SMOOTHING SEASONALLY ADJUSTED TIME SERIES

Estela Bee Dagum and Alessandra Luati, University of Bologna

Estela Bee Dagum, Department of Statistics, via Belle Arti 41, Bologna, Italy

KEY WORDS Nonparametric Smoothers, Fixed Bandwidth, Loess, Cubic Smoothing Spline, Gaussian Kernel, Henderson Filters.

Abstract

This paper deals with the properties of several nonparametric estimators in the context of seasonal adjustment. The smoothers discussed are Loess, the Cubic smoothing spline and the Gaussian kernel, all constrained to a fix span of 13 terms for comparison with the Henderson filter widely applied in time series decomposition. Because of this constraint, the smoothers statistical properties are affected, and no longer optimal as when their respective smoothing parameters are optimally estimated according to data-driven automated selection methods.

We perform the comparison by means of spectral techniques, deriving the symmetric and asymmetric weights of each smoother and calculating the corresponding gains and phase shift functions. These latter provide information on the type of "signal" passed and "noise" suppressed by each smoother.

1. Introduction

The presence of high levels of variability in seasonally adjusted data, often encountered in recent years, poses the serious problem of properly identifying the underlying short-term trend. One solution has been the calculation of variances of seasonally adjusted series (see, e.g. Wolter and Monsour, 1981; Burridge and Wallis, 1985; Hillmer, 1985; Pfeffermann, 1994). Since the dominant seasonal adjustment method adopted by statistical agencies is not model based, the variance calculation has proven to be a great challenge; still not fully solved. Similarly, for model based seasonal adjustment procedures, the fact that models and/or parameter values change with new observations, limits the validity of the results for most recent observations which are the most important.

Another alternative to the identification of the underlying short-term trend is that of smoothing the seasonally adjusted data as suggested by Dagum (1987). Traditionally, the most often applied trend-cycle predictor for monthly data is the 13-term Henderson filter (H13). This filter, however, has the two major limitations of producing: (1) a large number of 9-10 month cycles (unwanted ripples) leading to false turning points, and (2) large revisions when new observations are added to the series.

In this paper, we investigate if other nonparametric smoothers, subject to length restriction, can improve on H13 and, thus, offer an alternative to the nonparametric nonlinear method developed by Dagum (1996) that corrects for the two H13 shortcomings.

The smoothers considered are based on two classes of weight generating functions, local polynomials and probability distributions. We consider, within the first class, the locally weighted regression smoother (loess) of degree 2 (L2) and the cubic smoothing spline (CSS); and in the second class, the Gaussian kernel (GK). These estimators depend on a smoothing parameter that determines the degree of smoothness of the output. Traditionally, the smoothing parameter is estimated on the basis of procedures that minimize a loss function, e.g. the mean square error or prediction risk. Frequently, when the smoothing parameter is optimally estimated as a function of the characteristics of the input data, the span of the filter is rather large (see, e.g., Dagum and Capitanio, 1998, and Dagum and Luati, 2000). Therefore, a large number of end points can be estimated only with asymmetric weights; a serious limitation for seasonal adjustment. Asymmetric weights introduce revisions in the more recently estimated values as new observations are added and phaseshifts at cyclical turning points. To reduce the undesirable effects of a large number of asymmetric weights, we impose to each smoother the constraint of a fixed 13-term length in agreement with the H13 filter that will be used as a our "benchmark". With this a priori constraint, all the smoothers become linear, and consequently, no longer data dependent.

Section 2 introduces the definition and a brief description of each nonparametric smoother. Section 3 gives the symmetric and asymmetric (last data point) weights derived for each constrained smoother. Section 4 calculates their frequency response functions for symmetric and asymmetric weights and discusses their smoothing properties. Finally, section 5 gives the conclusions.

2. Nonparametric Estimators

The nonparametric estimators discussed are based on different assumptions of smoother building. The locally weighted regression smoother known as loess, fits local polynomials of a degree d where the parameters are estimated by ordinary or weighted least squares. Hence, it satisfies the property of best fit to the data. The cubic smoothing spline searches for an optimal solution between both fitting and smoothing of the data under the assumption that the signal locally follows a second degree polynomial. The Gaussian kernel is a locally weighted average where the weighting function follows the Gaussian standard probability distribution. Finally, the Henderson smoothing filter , derived from the graduation theory, minimizes smoothing with respect to a third degree polynomial within the span of the filter.

