

A Symmetric Linear Filter for Non Stationary Mean Prediction of Seasonally Adjusted Time Series

Un Filtro Lineare Simmetrico per la Stima della Media Non Stazionaria di Serie Storiche Destagionalizzate

Estela Bee Dagum, Alessandra Luati

Dipartimento di Scienze Statistiche

Università di Bologna

beedagum@stat.unibo.it; luati@stat.unibo.it

Riassunto: Un filtro simmetrico per la stima del trend-ciclo di serie storiche destagionalizzate viene introdotto ed analizzato. Il filtro consiste in un'approssimazione lineare del filtro non lineare proposto da Dagum (1996) per ovviare al problema dei falsi punti di svolta introdotti dal filtro di Henderson a 13 termini, in un contesto di stima non parametrica. Lo studio delle proprietà teoriche del nuovo filtro viene effettuato attraverso l'analisi spettrale e misure teoriche locali di distorsione, varianza ed errore quadratico medio. Il filtro viene poi applicato a serie storiche reali e le proprietà di fitting e smoothing vengono confrontate con quelle di altri stimatori quali loess, spline cubica e kernel gaussiano.

Keywords: non parametric estimators, smoothing parameters, local bias and variance.

1. Introduction

The problem of identifying the direction of the short term trend (non stationary mean) of seasonally adjusted series contaminated by high levels of variability has become of relevant interest in recent years. In fact, major financial and economic changes of global character have introduced a large amount of noise in time series data, particularly, in socioeconomic indicators used for the analysis of current economic conditions. Traditionally such indicators are seasonally adjusted. Alternative solutions to this problem consisted in calculating variances of seasonally adjusted series (see Cleveland, 2002, and references therein) and in smoothing the seasonally adjusted data through non parametric estimators, as proposed by Dagum (1987) and Kenny and Durbin (1982). The most often applied smoother by officially adopted seasonal adjustment methods is the 13-term trend Henderson filter (H13) that has the excellent property of a fast detection of true turning points. However, H13 has the limitations of (a) producing large revisions to the most recent estimates when new observations are added to the series and (b) introducing a large number of 10-month cycles (unwanted ripples) in the final trend-cycle curve. The former limitation concerns the asymmetric filter for the last observations whereas the latter is due to the symmetric filter applied to central observations. Such limitations were overcome by a *nonlinear* nonparametric predictor, developed by Dagum (1996): a modified version of the H13, consisting of (a) extending the seasonally adjusted series with *ARIMA* extrapolations, and then (b) applying a strict replacement of extreme values.

The aim of this paper is to provide a linear approximation of the symmetric part of the non linear Dagum estimator which offers a solution to the unwanted ripples problem. The main advantages of a linear approximation are easiness of application and feasibility

of studying the filter statistical properties independently of the data characteristics. Furthermore, a linear transformation to the original series does not destroy the relationship among the seasonally adjusted variables.

The symmetric linear filter then obtained is analysed here in terms of its theoretical properties of fitting and smoothing as well as applying it to a large sample of real time series, representative of different degrees of variability. The results are evaluated in terms of empirical fitting and smoothing measures and compared with those of other widely applied non parametric estimators such as loess of degree 1 and 2 (L1, L2), the Gaussian kernel (GK) and the Cubic smoothing spline (CSS), all constrained to be linear by a priori fixing their length equal to 13 terms as that of our benchmark, H13.

2. The Short Cascade Linear Filter (SCLF)

To obtain a full linear approximation of the non linear Dagum filter it is necessary to (a) select some parsimonious *ARIMA* models with a priori fixed parameter values and (b) define a method for a linear replacement of the extreme values. Since our aim here is to develop a linear approximation of only the symmetric Dagum filter, we do not consider the *ARIMA* extrapolations, but we concentrate on the linear replacement of extreme values. Hence, we selected a 5- and a 7-term moving averages applied sequentially.

