UNIVERSITÀ DELL'AQUILA
ISTITUTO DI ELETTROTECNICA

Research Reports

P. d'ALESSANDRO

OPTIMAL SMOOTHINGS

Report No 01-81 February 1981

Submitted for publication

PAOLO d'ALESSANDRO

OPTIMAL SMOOTHINGS

Paolo d'Alessandro - Istituto di Elettrotecnica -

Università dell'Aquila - Monteluco - 67100 L'Aquila (Italy)

1. INTRODUCTION

It is a customary in time series analysis to assume that a certain variable be represented by a stochastic process plus one or more deterministic components, that, for example, account separately for "cycles" and "trends". These deterministic components are removed by means of different and often fairly complex procedures, which, generally speaking, involve a presumed "smoothing" of the original time series. Such idea of smoothing, defined intuitively rather than formally, appears frequently through the discussion of time series analysis.

However smoothness can be given a precise formal meaning even for "continuous time" series, in which case it would be though not so easy to handle, whereas it is both a rigorous and a handy concept for time series. It is in fact possible to introduce some suitable functionals that give a measure of the degree of smoothness. Once the appropriate quantitative definition of smoothness is thereby obtained, it is possible to introduce optimal smoothings to a certain degree of a time series.

In this way, back to the starting point above, in lieu of a case study in which no variety of different models can entirely capture the complexity of reality and in lieu of a large amounts of methods and filters, a single general method is substituted, not dependent on disaggregation in components, nor on models for each component, and leading to an optimization problem, which yields the sought smoothed series.

2. THE OPTIMAL SMOOTHING APPROACH

Any finite time series is represented as a point of R^N where N is the number of samples. The space R^N is given the usual Hilbert space structure, using the symbols $(.,.)$ and $\|\cdot\|$ to denote the corresponding inner product and its norm respectively.

A large variety of functional on R^N representing smoothness can be defined. The choice depends on practical purposes and technical convenience.

Two example are given here: the first is a pseudonorm the second is a norm (recall that any pseudonorm or norm for R^N is necessarily continuous).

The pseudonorm v, the variation functional is defined as follows: for each f in R^N

$$v(f) = \sum_{i=1}^{N-1} |f(i+1) - f(i)|$$

A modification of v is the norm w defined by

$$w(f) = |f(1)| + v(f)$$

For the sake of brevity it is left to the reader to formulate the numerous other possibilities. For instance a finite linear combination of pseudonorms with positive coefficients is another pseudonorm or possibly a norm.

Once anyway a "smoothness functional" say p, which is either a norm or a pseudonorm, has been choosen then the corresponding optimal smoothing problem can be formulated as follows.

Given a signal, that is an elements of R^N, suppose it is required to reduct its variability, measured by p(s) of, say, c%. To do so it is natural to look for the signal, that is the element $\hat{s}$ of R^N, with $p(\hat{s})$ not exceeding $r = p(s)\, c/100$, which approximates at best (that is having minimum distance from) s.

Formally if S_p^r denotes the closed sphere of radius **r of p** about the origin the above amount to find $\hat{s}$ so as to minimize $\| s-\hat{s} \|$ subject to $\hat{s} \in S_v^r$

Existence and uniqueness of the solution of this convex programming problem is very well-known in view of the fact that S_p^r is closed and convex. It is also known that the solution belongs to the boundary of S_p^r so that the value of p at the solution is just r. (see e.g. /1/ and /2/). The solution is called the projection P(s) of s onto S_p^r.

In view of these result the relevant point is that of actually computing the solution of the above "optimal smoothing" problem, devising to this purpose methods as fast and simple as possible. This important area of research is based on concepts and results of convex analysis.

Various techniques are offered by the literature. The author himself has developed a finite convergence method (that is requaring at most N steps) solving a certain class of convex programming problems, that in particular has allowed the implementation of the smoothing filter based on the functional w. A detail exposition of the relevant theory is given in /3/.

In this latter work much use is made of the polytopic structure of spheres of certain norms. A polytope is the convex estension of a finite subset of R^N, and as such is a compact set. (The convex extension of a set A is denoted by $\mathscr{C}(A)$).

The section is a concluded illustrating the polytopic structure of the closed spheres defined by v and w. It is clearly sufficient to consider spheres of radius 1. Consider the norm p_1 on R^N defined by

$$p_1(f) = \sum_{i=1}^{N} |f(i)| \qquad \text{for each } f \text{ in } R^N$$

The sphere $S^1_{p_1}$ is the convex extension of the set (the notation is self explanatory).

$$E = \{(1,0,\dots,0),\ (-1,0,\dots,0),\dots,(0,0,\dots,-1)\}$$

which is precisely the set of its extreme points.

