UPPSALA UNIVERSITET

Identification of jump Markov linear models using particle filters

Andreas Svensson

Department of Information Technology
Uppsala University, Sweden

CDC 2014, December 17

Problem

Jump Markov Linear System

$$
\begin{aligned}
s_{t+1} &\sim p(s_{t+1}|s_t),\\
z_{t+1} &= A_{s_{t+1}} z_t + B_{s_{t+1}} u_t + w_t,\\
y_t &= C_{s_t} z_t + D_{s_t} u_t + v_t,\\
s_t &\in \{1, \dots, K\}
\end{aligned}
$$

Maximum-Likelihood identification of θ from $y_{1:T}$ (and $u_{1:T}$)

$$
\widehat{\theta}_{\mathrm{ML}} = \underset{\theta \in \Theta}{\operatorname{argmax}}\ p_\theta(y_{1:T})
$$

$\theta \triangleq \{A_n, B_n, C_n, D_n, \mathbb{E} w_t w_t^T, \mathbb{E} e_t e_t^T, p(s_{t+1}|s_t)\}$ (not K and dimensions of matrices)

UPPSALA UNIVERSITET

Idea

Expectation Maximization for Sys Id

Model

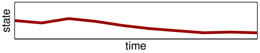
$$x_{t+1} = f_\theta(x_t)$$
$$y_t = g_\theta(x_t)$$

Expectation (E)

Maximization (M)

Guess θ_0

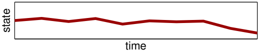
Run smoother with θ_0 to get state estimate $p_{\theta_0}(x_{1:T}|y_{1:T})$



Pick θ_1 as the one explaining the obtained state estimate the best;

$$\theta_1 = \underset{\theta}{\text{argmax}} \ \mathbb{E}_{\theta_0}\left[\log p_\theta(x_{1:T}|y_{1:T})|y_{1:T}\right]$$

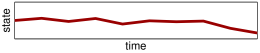
Run smoother with θ_1 to get new state estimate $p_{\theta_1}(x_{1:T}|y_{1:T})$



Pick θ_2 as the one explaining the new obtained state estimate the best;

$$\theta_2 = \underset{\theta}{\text{argmax}} \ \mathbb{E}_{\theta_1}\left[\log p_\theta(x_{1:T}|y_{1:T})|y_{1:T}\right]$$

Run smoother with θ_2 to get new state estimate $p_{\theta_2}(x_{1:T}|y_{1:T})$



⋮

Previous work

2005: Identification of linear systems using EM

E: Compute $\mathcal{Q}(\theta, \theta_{k-1}) = \mathbb{E}_{\theta_{k-1}}[\log p_\theta(x_{1:T}, y_{1:T}) \mid y_{1:T}]$ – *RTS smoother!*

M: Compute $\theta_k = \underset{\theta \in \Theta}{\text{argmax}}\ \mathcal{Q}(\theta, \theta_{k-1})$ – *Anlaytical expressions exist!*

S. Gibson and B. Ninness, **Robust maximum-likelihood estimation of multivariable dynamic systems**, *Automatica*, vol. 41, no. 10, pp. 1667-1682, October 2005.

2011: Identification of nonlinear systems using PS+EM

E: Compute $\mathcal{Q}(\theta, \theta_{k-1}) = \mathbb{E}_{\theta_{k-1}}[\log p_\theta(x_{1:T}, y_{1:T}) \mid y_{1:T}]$ – *Particle smoother!*

M: Compute $\theta_k = \underset{\theta \in \Theta}{\text{argmax}}\ \mathcal{Q}(\theta, \theta_{k-1})$ – *Anlaytical expressions sometimes exist!*

T. B. Schön, A. Wills and B. Ninness. **System identification of nonlinear state-space models**. *Automatica*, vol 47, no. 1, pp. 39-49, January 2011.

2013: More efficient id of nonlinear systems using PSAEM

E: Compute $\widehat{\mathcal{Q}}_k(\theta) = (1 - \gamma_k)\widehat{\mathcal{Q}}_{k-1}(\theta) + \gamma_k \log p_\theta(y_{1:T}, x_{1:T}[k])$ – *Conditional particle filter!*

M: Compute $\theta_k = \underset{\theta \in \Theta}{\text{argmax}}\ \widehat{\mathcal{Q}}_k(\theta)$ – *Anlaytical expressions sometimes exist!*

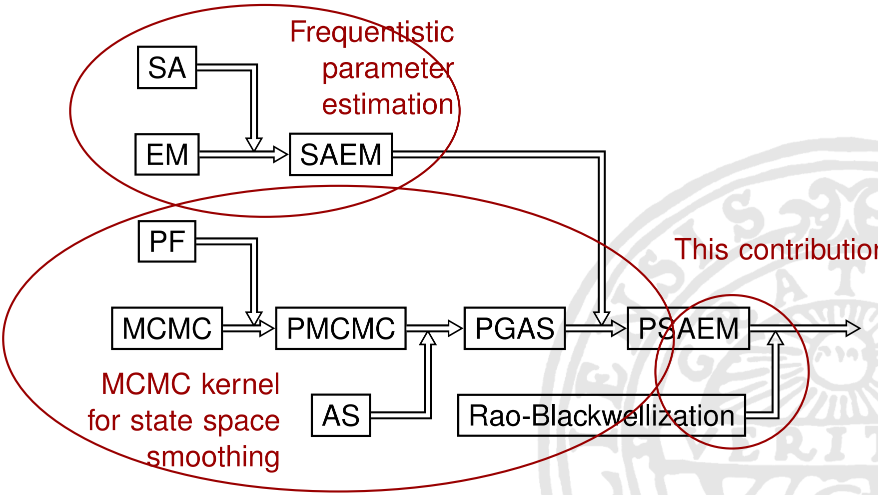
F. Lindsten, **An efficient stochastic approximation EM algorithm using conditional particle filters**, *Proc. 38th Int. Conf. Acoustics, Speech, and Signal Processing (ICASSP)*, Vancouver, Canada, May 2013, pp. 6274-6278.

This contribution

A version of PSAEM exploiting the (conditional) linear structure of the model by using exact methods (Kalman filter, RTS smoother, . . .).

I. e., a Rao-Blackwellized version of PSAEM.

Another way to view it



Algorithm

[0] Intitialize $\widehat{\theta}_0$, $s'_{1:T}[0]$ and $\widehat{\mathcal{Q}}_k(\theta) \equiv 0$

for $k \geq 1$ **do**

[1] Given $s'_{1:T}[k-1]$ and θ_{k-1}, run RB CPF-AS to get $\{s^i_{1:T}, w^i_T\}_{i=1}^N$

[2] Compute $\widehat{\mathcal{Q}}_k(\theta) = (1-\gamma_k)\widehat{\mathcal{Q}}_{k-1}(\theta) + \gamma_k \sum_i w^i_T \log p_\theta(y_{1:T}, s^i_{1:T})$

[3] Compute $\widehat{\theta}_k = \text{argmax } \widehat{\mathcal{Q}}_k(\theta)$

[4] Set $s'_{1:T}[k] = s^J_{1:T}$ with $\mathbb{P}(J = i) = w^i_T$

end for

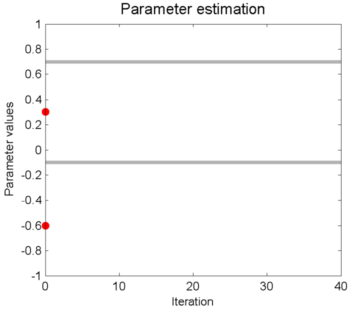
RB CPF-AS: Rao-Blackwellized Conditional Particle Filter with Ancestor Sampling.

Convergence

- PMCMC and Markovian stochastic approximation theory
- Convergence to a maximizer of $p_\theta(y_{1:T})$ for any *finite* (!) number of particles $N \geq 2$
- In practice, very few (e.g., 4) particles are needed

Illustration

One-dimensional system, 2 modes, only A_n unknown.



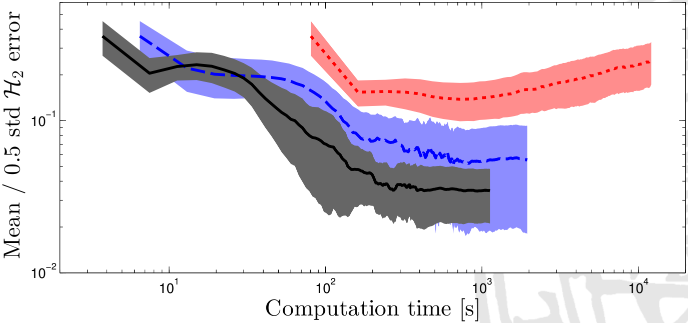
Comparison

Identification of 2-dimensional stable system with 3 modes.

Black: Rao-Blackwellized PSAEM, $N = 3$

Blue: Original PSAEM, $N = 20$

Red: PSEM, $N = 100$, $M = 20$



Summary

- Identification of Jump Markov Linear Models
  - Rao-Blackwellized version of PSAEM
  - Extensions to more general mixed linear/nonlinear models?
- Convergence results
  - Even for finite number of particles!
- Good numerical performance

www.it.uu.se/katalog/andsv164/research/RB-PSAEM