Given a time series $\{y_t\}_{t=1,\dots,N}$ that will be indicated as an N -dimensional vector $\mathbf{y} \in \mathbb{R}^N$ of generic element y_t , the smoothing matrix associated to the symmetric linear filter results

$$\mathbf{H} [\mathbf{H} + \mathbf{M}_{7(0.14)} (\mathbf{I}_N - \mathbf{H})] [\mathbf{H} + \mathbf{M}_{5(0.25)} (\mathbf{I}_N - \mathbf{H})]$$

where $\mathbf{H}$ is the N -dimensional matrix canonically associated to the 13-term Henderson filter, $\mathbf{M}_{5(0.25)}$ is the matrix representative to the 5-term moving average with weights (0.250, 0.250, 0.000, 0.250, 0.250) and $\mathbf{M}_{7(0.14)}$ is the matrix representative of a 7-term simple moving average with all weights equal to 0.143. It is evident that the linear Dagum filter is the result of the convolution between three filters, or, equivalently, the product of three matrices: $\mathbf{H}$, $[\mathbf{H} + \mathbf{M}_{5(0.25)} (\mathbf{I}_N - \mathbf{H})]$ and $[\mathbf{H} + \mathbf{M}_{7(0.14)} (\mathbf{I}_N - \mathbf{H})]$, representative of a Henderson trend estimation combined with two replacements of extreme values sequentially applied to preliminary Henderson trend estimations. This convolution gives a 31-term symmetric filter, which produces 15 asymmetric weights. The latter introduce phase shifts (temporal displacements of the points of maxima and minima of the filtered series), *i.e.* bias in the estimates of the last observations. For these reason, and being the ending weights of the 31-term symmetric filter very small relative to the central ones, the filter is truncated to 13 terms (length of H13) and normalized such that its weights add up to unity (constant mean preservation). To normalise the filter, the total weight discrepancy (the 13 truncated weights add up to 1.065) must be distributed over the 13 weights, w_j , $j = -6, \dots, 6$, according to a well defined pattern. This is a very critical transformation, because different distributions will produce linear filters with distinctive properties. The two most common normalisations procedures are the uniform and the proportional, both discussed in Dagum and Luati (2004a) together with a mixture of the two that is shown to give the best performance. The resulting weights of the short cascade linear filter (SCLF) are

$-0.027, -0.007, 0.031, 0.067, 0.136, 0.188, 0.224, 0.188, 0.136, 0.067, 0.031, -0.007, -0.027$.

This new filter belongs to the family of second-order kernel estimators since it satisfies the conditions $\sum_{j=-6}^6 w_j = 1$, $\sum_{j=-6}^6 jw_j = 0$ and $\sum_{j=-6}^6 j^2w_j \neq 0$, as well as L1 and GK. On the other hand, H13, L2 and CSS are higher-order kernels since they also satisfy $\sum_{j=-6}^6 j^2w_j = 0$ and $\sum_{j=-6}^6 j^pw_j \neq 0$ for some even $p > 2$. In general, second-order kernels produce smoother estimates relative to those given by higher-order kernels.

The theoretical properties of SCLF are studied by means of classical spectral analysis techniques and by local weight based measures of bias, variance and mean square error introduced and analysed by Dagum and Luati (2002, 2004b).

Figure 1. Gain functions of 13-term symmetric smoothers

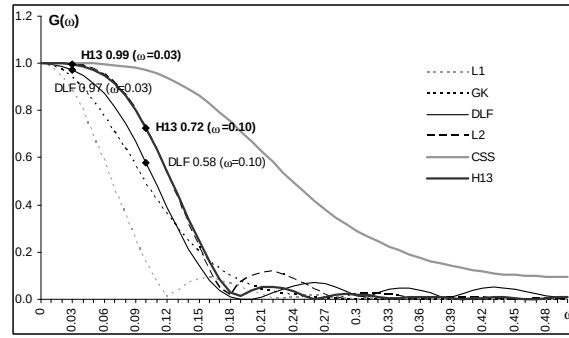


Figure 1 shows the gain functions of the symmetric estimators. It is evident that second order kernels tend to oversmooth the signal whereas higher order kernels leave a great amount of noise (CSS in particular loses all its optimal fitting and smoothing properties when its smoothing parameter is a priori fixed). SCLF constitutes a sort of compromise between a second and a higher order kernel. In fact, compared to H13, it suppresses more signal only in the frequency band pertaining to short signal cycles, ranging from 15 to 24 months periodicity ($0.03 < \omega \leq 0.06$), whereas it passes practically all three years cycles ($0 < \omega \leq 0.03$, where the value of the gain function is 0.99 for H13 and 0.97 for SCLF). Furthermore, respect to H13, SCLF reduces of the 15% the number of unwanted ripples ($\omega = 0.10$) and passes the same amount of noise, corresponding to the frequency band $0.10 \leq \omega \leq 0.50$ for seasonally adjusted series.