Next, we give a brief description of each nonparametric smoother and refer the reader to Dagum and Luati (2000) for more details.

(I) The locally weighted regression smoother that we use in this study, known in the current literature as loess, is the one developed by Cleveland (1979) originally called lowess, (LOcally WEighted Scatterplot Smoother). Loess is based on nearest neighbors weights and applied iteratively, which makes it a robust procedure. Given a series of equally spaced observations and corresponding target points

$$\{(y_j, t_j), j = 1, \ldots, N\}, \quad t_1 \leq \ldots \leq t_N$$

where t_j is the time the observation y_j is taken, loess produces a smoothed estimate $\widehat{y}_j$, as follows,

$$\widehat{y}_j = \mathbf{t}_j^T \widehat{\beta}_j$$

where $\mathbf{t}_j$ is the $(d+1)$-dimensional vector of generic component $t_j^p, p = 0, \ldots, d$; $d = 0, 1, 2, \ldots$ denotes the degree of the fitting polynomial, and $\widehat{\beta}_j$ is the $(d+1)$-dimensional least squares estimate of a weighted regression computed over a neighborhood of t_j constituting a subset of the full span of the series.

The weights depend on the distance between the target point t_j^* and any other point belonging to its neighborhood. Each neighborhood is made of the same number of points chosen to be nearest to t_j^*, and the ratio between the amplitude of the neighborhood, n, and the full span of the series, N, defines the smoothing parameter. It is sensible to choose an odd value for n in order to allow symmetric neighborhoods, at least for central observations.

Concerning the degree of the fitting polynomial, $d = 1$ or $d = 2$ are usually appropriate choices. The highest degree is more appropriate when the plot of the observations against the target points presents many maxima and minima, since the flexibility of a quadratic curve best fits highly noisy time series, therefore, in this study we use loess with $d = 2$. As far as it concerns the weighting function, we use the one based on the tricube function

$$W(x) = \left(1 - |x|^3\right)^3 I_{\{[0,1[\}}(x)$$

whose quasi-semicircular shape allows about 45% of the points belonging to any symmetric neighborhood to have considerable weight (greater than 0.8), the remaining 55% having weights decreasing to zero quite slowly.

(II) Kernel type smoothers are locally weighted averages. A kernel smoothing gives, at time $t_h^*, 1 \leq h \leq N$, the smoothed estimate

$$\widehat{y}_h = \sum_{j=1}^{N} w_{hj} y_j$$

where

$$w_{hj} = \frac{K_b\left(\frac{t_h^* - t_j}{b}\right)}{\sum_{i=1}^{N} K_b\left(\frac{t_h^* - t_i}{b}\right)}$$

are the weights from a parametric kernel (a nonnegative function that integrated over its domain gives unity), $b > 0$ denotes the smoothing parameter, and $K_b(x) = K_b(-x)$.

In this study, we consider the standard Gaussian kernel function given by

$$K_b\left(\frac{t_h^* - t_j}{b}\right) = \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}\left(\frac{t_h^* - t_j}{b}\right)^2\right\}.$$

(III) A spline function is the unique solution of the following optimization problem (Whittaker and Robinson, 1924; Schoenberg, 1964)

$$f_\lambda = \min_{f_\lambda \in \mathcal{C}^s} \frac{1}{N} \sum_{t=1} [y_j - f(t_j)]^2 + \lambda \int_a^b \left[f^{(s)}(u) \right]^2 du \tag{2.1}$$

where $\mathcal{C}^s$ is the class of functions with s continuous derivatives and $\lambda > 0$. The solution of (2.1) is a univariate piecewise polynomial of degree $2s-1$ with the pieces joined at the 'knots' $t_1, \cdots, t_N$. For $s = 2$ we have the cubic smoothing spline, that is

$$\min_{f_\lambda \in \mathcal{C}^2} \frac{1}{N} \sum_{t=1} [y_j - f(t_j)]^2 + \lambda \int_a^b \left[f''(u) \right]^2 du \tag{2.2}$$

where the smoothing parameter λ balances the trade off between the fit to the data and the smoothness of the estimates. The fitting is guaranteed by the least squares condition on the left hand of the (2.2), whereas the smoothing condition is imposed by balancing the curvature of the function on the definition domain with the smoothing parameter λ.

