

Properties of Estimators Based on Local Polynomial Functions and on Probability Distribution Functions in Time Series Analysis

Proprietà di Stimatori Basati su Polinomi Locali e su Distribuzioni di Probabilità nell'Analisi delle Serie Storiche

Estela Bee Dagum

Dipartimento di Scienze Statistiche
Università di Bologna
beedagum@stat.unibo.it

Alessandra Luati¹

Dipartimento di Scienze Statistiche
Università di Bologna
luati@stat.unibo.it

Riassunto: Vengono analizzate le proprietà statistiche, globali e locali, di alcune funzioni nonparametriche utilizzate per stimare il segnale di una serie storica. Gli stimatori considerati appartengono a due classi di funzioni, i polinomi locali e le distribuzioni di probabilità. Le proprietà globali e locali vengono studiate attraverso il sistema di pesi generato da ogni stimatore una volta fissato il parametro di smoothing (che regola il compromesso tra distorsione e varianza) in modo da ottenere filtri lineari di uguale lunghezza, in particolare pari a 9 e 13.

Keywords: signal estimation, bias, variance, smoothness, systems of weights.

1. Introduction

A basic assumption in time series analysis, is that a time series $\{y_t\}_{t=1,\dots,N}$, $N < \infty$, can be decomposed in the sum of a signal, $g(t)$, plus an erratic component, u_t , as follows: $y_t = g(t) + u_t$ where $g(t)$ is either deterministic or stochastic and u_t is a white noise or an *ARMA* stochastic process with zero mean and constant variance σ_u^2 .

The main purpose of this study is to analyse the **global** and **local** statistical properties of some nonparametric signal estimators, the weights of which are based on local polynomials and probability distribution functions.

To analyse the global properties, we assume that the signal is described **locally** by some polynomial function of degree $p \in N$ in t , that is

$$g(t) = a_0 + a_1t + a_2t^2 + \dots + a_pt^p. \quad (1)$$

The estimation of local polynomials is generally obtained by applying, in a **moving manner**, a finite number of weights to the observations, such that, for $t > m$, $t \leq N - m$,

$$\hat{g}(t; j) = \sum_{j=-m}^m w_j y_{t+j}. \quad (2)$$

Hence, $\hat{g}(t; j)$ can be interpreted equivalently as: (a) a local estimator of $g(t)$ for fixed values of m and p and, (b) a smooth estimate of the value y_t resulting from a $2m + 1$ -term moving average of symmetric neighboring observations. For the first and last m

¹Financing from MURST is gratefully acknowledged.

observations only asymmetric filters can be used, such that, *e.g.*, the last observation is estimated by $\hat{g}(N; j) = \sum_{j=-h}^0 w_j y_{N+j}$, for some h that depends on the estimating function.

The estimators here studied are based on two classes of weight generating functions, local polynomials and probability distributions. We consider, in the first class, the locally weighted regression smoother or loess of degree 1 and 2 (L1 and L2), the cubic smoothing spline (CSS) and the Henderson filter (H); and in the second class, the Gaussian kernel smoother (GK) (see Dagum and Luati, 2000 and 2001). These estimators depend on a smoothing parameter, η , that determines the degree of smoothness of the estimated curves, and once the smoothing parameter is fixed, all the estimators are linear as given in (2).

In general, estimators for $g(t)$ are chosen such that the smoothing parameter gives the best compromise between bias and variance. We follow here an entirely different approach, for we analyse the properties of the estimators when the smoothing parameter is obtained by **constraining** the smoothers length to be **fixed**. We select 9 and 13 terms in accordance with the most often applied Henderson filters for trend estimation, but our approach can be easily extended to any a-priori fixed number of terms. Since the smoothers are constrained to be of a fixed length, their statistical properties are no longer necessarily optimal as in the case when the optimal smoothing parameter is chosen. Therefore, we study their statistical properties within the context of their respective weighting systems. In particular, we study the global and local properties of the symmetric and asymmetric weights, concerning bias, variance, mean square error and degree of smoothing of the output.

