GENERALIZED ADAPTIVE EXPONENTIAL SMOOTHING OF ERGODIC MARKOVIAN OBSERVATION SEQUENCES

ULRICH HERKENRATH[1],

Gerhard-Mercator-Universität — Gesamthochschule Duisburg

Abstract

An exponential smoothing procedure applied to a homogeneous Markovian observation sequence generates an inhomogeneous Markov process as sequence of smoothed values. If the underlying observation sequence is moreover ergodic then for two classes of smoothing functions the strong ergodicity of the sequence of smoothed values is proved. As a consequence a central limit theorem and a law of large numbers hold true for the smoothed values. The proof uses general results for so-called convergent inhomogeneous Markov processes. In the literature a lot of time series are discussed to which the smoothing procedures are applicable.

SMOOTHING OF NON-LINEAR TIME SERIES; GENERALIZED EXPONENTIAL SMOOTHING; CONVERGENT INHOMOGENEOUS MARKOV PROCESSES

AMS 1991 SUBJECT CLASSIFICATION: PRIMARY 60J20
SECONDARY 62M10

---

[1]Postal address: Fachbereich 11 Mathematik, Gerhard-Mercator-Universität — GH Duisburg, Postfach 10 15 03,   D-47048 Duisburg

1. Introduction

Exponential smoothing procedures are well-known and widely used to construct sequentially prognoses and "means" $(W_n, n \geq 1)$ for a sequence of observations $(X_n, n \geq 1)$, especially time series, see e.g. Bonsdorff (1989), Gardner (1985). Two kinds of generalizations of the classical case are possible: First as to the smoothing function, the usual mean $(1 - \lambda)W_n + \lambda X_n$ of last smoothed value W_n and actual observation X_n with constant weight λ may be substituted by the value $u(W_n, X_n)$ of a much more general function u. Next as to the structure of the observation sequence, the case of i.i.d. variables may be extended. More general smoothing functions are dealt in Bonsdorff (1989) and Herkenrath (1994a), in the latter article moreover a dependence of the distribution of X_n from W_n is allowed. There so-called compact smoothing systems are introduced and dealt.

In this article general smoothing functions are regarded too, but furthermore they are applied to an observation sequence, which may constitute an ergodic homogeneous Markov process. This seems to be a reasonable extension with regard to a dependence within the sequence of observations. Many important examples of Markovian time series are e.g. presented and discussed in Meyn and Tweedie (1996) and Doukhan (1994). In particular the ARCH- and GARCH-time series have attracted much attention in the last ten years in the context of economics.

Whereas for i.i.d. observations $(X_n, n \geq 1)$ the corresponding sequence of smoothed values $(W_n, n \geq 1)$ turns out to be a homogeneous Markov process, in the case of a Markovian observation sequence $(X_n, n \geq 1)$ the sequence $(W_n, n \geq 1)$ constitutes an inhomogeneous Markov process. This kind of Markov process is more difficult to handle than the homogeneous one.

For a class of smoothing functions, which satisfy certain contraction conditions, the case of a Markovian observation sequence is treated in Herkenrath (1994 b). In that context the analysis of the sequence of smoothed values $(W_n, n \geq 1)$ is achieved by the study of the homogeneous Markov process $((W_n, X_n), n \geq 1)$. That kind of assumption on u allows to exploit a certain ergodic theorem according to the theory of "compact Markov processes" in the sense of Norman (1972). For different types of functions u, which do not obey a contraction condition, the method described in Herkenrath (1994 b) does not work. Therefore the sequence of smoothed values $(W_n, n \geq 1)$ has to be studied as an inhomogeneous Markov process.

Since stability properties for the sequence $(W_n, n \geq 1)$ are desired, we ask for appropriate concepts for inhomogeneous Markov processes. In this context there are some notions of ergodicity for inhomogeneous Markov processes which are discussed e.g. by Iosifescu (1980). The most essential property is the so-called strong ergodicity, which means that the distribution of the state of the process converges to a limiting distribution in the norm of total variation uniformly for all starting points.

This stability property of the sequence of smoothed values $(W_n, n \geq 1)$ is proved for two classes of smoothing functions u applied to ergodic homogeneous Markovian observation sequences. Additionally a law of large numbers and a central limit theorem hold true for

the sequence $(W_n, n \geq 1)$. The method to prove the results is to show that the sequence $(W_n, n \geq 1)$ represents a so-called convergent inhomogeneous Markov process with a mixing limit. Then the desired results for the sequence $(W_n, n \geq 1)$ follow according to Herkenrath (1998).

2. The smoothing procedure and associated stochastic processes

As mathematical model for our analysis of the smoothing procedure we refer to the so-called compact smoothing system (CSS) introduced in Herkenrath (1994 a). We modify that system to a compact Markovian smoothing system (CMSS).

Definition. A (time-homogeneous) CMSS is a tuple $\{W, X, u, P\}$, where

(i) $W = [a, b]$ is a compact real interval, $a, b \in \mathbb{R}$, endowed with the Borel-σ-algebra $\mathfrak{W}$,

(ii) X is a real interval (which may be unbounded) with the Borel-σ-algebra $\mathfrak{X}$,

(iii) $u : (W \times X, \mathfrak{W} \otimes \mathfrak{X}) \to (W, \mathfrak{W})$ is a measurable function, $\mathfrak{W} \otimes \mathfrak{X}$ denotes the product-σ-algebra,

(iv) $P : (X, \mathfrak{X}) \times \mathfrak{X} \to [0, 1]$ is a transition probability on $(X, \mathfrak{X})$.

X denotes the set of possible observations or events, W the range of possible smoothed values or states of the system, u represents the smoothing or transition function of the system, which transforms a given smoothed value w and an actual observation x into a new smoothed value $u(w, x)$. In contrast to the CSS in Herkenrath (1994 a) here P denotes a transition probability on the set of observations which generates a homogeneous Markov process of observations $(X_n, n \geq 1)$. This explains the notation CMSS.

According to the theorem of Ionescu Tulcea, for each $(w, x) \in (W \times X)$ there exist a probability measure $\mathbb{P}_{w,x}$ on $((W \times X)^{\mathbb{N}}, (\mathfrak{W} \otimes \mathfrak{X})^{\mathbb{N}}) = (\Omega, \mathfrak{K})$ and random variables

$$
\begin{aligned}
W_n : (\Omega, \mathfrak{K}) &\to (W, \mathfrak{W}), \\
X_n : (\Omega, \mathfrak{K}) &\to (X, \mathfrak{X}), \qquad n \geq 1
\end{aligned}
$$

such that for $n \geq 1$, $A \in \mathfrak{W}$, $B \in \mathfrak{X}$

$$
\begin{aligned}
\mathbb{P}_{w,x}(W_1 = w) &= 1, \\
\mathbb{P}_{w,x}(X_1 = x) &= 1,
\end{aligned}
$$

$$
\mathbb{P}_{w,x}(W_{n+1} \in A \mid W_1, X_1, W_2, \ldots, W_n, X_n) = \delta_{u(W_n, X_n)}(A),
$$

$$
\mathbb{P}_{w,x}(X_{n+1} \in B \mid W_1, X_1, W_2, \ldots, X_n, W_{n+1}) = P(X_n, B),
$$

where δ_v denotes the one-point measure on $\{v\}$. Furthermore $\mathbb{E}_{w,x}$ and $\mathbb{V}_{w,x}$ represent the mathematical expectation respectively the variance with respect to $\mathbb{P}_{w,x}$. So $(X_n, n \geq 1)$ is the Markovian observation sequence generated by P with starting point $X_1 = x$. $(W_n, n \geq 1)$ with starting point $W_1 = w$ constitutes the sequence of smoothed values which is sequentially generated by the function u, since $W_{n+1} = u(W_n, X_n)$, $n \geq 1$.

3

The two stochastic processes $(W_n, n \geq 1)$ and $(X_n, n \geq 1)$ are linked together as shown in Figure 1.
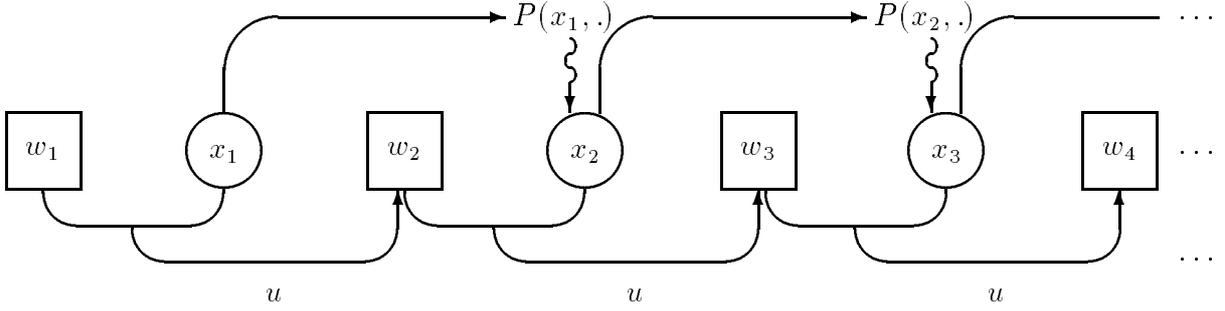

Figure 1.

$(X_n, n \geq 1)$ is a homogeneous Markov process by definition, $((W_n, X_n), n \geq 1)$ has this property too where its transition probability is given by

$$R((w,x), A \times B) = \delta_{u(w,x)}(A) \cdot P(x,B).$$

$(W_n, n \geq 1)$ turns out to constitute an inhomogeneous Markov process with one-step transition probability $Q^{m,m+1}$ at time m, where $Q^{m,m+1}$ is determined by

$$\mathbb{P}_{w,x}(W_{m+1} \in A \mid W_m = v) = \mathbb{P}_{w,x}(X_m \in u_v^{-1}(A)) = P^{m-1}(x, u_v^{-1}(A)) = Q^{m,m+1}(v,A)$$

for $m \geq 1, A \in \mathfrak{W}, v \in W, P^{m-1}$ the $(m-1)$-step transition probability induced by P and where $u_v : (X, \mathfrak{X}) \to (W, \mathfrak{W})$ is defined by $u_v(x) = u(v,x)$.

This representation of $Q^{m,m+1}$ shows: If the sequence of iterates $(P^m, m \geq 1)$ of P converges e.g. in the norm of total variation to some probability π, then the sequence of one-step transition probabilities $(Q^{m,m+1}, m \geq 1)$ converges, too. This leads to the concept of so-called convergent inhomogeneous Markov processes, which was introduced by Mott (1959) and is dealt in Herkenrath (1998) for a general state space.

The essential concept for stability of an inhomogeneous Markov process $(W_n, n \geq 1)$ with one-step transition probabilities $(Q^{m,m+1}, m \geq 1)$ on a state space $(W, \mathfrak{W})$ is the so-called strong ergodicity (see e.g. Iosifescu (1980)), which claims that there exists a probability measure π on $(W, \mathfrak{W})$ such that

$$\lim_{n \to \infty} \sup_{w \in W} \parallel Q^{m,m+n}(w, \cdot) - \pi(\cdot) \parallel = 0$$

for all $m \geq 1$. Here, for a signed measure ν on $(W, \mathfrak{W})$, $\parallel \nu \parallel = \frac{1}{2}(\sup_{A \in \mathfrak{W}} \nu(A) - \inf_{A \in \mathfrak{W}} \nu(A))$ denotes $\frac{1}{2} \times$ total variation of ν. Furthermore $Q^{m,m+n}$ is the n-step transition probability of the process starting at time m, i.e.

$$Q^{m,m+n}(v,A) = \mathbb{P}_{w,x}(W_{m+n} \in A \mid W_m = v),$$

$v, w \in W, x \in X, A \in \mathfrak{W}$. In particular $Q^{1,1+n}(w, \cdot)$ represents the distribution of W_{n+1}, if the sequence of smoothed values starts in w.

4

For a homogeneous Markov process $(X_n, n \geq 1)$ on the measurable space $(X, \mathfrak{X})$ with transition probability P strong ergodicity amounts to

$$\lim_{n \to \infty} \sup_{x \in X} \| P^n(x, \cdot) - \pi(\cdot) \| = 0$$

for a probability measure π. This property of a homogeneous Markov process respectively a transition probability is named in the literature strong ergodicity or uniform ergodicity. In accordance with e.g. Iosifescu and Grigorescu (1990), Meyn and Tweedie (1996) we call it uniform ergodicity and a corresponding process uniformly ergodic.

The weaker concept of ergodicity means according to Meyn and Tweedie (1996) that for all $x \in X$

$$\lim_{n \to \infty} \| P^n(x, \cdot) - \pi(\cdot) \| = 0$$

holds true for a unique invariant probability measure π. If moreover the above convergence of the $(P^n, n \geq 1)$ to π is of the kind

$$\| P^n(x, \cdot) - \pi(\cdot) \| = M(x) r^n$$

for some positive function M with $\int_X M(x)\, \pi(dx) < \infty$ and some $0 < r < 1$ then Meyn and Tweedie (1996) call it geometric ergodicity. A lot of examples of ergodic respectively geometrically ergodic Markov processes are given by Meyn and Tweedie (1996), Doukhan (1994), Tong (1990), Tjostheim (1990).

In Meyn and Tweedie (1996, p. 384) several equivalent conditions for uniform ergodicity are presented. One of them, namely (iii) there, is formulated by means of the so-called ergodicity coefficient of the transition probability P, see e.g. Iosifescu and Theodorescu (1969, p. 60). P is called mixing iff there exists $n_0 \in \mathbb{N}$ such that for the ergodicity coefficient $\alpha(P^{n_0})$ of the iterate P^{n_0} it holds true

$$\alpha(P^{n_0}) = 1 - \sup_{x', x'' \in X, B \in \mathfrak{X}} | P^{n_0}(x', B) - P^{n_0}(x'', B) | > 0 \,.$$

As mentioned above uniform ergodicity of the homogeneous Markov process is equivalent to mixing of its transition probability. The context of the ergodicity coefficient and different concepts of ergodicity is extensively treated in the books cited above.

An important example for a mixing transition probability is given by the following Lemma.

Lemma 1. Let P be a transition probability on $(X, \mathfrak{X})$, where $X \subset \mathbb{R}$ and $\mathfrak{X}$ is the Borel-σ-algebra, with a Lebesgue density p. If there is a finite interval $[\alpha, \beta] \subset X$ such that

$$\exists\, \gamma > 0 \quad \forall\, x \in X, y \in [\alpha, \beta] \ : \ 0 < \gamma \leq p(x, y) \,,$$

then the transition probability P is mixing.

Proof. For $B \in \mathfrak{X}$ let denote $B_I = B \cap [\alpha, \beta]$, $B_0 = B \setminus B_I$, $B_c = [\alpha, \beta] \setminus B_I$, and ℓ the Lebesgue measure.
Now, in case $\ell(B_c) > \frac{1}{2}\, \ell([\alpha, \beta])$ it follows:

$$P(x, B) \leq 1 - \gamma\, \ell(B_c) \leq 1 - \frac{1}{2} \gamma\, \ell([\alpha, \beta]),$$

which implies for $\delta = \frac{1}{2}\gamma\,\ell([\alpha,\beta])$ the estimate $\sup\limits_{x',x''\in X}\mid P(x',B) - P(x'',B)\mid\leq 1 - \delta$.

If $\ell(B_c) \leq \frac{1}{2}\ell([\alpha,\beta])$, one can estimate

$$
\begin{aligned}
P(x,B) \;&\geq\; P(x,B_I) = P(x,[\alpha,\beta]\setminus B_c)\\
&\geq\; \gamma\,\ell([\alpha,\beta]\setminus B_c) = \gamma\{\ell([\alpha,\beta]) - \ell(B_c)\} \geq \tfrac{1}{2}\gamma\ell([\alpha,\beta]) = \delta,
\end{aligned}
$$

which also implies $\sup_{x',x''\in X}\mid P(x',B) - P(x'',B)\mid\leq 1 - \delta$.

Therefore $\alpha(P) \geq \delta > 0$. $\square$

Now we come back to the study of the CMSS.

A first important consequence of ergodicity of the observation sequence for the sequence of smoothed values $(W_n, n \geq 1)$ is:

Lemma 2. Consider a CMSS with an ergodic observation sequence $(X_n, n \geq 1)$ with limiting distribution π.

(i) Then the inhomogeneous Markov process $(W_n, n \geq 1)$ of smoothed values is convergent in the following sense: Its one-step transition probabilities $Q^{m,m+1}$ converge to the transition probability Q^∞ on $(W, \mathfrak{W})$ defined by

$$
Q^\infty(w, A) = \pi(u_w^{-1}(A)), \quad w \in W, A \in \mathfrak{W}
$$

in the sense

$$
\lim_{m \to \infty}\ \sup_{w \in W}\ \|\, Q^{m,m+1}(w,\cdot) - Q^\infty(w,\cdot)\,\| = 0.
$$

(ii) If moreover Q^∞ is mixing, then the sequence $(W_n, n \geq 1)$ is strongly ergodic with a limit ρ, which is determined as the unique stationary distribution corresponding to Q^∞.

(iii) In the case of a mixing limit Q^∞ in addition a law of large numbers and a central limit theorem hold true for a sequence $(Y_n = F_n(W_n), n \geq 1)$ of real-valued measurable functions of W_n:

$$
\sum_{n=1}^{\infty} \frac{\mathbb{V}_{w,x}(Y_n)}{n^2} < \infty \quad \Longrightarrow \quad \frac{S_n - \mathbb{E}_{w,x}(S_n)}{n} \xrightarrow[n\to\infty]{} 0 \quad \mathbb{P}_{w,x}\text{-}a.s.
$$

with $S_n = \sum_{i=1}^{n} Y_i, n \geq 1$, for each $(w,x) \in W \times X$ as starting points of the associated processes and,
under the conditions

$$
\mid Y_n \mid\leq C < \infty
$$

$$
n^{-2/3}\,(\min_{1 \leq i \leq n}\ \alpha\,(Q^{i,i+1}))\sum_{i=1}^{n}\mathbb{V}_{w,x}(Y_i) \xrightarrow[n\to\infty]{} \infty,
$$

it holds true

$$
\lim_{n \to \infty} \mathbb{P}_{w,x}\left(\frac{S_n - \mathbb{E}_{w,x}(S_n)}{\sqrt{\mathbb{V}_{w,x}(S_n)}} < a\right) = \phi(a) \quad \text{for all } a \in \mathbb{R},
$$

where ϕ denotes the distribution function of the standard normal distribution.

Proof. Since for each starting point $x \in X$ of the Markov process $(X_n, n \geq 1)$ it holds true

$$\lim_{m \to \infty} \| P^m(x, \cdot) - \pi(\cdot) \| = 0 \,,$$

the convergence of the sequence $(Q^{m,m+1}, m \geq 1)$ to Q^∞ uniformly in $w \in W$ follows from

$$\sup_{w \in W} \sup_{A \in \mathfrak{W}} | \, Q^{m,m+1}(w, A) - Q^\infty(w, A) \, | = \sup_{w \in W} \sup_{A \in \mathfrak{W}} | \, P^{m-1}(x, u_w^{-1}(A)) - \pi(u_w^{-1}(A)) \, |$$

$$\leq \sup_{B \in \mathfrak{X}} | \, P^{m-1}(x, B) - \pi(B) \, | = \| \, P^{m-1}(x, \cdot) - \pi(\cdot) \, \| \,.$$

The other statements are valid according to Theorem 6 in Herkenrath (1998). □

The above concept of convergent inhomogeneous Markov processes was introduced by Mott (1959) for the case of a finite state space and is dealt by Herkenrath (1998) for a general state space. In the latter article also further consequences of strong ergodicity and relations to different ergodicity properties are shown.

In order to ensure desirable properties for the sequence of smoothed values $(W_n, n \geq 1)$ of the CMSS $\{W, X, u, P\}$, according to the above Lemma we look for such conditions on u and P, which induce a convergent inhomogeneous Markov process $(W_n, n \geq 1)$ with a mixing limit Q^∞. Because of the representation of Q^∞ by means of u and π the problem is reduced to the study of the CSS $\{W, X, u, \pi\}$:

Corollary. Consider a CMSS $\{W, X, u, P\}$ with a P, that generates an ergodic observation sequence with a corresponding unique stationary distribution π, next the CSS $\{W, X, u, \pi\}$. If the smoothing function u induces a uniformly ergodic Markov process $(\widetilde{W}_n, n \geq 1)$ associated to the CSS $\{W, X, u, \pi\}$, then (i), (ii) and (iii) of Lemma 2 are valid.

Proof. (i) is obviously valid. Let Q be the one-step transition probability of the homogeneous Markov process $(\widetilde{W}_n, n \geq 1)$ associated to the CSS $\{W, X, u, \pi\}$. Since $Q(w, A) = \pi(u_w^{-1}(A)) = Q^\infty(w, A)$, $w \in W, A \in \mathfrak{W}$, the mixing property of Q^∞ is the one for Q. Again a mixing Q is equivalent to a uniformly ergodic Markov process $(\widetilde{W}_n, n \geq 1)$. Therefore the statements of (ii) and (iii) hold true too. □

Due to Lemma 2 and its Corollary the analysis of the sequence of smoothed values $(W_n, n \geq 1)$ of a CMSS $\{W, X, u, P\}$ with an ergodic observation sequence with a limiting distribution π can be reduced to that of the CSS $\{W, X, u, \pi\}$.

3. Strongly ergodic smoothed values

For a CSS sufficient conditions for a uniformly ergodic associated Markov process of smoothed values are given in Herkenrath (1994a). Thus we ask for conditions on u and P to meet the theorems in that paper.

A first class of smoothing functions which induce a strongly ergodic sequence of smoothed values $(W_n, n \geq 1)$ is characterized by an only "linear influence" of the actual observation x on the following smoothed value.

Theorem 1. Consider a CMSS with $W \subset X$ and let the smoothing function u be given by

$$u(w, x) = Proj_W\left[(1 - \lambda(w))w + \lambda(w)x\right],$$

where $Proj_W[w']$ denotes the projection of $w' \in \mathbb{R}$ onto W and

$$\lambda : W \rightarrow [p, 1], \quad 0 < p < 1,$$

is continuous.

The transition probability P should induce an ergodic observation sequence $(X_n, n \geq 1)$ with a limiting distribution π and moreover satisfy:

$$W \subset \operatorname{supp} P(x, \cdot) \quad \text{for all } x \in X,$$

where supp denotes the support of a measure, and

$$\exists \, \gamma > 0 \quad \forall x \in X, B \in \mathfrak{X} : \quad P(x, B) \leq \gamma \ell(B),$$

where ℓ denotes the Lebesgue measure on X.

Under these conditions the induced inhomogeneous Markov process $(W_n, n \geq 1)$ of smoothed values is strongly ergodic with a unique limiting stationary distribution ρ given by

$$\rho(A) = \int\limits_W \pi(u_w^{-1}(A)) \, \rho(dw), \quad A \in \mathfrak{W}.$$

Moreover the law of large numbers and the central limit theorem described in Lemma 2 hold true.

Proof. According to the Corollary of Lemma 2 it only remains to show that the CSS $\{W, X, u, \pi\}$ induces a uniformly ergodic associated Markov process $(\widetilde{W}_n, n \geq 1)$. Since $W \subset \operatorname{supp}P(x, \cdot)$ for all $x \in X$,

$$\pi(W) = \lim_{n \to \infty} P^n(x, W) = \lim_{n \to \infty} \int\limits_X P(x', W) P^{n-1}(x, dx') > 0$$

and

$$\pi(B) = \lim_{n \to \infty} P^n(x, B) = \lim_{n \to \infty} \int\limits_X P(x', B) P^{n-1}(x, dx') \leq \gamma \ell(B)$$

for all $x \in X, B \in \mathfrak{X}$, π satisfies the conditions of Theorem 1 in Herkenrath (1994a) as u does too. That Theorem in turn ensures the uniform ergodicity of $(\widetilde{W}_n, n \geq 1)$. $\quad \square$

A different class of smoothing functions u is covered by the following Theorem.

Theorem 2. Let $\{W, X, u, P\}$ be a CMSS with the following properties:
For each $w \in W$, X is the direct sum of countably many bounded open sets $X_a(w), X_b(w)$, $X_1(w), X_2(w), \ldots$ and a set of Lebesgue measure 0 such that $u(w, \cdot)$ is continuously differentiable with a strictly positive finite derivative on all $X_i(w), i \in \mathbb{N}$, i.e.

$$0 < c \leq \frac{d}{dx} u_w(x) < \infty \quad \text{for some } c \in \mathbb{R},$$

8

$X_a(w), X_b(w)$ *are the open kernels of the sets of all* $x \in X$, *for which* $u(w, x) = a$
respectively $u(w, x) = b$. *These sets may be non-empty, if e.g. u contains a projection
onto W. Moreover all mappings $u(w, \cdot)$ are continuous and increasing.*

Furthermore there exists a closed interval $[\alpha, \beta] \subset X$ *such that:*

$$u(\cdot, \alpha) \text{ is continuous}, \quad u(a, \alpha) = a, \quad u(w, \alpha) < w \quad \text{for } w > a,$$
$$u(\cdot, \beta) \text{ is continuous}, \quad u(b, \beta) = b, \quad u(w, \beta) > w \quad \text{for } w < b.$$

*The transition probability P should induce an ergodic observation sequence $(X_n, n \geq 1)$
with a limiting distribution π and moreover have a bounded Lebesgue density p which is
positive on $[\alpha, \beta]$ for all $x \in X$, i.e.*

$$P(x, B) = \int_B p(x, y) \, \ell(dy), \quad B \in \mathfrak{X},$$

with

$$p(x, y) \leq C < \infty$$

and

$$\forall x \in X, y \in [\alpha, \beta] : \quad 0 < p(x, y).$$

*Then the inhomogeneous Markov process $(W_n, n \geq 1)$ of smoothed values is strongly
ergodic with a limiting distribution ρ as in Theorem 1. Moreover the law of large numbers
and the central limit theorem described in Lemma 2 hold true.*

*This result remains unchanged if in the case of a finite interval $X = [\alpha, \beta]$ the $P(x, \cdot)$
are determined by a bounded Lebesgue density only in the interior of X and give positive
mass to the boundary points α and β.*

Proof. For the CSS $\{W, X, u, \pi\}$ the conditions of Lemma 2 and Theorem 3 in Herkenrath
(1994 a) are satisfied:
$[\alpha, \beta]$ is contained in the support of all $P^n(x, \cdot)$ for all $n \in \mathbb{N}, x \in X$, therefore in the
support of the limit π too and π has a bounded Lebesgue density, since all $P^n(x, \cdot)$ have
it. According to that Theorem the associated Markov process of the CSS $\{W, X, u, \pi\}$
is uniformly ergodic, which proves the statement due to the Corollary.
The supplement to the statement is true, because the above cited Theorem 3 remains
true. $\qquad\square$

For a discussion of the conditions on u, see Herkenrath (1994a).

An important question on a CMSS is of course, whether the smoothing function u induces
such smoothed values $(W_n, n \geq 1)$ which approach "in the mean" the corresponding mean
of observations, i.e. for which $\lim_{n\to\infty} \mathbb{E}_{w,x}[W_n] = c = \int_X x\pi(dx) = \mu$. For CMSS with a
strongly ergodic sequence $(W_n, n \geq 1)$ we prove:

*Theorem 3. Let $\{W, X, u, P\}$ be a CMSS which satisfies the conditions of one of the
Theorems 1 or 2, i.e. induces a strongly ergodic sequence of smoothed values $(W_n, n \geq 1)$.
Let the function $F : W \to W$ be defined by*

$$F(w) = \int_X u(w, x) \, \pi(dx)$$

π being the unique stationary distribution of P.

The following statements are valid:

(i) $F(a) > a, \quad F(b) < b.$

(ii) If F is linear then there is a unique fixed point w_F of F and $\lim\limits_{n \to \infty} \mathbb{E}_{w,x}[W_n] = c = w_F$ for all $(w,x) \in (W \times X)$. Therefore the condition $F(\mu) = \mu$ implies

$$\lim_{n \to \infty} \mathbb{E}_{w,x}[W_n] = c = w_F = \mu = \int_X x\pi(dx) \; .$$

(iii) If F is convex, then there is a unique fixed point w_F and $c \geq w_F$.

(iv) If F is concave then there is a unique fixed point w_F and $c \leq w_F$.

Proof. Since according to Theorem 1 respectively 2 there holds true

$$\lim_{n \to \infty} \sup_{w \in W} ||Q^{1,1+n}(w, \cdot) - \rho(\cdot)|| = 0$$

and W is bounded, $\mathbb{E}_{w,x}[W_n]$ converges to a limit $c \in W$ independently of the starting points $(w,x) \in (W \times X)$:

$$\mathbb{E}_{w,x}[W_{n+1}] = \int_W w' \, Q^{1,1+n}\,(w, dw') \xrightarrow[n \to \infty]{} \int_W w'\rho(dw') = c.$$

To be able to relate the limit c to the "asymptotic observation mean" $\mu = \int_X x\,\pi(dx)$, we again consider the CSS $\{W, X, u, \pi\}$ with the associated Markov process $(\widetilde{W}_n, n \geq 1)$, $\widetilde{W}_1 = w$ and corresponding probability $\mathbb{P}^\pi_w$ and expectation $\mathbb{E}^\pi_w$.

Because of the uniform ergodicity of the sequence $(\widetilde{W}_n, n \geq 1)$, which is guaranteed by Theorem 1 respectively 2, the sequence $(\mathbb{E}^\pi_w[\widetilde{W}_n], n \geq 1)$ converges too, even to the same limit c. This follows from

$$|\mathbb{E}_{w,x}[W_{n+1}] - \mathbb{E}^\pi_w[\widetilde{W}_{n+1}]| = |\int_W w' \, Q^{1,1+n}(w, dw') - \int_W w' \, Q^{\infty n}(w, dw')|$$

$$\leq b \; ||Q^{1,1+n}(w, \cdot) - Q^{\infty n}(w, \cdot)|| \leq b \; \{||Q^{1,1+n}(w, \cdot) - \rho(\cdot)|| + ||\rho(\cdot) - Q^{\infty n}(w, \cdot)||\},$$

taking into account that the right side converges to 0 for $n \to \infty$, since ρ is the limit of the iterates $(Q^{\infty n}, n \geq 1)$ of the mixing transition probability Q^∞.

As a consequence c can be analyzed within the CSS $\{W, X, u, \pi\}$ as $\lim\limits_{n \to \infty} \mathbb{E}^\pi_w[\widetilde{W}_n]$. Therefore the statements immediately follow from Theorem 4 in Herkenrath (1994 a). $\qquad \square$

4. Applications and examples

If one wants to apply the CMSS to smoothen an ergodic Markovian observation sequence one should take into account: First X as set of possible observations is given by the

problem, maybe as an unbounded interval, e.g. as $I\!R_+$. Next the statistician has to choose a compact interval $W = [a, b]$ as domain of possible smoothed values. With respect to the problem, the experience with the observed data, the aims or the use of the smoothed values the statistician should be able to choose a suitable W, which covers the essential range of observations. If W is set too narrow too much information and variation of the observations gets lost. In case the smoothing function u contains a projection operator as for $u(W_n, X_n) = Proj_W [(1 - \lambda) W_n + \lambda X_n]$ and this projection becomes active in many steps, the influence of the observations on the smoothed values becomes even more difficult. A large W and a smoothing function without a projection operator are preferable. Occasionally X should be modeled as compact interval and then W chosen equal to X with the device that an observation value exceeding X is regarded as an outlier and truncated to the corresponding boundary point of X.

Time series which represent ergodic Markov processes are dealt by several authors within different contexts, see Tong (1990),Tjostheim (1990), Doukhan (1994), Meyn and Tweedie (1996). There various types of sufficient conditions for ergodicity or even geometric ergodicity are presented. Of special interest within this class of time series are the so-called ARCH- and GARCH-processes which arise in modern finance theory. These processes show extreme variations of their values and thus claim for smoothing techniques to learn something about their essential parameters.

As for examples of smoothing functions u we refer to Herkenrath (1994 a). Here we list the most important ones for the case $W = X = [0, 1]$:

(SF1) *Classical exponential smoothing*

$$u(w, x) = (1 - \lambda)w + \lambda x, \quad 0 < \lambda < 1$$

(SF2) *Weight* $\quad \lambda(w) = e^{-\lambda w}$ *for x*

$$u(w, x) = (1 - e^{-\lambda w}) w + e^{-\lambda w} x, \quad 0 < \lambda \leq 1$$

(SF3) *Weight for x depending on* $|w - x|$

$$u(w, x) = (1 - e^{-\lambda |w - x|})w + e^{-\lambda |w - x|} x, \quad 0 < \lambda \leq 1$$

(SF4) *Quadratic smoothing*

$$u(w, x) = w + \lambda(x - \tfrac{1}{2})(1 - w)^2 - \lambda(w - \tfrac{1}{2})(1 - x)^2, \quad 0 < \lambda \leq 1$$

Concerning the assumptions on u Theorem 1 covers (SF1) and (SF2), Theorem 2 (SF1), (SF2), (SF3) and (SF4). An analysis of the associated function F defined in Theorem 3 is done in Herkenrath (1994 a) as well as a discussion of the various types of u. The most convenient procedure under theoretical aspects is Classical exponential smoothing (SF1) for which F is linear and moreover $'F(\mu) = \mu'$ holds true which implies $\lim_{n \to \infty} I\!E_{w,x}[W_n] = \mu$.

The idea for the selection or construction of smoothing functions u is to prescribe certain structural properties for u, e.g. for the mappings $u_x(\cdot)$ like convexity or concavity. Such properties and parameters of a smoothing function, as e.g. λ in (SF1) to (SF4), have to be chosen adapted to the problem.

References

BONSDORFF, H. (1989) A comparison of the ordinary and a varying parameter exponential smoothing. *J. Appl. Prob.* 26, 784–792.

DOUKHAN, P. (1994) *Mixing: Properties and Examples.* Lecture Notes in Statistics. Springer, New York.

GARDNER, E. S. JR. (1985) Exponential smoothing: the state of the art. *J. Forecasting* 4, 1–28.

HERKENRATH, U. (1994 a) Generalized adaptive exponential smoothing procedures. *J. Appl. Prob.* 31, 673–690.

HERKENRATH, U. (1994 b) Premium adjustment by generalized adaptive exponential smoothing. *Insurance: Mathematics and Economics* 15, 203–217

HERKENRATH, U. (1998) On ergodic properties of inhomogeneous Markov processes, *to appear in Rev. Roumaine Math. Pures Appl. No. 5–6.*

IOSIFESCU, M. (1980) *Finite Markov Processes and Their Applications.* Wiley, New York.

IOSIFESU, M. AND GRIGORESCU, S. (1990) *Dependence with Complete Connections and its Applications.* Cambridge University Press.

IOSIFESCU, M. AND THEODORESCU, R. (1969) *Random Processes and Learning.* Springer, Berlin.

MEYN, S.P. AND TWEEDIE, R.L. (1996) *Markov Chains and Stochastic Stability.* Springer, London.

MOTT, J.L. (1959) The central limit theorem for a convergent nonhomogeneous finite Markov chain. *Proc. Roy. Soc. Edinburgh Sect. A* 65, 109–120.

NORMAN, M.F. (1972) *Markov Processes and Learning Models.* Academic Press, New York.

TJOSTHEIM, D. (1990) Non-linear time series and Markov chains. *Adv. Appl. Prob.* 22, 587–611.

TONG, H. (1990) *Non-linear Time Series: A Dynamical System Approach.* Oxford University Press, Oxford.