It follows from (2.2) that if $\lambda \to 0$, then the solution f tends to the univariate natural polynomial spline which interpolates the data, while if $\lambda \to \infty$, the solution tends to the second degree polynomial best fitting the data in the least squares sense.

The cubic smoothing spline estimates are given by

$$\hat{y}_j = f(t_j)$$

with $f(t_j)$ being a smooth function on the interval $[a, b]$, $a \leq t_1 \leq \ldots \leq t_N \leq b$.

(IV) The Henderson smoothing filters often applied for trend estimation in time series decomposition are based on summation formulae mainly used by actuaries. The basic principle for a summation formula is the combination of operations of differencing and summation in such a manner that, when differencing above a certain order is ignored, they will reproduce the functions operated on.

If the length of the filter is $2k-3$, Henderson (1916) showed that the general expression for the j -th term of the filter that minimizes $\sum_j \left(\Delta^3 w_j\right)^2$ is $w_j = \frac{315\left[(k-1)^2-j^2\right]\left(k^2-j^2\right)\left[(k+1)^2-j^2\right]\left(3k^2-16-11j^2\right)}{8k(k^2-1)(4k^2-1)(4k^2-9)(4k^2-25)}$. Making $k = 8$ and j from -6 to 6 we obtain a symmetric set of 13 weights.

Kenny and Durbin (1982) and Gray and Thomson (1996) proved that the symmetric weights w_j of the Henderson smoothing filters can also be obtained as the solution of a polynomial in j of degree 8 for the 13-term filter, subject to the unbiasedness conditions,

$$\sum_{j=-m}^{m} w_j = 1, \qquad \sum_{j=-m}^{m} j w_j = 0$$

and $\sum_{j=-m}^{m} j^2 w_j = 0$ and to the smoothness restriction S given by,

$$S = \sigma_u^2 \sum_{j=-m}^{m+3} \left(\Delta^3 w_j\right)^2$$

where the w_js satisfy the boundary condition of $w_j = 0$ for $j = \pm(m+1), \pm(m+2), \pm(m+3)$.

On the contrary, the weights of the asymmetric Henderson filters available in software such as $CENSUSX11$, $X11ARIMA$, and $X12ARIMA$, were developed by Musgrave (1964) on the basis of the minimization of the mean squared revision between the final estimates (obtained by the application of the symmetric filter) and the preliminary estimates (obtained by the application of an asymmetric filter) subject to the constraint that the sum of the weights is equal to one. The assumption made is that the most recent values of the series (where seasonality has been removed, if present in the original observations) follow a linear trend plus an erratic component ε_t such that $\varepsilon_t \sim N\left(0, \sigma_\varepsilon^2\right)$ (see Laniel, 1985, and Doherty, 1992).

3. Symmetric and Asymmetric Weights of Fixed 13-term Length Smoothers

Setting the length of the smoothers equal to 13, we derive the following symmetric and asymmetric filters to be applied to central and last available observations, respectively.

L2	-0.015	-0.036	-0.004	0.074	0.157	0.210	0.227*
CSS	0.001	-0.001	-0.010	-0.023	0.022	0.250	0.522*
GK	0.001	0.007	0.023	0.060	0.121	0.183	0.210*
H13	-0.019	-0.028	0.000	0.066	0.147	0.214	0.240*

Table 1. Symmetric weights of fixed 13-term smoothers (*central weight).

L2	-0.099	-0.085	-0.029	0.067	0.201	0.371	0.574*
CSS	0.002	0.002	-0.008	-0.037	-0.038	0.201	0.879*
GK	0.002	0.011	0.038	0.100	0.199	0.302	0.347*
H13	-0.092	-0.058	0.012	0.120	0.244	0.353	0.421*

Table 2. Asymmetric weights of fixed 13-term smoothers (*end point weight).

