Analysis of Recursive MAP Algorithm for State Estimation of Bilinear Systems[1]

Vikram Krishnamurthy
Department of Electrical and Electronic Engineering
University of Melbourne, Parkville, Victoria 3052, Australia.
vikram@mullian.ee.mu.oz.au

George Yin
Department of Mathematics,
Wayne State University, Detroit, MI 48202, U.S.A.
gyin@math.wayne.edu

Abstract

This paper derives and analyses a recursive algorithm for maximum aposteriori (MAP) state estimation of partially observed bilinear systems. The recursive algorithm is based on cross-coupling two Kalman filters, one for each component of the bilinear system.

1 Introduction

Bilinear models [3] are widely used to model nonlinear processes in signal and image processing and communication systems modelling. Unfortunately, the optimal filter for reconstructing conditional mean estimates of the state of a partially observed bilinear system is infinite dimensional. Thus practical filtering algorithms for bilinear systems are necessarily suboptimal. For example, the extended Kalman filter (EKF) is an approximate filter that linearizes around conditional mean state estimates at each time instant.

In recent work [1], rather than computing approximations of the conditional mean estimates, iterative finite dimensional algorithms were presented for computing the optimal MAP state sequence estimate for a bilinear system. In particular, the Expectation Maximization (EM) algorithm was used to compute these MAP state sequence estimates. The EM algorithms in [1] for estimating the state of the bilinear system, involved cross-coupling two Kalman smoothers, each estimating one of the two component signals of the bilinear system.

This paper presents and prove convergence of a recursive version of the EM algorithm for state estimation of bilinear systems. In analogy to the off-line EM algorithm presented in [1], the recursive algorithm we present exactly cross-couples two Kalman filters, each estimating one of the two component signals of the bilinear system. We show that for scalar bilinear models, providing the noise variances of the state processes are a sufficiently small value ϵ, the estimation errors are $O(\epsilon)$.

[1]This research was partially supported by the National Science Foundation under DMS-9877090.

2 Formulation and Recursive Algorithm

The underlying model considered is a scalar discrete-time system. For each $n = 0, 1, \ldots$, x_n, s_n, $y_n \in \mathbb{R}$ are the state, the signal, and the observation, respectively. The quantities c, b, $\varepsilon \in \mathbb{R}$ satisfy $c \neq 0$, $|b| < 1$, $0 < \varepsilon \ll 1$. Denote the system noise, the disturbance in the signal, and the observation noise by w_n, v_n, and e_n, respectively. The bilinear system is given by

$$\left\{ \begin{array}{l} s_{n+1} = bs_n + \varepsilon v_n \\ x_{n+1} = s_n x_n + \varepsilon w_n \\ y_n = cx_n + e_n. \end{array} \right. \tag{2.1}$$

Basic Assumption. The noise sequences $\{v_n\}$, $\{w_n\}$, and $\{e_n\}$ are independent martingale difference sequences such that $\{v_n\}$ is a sequence of uniformly bounded random variables,

$$\begin{array}{l} Ev_n = Ew_n = Ee_n = 0, \quad \text{and} \\ \sup_n E|w_n|^2 < \infty, \ \sup_n E|e_n|^2 < \infty. \end{array}$$

2.1 Motivation – MAP State Estimation

Denote the T–point sequences $S_T = [s_1 \ldots s_T]$, $X_T = [x_1 \ldots x_T]$ and $Y_T = [y_1 \ldots y_T]$ where T is a positive integer. Consider now the following MAP sequence estimation problem for S_T and X_T: Compute

$$(\hat{S}_T, \hat{X}_T)^{MAP} = \arg \max_{S_T, X_T} f(S_T, X_T | Y_T) \tag{2.2}$$

where $f(S_T, X_T|Y_T)$ denotes the joint conditional probability density function of S_T and X_T, conditioned on the measurements Y_T. In [1], this optimization was performed iteratively by using the a coordinate descent algorithm involving cross-coupling two Kalman filters.

2.2 Recursive Algorithms

Motivated by the iterative coordinate descent algorithm in [1], consider the following recursive (online) algorithm for state estimation: Replace the Kalman smoothers by Kalman filters. The resulting recursive algorithm has a clear intuitive interpretation: Knowledge of the signal states s_k results in a Kalman filter achieving optimal estimates of x_k. Conversely, knowledge of the states x_k means a Kalman filter achieves optimal estimates of s_k. Simulation studies in [1] show that the resulting algorithm performs significantly better than the Extended Kalman filter.

Let $\mathcal{F}_n$ denote the sigma algebra that measures at least $\{x_0, s_0, v_j, w_j, e_j,\ j < n\}$. With given estimate $\widehat{s}_n$ of s_n, we construct the recursive estimate of x_n by choosing a square integrable x_n (i.e., $\sup_n E|x_n|^2 < \infty$) such that $E|y_n - cx_n|^2$ is minimized subject to the constraints of x_n and y_n given in (2.1). Likewise, with given $\widehat{x}_{n+1}$ and $\widehat{x}_n$, choose a square integrable s_n (i.e., $\sup_n E|s_n|^2 < \infty$) that minimizes $E|\widehat{x}_{n+1} - s_n\widehat{x}_n|^2$ subject to s_n and x_n given in (2.1). We then have the following algorithm

$$\begin{cases} \widehat{x}_{n+1} = \Pi_G\left(\widehat{s}_n\widehat{x}_n + \alpha_n c(y_n - c\widehat{x}_n)\right) \\ \widehat{s}_{n+1} = b\widehat{s}_n + \beta_n\widehat{x}_n(\widehat{x}_{n+1} - \widehat{s}_n\widehat{x}_n), \end{cases} \tag{2.3}$$

where Π_G denotes the projection of the iterates on a pre-selected bounded and convex region G, and where α_n and β_n are step-size sequences.

Step Sizes. Assume either $\alpha_n = \alpha_\varepsilon$, and $\beta_n = \beta_\varepsilon$ such that $0 < \alpha_\varepsilon \to 0$ and $0 < \beta_\varepsilon \to 0$ as $\varepsilon \to 0$, or $\{\alpha_n\}$ and $\{\beta_n\}$ are decreasing sequences of positive real numbers such that $\alpha_n \to 0$ and $\beta_n \to 0$ as $n \to \infty$, or they are both $\mathcal{F}_n$ measurable random sequences satisfying $\alpha_n \to 0$ w.p.1 and $\beta_n \to 0$ w.p.1 uniformly.

Remark 2.1. For the scalar problem, a simplest projection region is an interval. The algorithm then can be written in an alternative form as

$$\widehat{x}_{n+1} = \left[\widehat{s}_n\widehat{x}_n + \alpha_n c(y_n - c\widehat{x}_n)\right]_l^u,$$

where $l, u \in \mathbb{R}$ are some prescribed numbers serving as the lower and upper bounds, respectively.

Proposition 2.2. *Under the basic assumption, assuming that s_0 is bounded w.p.1, then $\{s_n\}$ is bounded w.p.1. Moreover,* $\limsup_n |s_n| = O(\varepsilon)$ *w.p.1..*

Remark 2.3. Note that for sufficiently small $\varepsilon > 0$ the bound given above satisfies $O(\varepsilon) < 1$. Then for any $\eta > 0$ sufficiently small such that $0 < O(\varepsilon) + \eta < 1$, there can be only finitely many n such that $||s_n| - O(\varepsilon)| > \eta$. Thus there is an n_0 (the boundedness of $\{v_n\}$ and s_0 implies that n_0 can be independent of ω) such that for all $n \geq n_0$, $|s_n| \leq O(\varepsilon) + \eta$. As a consequence,we have:

Corollary 2.4. *Under the condition of Proposition* 2.2, *for sufficiently small $\varepsilon > 0$,*

$$\sum_{j=0}^{n}\left|\prod_{i=j+1}^{n} s_i\right| \text{ converges w.p.1, } \limsup_n \sum_{j=0}^{n} \sup_\omega \left|\prod_{i=j+1}^{n} s_i\right| < \infty,$$
$$\text{and } \limsup_n \sum_{j=0}^{n} E^{1/2}\left|\prod_{i=j+1}^{n} s_i\right|^2 < \infty. \tag{2.4}$$

Proposition 2.5. *Under the conditions of Proposition* 2.2, *for sufficiently small $\varepsilon > 0$, $\{x_n\}$ is bounded w.p.1, and* $\limsup_n |x_n| = O(\varepsilon)$ *w.p.1..*

3 Error Bounds

Consider the tracking errors $s_n - \widehat{s}_n$ and $x_n - \widehat{x}_n$. Define $\widetilde{s}_n = s_n - \widehat{s}_n$, $\widetilde{x}_n = x_n - \widehat{x}_n$, $\widetilde{s}_0 = 0$, and $\widetilde{x}_0 = 0$. Following the approach in [2], we take an "expansion" of $\widehat{x}_n$ as follows. Define a reflection term z_n by

$$\alpha_n z_n = \widehat{x}_{n+1} - \widehat{s}_n\widehat{x}_n - \alpha_n c(y_n - c\widehat{x}_n), \tag{3.1}$$

where z_n is a real number with shortest Euclidean length needed to bring $\widehat{s}_n\widehat{x}_n + \alpha_n c(y_n - c\widehat{x}_n)$ back to the constraint set G. Under (3.1), the algorithm for $\widehat{x}_n$ can be rewritten as

$$\widehat{x}_{n+1} = \widehat{s}_n\widehat{x}_n + \alpha_n c(y_n - c\widehat{x}_n) + \alpha_n z_n. \tag{3.2}$$

Proposition 3.1. *Under the basic condition, for sufficiently small $\varepsilon > 0$, the following assertions hold:*
A. If α_ε and β_ε are constant step sizes then $\limsup_n E|\widetilde{s}_n| = O(\varepsilon + \beta_\varepsilon)$, $\limsup_n E|\widetilde{x}_n| = O(\varepsilon + \alpha_\varepsilon)$.
B. If α_n and β_n are decreasing sequences satisfying the conditions on step size, then $\limsup_n E|\widetilde{s}_n| = O(\varepsilon)$, *and* $\limsup_n E|\widetilde{x}_n| = O(\varepsilon)$.

References

[1] L. Johnston and V. Krishnamurthy. Finite dimensional smoothers for MAP state estimation of bilinear systems. *IEEE Transactions on Signal Processing*, 47(9):2444–2459, September 1999.

[2] H.J. Kushner and G. Yin. *Stochastic Approximation Algorithms and Applications.* Springer-Verlag, New York, 1997.

[3] M.B. Priestley. *Non-linear and Nonstationary Time Series Analysis.* Academic Press, 1991.