Let T be the linear isomorphism of R^N onto itself represented by the matrix

$$\hat{T} = \begin{pmatrix} 1 & 0 & \dots & 0 & 0 \\ -1 & 1 & \dots & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \dots & -1 & 1 \end{pmatrix}$$

Then clearly

$$w = p_1 \circ T$$

and therefore

$$S_w^1 = w^{-1}([0,1]) = T^{-1}(p_1^{-1}([0,1])) = T^{-1}(S_{p_1}^1) =$$

$$= T^{-1}(\mathcal{C}(E)) = \mathcal{C}(T^{-1}(E))$$

and again it is quickly proved that the finite set $T^{-1}(E)$ is the set of extreme points of S_w^1. This shows that S_w^1 is a polytope. The trivial explicit computations are left to the reader.

As to the pseudonorm v note that the linear subspace $v^{-1}(\{0\})$ is given by the set of all constant function in R^N. Denote this subspace by F and observe that its orthogonal complement $F^{\perp}$ is given by the linear subspace $\tilde{N}$ of all zero-mean functions. Let $V_{\tilde{N}}$ be the restriction of v on $\tilde{N}$, and note that it is a norm for $\tilde{N}$.

Then if $S^1_{v_{\tilde{N}}}$ is the closed unit-sphere in $\tilde{N}$ defined by $v_{\tilde{N}}$:

$$S_v^1 = N + S^1_{v_{\tilde{N}}}$$

The, so to speak, bounded part of the sphere, i.e. $S^1_{v_{\tilde{N}}}$ is again a polytope for if $L_{|\tilde{N}}$ is the linear isomorphism of $\tilde{N}$ into $\mathbb{R}^{N-1}$, where L is represented by the matrix

$$\hat{L} = \begin{pmatrix} -1 & 1 & \dots & 0 & 0 \\ 0 & -1 & \dots & 0 & 0 \\ \vdots & & & & \\ 0 & 0 & \dots & -1 & 1 \end{pmatrix}$$

then

$$v|_{\tilde{N}} = \tilde{p}_1 \circ L_{|\tilde{N}}$$

where $\tilde{p}_1$ is the norm defined for R^{N-1} in the same way as p_I for R^N.

The remaining trivial details can be left to the reader. Note that in view of the results given in /3/ the filter based on v leaves unchanged the time average of the signal.

3. COMMENTS

In view of existence and uniqueness of solution it is well defined a projection transformation taking the space R^N onto the closed convex subset in question. Denote this map by P. The implementation of P is a filter for finite time series. A major question that arises at this point regards the possibility of studying the stochastic performance of the filter. This requires P to a measurable transformation of R^N (endowed with its Borel σ-algebra) into itself. But it is very well known (see /1/ and /3/) that P is actually (and even in Hilbert space environment) a uniformly continuous transformation, thus measurability is always and largely guaranteed. More than that the filter thereby obtained retains a feature that is typical of all linear filters. If R^N is equiped with a probability measure say p on its Borel σ-algebra the performance of the filter may be studied, for example in case the functional v is adopted, comparing the probability distributions $p \circ v^{-1}$ and $p \circ P^{-1} \circ v^{-1}$.

As to applications, the technique suggested here is concurrent in the first place with the various rather empirical method whereby routine smoothing of time series are performed for a number of purposes (e.g. moving average seasonal adjustements). In addition to that, many other application, like the derivation of new forcasting methods, are possible, and it is expected (and hoped) that interest on both practical and theoretical sides of the research regarding this new family of methods will grow up in the near future.

It is interesting to note, as further comment, that of course the above filters should be considered non-dynamical. However it is questionable that a dynamical filter is always the best solution. In fact sometimes its advantages may not be enough compensation of the drawback, for example, of being forced to pay the smoothing in terms of lags.

Work at theory and implementation of optimal smoothing methods started in the second half of 1979: a first implementation of the filter based on the functional $\mathbf{w}$ goes back to the beginning of 1980, whereas theory and implementation (see next section) of finite convergence method has been completed in september 1980.

4. NUMERICAL RESULTS

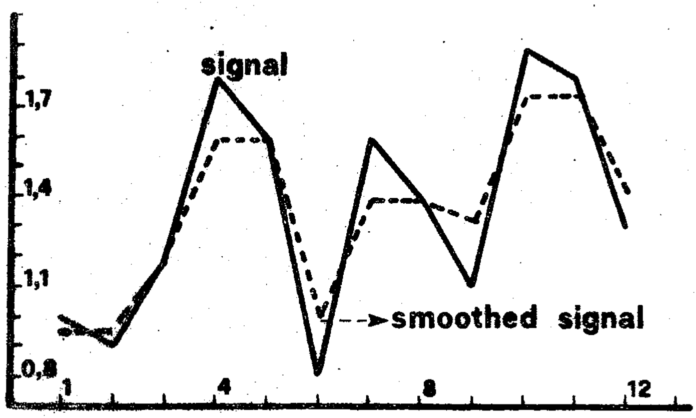
Just to give an idea of how the smoothing works the results on a couple of time series used during development of software are given. The finite convergence method has lead to a fairly small program (less then three hundred lines of Pascal) memory trifty, and amazingly fast even though it was run on a popular "personal computer."

In the tables Filsig stands for filtered signal. The percent refers to the cut in the variation term of the functional w.