It should be noted that the CSS weights reported here are different from those given in Dagum and Capitanio (1999) where the value of $\lambda = 1.0$ was selected to approximate the H13 and led originally to a number of weights equal to 31 which were truncated to 13. We obtained the value of $\lambda = 0.2$ as the one that corresponds to a span of 13 weights.

Figures 1 and 2 show the symmetric and asymmetric weights, respectively, for the various smoothers. It is apparent that the symmetric weights of CSS have the smallest dispersion around the central value and that those of L2 and H13 are very similar one another. On the other hand, the asymmetric last-point weights are quite different for all the filters.

4. Gains and Phaseshifts of the Symmetric and Asymmetric Smoothers Filters

We study the smoothing properties of the fixed 13-term nonparametric estimators by means of the traditional spectral analysis approach, and, thus, calculate their frequency response functions $H(\omega)$. The latter is decomposed into gain and phaseshift defined by $H(\omega) = G(\omega)\, e^{-i\phi(\omega)}$ where $G(\omega)$ is the gain function of the filter and $\phi(\omega)$ is the phaseshift. The phaseshift is given in months by $\phi^*(\omega) = \frac{\phi(\omega)}{2\pi\omega}$. The formula $G(\omega) = \sqrt{A^2(\omega) + B^2(\omega)}$, where $A(\omega) = \sum_j w_j \cos(\omega j)$ and $B(\omega) = \sum_j w_j \sin(\omega j)$, has been used to compute and plot the gain functions of the symmetric and asymmetric filters and $\phi^*(\omega) = \frac{1}{2\pi\omega} \arctan \left\{ \frac{B(\omega)}{A(\omega)} \right\}$ has been used for the phaseshifts of the last point asymmetric filters.

Figure 3 shows the gains of the symmetric filters for central observations. In the context of smoothing seasonally adjusted data it is useful to divide the total range of $\omega \in [0, 0.50]$ in two major intervals: (1) $0 \le \omega \le 0.06$, associated with cycles of 16 months or longer attributed to the "signal" (trend-cycle) of the seasonally adjusted series, and (2) the frequency band $0.10 \le \omega \le 0.50$ corresponding to short cyclical fluctuations attributed to the "noise". In this latter interval, it is of great interest to see how much of the power is not suppressed at $\omega = 0.10$, corresponding to 10-month cycles, known as "unwanted ripples", which can be wrongly interpreted as true turning points. An optimal smoothing filter should have a gain as close as possible to one for $0 \le \omega \le 0.06$ and near to zero for $0.10 \le \omega \le 0.50$. It should be noted that most seasonal adjustment methods will suppress almost all the power already present in the frequency band around the fundamental seasonal frequency, i.e. $0.6 < \omega < 0.10$.

For the symmetric filters, the largest reduction of noise is produced by GK which possesses the good property of suppressing a large amount of power at $\omega = 0.10$ but has the disadvantage of also suppressing power associated to the trend-cycle. On the contrary, CSS leaves untouched the signal but passes a lot of noise. In fact, restricting CSS to be of a rather short length has destroyed its optimal smoothing property. Another important observation is that the restricted loess of degree 2 gives almost identical results to those of the classical H13. Both filters have the good properties of passing almost all the signal without modification, and suppressing a large amount of noise. But both also have the disadvantage of leaving too much power at $\omega = 0.10$ and, therefore, will produce a large number of unwanted ripples in the output.

Figure 4 exhibits the gains of the asymmetric filters for the last observation. It is apparent that all the asymmetric filters pass a much larger amount of noise than the symmetric ones.

The L2 gain is now shifted up by almost a constant amount with respect to H13, simultaneously introducing a large amplification of the power attributed to the trend and suppressing less noise. CSS shows a good performance for the signal frequency band but at the expense of a very small noise reduction. The GK gain is close to that of H13 in the noise frequency band, but very different in the signal part where GK is shown to suppress too much power.

Figure 5 shows the various phaseshifts introduced by the asymmetric filters. Only CSS has practically no phaseshift all along the frequency range. It can also be seen that L2 produces almost half the phaseshift introduced by H13 at the signal frequency band, which is the most critical given the large amount of power passed by both filters.

