Recent Developments in Short-term Trend Prediction for Real Time Analysis

Estela Bee Dagum

Department of Statistical Sciences - University of Bologna
and
Time Series Research and Analysis Center - Statistics Canada

estela.beedagum@unibo.it

Outlook

Large variety of nonparametric trend-cycle estimators, based on different smoothing assumptions:

- density functions (kernel estimators)
- local polynomial fitting (Loess smoother)
- graduation theory (Henderson filter)
- smoothing spline regression

Aim: *to introduce a unified approach by means of the Reproducing Kernel Hilbert Space (RKHS) methodology*

Main advantages

(1) It enables the comparison of TC nonparametric predictors in the time domain. The nonparametric predictors are transformed into density functions which provide the initial weighting shape for near neighborhood observations.

(2) From the density function, a hierarchy of higher order kernels is derived.

(3) It enables to find TC nonparametric predictor kernel representations for relatively short bandwidths mainly used in real time analysis.

(4) It enables asymmetric filters for most recent observations which are coherent with the corresponding symmetric ones . In particular, these asymmetric filters have superior properties of signal passing, noise suppression, and revisions relative to the classical ones.

- A RKHS is a Hilbert space characterized by a kernel that reproduces, via an inner product, every function of the space.
- The concept was first introduced by Aronszajn (1950) and Bergman (1950) and it was Parzen (1959) the first to applied it in time series. Parzen used a parametric approach
- The RKHS approach followed in our study is strictly nonparametric.
- It basically consists in transforming all these filters into kernel functions of order two which are densities, and from which hierarchies of estimators are derived.

A Hilbert space is a linear finite or infinite complete space with an inner product. In particular, we will consider the space $L^2(f_0)$ of square integrable functions with respect to a density function f_0 defined on T (e.g. $T = [-1,1]; [0,1), or (-\infty,\infty))$.

That is, $Y(t)$ belongs to $L^2(f_0)$ if $\int_T Y(t)^2 f_0(t)dt<\infty$.

For Y_1 and Y_2 in $L^2(f_0)$,

$$< Y_1(t), Y_2(t) >_{L^2(T)} = \int_T Y_1(t) Y_2(t) f_0(t) dt$$

Assumptions

1)

$$y_t = g_t + u_t \qquad t = 1,2,..., N$$

where

y_t: input seasonally adjusted series;

g_t : signal or nonstationary mean (trend-cycle);

u_t: noise assumed to be either a white noise, $WN(0,\sigma^2{}_u)$, or more generally to follow a stationary and invertible ARMA process.

2) $\{y_t, t=1,2,\ldots,N\}$ is a finite realization of a stochastic process, whose trajectories belong to the Hilbert space $L^2(f_0)$.

Assumptions – continue

3) The signal ***g* is a smooth function of time,** hence g can be **locally approximated by a polynomial function** of the time distance j between y_t and the neighboring observations y_{t+j}, *that is,*

$$g_{t+j} = a_0 + a_1 j + a_2 j^2 + \ldots + a_p j^p + \varepsilon_{t+j} \qquad j = -m, \ldots, m$$

$a_0, a_1, \ldots, a_p$ are real and ε_t is assumed to be purely random and mutually uncorrelated with u_t.

This implies that the analysis of the signal can be performed in the space $\boldsymbol{P_p}$ of polynomials of degree at most p, being p a nonnegative integer.

$a_0, a_1, \ldots, a_p$ are estimated by projecting the observations in a neighborhood of y_t on the subspace P_p, or equivalently by minimizing the weighted least square fitting criterion

$$\min_{g \in P_p} \| y - g \|^2_{P_p} = \int_T (y(t-s) - g(t-s))^2 f_0(s) ds \qquad (1)$$

$\| \cdot \|^2_{P_p}$ denotes the P_p-norm.

The **solution** for a_0 provides the trend-cycle estimate g_t, for which a general characterization and explicit **representation** can be provided **by means of the RKHS methodology**.

$$R : T \times T \to \Re$$

Reproducing kernel

$$1. R(t,.) \in H, \forall t \in T;$$

$$2. < g(.), R(t,.) >= g(t), \forall t \in T \; and \; g \in H$$

Fundamental results

The space $\boldsymbol{P_p}$ is a reproducing kernel Hilbert space of polynomials on some domain T, that is there exists an element $R_p(t,.)$ belonging to $\boldsymbol{P_p}$, such that

$$P(t) = < P(.), R_p(t,.) > \quad t \in T, P \in P_p$$

Under assumptions 1,2, and 3, the minimization problem (1) has a unique and explicit solution given by

$$\hat{g}_t = \int_T y(t-s) R_p(s,0) f_0(s) ds$$

Hence, the estimate g_t can be seen as a local weighted average of the observations, where the weights are derived by a kernel function of order $p+1$

$$K_{p+1}(t) = R_p(t,0) f_0(t)$$

where p is the degree of the fitted polynomial.

Kernels of order p

Given $p \geq 2$, K_p is said to be of order p if

$$\int_T K_p(s) ds = 1$$

and

$$\int_T s^i K_p(s) ds = 0 \qquad i = 1,2,\ldots., p-1$$

In other words, it will reproduce a polynomial trend of degree p-1 without distortion.