FIRST TIME SERIES. N=12.

SIGNAL	FILSIG 10%	FILSIG 20%	FILSIG 30%
1	0.967586	0.950000	0.950000
0.9	0.932414	0.950000	0.950000
1.2	1.200000	1.200000	1.200000
1.8	1.767590	1.729520	1.683160
1.6	1.600000	1.600000	1.600000
0.8	0.832414	0.870476	0.916842
1.6	1.567590	1.529520	1.483160
1.4	1.400000	1.400000	1.400000
1.1	1.132410	1.170480	1.216840
1.9	1.867590	1.829520	1.791580
1.8	1.800000	1.800000	1.791580
1.3	1.316210	1.335240	1.358420

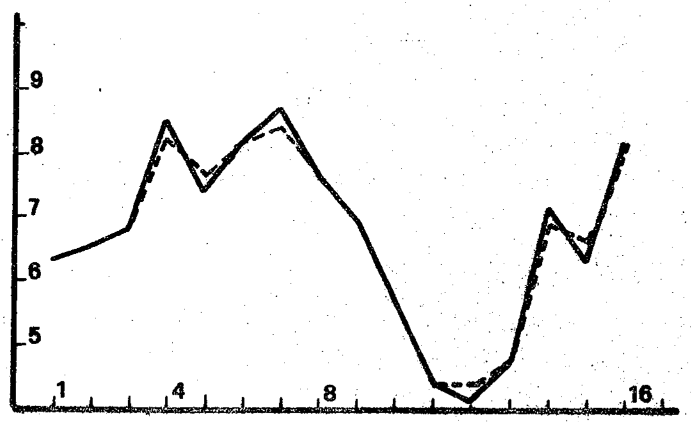
FILSIG 40%	FILSIG 50%	FILSIG 60%	FILSIG 70%
0.950000	0.950000	0.950000	0.950000
0.950000	0.950000	0.950000	0.950000
1.200000	1.200000	1.200000	1.200000
1.633680	1.590000	1.553330	1.501110
1.600000	1.590000	1.553330	1.501110
0.966316	1.020000	1.093330	1.197780
1.433680	1.390000	1.366670	1.366670
1.400000	1.390000	1.366670	1.366670
1.266320	1.320000	1.366670	1.366670
1.766840	1.740000	1.703330	1.651110
1.766840	1.740000	1.703330	1.651110
1.383160	1.410000	1.446670	1.498890



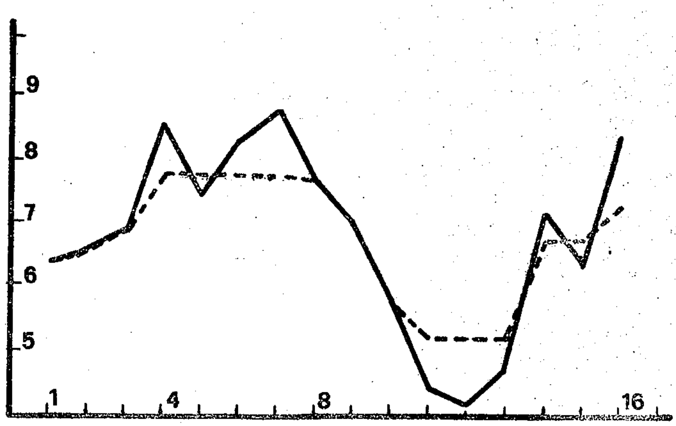
50% of variation cut

SECOND TIME SERIES. N=16

SIGNAL	FILSIG 20%	FILSIG 40%	FILSIG 60%
6.31680	6.31680	6.31680	6.31680
6.53190	6.53190	6.53190	6.53190
6.82000	6.82000	6.82000	6.82000
8.56190	8.32289	8.01310	7.70839
7.46430	7.70331	8.01310	7.70839
8.27670	8.27670	8.21099	7.70839
8.71730	8.47829	8.21099	7.70839
7.68960	7.68960	7.68960	7.68960
6.97510	6.97510	6.97510	6.97510
5.75990	5.75990	5.75990	5.75990
4.44870	4.44870	4.60826	5.19581
4.19580	4.43481	4.60826	5.19581
4.75630	4.75630	4.75630	5.19581
7.18720	6.94819	6.77310	6.77310
6.35900	6.59801	6.77310	6.77310
8.36010	8.24060	8.07409	7.26679



20% of variation cut



60% of variation cut

REFERENCES

/1/ J.L. KELLEY, I. NAMIOKA et AL., "Linear topological spaces", Springer Verlag, 1963.

/2/ A.V. BALAKRISHNAN, "Applied Functional Analysis", Springer Verlag, 1976.

/3/ P. d'ALESSANDRO, "Finite convergence method to solve certain projection problems". Report No 02-81, Feb. 1981, Istituto di Elettrotecnica, Università dell'Aquila.

Finito di stampare nel mese di febbraio 1981
dall'Artigiana Multistampa Snc
Via Ruggero Bonghi, 36 - Roma