5. Conclusions

We analyzed the properties of several non-parametric smoothers, restricted to fixed 13-term length, for the smoothing of seasonally adjusted series. Our benchmark was the traditional 13-term Henderson filter. We consider the locally weighted regression smoother (loess) of degree 2 (L2), the cubic smoothing spline (CSS) and the Gaussian kernel (GK). We derived their symmetric and asymmetric (end point) sets of weights filters and studied their properties by means of their corresponding gains and phaseshift functions.

The constraint of a 13-term span destroyed the optimal smoothing properties of the CSS that passed the largest amount of noise for both symmetric and asymmetric filters.

The L2 gain closely approximated that of H13 and, thus, it shared the disadvantage of passing a large amount of power at $\omega = 0.10$ implying a large number of unwanted ripples (short cycles of 10-month periodicity). Finally, the results from the restricted GK indicated an oversmoothing for it suppressed not only most of the noise power but also that of the signal.

Concerning the fixed-length asymmetric filters for the last point, they all passed much more noise than the symmetric ones. The asymmetric CSS filter was the only one not introducing phaseshift (a measure of bias for points of maxima and minima) all along the frequency range but it had the serious limitation, shared by its symmetric counterpart, of leaving all the noise almost unchanged.

The constrained asymmetric L2 no longer approximated closely the H13 filter, showing both power amplification of the signal and less phaseshift. Finally, the GK last point asymmetric filter introduced the largest amount of phaseshift.

In summary, none of the 13-term linearized smoothers improve on the limitations of H13.

In view of these results, we intend to linearize the Dagum filter (1996) and study its spectral properties.

Acknowledgment: Financing from MURST is gratefully acknowledged.

REFERENCES

Burridge, P. And Wallis, K.F. (1985): "Calculating the Variance of Seasonally Adjusted Series", JASA, 541-552.

Cleveland, W.S. (1979): "Robust Locally Weighted Regression and Smoothing Scatterplots", JASA, 74, 829-836.

Dagum, E.B. (1987): "Current Issues on Seasonal Adjustment", Survey Methodology , 13, n.1, 63-74.

Dagum. E. B. (1996): "A New Method to Reduce Unwanted Ripples and Revisons in Trend-cycle Estimates from X11ARIMA", Survey Methodology, 22, n.1, 77-83.

Dagum, E.B. and Capitanio, A. (1998): "Smoothing Methods for Short-Term Trend Analysis: Cubic Splines and Henderson Filters", Statistica, anno LVIII, n.1, 5-24.

Dagum, E.B. and Capitanio, A. (1999): "Cubic Spline Spectral Properties for Short Term Trend-cycle Prediction". Proceedings of the Business and Economic Statistics Section of the A.S.A, 100-106.

Dagum, E. B. and Luati, A. (2000): "Predictive Performance of Some Linear and Nonlinear Smoothers for Noisy Data". Statistica, anno LX, 4, 635-654.

Doherty, M. (1992): "Surrogate Henderson Filters in X-11". Technical Report, NZ Department of Statistics, Wellington, New Zealand.

Gray, A. and Thomson, P. (1996): "Design of Moving-Average Trend Filters Using Fidelity and Smoothness Criteria" in Time Series Analysis in Memory of E.J. Hannan (P.M. Robinson and M. Rosenblatt Editors), Vol. II, Springer Lecture Notes in Statistics, 115, 205-219.

Henderson, R. (1916): "Note on Graduation by Adjusted Average". Transaction of the Actuarial Society of America, 17, 43-48.

Hillmer, S.C. (1985): "Measures of Variability for Model-based Seasonal Adjustment Methods", JBES, 3,60-68.

Kenny, P.B. and Durbin, J. (1982): "Local Trend Estimation and Seasonal Adjustment of Economic and Social Time Series". JRSS A, 145, 1-41.

Laniel, N. (1985): "Design Criteria for the 13-Term Henderson End-Weights", Working Paper, Methodology Branch, Statistics Canada, Ottawa.

Musgrave, J. (1964): "A Set of End Weights to End all End Weights". Working Paper, U. S. Bureau of Census, Washington D.C.

Pfeffermann, D. (1994): "A General method for Estimating the Variances of X-11 Esti-

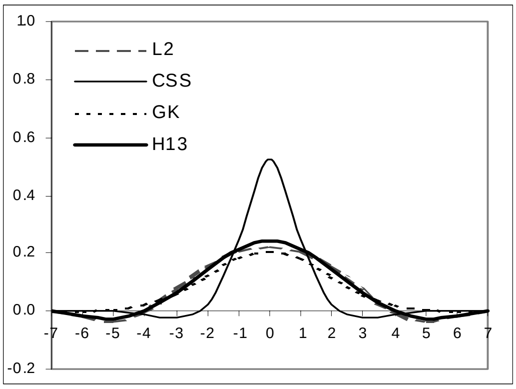

Figure 1: Symmetric weights of fixed 13-term smoothers for central observations.

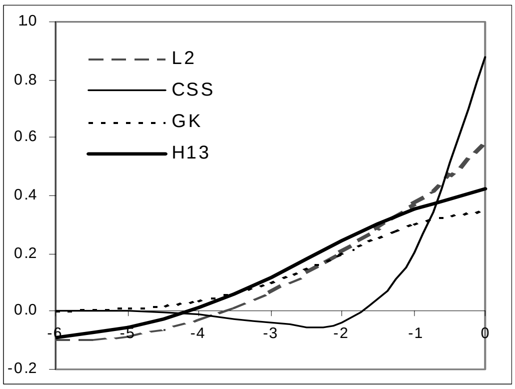

Figure 2: Asymmetric weights of fixed 13-term smoothers for the last data point.

mators". Journal of Time Series Analysis, 15, 85-116.

Schoenberg, I.J. (1964): "Spline Functions and the Problems of Graduation". Proceedings of the National Academy of Sciences, Usa, 52, 947-950.

Walter, K.M. and Monsour, N.J. (1981): "On the Problem of Variance Estimation for a Deseasonalized Series" in Current Topics in Survey Sampling: An Internatonal Symposium, (D.Krewski, R. Platek and J.N.K. Rao Editors), New York :Academic Press Inc., 367-403.

Whittaker E.T. and Robinson G. (1924): Calculus of Observations: a Treasure on Numerical Calculations, Blackie and Son, London.

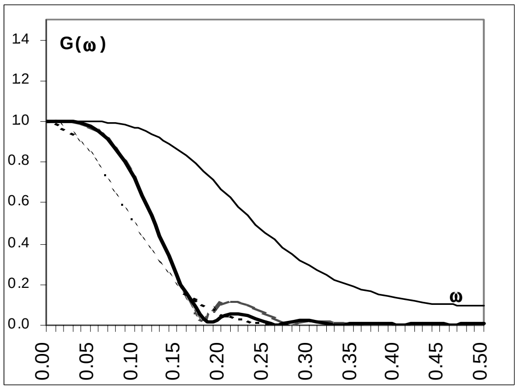

Figure 3: Gain functions of the symmetric filters for central observations.

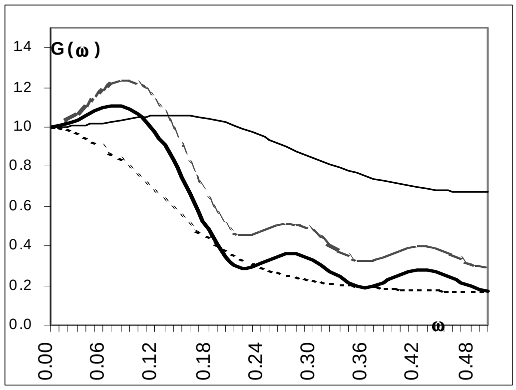

Figure 4: Gain functions of the asymmetric filters for the last data point.

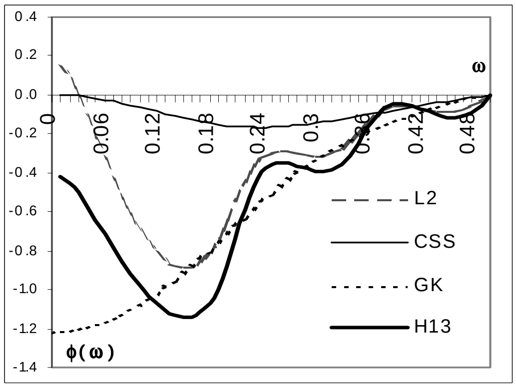

Figure 5: Phaseshifts of the asymmetric filters for the last data point.