreas de conocimiento.

Referencias

[1] Akaike, H. (1973). Maximum likelihood identification of Gaussian autorregresive moving average models. *Biometrika*, **60**, 255-265.

[2] Chu, C.-K. y Marron, J.S. (1992). Choosing a kernel regression estimator (with discussion). *Stat. Sci.*, **6**, 404-436.

[3] Currie, I. y Durbán, M. (2002). Flexible smoothing with Psplines: a unified approach. *Stat. Model.*, **2**, 333-349.

[4] Currie, I., Durbán, M. y Eilers, P. (2004). Smoothing and forecasting mortality rates. *Stat. Model.*, **4**, 279-298.

[5] Currie, I., Durbán, M. y Eilers, P. (2006). Generalized array models with applications to multidimensional smoothing. *J. R. Stat. Soc., B*, **68**, 259-280.

[6] De Boore, C. (1977). Package for calculating B-splines. *J. Numer. Anal.*, **14**, 441-472.

[7] Dierckx, P. (1993). *Curve and Surface Fitting with Splines*, Clarendon, Oxford (UK).

[8] Durbán, M., Harezlak, J., Wand, M. P. y Carroll, R. J. (2005). Simple fitting of subject-specific curves for longitudinal data. *Stat. Med.*, **24**, 1153-1167.

[9] Eilers P. y Marx B. (1996). Flexible smoothing with B-splines and penalties. *Stat. Sci.*, **11**, 89-121.

[10] Green, P.J. y Silverman B.W. (1994). *Nonparametric Regression and Generalized Linear Models*, Chapman & Hall, London (UK).

[11] Hardle, W. (1990). *Applied Nonparametric Regression*, Camdridge University Press, Cambridge (UK).

[12] Lee, D-J. y Durbán M. (2009). Smooth-CAR mixed models for spatial count data. *Comput. Stat. Dat. Anal.*, **53**, 2968-2979.

[13] Marx, B. y Eilers, P. (1998). Direct generalized additive modelling with penalized likelihood. *Comput. Stat. Dat. Anal.*, **28**, 193-209.

[14] O'Sullivan, F. (1986). A statitical perspective on ill-posed inverse problems. *Stat. Sci.*, **1**, 505-527.

[15] Rice, J.A. y Wu, C.O. (2001). Nonparametric mixed effects models for unequally sample noisy curves. *Biometrics*, **57**, 253-259.

[16] Ruppert, D., Wand, M.P. and Carroll, R.J. (2003). *Semiparametric Regression*, Cambridge University Press, Cambridge (UK).

[17] Speed, T. (1991). Comment on “blup is a good think: estimation of random effects”, by Robindon, G.K. *Stat. Sci.*, **6**, 15-51.

[18] Schwarz, G.E. (1978). Estimating the dimension of a model *Ann. Stat.*, **6**, 461-464.

[19] Verbyla, A., Cullis, B., Kenward, M. y Welham, S. J. (1999). The analysis of designed experiments and longitudinal data usinf smoothing splines. *Appl. Stat.*, **48**, 269-312.

[20] Wand, M.P. (2003). Smoothing and mixed models. *Comput. Stat.*, **18**, 223-249

[21] Wood, S. (2003). Thin plate regression splines. *J. R. Stat. Soc., B*, **65**, 95-114.

Acerca de la autora

María Durbán es profesora titular en el Departamento de Estadística de la Universidad Carlos III de Madrid. Es miembro electo del Council de la International Biometric Society, y Fellow de la Royal Statistical Society. Tiene numerosas artículos en el tema de P-splines y ha impartido cursos en esta materia. Sus líneas de investigación actuales se centran en los modelos de suavizado con P-splines, en particular, en el desarrollo de modelos para datos multidimensionales y sus aplicaciones en el campo de la Bioestadística.