On-line Identification of Non-minimum Phase Finite Impulse Response Linear Systems with Discrete Inputs via Hidden Markov Models

Jason J. Ford[1]

Abstract

This paper considers adaptive estimation of noisy finite impulse response linear systems driven by stationary inputs from a discrete set. An algorithm based on an equivalent hidden Markov model representation is presented. The algorithm presented is global convergent while previous approaches have only been locally convergent.

1 Introduction

The problem of blind identification and filtering of linear systems driven by inputs from a discrete set has received some attention in recent years[1]. It has received this attention because of its importance for digital communications. The discrete nature of the input signals adds additional *a priori* information that traditional deconvolution approaches for linear systems do not exploit (for example, the probability density of the input sequence is point concentrated rather than a smooth normal distribution). Although traditional identification techniques can be used, they are far from optimal.

Recently, several methods have been proposed that exploit this discrete nature to develop blind deconvolution algorithms with improved performance[1] . One notable approach is the algorithm of Gassiat and Gautherat[1] which requires only weak assumptions on the input sequence including that it is ergodic and stationary (they also suggest that the inputs need not be stationary but no theoretical results are given).

The key contribution of this paper is the proposal of a new globally convergent on-line algorithm for adaptive estimation via hidden Markov models (HMMs) of finite impulse response (FIR) linear systems (even non-minimum phase systems) driven by inputs from a discrete set. The algorithm presented in this paper is superior to previous approaches such as the "tuning the inverse" method of [1] and the recursive expectation-maximisation (EM) algorithm approach of [2], because the new algorithm is recursive and globally convergent.

[1]Weapons Systems Division, Defence Science Technology Organisation, P.O. Box 1500, Salisbury SA, 5108, Australia. Email: Jason.Ford@dsto.defence.gov.au

2 Linear Systems as Hidden Markov Models

This section introduces a model for a linear system with input signals coming from a discrete set. Consider a probability space $(\Sigma, \mathcal{F}, P)$; suppose that x_k is a scalar *iid* stationary ergodic stochastic process on this probability space. Further suppose that x_k takes on only one of the n (which is known) possible values from the set $\{M^1, \dots, M^n\}$. Let us denote the vector of possible input values as $\mathcal{M} := [M^1, \dots, M^n]$. Here $:=$ is the 'defined' by symbol. Then assume without loss of generality (*wlog*) that $M^i \neq M^j$ for $i \neq j$ and $M^i \neq 0$ for some i. Now suppose that a sequence of scalar outputs $\{y_k\}$ from an unknown linear time-invariant system that is driven by the process $\{x_k\}$ are observed as follows:

$$y_k = \sum_{j=0}^{K-1} h^j x_{k-j} + w_k \tag{2.1}$$

where h^j are coefficients (the input response of the linear system) and $\{w_k\}$ is a sequence of independent and identically distributed (*iid*) $N(0, \sigma_w^2)$ (Gaussian distribution with zero mean and σ_w^2 variance) scalar random variables on the same probability space $(\Sigma,\mathcal{F},P)$. The stochastic processes $\{w_k\}$ and $\{x_k\}$ are assumed mutually independent. Let us assume that K is finite and known, and *wlog* that h^0 and $h^{K-1} \neq 0$ and let us denote the vector of coefficients as $\mathcal{H} := [h^0, \dots, h^{K-1}]$. Then the system (2.1) is a (causal time-invariant) finite impulse response (FIR) linear system driven by a discrete input.

To complete the system description, let us denote the probability that x_k takes on the value M^i by $p^i := P(x_k = M^i)$. Also, let us denote the statistics of the input sequence by the vector of probabilities $\mathcal{P} := [p^1, \dots, p^n]$. In addition we assume that the gain of (2.1) is known, ie. *wlog* assume $\sum_{j=0}^{K-1} h^j = 1$. It is important to note that (2.1) is not assumed to be minimum phase.

The idea of representing the system (2.1) as a hidden Markov model (HMM) appears previously in [2]. The

HMM formulation of the problem follows by considering the linear system (2.1) as not only having inputs from a discrete set but also having a system state (a memory of the last K inputs) that can only take on one of $N = n^K$ possible values (a K selection from a set of size n).

It can be shown that there is a one-to-one correspondence between the linear system given above and an HMM of a particular structure (see upcoming paper). Hence, a solution to the adaptive estimation problem for the HMM formulation immediately provides a solution to the linear system adaptive estimation problem (see [4, 5] for previous work).

3 Adaptive Estimation

The main problem of interest in this paper is estimation of $\mathcal{H}$, $\mathcal{M}$, $\mathcal{P}$ and the input sequence given an observation sequence $\mathcal{Y}_k$. This adaptive estimation problem can be posed as an HMM parameter estimation problem.

Let $\hat{A}_k$ and $\hat{C}_k$ denote estimates for the HMM parameters (equivalent to the linear system). Let $\hat{\lambda}_k = \lambda(\hat{A}_k, \hat{C}_k, \sigma_w^2, X_0)$ denote an estimate of the model at time k. In [4] it is shown that estimates for the HMM parameters can be obtained as follows:

$$\begin{aligned} \hat{A}_k &= \hat{J}_{k|\hat{\lambda}_{k-1}} \left(\hat{O}_{k|\hat{\lambda}_{k-1}}\right)^{-1} \quad \text{and} \\ \hat{C}_k &= \hat{\mathcal{T}}_{k|\hat{\lambda}_{k-1}} \left(\hat{O}_{k|\hat{\lambda}_{k-1}}\right)^{-1}. \end{aligned} \tag{3.1}$$

Here

$$\begin{aligned} \hat{J}_{k|\hat{\lambda}_{k-1}} &= E[J_k|\mathcal{Y}_k, \hat{\lambda}_0, \ldots, \hat{\lambda}_{k-1}], \\ \hat{O}_{k|\hat{\lambda}_{k-1}} &= E[O_k|\mathcal{Y}_k, \hat{\lambda}_0, \ldots, \hat{\lambda}_{k-1}] \quad \text{and} \\ \hat{\mathcal{T}}_{k|\hat{\lambda}_{k-1}} &= E[\mathcal{T}_k|\mathcal{Y}_k, \hat{\lambda}_0, \ldots, \hat{\lambda}_{k-1}]. \end{aligned} \tag{3.2}$$

where the ijth element of J_k is the number of transitions from state j to state i; the ith diagonal element of O_k is the number of times in state i (O_k is a diagonal matrix); and the ith element of T_k is the summation of y_k when in state i (finite dimensional filters are given in [3]).

Under certain persistence of excitation conditions (see [4] for details) the following theorem holds (which is a restatement of a result established in [4]).

Theorem 1 *Consider the HMM representation of the linear system and assume the system is persistently exciting.*

$$\lim_{k\to\infty} \hat{A}_k, \hat{C}_k = A, C \text{ almost surely} \tag{3.3}$$

or to a permutation (relabelling of states).

This estimation algorithm does not provide explicit estimates of the quantities $\mathcal{H}$, $\mathcal{M}$ and $\mathcal{P}$. When the mapping from HMM to the linear system is used the following theorem holds.

Theorem 2 *Under the same conditions as Thm. 1*

$$\lim_{k\to\infty} \hat{\mathcal{P}}_k, \hat{\mathcal{H}}_k, \hat{\mathcal{M}}_k = \mathcal{P}, \mathcal{H}, \mathcal{M} \text{ almost surely,} \tag{3.4}$$

or to a permutation.

Proof: Follows from Thm. 1 and the one-to-one correspondence between the linear model and the HMM.

4 Discussion

Theorem 2 states that the actual linear system is identified (not the minimum phase system equivalent to the true linear system) and this global convergence property has been demonstrated in simulations studies.

In addition the HMM representation of this problem aids understanding of the effect of non-stationary statistics considered in [1]. For example, a HMM formulation suggests the conjecture of [1] that their algorithm can handle inputs sequences with non-stationary statistics is unfounded because their simulation studies are special cases of the estimation problem.

The key advantage of the algorithm presented here over the approach of [2] is the use of new globally convergent HMM parameter estimators. In fact, this problem is an interesting example of a situation where many HMM estimators are locally convergent.

References

[1] E. Gassiat and E. Gautherat, "Identification of Noisy Linear Systems with Discrete Random Input," *IEEE Transactions on Information Theory*, Vol. 44, No. 5, pp. 1941-1952, Sept. 1998.

[2] L.B. White and V. Krishnamurthy, "Adaptive Blind Equalisation of FIR channels using HMMs," in *IEEE Conf. on Communications*, Geneva, Switzerland, pp. 1128-1133, May 1993.

[3] R.J. Elliott, L. Aggoun, and J.B. Moore, *Hidden Markov Models: Estimation and Control*, Springer, New York, 1995.

[4] J.J. Ford, *Adaptive hidden Markov model estimation and applications*, PhD thesis, Australian National University, 1998.

[5] J.J. Ford and J.B. Moore, "Adaptive Estimation of HMM Transition Probabilities," *IEEE Transactions on Signal Processing*, Vol. 46, No. 5, pp. 1374-1385, May 1998.