Whittaker Smoother
Neal B. Gallagher

Key words: Data fitting, smoothing penalty

Introduction: Data are often fit to a curve such as a low order polynomial or other basis function. However, at times the fitting functions are too restrictive and don’t provide a good representation of the original data. A more flexible approach was developed by Eilers[1] and is further modified to provide additional flexibility as demonstrated by the 
www.eigenvector.com

The Whittaker Smoother: Eiler’s paper[1] introduces the following objective function

\[ O(z) = (y - z)^T W_0 (y - z) + \lambda_s z^T D_s^T D_s z \]

where \( y \) is a \( N \times 1 \) vector of measured data, \( z \) is smooth curve to be fit to the data, \( W_0 \) is a diagonal matrix of weights (typically \( 0 \leq w_{0,n} \leq 1 \) for \( n = 1, \ldots, N \), \( D_s \) is a second derivative operator (e.g., \( D_s z \) is the second derivative of \( z \)) and \( \lambda_s \) is a scalar penalty on the smoothing term. When data are missing, the corresponding weight, \( w_{0,n} \), can be set to zero. Once that \( W_0 \) and \( \lambda_s \) are given (set by default or provided by the user) the corresponding estimate of \( z \) is given by

\[ \hat{z} = \left( W_0 + \lambda_s D_s^T D_s \right)^{-1} W_0 y. \]

For example, an optical emission (OES) spectrum is plotted Figure 1 along with two smoothed versions shown for \( W_0 = I \) and \( \lambda_s = 0.1 \) (low smoothing) and \( \lambda_s = 10 \) (stronger smoothing). The optical emission spectra are available at www.eigenvector.com and discussed in References [3,4]. The stronger smoothing (green curve) appears to suppress the noise better than the weak smoothing but it also suppresses the peaks more. With a slight modification to the objective function, the best of both worlds can be obtained.

The Modified Whittaker Smoother: The original smoother can be modified to allow for different smoothness weighting on each of the channels using

\[ O(z) = (y - z)^T W_0 (y - z) + \lambda_s z^T D_s^T W_s D_s z \]

where \( W_s \) is a diagonal matrix of weights with entries \( 0 \leq w_{s,n} \leq 1 \) and corresponding estimator given by

\[ \hat{z} = \left( W_0 + \lambda_s D_s^T W_s D_s \right)^{-1} W_0 y. \]

For the example, lowering the weights in \( W_s \) from 1 to 0.1 for the wavelengths (\( \lambda \)) with peaks corresponding to \( 258 < \lambda < 280 \), \( 307 < \lambda < 311 \) and \( 393 < \lambda < 397 \), the smoother gives the final smoothed spectrum in Figure 2 (the black curve). The smooth black curve follows the green curve outside the peak ranges and the red curve with the peak ranges.

![Figure 1: Uncorrected OES spectrum (blue), smoothed spectrum \( \lambda_s = 0.1 \) (red) and strongly smoothed spectrum \( \lambda_s = 10 \) (green). Zoom in on a small peak (middle) and significant sharp peaks (bottom).](image)
shown in Figure 3 for the original data and the three different smoothing approaches. The eigenvalue distribution for principal components (PCs) ≥6 are mostly attributable to noise. A proxy for S/N is the ratio of the sum of the eigenvalues 1 to 5 to the sum of eigenvalues ≥6 shown in the figure. The original spectra and the low smoothing appear to have the lowest S/N while the higher smoothing has the biggest S/N. Interestingly, relaxing the smoothing on the peaks appears to lower the S/N slightly compared to including smoothing over all the wavelengths.

Figure 2: Uncorrected OES spectrum (blue), smoothed spectrum $\lambda_s = 0.1$ (red) and strongly smoothed spectrum $\lambda_s = 10$ (green). Strong smoothing except for the peaks (black). Zoom in on a small peak (middle) and significant sharp peaks (bottom).

Conclusions: Smoothing is a useful tool for providing interpretable trends given by the smoothed signal and has the potential to improve signal-to-noise.

References:


