ESTIMATION OF A CERTAIN CLASS OF CHAOTIC SIGNALS: AN EM-BASED APPROACH

Carlos Pantaleón, Luis Vielva, David Luengo and Ignacio Santamaría

Dpto. Ing. Comunicaciones, ETSII y Telecom., University of Cantabria
Avda. Los Castros, 39005 Santander, Spain
e-mail: carlos@gtas.dicom.unican.es

ABSTRACT

Maximum-likelihood estimation of chaotic signals generated by iterating piecewise-linear maps on the unit interval exhibits an exponential increase in computational cost with the register length. This paper considers iterative estimation algorithms based on the Expectation-Maximization (EM) algorithm and related space alternating methods. This approach is inspired in the parallelism that may be drawn between chaotic estimation and multiuser detection, which also becomes prohibitively complex as the number of users increases. The resulting algorithms are based on an iterative updating of estimates of the chaotic signal itinerary. Computer simulations show that the proposed algorithms achieve the performance of the ML estimator for short data registers and improve the computationally feasible (suboptimal) estimators for long records.

1. INTRODUCTION

Chaotic signals, i.e. signals generated by a non-linear dynamical system in chaotic state, are potentially atractive in a wide range of signal processing applications. Classical signal processing techniques are not adequate for this class of signals that, while deterministic, exhibit a noise-like behaviour. Therefore there is a need for robust and efficient algorithms for the estimation of these signals in noise.

Several authors have proposed signal estimation algorithms for chaotic signals [1, 2, 3]. These methods are usually based on the connection between the symbolic sequence associated to a particular chaotic signal and its initial condition, and are, in general, suboptimal. A dynamical programming approach has also been proposed [2]. Maximum-likelihood (ML) estimators have been developed for the tent map dynamics [4], and generalized to all piecewise linear (PWL) maps [5]. In [6] a Bayesian estimator for chaotic signals generated by the tent map has been proposed.

The exponential increase in complexity that arises in chaotic signal estimation —since estimates have to be computed for any possible itinerary [5]— motivates the search for more efficient estimation approaches. In this sense, it is interesting to notice that the same problem arises in multiuser detection [7] where optimal detectors have to consider all possible transmitted binary sequences in the search for the optimal demodulated sequence. This parallelism justifies the application of good performing techniques in the multiuser detection field to chaotic signal processing. In particular in this paper we consider the application of the Expectation-Maximization (EM) algorithm to chaotic signal estimation. The EM algorithm provides an iterative approach to ML parameter estimation when direct maximization of the likelihood function is not feasible or very time consuming. Although convergence to the ML estimator is not guaranteed, the EM algorithm produces estimates that monotonically increase in likelihood. In some cases, however, better performance is obtained by the Space Alternating Generalized EM (SAGE) algorithm, that updates only a subset of the parameter components [8]. In [9] several receiver structures based on EM, SAGE, and related algorithms are developed based on this philosophy.

In this paper we develop estimators for chaotic signals based on the EM and SAGE algorithms, following the same approach applied in [9]. The fact that chaotic signal estimation is a detection problem (the detection of the sequence of signs of the chaotic sequence) and the parallelism with multiuser detection justifies this approach. Although in this paper we focus on a particular class of maps, the method can be easily applied to general PWL maps [5].

This work has been partially supported by the European Commission and the Spanish Government under contract 1FD97-1066-C02-01

2. SKEW-TENT MAPS

The signals $x[n]$ that we consider in this work are generated according to

$$x[n+1] = F(x[n]), \tag{1}$$

where $F(\cdot)$ is the so called skew-tent map:

$$F(x) = \begin{cases} x/a, & 0 \le x \le a, \\ (1-x)/(1-a), & a < x \le 1; \end{cases} \tag{2}$$

for some parameter $0 < a \le 1$. This map produces sequences that are chaotic with invariant density uniform in the range $[0, 1]$.

The phase space of non-linear maps can be divided in a collection of non-overlapping regions E_i. If a symbol from a known alphabet is assigned to each of the regions, the dynamics of the map may be characterized by following the different regions that the map visits during its dynamical evolution. In the particular case of the skew-tent map, we divide the phase space in two regions $E_1 = [0, a]$ and $E_2 = [a, 1]$ and associate a symbol $s[n]$ to each $x[n]$ according to

$$s[n] = \text{sign}(x[n] - a).$$

The sequence $\mathbf{s} = \{s[0], \ldots, s[N-1]\}$ associated with a length $N+1$ chaotic signal $x[n]$ is called itinerary, and it can be considered a symbolic coding of the chaotic signal. We can define another partition of the phase space in a collection of P intervals R_j composed of the points in $[0, 1]$ that share a common

length N symbolic sequence. Given a certain itinerary $\mathbf{s}_j$ we define $R_j = \{x \in [0,1] : S_N(x) = \mathbf{s}_j\}$, where $S_N(\cdot)$ is the map that associates an initial condition in $[0,1]$ to its corresponding symbolic sequence. If all the linear components are onto, as it is the case for the skew-tent maps, all the symbolic sequences are admissible and there are $P = 2^N$ regions R_j [5].

Considering the two cases of (2), it is easy to verify that $F(x[n])$ can be expressed as

$$F(x[n]) = \frac{x[n](1-2a-s[n]) + a(1+s[n])}{b}, \qquad (3)$$

with $b = 2a(1-a)$. As it is shown in [5], a closed form expression for n-fold composition of $F(\cdot)$ as a function of the initial condition and the itinerary can be obtained by iterating (3):

$$F_{\mathbf{s}}^n(x[0]) = a\sum_{i=0}^{n-1} b^{i-n}(1+s[i])S_{n-i-1}^n + b^{-n}S_n^n x[0], \qquad (4)$$

where $S_0^n = 1$ and

$$S_i^n = \prod_{j=n-i}^{n-1}(1-2a-s[j]) \qquad i = 1,\ldots,n.$$

The subindex $\mathbf{s}$ in $F_{\mathbf{s}}^n(x[0])$ stresses the fact that this expression is only true if $x[0]$ has a sign sequence given by $\mathbf{s}$.

3. EM AND SAGE ALGORITHMS

Given a realization of a random vector $\mathbf{Y}$, that we will denote $\mathbf{y}$, the ML estimator of a parameter θ is defined as

$$\hat{\theta}(\mathbf{y}) = \arg\max_{\theta} \log f(\mathbf{y};\theta). \qquad (5)$$

In many situations direct maximization of (5) may produce algorithms that increase exponentially in complexity with the register length, as it is the case of multiuser detection. In these situations the EM algorithm provides an attractive alternative [9]. This algorithm is based on the existence of some missing data $\mathbf{Z}$ that would help to estimate θ but that are not observable. We will refer to $\mathbf{X} = \{\mathbf{Y},\mathbf{Z}\}$ as the complete data. The basic approach is to alternate between an expectation step, where sufficient statistics of the complete data are obtained based on the incomplete data and current parameter estimates, and a maximization step where the parameters are reestimated based on the sufficient statistics of the complete data. The key point of the EM algorithm is to iteratively maximize the new objective function

$$Q(\theta;\theta^i) = E\{\log f(\mathbf{Y},\mathbf{Z};\theta)|\mathbf{Y}=\mathbf{y};\theta^i\}.$$

Given an initial estimate of the parameter θ^0, the EM algorithm is described by

- E(xpectation)-step: Compute $Q(\theta;\theta^i)$,
- M(aximization)-step: $\theta^{i+1} = \arg\max_{\theta} Q(\theta;\theta^i)$.

The EM algorithm produces estimates that monotonically increase in likelihood; and, in addition, ML estimates are fixed points of the algorithm. As for any deterministic iterative maximization, the ability of the approach to find the global maximum is heavily dependent on the initialization step.

As an alternative, the SAGE algorithm updates only a subset of the parameter components at each iteration [8, 9]. If we denote by θ_S the parameter components indexed by the set S, the i-th iteration of the SAGE algorithm is given by:

- Defn-step: Choose the parameter index set S and the missing data $\mathbf{Z}^S$.
- E-step: Compute $Q^S(\theta_S;\theta^i)$.
- M-step: $\begin{cases} \theta_S^{i+1} = \arg\max_{\theta} Q^S(\theta_S;\theta_S^i), \text{ and} \\ \theta_{\tilde{S}}^{i+1} = \theta_{\tilde{S}}^i, \end{cases}$

where $\tilde{S}$ is the complement of S. Similar to EM, the SAGE estimates monotonically increase in likelihood and the maxima of the likelihood function are fixed points of the algorithm. When an *a priori* distribution for the parameter θ is available, both algorithms may be easily extended to work with conditional densities.

4. EM-BASED ESTIMATION OF CHAOTIC SIGNALS

4.1. Problem Statement

The signal model we are considering is

$$y[n] = x[n] + w[n], \qquad n = 0,1,\ldots,N; \qquad (6)$$

where $x[n]$ is generated using (2) by iterating some unknown initial condition $x[0] \in [0,1]$ according to (1), and $w[n]$ is a stationary, zero-mean, white Gaussian noise with variance σ^2.

4.2. ML estimation of chaotic signals

The observations $\mathbf{y} = \{y[0], y[1], \ldots, y[N]\}$ in (6) are a collection of independent Gaussian random variables with equal variance. The conditional density then may be expressed as

$$f(\mathbf{y}|x[0]) = \frac{1}{(\sqrt{2\pi}\sigma)^{N+1}} \exp\left(-\frac{J(x[0])}{2\sigma^2}\right), \qquad (7)$$

where $J(x[0])$ is given by

$$J(x[0]) = \sum_{k=0}^{N}(y[k] - F^k(x[0]))^2. \qquad (8)$$

Using (4), we can express (8) in a certain region R_j as

$$J_j(x[0]) = \sum_{k=0}^{N}(y[k] - F_{\mathbf{s}_j}^k(x[0]))^2. \qquad (9)$$

If we define the indicator or characteristic function

$$\chi_j(x) = \begin{cases} 1, & \text{if } x \in R_j, \\ 0, & \text{if } x \notin R_j; \end{cases} \qquad (10)$$

the equation (8) can be written as

$$J(x[0]) = \sum_{j=1}^{P}\chi_j(x[0])J_j(x[0]). \qquad (11)$$

This cost function is quadratic within the limits of each region. Differentiating and solving for the unique minimum we obtain

$$\hat{x}_j[0] = \frac{\sum_{k=0}^{N} b^{-k}S_k^k\gamma[k]}{\sum_{k=0}^{N}\left(b^{-k}S_k^k\right)^2}, \qquad (12)$$

where

$$\gamma[k] = y[k] - a\sum_{i=0}^{n-1} b^{i-k}(1+s[i])S_{k-i-1}^{k}. \qquad (13)$$

The ML estimate of $x[0]$ for a given itinerary is $\hat{x}_j[0]$ only if $\hat{x}_j[0] \in R_j$; otherwise the minimum of $J(x[0])$ in the region R_j is given by the closest value to $\hat{x}_j[0]$ in R_j

$$\hat{x}_{\mathrm{ML}}^{j}[0] = \min(R_j^u, \max(R_j^l, \hat{x}_j[0])), \qquad (14)$$

where R_j^l and R_j^u are the lower and upper limits of the region R_j. To simplify the notation the dependence on j, that has been used to indicate a particular itinerary, will we droped from now on.