(Berlinet, 1993). Kernels of order $p+1$, $p \geq 1$, can be written as products of the reproducing kernel $R_p(t,.)$ of the space $\boldsymbol{P_p}$ and a density function f_0 with finite moments up to order $2p$. That is

$$K_{p+1}(t) = R_p(t,0) f_0(t) = \sum_{i=0}^{p} P_i(t) P_i(0) f_0(t)$$

Remark 1 (Christoffel-Darboux formula). For any sequence $(P_i)_{0 \leq i \leq p}$ of $p+1$ orthonormal polynomials in $L^2(T)$

$$R_p(t,0) = \sum_{i=0}^{p} P_i(t) P_i(0)$$

Applied to real data,

$$\hat{g}_t = \sum_{j=-m}^{m} w_j y_{t-j}$$

where w_j, $j = -m,\ldots,m$, depend on the degree p of the polynomial, the b bandwidth length and the shape of the density function f_0.

When f_0 is defined on $T = [-1, 1]$, symmetric weights for a filter length $2m+1$ are derived by fixing $b = m+1$, that is

$$w_j = \frac{K_{p+1}\left(\frac{j}{(m+1)}\right)}{\sum_{j=-m}^{m} K_{p+1}\left(\frac{j}{(m+1)}\right)}, \qquad j = -m,\ldots,m$$

Several nonparametric estimators developed in the literature for smoothing functional data. Two main approaches: (a) least squares (kernel estimation, local polynomial fitting, graduation theory), and (b) roughness penalty (smoothing spline regression).

Unified perspective by means of RKHS: different nonparametric estimators are transformed into kernel functions and grouped into hierarchies with the following property: each hierarchy is identified by a density f_0 and contains estimators of order 2,3,4,… which are products of orthonormal polynomials with f_0.

Polynomial kernel regression

Problem: fitting a polynomial trend to the observations y_{t+j}, j=-m,…,m, by minimizing

$$\sum_{j=1}^{N} K_0\left(\frac{t-j}{b}\right)\left[y_t - a_0 - a_1(t-j) - \ldots - a_p(t-j)^p\right]^2$$

where b determines the bandwidth of the symmetric and nonnegative weighting function, since $K_0(z)=0$, if $|z|\geq 1$.

Kernel estimators, local polynomial regression smoothers and filters derived in the graduation theory differ in:

- degree of the fitted polynomial;
- shape of the weighting function;
- neighborhood of observations taken into account.

To derive the corresponding kernel hierarchy by means of the RKHS methodology, the density function corresponding to K_0 and its orthonormal polynomials have to be determined.

Polynomial kernel regression : (1) Kernel estimators

Kernel estimates are obtained by locally fitting linear (p=1) polynomial trends weighting the observations in a neighborhood of the target point t using a density function.

Gaussian kernel family well-known in the literature as Gram-Charlier hierarchy (Deheveuls, 1977; Wand and Schucany, 1990; Granovsky and Muller, 1991)

Corresponding density function $f_{0G}(t) = \frac{1}{\sqrt{2\pi}} exp\left(-\frac{t^2}{2}\right)$, with Hermite orthonormal polynomials.

Third order kernel within the hierarchy $\frac{1}{\sqrt{2\pi}} exp\left(-\frac{t^2}{2}\right) \times \left(\frac{3-t^2}{2}\right)$

Clear relationship between different order estimators within the hierarchy.

Polynomial kernel regression : (2) Loess smoother

The Loess estimator is based on nearest neighbor weights and is applied in an iterative manner for robustification. It consists of locally fitting a polynomial of degree p by weighted least squares, where the weighting function proposed by Cleveland (1979) is the tricube one

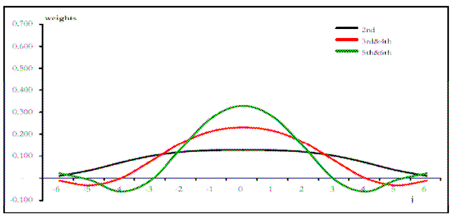
$$K_{0T}(t) = \left(1 - |t|^3\right)^3 I_{[-1,1]}(t)$$

Dagum and Bianconcini (2006) derived the **Loess kernel hierarchy** based on the tricube density

$$f_{0T}(t) = \frac{70}{81}\left(1 - |t|^3\right)^3$$

Third order kernel

$$f_{0T}(t) = \frac{70}{81}\left(1 - |t|^3\right)^3 \left(\frac{539}{293} - \frac{3719}{638} t^2\right)$$



Symmetric weights of the 13-term Loess kernels.

Polynomial kernel regression : (3) Henderson filters

Henderson's starting point was the requirement that the filters should reproduce a cubic polynomial trend without distortion.