These results are in agreement with those inferred by the theoretical weight-based measures of local bias $1 - \sum_{j=-2}^2 w_j$, local variance $\sum_{j=-6}^6 w_j^2$ and local mean square error ($LMSE$), shown in Table 1 together with the empirical root mean square error and number of false turning points, calculated over a sample of real series (mean values standardized by those of H13). In fact, the $LMSE$ of SCLF is very similar to that of H13, but SCLF is more biased than H13 (*i.e.* less signal passed) and with smaller variance (*i.e.* less number of unwanted ripples). In general, symmetric smoothers where the $LMSE$ is mainly due to the bias produce an oversmoothed output and the converse holds for smoothers where the variance is dominant (see, respectively, the empirical measures of L1 and CSS compared to those of H13). One major advantage of the weight-based statistical measures relative to the classical spectral ones is their easiness of calculation. Furthermore, it is possible to have a local measure of bias for symmetric filters which

cannot be obtained by means of spectral techniques. Finally, we compare the empirical values of SCLF with those of H13. As expected, being SCLF a second order kernel with local variance slightly dominant over the local bias, its *RMSE* is higher than that of H13 but it reduces by 20% the number of false turning points in the final trend. Respect to L1 and GK, SCLF shows a smaller mean square error, *i.e.* it does not oversmooth the data.

Table 1. *Theoretical and empirical measures for 13-term symmetric smoothers*

Theoretical Weight-based Measures	SCLF	L1	GK	L2	CSS	H13
Local Bias	0.13	0.40	0.18	0.04	-0.07	0.04
Local Variance	0.17	0.10	0.15	0.20	0.40	0.20
LMSE	0.19	0.26	0.18	0.20	0.40	0.20
Empirical Data-based Measures						
RMSE/RMSEH13	1.05	1.35	1.08	1.04	0.58	1
$ftp/ftpH13$	0.80	0.42	0.80	1.03	2.29	1

3. Concluding Remarks

A linear symmetric estimator for the trend-cycle of seasonally adjusted time series was introduced. The new filter approximated the sA linear symmetric estimator for the trend-cycle of seasonally adjusted time series was introduced. The new filter approximated the symmetric part of the non linear Dagum estimator, shown to possess the good property, respect to the widely applied Henderson filter, of reducing the number of false turning points in the final estimates. Applied to real time series, the linear estimator obtained reduced by 20% the number of false turning point produced by the 13-term Henderson filter and did not oversmooth the data, as generally second-order kernels do, representing a good compromise between a second and a higher-order kernel. It possesses the smoothing properties of the former and the goodness of fit typical of the latter, without introducing a large number of false turning points.

References

- Cleveland (2002), *Estimating Variance of Seasonally Adjusted Series*, Finance and Economics Discussion Series, Board of Governors of the Federal Reserve System (U.S.), 15.
- Dagum E. B. (1987), *Current Issues on Seasonal Adjustment*, Survey Methodology, 13, 1, pp. 63-74.
- Dagum E. B. (1996), *A New Method to Reduce Unwanted Ripples and Revisions in Trend-Cycle Estimates from X11-ARIMA*, Survey Methodology, 22, 1, pp. 77-83.
- Dagum E. B., Luati A. (2002), *Global and Local Statistical Properties of Fixed Length Non Parametric Smoothers*, Statistical Methods and Applications, 11, 3, pp. 313-333.
- Dagum E. B., Luati A. (2004a), *A Robust Linear Filter for Non Stationary Mean Prediction*, Computational Statistics and Data Analysis, (forth).
- Dagum E. B., Luati A. (2004b), *Relationship between Local and Global Measures of Fitting and Smoothing of Nonparametric Function Estimators*, Studies in Nonlinear Dynamics and Econometrics, (forth).
- Kenny P. B., Durbin J. (1982), *Local Trend Estimation and Seasonal Adjustment of Economic and Social Time Series*, JRSS A, 145, pp. 1-41.