2. Asymptotic and Local Measures of Bias and Variance

An estimator of the form (2) for a polynomial of degree $p > 0$ of the form (1) is globally unbiased if $E[\hat{g}(t; j)] = g(t)$ for all t , which implies $\sum_{j=-m}^m j^p w_j = 0$. Hence, we consider $\sum_{j=-m}^m j^p w_j$ as a measure of **global bias for symmetric weights**, for $m = 6$ and $m = 4$; $p = 1, 2$. Since the weights always add to one, a local constant ($p = 0$) is always reproduced. Concerning the asymmetric weights, we analyse those for the last observation and, thus, the global bias is given by $\sum_{j=-h}^0 j^p w_j$, where $h = m$ for all the estimators except for loess where $h = 2m + 1$ (since neighborhoods of $2m + 3$ points need to be considered in order to have $2m + 1$ nonnull symmetric weights; these produce asymmetric filters of length $2m + 2$).

The variance of $\hat{g}(t, j)$ is $V\left(\sum_{j=-m}^m w_j y_{t+j}\right) = \sum_{j=-m}^m w_j^2 \sigma_u^2 + \sum_{i,j=-m; i \neq j}^m w_i w_j \gamma_{ij}$, where γ_{ij} is the covariance between y_{t+i} and y_{t+j} . It follows that we cannot draw any conclusion about the global variance unless we know the behavior of the autocovariance function of the observed series.

As **local statistical measures** we propose the following:

(1) The local bias for the symmetric weights is given by $1 - \sum_{j=-k}^k w_j$ with $k = 2$ when $m = 6$ and $k = 1$ when $m = 4$. In this way, we take into account the amount of the total weight (equal to 1) assigned by 13- and 9-term estimators to respectively the

38% and 33% of the observations surrounding the point to be estimate. It is apparent that the bias will be null if the total weight given to near the 40% and 30% of neighboring observations adds to one. On the other hand, the greater or lesser than zero, the higher the systematic overestimation or underestimation performed by the smoother, respectively. A negative value for this measure of local bias is not possible for systems of weights based on probability distributions, that are necessarily all nonnegative. For the last point, the local bias indicator is $1 - \sum_{j=-k}^0 w_j$, with k as given above for all the smoothers except loess where $k = m - 1$, for the 40% of the total weight to be distributed on ending observations.

(2) As a measure of local variance, we use the sum of the squares of the weights. This measure gives exactly the amount of variance reduction of the output only if the input is a white noise. For most socio-economic time series this is not the case, but we can still use it as a good indicator of the total variance. In fact, if $\sum_{j=m}^m w_j = 1$, then $\left(\sum_{j=m}^m w_j\right)^2 = 1$ and, equivalently, $\sum_{j=m}^m w_j^2 + \sum_{i,j,i \neq j} w_i w_j = 1$. This means that the smaller $\sum_{j=1}^n w_j^2$, the larger is its complement to unity given by $\sum_{i,j,i \neq j} w_i w_j$. Therefore, on the basis of the assumption that, for the majority of economic time series, γ_{ij} is a linearly decreasing and positive function of the time lag, most of the variance of the output is expected to be accounted for by $\sum_{j=m}^m w_j^2$. For the last observation, the measure of the variance is given by $\sum_{j=-h}^0 w_j^2$.

(3) A measure of the compromise between bias and variance is given by the mean square error (MSE) $\left(1 - \sum_{j=-k}^k w_j\right)^2 + \sum_{j=m}^m w_j^2$ and $\left(1 - \sum_{j=-k}^0 w_j\right)^2 + \sum_{j=-h}^0 w_j^2$.

(4) Finally, a measure of smoothness of the estimator (and equivalently of the estimated signal) is given by $\sum_{j=-m}^{m+3} (\Delta_B^3 w_j)^2$ and $\sum_{j=-h}^0 (\Delta_B^3 w_j)^2$, for symmetric and asymmetric weights, respectively, where $\Delta_B = 1 - B$, B the backshift operator, which measures the difference between a parabola and the curve described by the weights. The higher it is (the sharper is the distribution of the weights), the lesser smoothed the estimated curve.