Three alternative smoothing criteria give the same weight diagram (Kenny and Durbin, 1982; Gray and Thomson, 1996):

- minimization of the variance of the third differences of the smoothed series;
- minimization of the sum of squares of the third differences of the coefficients of the moving average formula;
- fitting a cubic polynomial by weighted least squares, with weighting function given by

$$K_{0H}(t) \propto \{(m+1)^2 - j^2\}\{(m+2)^2 - j^2\}\{(m+3)^2 - j^2\} \quad (2)$$

Dagum and Bianconcini (2008) showed that the weight diagram of the Henderson smoother is well-reproduced by two different density functions:

- the exact density f_{0H} derived by the penalty function K_{0H};
- the **biweight density**

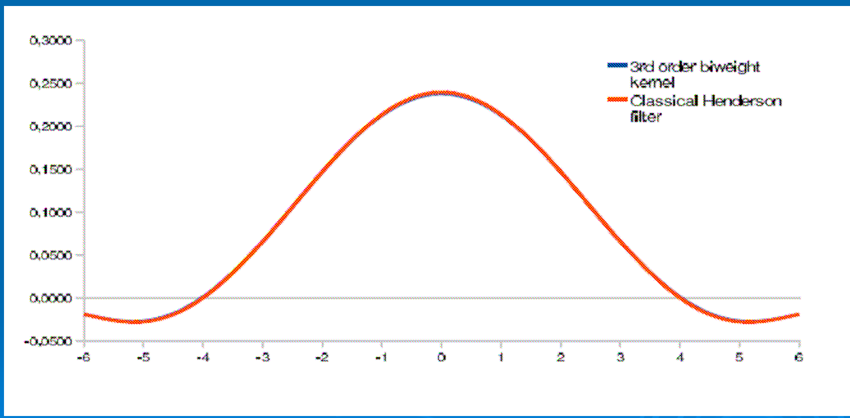
$$f_{0B}(t) = \frac{15}{16}\left(1-t^2\right)^2 I_{[-1,1]}(t)$$

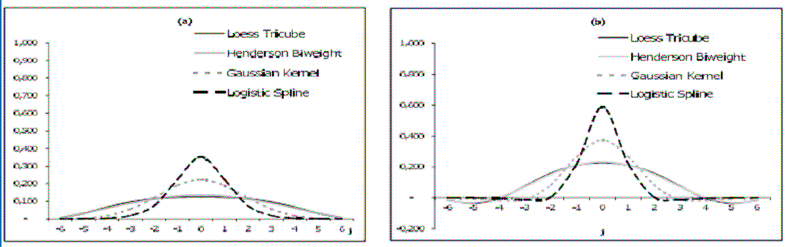
Advantages f_{0B}

- the biweight density function, and the corresponding hierarchy, does not need to be calculated any time that the length of the filter changes, as happens for f_{0H};
- it belongs to the well-known Beta distribution family;
- the corresponding orthonormal polynomials are the Jacobi ones, for which explicit expressions for computation are available.

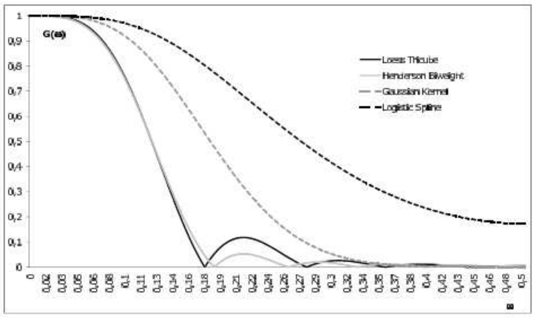
Third order kernel $\frac{15}{16}\left(1-t^2\right)^2 \times \left(\frac{7}{4} - \frac{21}{4}t^2\right)$

Symmetric Kernel and Classical 13 Term Filters





13-term (a) Second and (b) Third Order Kernels.



ain Functions of Symmetric 13-term Third Order Kernels.

Smoothing spline regression

Problem: search for an optimal solution between fitting and smoothing of the data, under the assumption that the signal follows locally a polynomial of degree *p*.

Schoenberg (1964) showed that natural smoothing spline estimator of order ℓ arises as the solution of the minimization problem

$$\min_{g \in W_2^{\ell}(T)} \|y - g\|^2_{W_2^{\ell}} \qquad (3)$$

where $\|\cdot\|_{W_2^{\ell}}$ denotes the W_2^{ℓ}, defined

$$\|y - g\|^2_{W_2^{\ell}} = \int_T (y(t) - g(t))^2 dt + \lambda \int_T \left(g^{(\ell)}(t)\right)^2 dt$$

Where p=2ℓ-1. For ℓ=2, hence p=3, (Wahba, 1990; Green and Silverman, 1994)

$$\hat{\mathbf{g}} = \mathbf{A}(\lambda)\mathbf{y}$$

- $\hat{\mathbf{g}} = (\hat{g}_1, \hat{g}_2, \ldots, \hat{g}_N)'$ *and* $\mathbf{y} = (y_1, y_2, \ldots, y_N)'$

Influential matrix

Equivalent kernel representation in Sobolev spaces

For each y_t belonging to $L^2(T)$, it can be shown that the solution to the minimization problem (3) exists and is unique. It is determined by the unique **Green's function** $G_\lambda(t,s)$, such that

$$\hat{g}(t) = \int_T G_\lambda(t,s)\, y(s)\, ds$$

The derivation of $G_\lambda(t,s)$ corresponding to a smoothing spline of order ℓ requires the solution of a $(2p+2)x(2p+2)$ system of linear equations for each value of λ.

A simplification is provided by studying $G_\lambda(t,s)$ as the reproducing kernel $R_{\ell,\lambda}(t,s)$ of the Sobolev space, where T is an open subset of the real space.

When T is the real space, Thomas-Agnan (1991) provided a general formula for $R_{\ell,\lambda}(t,s)$

Corollary 2

$$R_{1,1}(t) = \frac{1}{2} exp(-|t|)$$

$$R_{2,1}(t) = \frac{1}{2} e^{\frac{t}{\sqrt{2}}} sin\left(|t| \frac{\sqrt{2}}{2} + \frac{\pi}{4} \right)$$

$$R_{3,1}(t) = \frac{1}{6} \left\{ e^{|t|} + 2e^{\frac{t}{2}} sin\left(|t| \frac{\sqrt{3}}{2} + \frac{\pi}{6} \right) \right\}$$

Equivalent kernel representation in the polynomial space

Sobolev minimization problem

$$\min_{g \in W_2^\ell} \|y-g\|^2_{W_2^\ell} = \int_T (y-g)^2 dt + \lambda \int_T \left(g^{(\ell)}(t) \right)^2 dt$$

Weighted least square criterion

$$\min_{g \in P_p} \|y-g\|^2_{P_p} = \int_T (y(t-s) - g(t-s))^2 f_0(s) ds$$

Problem: find the density function f_0 according to which the spline estimates are obtained by minimizing the weighted least squares fitting criterion.

Considered densities

- Standard Laplace density

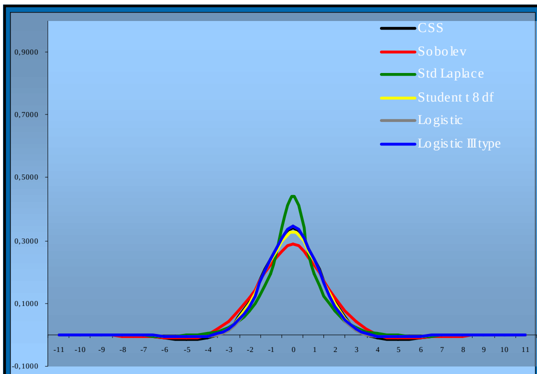
$$f_{0Lap}(t)=\frac{1}{2}exp(-|t|)$$

- Student's *t* with 8 degrees of freedom

$$f_{0Stu}(t)=\frac{\Gamma\left(\frac{9}{2}\right)}{\sqrt{8\pi}\,\Gamma(4)}\left(1-\frac{t^2}{8}\right)^{-\frac{9}{2}}$$

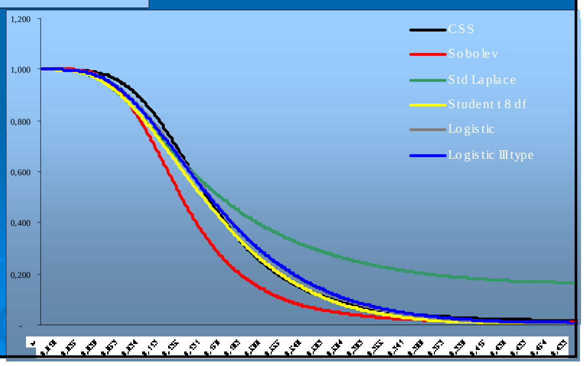
- Logistic density with mean α=0 and dispersion parameter β=0.2

$$f_{0Log}(t)=\left(\frac{1}{4\beta}\right)^{-1}sec\,h^2\left[\frac{1}{2}\left(\frac{t-\alpha}{\beta}\right)\right]$$



Symmetric weights of 23-term equivalent kernels and classical cubic splines.

Gain functions of symmetric 23-term equivalent kernels and classical cubic spline.



$$\Delta_{weight} = \left(\sum_{j=-m}^{m} \left| \kappa_{CSS}(j) - \kappa_K(j) \right|^2 \right)^{1/2}$$

$$\Delta_{gain} = \left(\sum_{\omega=0}^{1/2} \left| G_{CSS}(\omega) - G_K(\omega) \right|^2 \right)^{1/2}$$

3rd order kernel	Filter length					
	9		*13*		*23*	
	weights	**gain**	**weights**	**gain**	**weights**	**gain**
Sobolev space	**0,144**	**3,213**	**0,428**	**1,925**	**0,482**	**1,709**
Std Laplace	**0,049**	**1,088**	**0,609**	**3,863**	**0,583**	**2,916**
Student's t	**0,028**	**0,626**	**0,287**	**0,678**	**0,303**	**0,619**
Logistic α=0, β=0.2	**0,021**	**0,480**	**0,280**	**0,592**	**0,260**	**0,471**

Behavior at the boundaries – Polynomial kernel regression

The kernel derived by means of the RKHS methodology provide a new and unified way to represent nonparametric estimators based on different assumptions of fitting and smoothing.

For Loess and Henderson filter, Dagum and Bianconcini (2006 and 2008) showed how this has important consequences in the derivation of the asymmetric weights.

The third order kernel in the tricube and biweight hierarchies are continuous versions of the classical Loess of degree 2 (LOESS 2) and Henderson filters respectively. On the other hand, no comparisons can be made for the third order Gaussian kernel which is already a kernel function, and for which no counterpart exists in the literature.

In the RKHS approach, all the filters are transformed into kernel functions and applied as local weighted averages to the data. At the boundary of the observation interval the local averaging process get asymmetric, that is, half of the weights are non defined and outside the boundary.

Behavior at the boundaries in RKHS

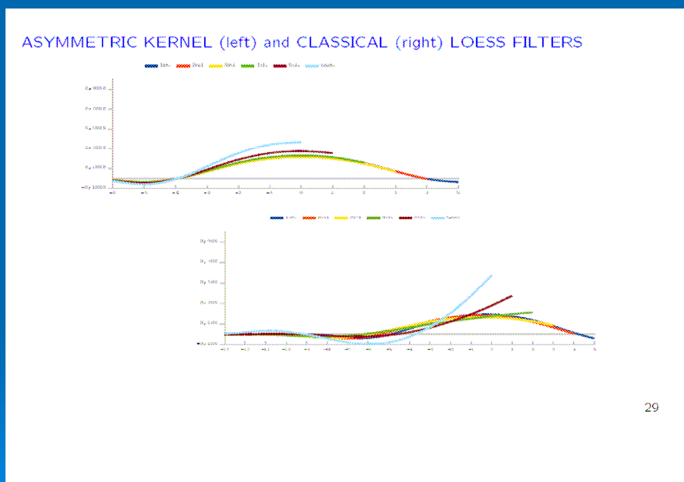
The third order kernels are unbiased estimators of a local quadratic/cubic polynomial trend when applied in the middle of the observation interval ($m+1 \leq t \leq N-m+1$). However, when applied to the first and last m observations, the unbiasedness condition is not fulfilled.

Common approach is to compute the asymmetric weights by applying the so called *"cut-and-normalized"* method.

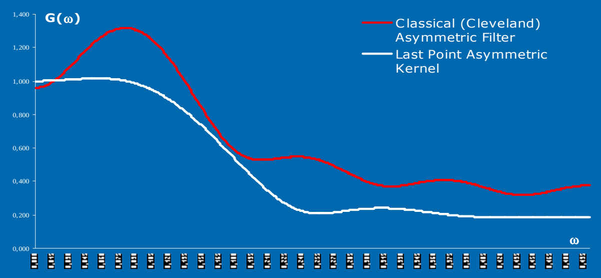
$$w_j = \frac{K_3\left(\frac{j}{b}\right)}{\sum_{j=-m}^{q} K_3\left(\frac{j}{b}\right)} \qquad j = -m, \ldots, q; q = 0, \ldots, m-1$$

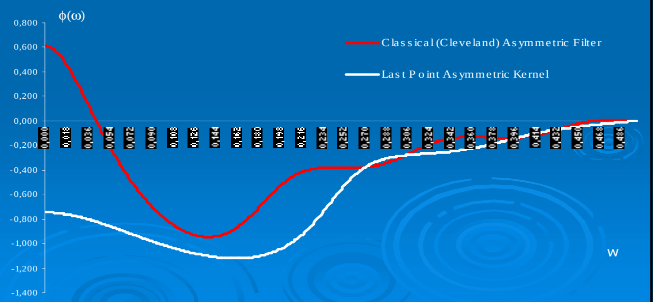
where

- j: distance to the target point t ($t = N-m+1, \ldots, N$);
- b: bandwidth parameter ensuring a symmetric filter of length $2m+1$
- $m+q+1$: asymmetric filter length.

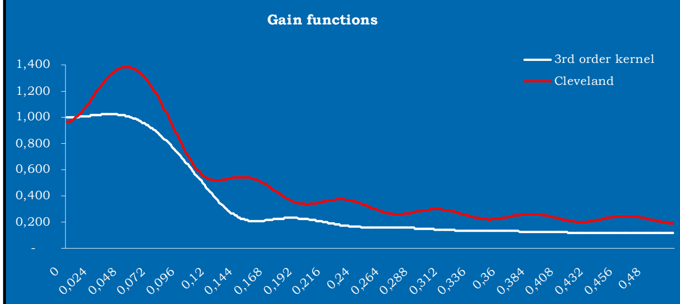


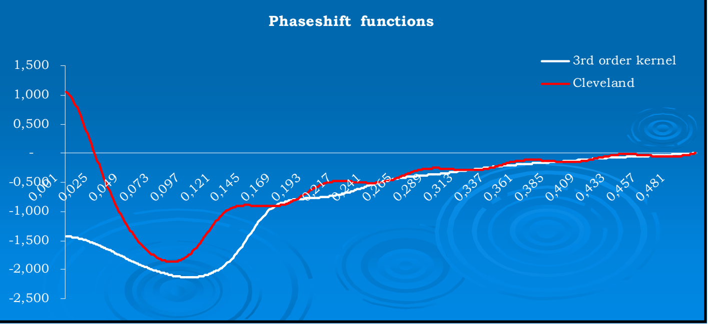
Behavior at the boundaries – 13-term Loess filters



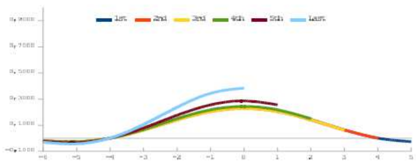


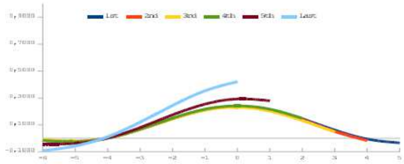
Behavior at the boundaries – 23-term Loess filters





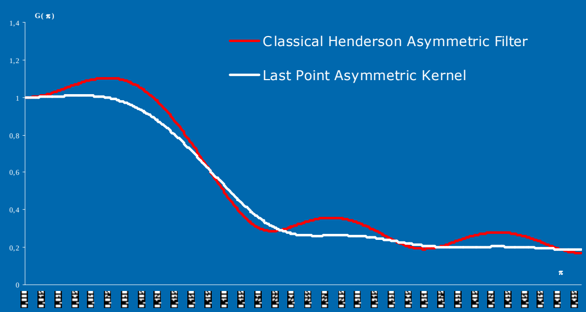
ASYMMETRIC KERNEL (left) and CLASSICAL (right) 13-TERM HENDERSON FILTERS

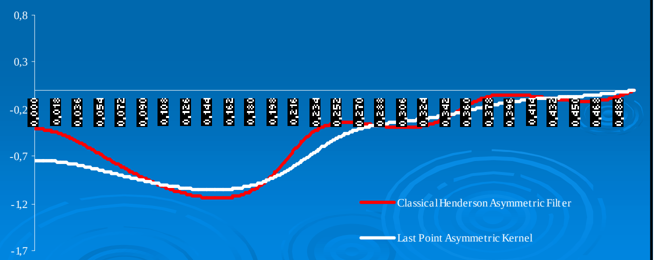




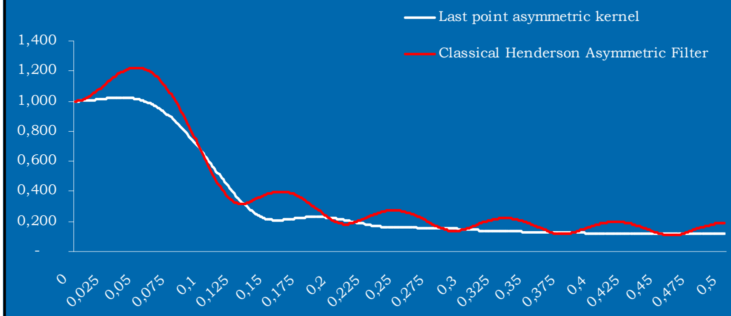
32

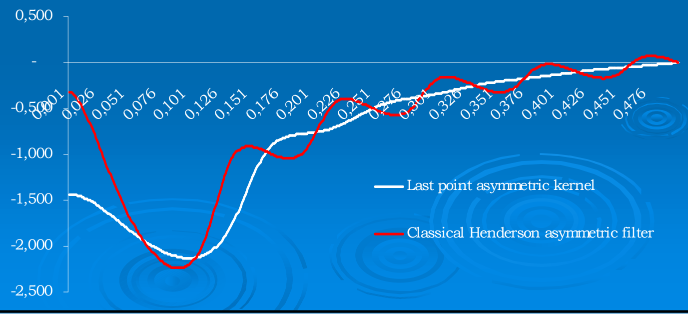
Behavior at the boundaries – 13-term Henderson filters





Behavior at the boundaries – 23-term Henderson filters





Behavior at the boundaries – Kernel splines

Third order kernel within the logistic hierarchy

$$K_3(t) = \frac{5}{4} \sec h^2 \left(\frac{5}{2} t \right) \left(\frac{21}{16} - \frac{2085}{878} t^2 \right)$$

K_3 is an unbiased estimator of a local cubic polynomial trend when applied in the middle of the observation interval ($m+1 \leq t \leq N-m+1$). However, when applied to the first and last m observations the unbiasedness condition are not still fulfilled.

As before, in the RKHS

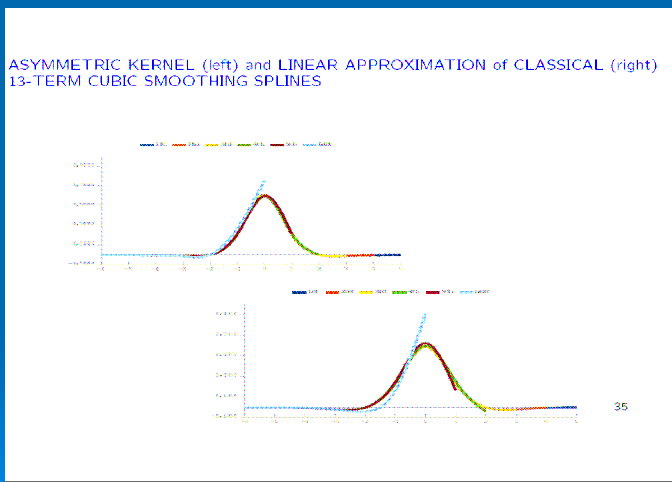
$$w_j = \frac{K_3\left(\frac{j}{b}\right)}{\sum_{j=-m}^{q} K_3\left(\frac{j}{b}\right)} \qquad j = -m, \ldots, q; q = 0, \ldots, m-1$$

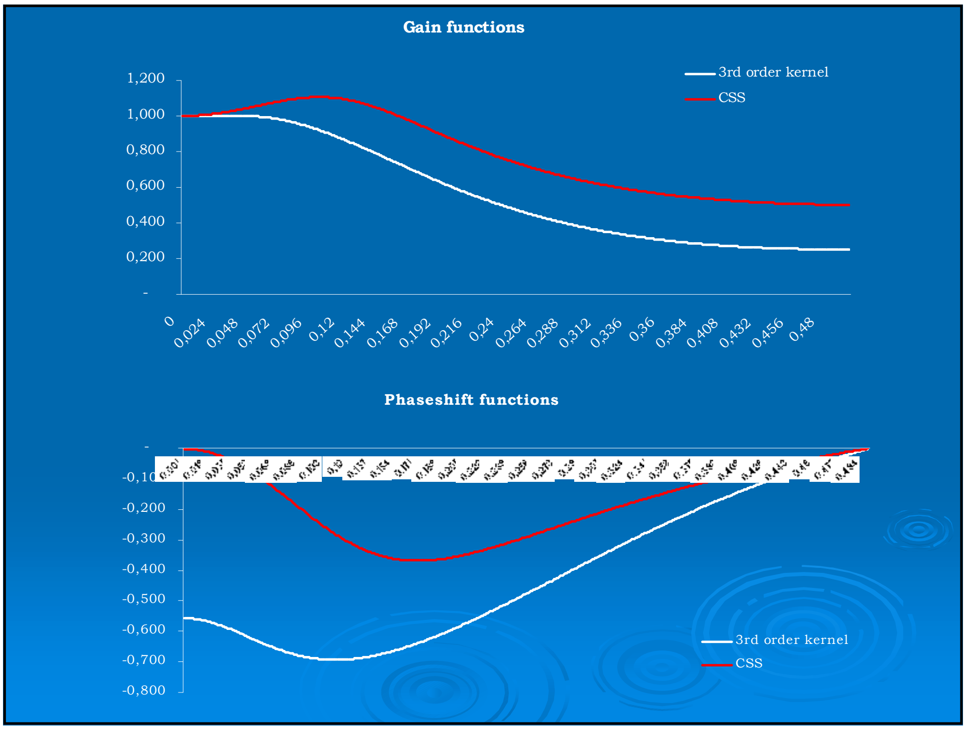
Behavior at the boundaries – Natural splines

The problem of an erratic polynomial behavior near the boundaries is exacerbated with natural cubic smoothing splines.

Natural cubic smoothing splines add additional constraints, ensuring that the function is of degree 1 beyond the boundary knots.

In this study, the asymmetric classical splines are obtained by fixing the λ parameter in view of ensuring a $2m+1$-term symmetric filter, and then selecting the last m rows of the influential matrix $A(\lambda)$.