4.3. EM chaotic estimation

In this paper we apply iterative algorithms to find the ML estimate of chaotic signals observed in noise. It might be argued that the highly irregular shape of the likelihood function with many local maxima [3] precludes the application of iterative approaches due to the fact that these methods would always fall in suboptimum solutions. Although this is true, there is another fact: it is well known that the itinerary obtained by hard limiting the noisy chaotic sequence is a good estimate of the itinerary [1, 2, 3]. Hence we have a good starting point for our iterative approach, and this will compensate, to a certain extent, the irregular shape of the likelihood function.

In this section we derive EM-based estimators that consider all of the signs in the itinerary but one as missing data. Let $\mathbf{s}_{\bar{k}}$ denote all the signs in the itinerary $\mathbf{s}$ except s_k. Application of the EM to the detection of s_k considering $\mathbf{s}_{\bar{k}}$ as missing data is derived below. The complete-data log-likelihood function is given by

$$\log f(\mathbf{s}_{\bar{k}}, \mathbf{y}|s_k) = \log f(\mathbf{s}_{\bar{k}}|s_k) + \log f(\mathbf{y}|\mathbf{s}).$$

Since the itinerary components are independent in the case of chaotic signals generated by skew tent maps [10], we only have to consider the second density.

Given a certain itinerary $\mathbf{s}$, and taking into account (7) and (11), the density function of the observations is

$$f(\mathbf{y}|x[0], \mathbf{s}) = \frac{1}{(\sqrt{2\pi}\sigma)^{N+1}}\chi(x[0])\exp\left(-\frac{J(x[0])}{2\sigma^2}\right).$$

To eliminate the dependence with $x[0]$, we integrate this expression,

$$\int_R f(\mathbf{y}|x[0], \mathbf{s}) f(x[0]|\mathbf{s})\, dx[0];$$

and because the density function of $x[0]$ given a certain itinerary is uniform in the region associated with the given itinerary, we obtain

$$f(\mathbf{y}|\mathbf{s}) = \frac{1}{(\sqrt{2\pi}\sigma)^{N+1}}\exp\left(-\frac{J(\hat{x}[0])}{2\sigma^2}\right)\frac{B}{L}.$$

where L is the length of the region R and B is given by

$$B = \int_R \exp\left(-\frac{(x[0]-\hat{x}[0])^2}{2\sigma^2}\sum_{k=0}^{N}\left(b^{-k}S_k^k\right)^2\right) dx[0].$$

Finally we obtain

$$Q(s_k; s_k^i) = \sum_{\mathbf{s}_{\bar{k}}\in\{\pm1\}^{N-1}} f(\mathbf{s}_{\bar{k}}, s_k^i) f(\mathbf{y}|\mathbf{s}_{\bar{k}}, s_k^i)\log f(\mathbf{y}|\mathbf{s}_{\bar{k}}, s_k).$$

The summation index $\mathbf{s}_{\bar{k}} \in \{\pm1\}^{N-1}$ in the final result indicates that it is an expression composed of 2^{N-1} terms. Consequently it is not useful to reduce the exponential increase in complexity of direct approaches. The difference with multiuser detection, where the complexity is indeed reduced, is that in our problem the dependence on the itinerary is highly nonlinear, while in multiuser detection there is a linear dependence with the bit sequence. Nonetheless we may use the preceding expressions for developing more efficient algorithms in the next section.

4.4. SAGE chaotic estimation

The SAGE algorithm can be applied to our problem with $\theta = \mathbf{s}$ and the index sets cycling through $0, \ldots, N-1$. In this way the algorithm may be implemented without any missing data. Consequently the E-step is trivial and may be ignored. The M-step is simply

$$s_k^{i+1} = \arg\max_{s_k}\left[\frac{B_{\bar{k}}}{L_{\bar{k}}}\exp\left(-\frac{J(\hat{x}_{\bar{k}})}{2\sigma^2}\right)\right],$$

where $\hat{x}_{\bar{k}}$ is the estimate obtained according to (12), when all the itinerary components are fixed except s_k (the same idea applies to $B_{\bar{k}}$ and $L_{\bar{k}}$). So the algorithm becomes:

- Defn-step: Let $k = 0 + i \bmod (N-1)$.
- M-step: $\begin{cases} s_k^{i+1} = \arg\max_{s_k}\left[\frac{B_{\bar{k}}}{L_{\bar{k}}}\exp\left(-\frac{J(\hat{x}_{\bar{k}})}{2\sigma^2}\right)\right], \\ s_m^{i+1} = s_m^i, \text{ for all } m \neq k. \end{cases}$

This algorithm, that we will denote SAGE-1, selects the most probable value of s_k, and is a coordinate descendent or greedy algorithm.

As an alternative, another SAGE algorithm may be implemented by defining $\theta = \{x[0], \mathbf{s}\}$ and the index set including always the component in $x[0]$ and cycling through the components of the itinerary, one each iteration. The E-step again may be ignored and the M-step is

$$\left\{x[0]^{i+1}, s_k^{i+1}\right\} = \arg\min_{s_k, x[0]} J(x[0]). \qquad (15)$$

so the final algorithm, denoted as SAGE-2, becomes:

- Defn-step: Let $k = 0 + i \bmod (N-1)$.
- Mstep: $\begin{cases} s_k^{i+1} = \arg\min_{s_k} J(\hat{x}_{\mathrm{ML}}^{\bar{k}}[0]), \\ s_m^{i+1} = s_m^i, \text{ for all } m \neq k. \end{cases}$

where $\hat{x}_{\mathrm{ML}}^{\bar{k}}[0]$ is the ML estimate, obtained from (12) and (14), with all itinerary components fixed except s_k. In this case we search for the most likely of the possible values of the initial condition.

Further improvement may be expected by modeling s_m as taking values in a continuous set, for example in the range $[-1, 1]$, producing fractional itinerary values. In any case the boundness and monotonicity of the log-likelihood function guarantees the convergence of the SAGE algorithm.

5. SIMULATION RESULTS

In this section we analyze the performance of the different signal estimation algorithms. We consider a skew-tent map with parameter $a = 0.9$. We first study short sequences, namely with $N = 6$.

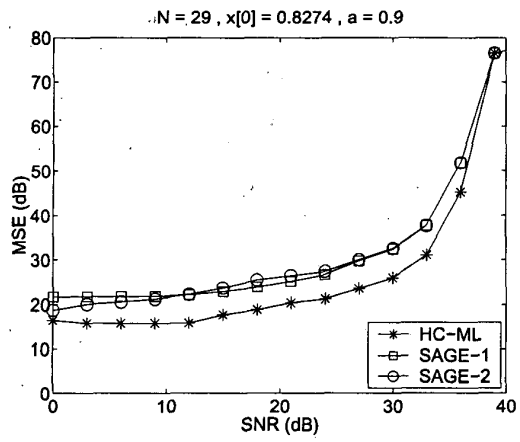

Fig. 1. Mean square error (MSE) for the three estimators: HC-ML, SAGE-1, and SAGE-2.

One thousand initial conditions have been selected according to the uniform invariant density of the map. For each initial condition and SNR, one thousand data registers have been generated. We compare four methods: ML estimator, SAGE-1, SAGE-2, and a method similar to the those proposed in [1, 2, 3] (where the itinerary is obtained by hard-censoring the noisy data) that we will denote as HC-ML. Table 1 summarizes the results. The performance of both SAGE algorithms is similar to that of the ML estimator, with a much lower computational cost, and much better than that provided by the HC-ML. It is also obvious that no much room for improvement is left, so it is arguable whether a soft decision SAGE algorithm would be useful for short data records. In a second set of tests, we compare the performance of the proposed algorithms on longer sequences. Specifically, we have generated noisy registers with $N + 1 = 30$ and $a = 0.9$. Figure 1 shows an example for $x[0] = 0.8274$. Again both approaches improve the performance of the HC-ML estimator. It should be noted that in this case it is unfeasible to compute the ML estimate.

	MSE (dB)			
SNR (dB)	HC-ML	SAGE-1	SAGE-2	ML
0	15.6	15.7	15.0	14.7
5	18.7	18.9	19.0	18.7
10	22.7	24.0	24.6	24.3
15	27.2	30.0	30.8	30.4
20	31.9	36.2	36.8	36.8
25	38.3	42.0	42.6	43.3
60	88.6	89.0	89.0	89.0

Table 1. MSE of the four estimators as a function of the SNR. One thousand noisy sequences for each of the one thousand initial conditions have been averaged.

6. CONCLUSIONS

ML estimation of chaotic signals generated by iterating piecewise-linear maps on the unit interval exhibits an exponential increase in computational cost with the register length. In this paper we have developed efficient estimators for chaotic signals based on the EM and SAGE algorithms that are closely related to similar estimators developed for multiuser detection, where the same complexity problem arises. The EM-based approach produces an algorithm that mantains the exponential complexity and therefore it is not useful by itself. However, it allows to develop SAGE algorithms that achieve a quasi-optimal performance for short register lengths and that improve the performance of known estimation algorithms for long registers. Future lines of research include the extension of these algorithms to chaotic signals generated by any Markov map and to further exploit the connection between chaotic signal estimation and multiuser detection ir order to develop new methods in the chaos signal processing field.

7. REFERENCES

[1] S. Wang, P.C. Yip and H. Leung, "Estimating Initial Conditions of Noisy Chaotic Signals Generated by Piecewise-Linear Markov Maps Using Itineraries", *IEEE Trans. Signal Processing.*, vol. 47, no 12. pp. 3289–3302, 1999.

[2] S. M. Kay, "Methods for Chaotic Signal Estimation", *IEEE Trans. Signal Processing.*, vol. 43, no 8. pp. 2013–2016, 1995.

[3] L. Cong, W. Xiafou and S. Songgeng, "A General Efficient Method for Chaotic Signal Estimation",*IEEE Trans. Signal Processing.*, vol. 47, no 5. pp. 1424–1428, 1999.

[4] H. C. Papadopoulos, G. W. Wornell, "Maximum Likelihood Estimation of a Class of Chaotic Signals", *IEEE Trans. Information Theory*, vol. 41, no 1. pp. 312–317, 1995.

[5] C. Pantaleón, D. Luengo, I. Santamaría, "Optimal Estimation of Chaotic Signals Generated by Piecewise-Linear Maps", *IEEE Signal Processing Letters*, vol. 7, no. 8, pp. 235-237, Aug. 2000.

[6] C. Pantaleón, D. Luengo, I. Santamaría, "Bayesian Estimation of a Class of Chaotic Signals", Proc. of the 2000 IEEE Int. Conf. on Acoust., Speech, and Signal Processing (ICASSP 2000), vol. I, pp. 197-200, Istanbul, Turkey, June 2000.

[7] S.Verdu, *Multiuser detection*, Cambridge, UK: Cambridge Univ. Press, 1998.

[8] J. A. Fessler and A. O. Hero, "Space-alternating geeralized EM algorithm", *IEEE Trans. Signal Processing*, vol. 42, no 10. pp. 2664-2667,1994.

[9] L. B. Nelson, H. V. Poor,"Iterative Multiuser Receivers for CDMA Channels: An EM-Based Approach", *IEEE Trans. Communications*, vol. 44, no 12, pp. 1700–1710,1996.

[10] S. H. Isabelle and G. W. Wornell, "Statistical Analysis and Spectral Estimation Techniques for One-Dimensional Chaotic Signals", *IEEE Trans. Signal Processing*, vol. 45, no 6. pp. 1495–1506,1997.