3. Analysis of the Results and Main Conclusions

The global bias for $p = 1$ is null for all symmetric estimators (not shown for space reasons) whereas for $p = 2$ only L2, CSS, H13 and H9 can be considered globally unbiased. On the other hand, L1 and GK will introduce large bias if the signal follows a second degree polynomial.

Concerning the **symmetric filters local statistical properties**, the bias measure is not invariant to their length but consistent with the global results. Among the estimators with nearly zero bias, the ones with smallest variance are L2, H13 and H9 whereas CSS shows a very large value. This is reflected in the corresponding MSE values. Finally, all estimators are excellent smoothers with respect to a second degree parabola, except CSS. In fact, by fixing the smoothing parameter, the optimal properties of CSS as the best compromise between fitting and smoothing seem to be lost. On the other hand, the polynomial based estimators L2 and the Henderson filters can be considered as those being best linear unbiased with respect to the local properties. Furthermore, these two

SYMM.	<i>13-term estimators</i>					<i>9-term estimators</i>				
Global	L1	L2	CSS	H13	GK	L1	L2	CSS	H9	GK
<i>Bias p=2</i>	7.06	0	-0.04	0.02	3.58	3.61	0	0.01	-0.01	1.44
Local	<i>13-term estimators</i>					<i>9-term estimators</i>				
<i>Bias</i>	0.40	0.04	-0.07	0.04	0.18	0.49	0.14	-0.07	0.14	0.20
<i>Variance</i>	0.10	0.20	0.40	0.20	0.15	0.14	0.28	0.71	0.28	0.24
<i>MSE</i>	0.26	0.21	0.40	0.21	0.18	0.38	0.30	0.71	0.30	0.27
<i>Smooth.</i>	0	0.01	0.77	0	0.01	0.01	0.09	7.92	0	0.08
ASYMM.	<i>13-term estimators</i>					<i>9-term estimators</i>				
Global	L1	L2	CSS	H13	GK	L1	L2	CSS	H13	GK
<i>Bias p=1</i>	0	0	0.01	-0.41	-1.22	0	0	0	-0.07	-0.68
<i>Bias p=2</i>	-13.81	0	-0.32	-2.16	2.96	-6.45	0	-0.04	-1.64	1.08
Local	<i>13-term estimators</i>					<i>9-term estimators</i>				
<i>Bias</i>	-0.22*	-0.13*	-0.04	-0.02	0.15	-0.17*	-0.16*	-0.03	-0.01	0.15
<i>Variance</i>	0.31*	0.55*	0.82	0.39	0.26	0.41*	0.68*	0.95	0.58	0.39
<i>MSE</i>	0.36*	0.57*	0.82	0.39	0.29	0.44*	0.70*	0.95	0.58	0.41
<i>Smooth.</i>	0*	0*	0.09	0.07	0.01	0*	0.01*	0.54	0.26	0.05

Table 1: Global and local measures of bias and variance of symmetric and asymmetric weights of various 13- and 9-term estimators. * Note that asymmetric weights for loess are $2(m+1)$.

estimators have excellent smoothing properties.

For the **asymmetric weights**, the global bias of the Henderson filters differs very much from those of the symmetric version because they are based on different signal assumptions. However, concerning the local properties, only H13, H9 and CSS show small biases being the Henderson filters the ones to be preferred in terms of variance and smoothing.

We intend to pursue further research by testing the appropriateness of these local statistical measures using a large sample of both simulated and real-life time series.

References

- Dagum E.B. and Luati A. (2000), *Predictive Performances of Some Linear and Nonlinear Smoothers for Noisy data*, Statistica, anno LX, 4, pp. 635-654.
- Dagum E. B. and Luati A. (2001), *A Study of the Asymmetric and Symmetric Weights of Kernel Smoothers and their Spectral Properties*, Estadistica, 53, 160, pp. 215-258.