MSE REVISION RATIO BETWEEN KERNEL AND CLASSICAL LAST POINT PREDICTORS

Closings of the Dow-Jones industrial index

-LOESS 0.383

-HENDERSON 0.886

-SPLINE 0.860

Temperature, coppermine

-LOESS 0.539

-HENDERSON 0.663

-SPLINE 0.609

U.S. male (20 years and over) unemployment

-LOESS 0.478

-HENDERSON 0.789

-SPLINE 0.721

U.S. female (20 years and over)unemployment

-LOESS 0.540

-HENDERSON 0.893

-SPLINE 0.867

Macro-area	Series	LOESS	Henderson	Spline
Crime	Minneapolis public drunkenness	0,572	0,881	0,867
Finance	Monthly return on the S&P 500 index	0,722	0,610	0,232
	Return to an investment strategy based on the paper rate	0,225	0,490	0,333
	Commercial paper rate, expressed by the annual percentage rate	0,483	0,787	0,562
	Railroad bond yields (% x 100)	0,664	0,760	0,622
	Mutual savings bank data end-of-month balance	0,561	0,789	0,660
	Interest rates Government Bond, Reserve Bank of Australia	0,443	0,789	0,714
	Closings of the Dow-Jones industrial index	0,282	0,886	0,860
Health	Number of cases of measles, New York city	0,806	0,853	0,482
	Number of cases of measles, Baltimore	0,512	0,777	0,552
	Bodyweight of rats	0,544	0,760	0,674
	Number of chickenpox, New York city	0,597	0,879	0,837
Hydrology	Temperature, coppermine	0,535	0,663	0,609
	Flows, Colorado River Lees Ferry	0,580	0,856	0,368
	Lake Erie Levels	0,568	0,787	0,807
	Flows, chang jiang at han kau	0,658	0,877	0,874
Labour Market	Wisconsin employment time series, fabricated metals	0,502	0,774	0,712
	U.S. male (20 years and over) unemployment figures	0,478	0,785	0,721
	Unemployment Benefits in Australia	0,318	0,702	0,722
	Woman unemployed UK	0,487	0,786	0,801
	Canadian total unemployment figures	0,362	0,788	0,856
	Sutter county workforce	0,618	0,893	0,854
	U.S. female (20 years and over) unemployment figures	0,540	0,892	0,867
	Number of employed persons in Australia	0,318	0,702	0,851
	Passenger miles flow domestic UK	0,702	0,911	0,896
Utilities	Av. residentail gas usage Iowa	0,633	0,877	0,830
	Total number of consumers	0,615	0,715	0,894

Macro-area	Series	LOESS	Henderson	Spline
Macro-Economics	Consumer price index	0,346	0,719	0,736
Meteorology	Degree days per heating in Chicago	0,541	0,861	0,821
Micro-Economics	Gambling expenditure in Victoria, Australia	0,584	0,902	0,827
	Logged flour price indices over the 9-years	0,558	0,833	0,842
Miscellaneous	Average daily calls to directory assistance	0,607	0,803	0,727
Physics	Zurich sunspot numbers	0,546	0,806	0,274
	Critical radio frequencies in Washington D.C.	0,621	0,887	0,761
	Mean thickness (Dobson units) ozone column Switzerland	0,685	0,791	0,818
Production	Basic iron production in Australia	0,696	0,788	0,035
	Production of chocolate confectionery in Australia	0,638	0,782	0,335
	Production of Portland cement	0,687	0,827	0,261
	Electricity production in Australia	0,783	0,665	0,388
	Production of blooms and slabs in Australia	0,666	0,719	0,605
	Production of blooms and slabs	0,677	0,697	0,827
Sales	Sales of Tasty Cola	0,459	0,774	0,674
	Unit sales, Winnebago Industries	0,594	0,892	0,777
	Sales of new one-family houses sold in US	0,663	0,883	0,834
	Sales for a souvenir shop in Queensland, Australia	0,570	0,888	0,841
	Demand for carpet	0,502	0,778	0,846
Transport and Tourism	Portland Oregon average monthly bus ridership	0,511	0,777	0,712
	U.S air passenger miles	0,546	0,908	0,771
	International airline passengers	0,473	0,768	0,772
	Weekday bus ridership, Iowa city, Iowa (monthly averages)	0,567	0,916	0,873
	Passenger miles flow domestic UK	0,702	0,911	0,896
Utilities	Av. residentail gas usage Iowa	0,633	0,877	0,830
	Total number of consumers	0,615	0,715	0,894

Conclusions

A unified approach for studying different nonparametric smoothers was found within the context of RKHS.

We identified the density function or kernel of order two for LOESS, Henderson filter and the cubic smoothing spline. It provides the "*initial weighting shape*" from which the higher order kernels inherit their properties.

Hierarchies of higher order kernels have been generated via the multiplication of the density functions by their orthonormal polynomials.

Advantages:

- if f_0 is optimal according to a specific smoothing criteria, each kernel of the hierarchy inherits the optimality property at its own order;
- kernel functions can be compared by considering smoothers of different order within the same hierarchy as well as kernels of the same order, but belonging to different hierarchies;
- filters of any length, including the infinite ones.

In real cases the most often applied are estimators of order three, and we calculated their asymmetric last point kernels. A comparison was made with the corresponding classical smoothers. The results showed that the former are superior to the latter in terms signal passing, noise suppression and revisions.

References

Nonparametric estimators

Cleveland W. (1979), Robust Locally Regression and Smoothing Scatterplots, Journal of the American Statistical Association, 74, 829-836.

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

Wahba G. (1990), Spline Models for Observational Data, Philadelphia: SIAM.

RKHS methodology

Parzen E. (1961), An Approach to Time Series Analysis, The Annals of Mathematical Statistics, Vol. 32, No. 4. (Dec., 1961), pp. 951-989.

Berlinet A. and C. Thomas-Agnan (2003), Reproducing Kernel Hilber Spaces in Probability and Statistics, Kluwer Academic Publishers.

Nonparametric estimators in RKHS

Bianconcini S. (2006), Trend-cycle estimation in reproducing kernel Hilbert spaces, Ph.D. Thesis, Department of Statistics, University of Bologna, pp. 147.

Bee Dagum E. and S. Bianconcini (2006), Local Polynomial Trend-Cycle Predictors in Reproducing Kernel Hilbert Spaces for Current Economic Analysis, Anales de Economia Aplicada, pp. 1-22.

Bee Dagum E. and S. Bianconcini (2008), The Henderson Smoother in Reproducing Kernel Hilbert Space, Journal of Business and Economic Statistics, 26 (4), 536-545.