

# Anexo I

## Suavizado con splines

Este tipo de suavizado se sustenta en la aproximación que se suele hacer de funciones polinómicas complejas a través de *splines*. Este tipo de técnicas dividen la función a aproximar en intervalos para los cuales se ajusta individualmente una función del orden deseado, dependiendo del grado de ajuste que queramos.

Para aplicar esta técnica a nuestra serie temporal primero dividiremos el tiempo  $t$  en  $k$  intervalos iguales:

$$[t_1 = 1, t_2], [t_2 + 1, t_3], \dots, [t_{k-1} + 1, t_k = n] \quad (1)$$

En los que los valores  $t_1, t_2, \dots, t_k$  reciben el nombre de nudos. Hecho esto a cada intervalo se le ajusta una función de regresión dependiente del tiempo de la forma:

$$\beta_0 + \beta_1 t + \dots + \beta_k t^p \quad (2)$$

Normalmente con un  $p = 3$  suele ser suficiente para captar el comportamiento de la serie. Hasta ahora lo único que hemos hecho ha sido realizar un suavizado un poco más elaborado a través de funciones polinómicas. El suavizado con *splines* va un poco más allá e intenta encontrar la relación óptima entre el ajuste y el grado del spline dada por:

$$\sum_{t=1}^n (Y_t - f_t)^2 + \lambda \int (f_t'')^2 dt \quad (3)$$

Donde  $f_t$  es un *spline* cúbico con nudos en todo  $t$ . Si  $\lambda > 0$  entonces el suavizado que se logra es mayor eliminando la gran parte de fluctuaciones propias de la serie. El ajuste de este parámetro determinará el grado de ajuste de la serie suavizada a la original.

## Suavizado con kernel

El suavizado a través de kernels es una técnica utilizada para suavizar series temporales muy influenciadas por la aleatoriedad, es decir, con mucho ruido. Es una técnica no paramétrica que nos permite una muy buena suavización de series complejas debido a su facilidad de ajuste a procesos no lineales.

Este tipo de técnicas suavizan la serie en un punto  $t$  a partir de observaciones pasadas y futuras dentro de un entorno delimitado por el ancho de banda  $b$ . Este parámetro determinará el grado de suavizado de la serie, ya que si nuestro entorno es amplio se tendrán en cuenta más observaciones haciendo que el suavizado sea mayor. Si el ancho de banda es pequeño las fluctuaciones de la serie tendrán una mayor presencia en la suavización.

La estimación de  $Y_t$  es una media ponderada de los valores pasados y futuros dentro del entorno seleccionado. Esta ponderación se realiza a través de una función kernel  $K$ :

$$Y'_t = \frac{\sum_{i=1}^n K\left(\frac{t-i}{b}\right) Y_i}{\sum_{j=1}^n K\left(\frac{t-j}{b}\right)} \quad (4)$$

siendo  $1, \dots, n$  las observaciones del entorno generado por  $b$ . Existen una gran variedad de funciones kernel aunque quizás las que más se suele usar es la Gaussiana:

$$K(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2} \quad (5)$$