The Enting spline

December 5, 2022

This splining method is motivated by the trade-off between staying close to the data (small deviation of the spline from the original data, e.g., small sum of squared errors "S") and obtaining a smooth spline (little curvature, e.g., small integrated second derivative of the spline Q). As such, it minimizes

$$\Theta(\lambda) = S + \lambda \cdot Q, \tag{1}$$

with $\lambda > 0$ controlling whether good data fit (λ small) or smoothness (λ large) is prioritized. Enting [Enting, 1987] showed that a spline function that minimizes Eq. (1) behaves like a low-pass filter, removing periodicities below a cut of cut-off period $\mathrm{T}_{0.5}$ (corresponding to higher frequencies) from the signal. The signal at the cut-off period is attenuated by 50 %. He showed that $\mathrm{T}_{0.5}$ is related to λ by

$$\mathrm{T}_{0.5} = 2\pi\,(\lambda\Delta t)^{1/4} \quad \text{or} \quad \lambda = (\mathrm{T}_{0.5}/2\pi)^4\,/\Delta t. \tag{2}$$

Δt denotes the temporal resolution of the data (the spacing of the x values $\Delta t = x_{i+1} - x_i$). For annual mean temporal resolution and $\mathbf{x}$ given in units of years, $\Delta t = 1$y. Eq. (2) allows to determined λ based on a desired cut-off frequency.

So far, the spline function to minimize Eq. (1) has not been specified. Enting there relied on an approach by de Boor [1978] (Chapter XIV therein, p. 207ff) where a separate third-order polynomial function p_i is fitted to each interval between two adjacent data points in the input data: $[x_i, x_{i+1})$ for the x values (*years*) and $[y_i, y_{i+1})$ for the y values (*global mean temperature*);

$$p_i(x) = a_i + b_i \cdot (x - x_i) + c_i \cdot (x - x_i)^2 + d_i \cdot (x - x_i)^3. \tag{3}$$

It is then required that the overall spline function over the full interval $[x_1, x_n]$ is smooth in the break points (the x_i's where the polynomials are knit together), by requiring continuity of the values as well as the first and second derivatives in the break points,

$$p_{i-1}(x_i) = p_i(x_i) \quad p'_{i-1}(x_i) = p'_i(x_i) \quad \text{and} \quad p''_{i-1}(x_i) = p''_i(x_i). \tag{4}$$

Furthermore, it is required that the second derivative vanishes at the beginning (first year of input data) and the end of the time series (last year of input data). Based on these boundary conditions and Eq. (1), one obtains a system of equations for the coefficients in Eq.(3) that can be solved such that Eq. (1) is minimized. This way, one obtains an analytical spline that can be evaluated at any x value in principal. We are, however, only interested in the values of the spline at the original x_i to obtain the spline for the annual average global-mean temperature.

Bibliography

Carl de Boor. *A Practical Guide to Splines*. Springer New York, NY, 1978.

I. G. Enting. On the use of smoothing splines to filter co2 data. *Journal of Geophysical Research: Atmospheres*, 92(D9):10977–10984, 1987. doi: https://doi.org/10.1029/JD092iD09p10977.