Parameter inference in AR processes with missing data

J. Dahlin, F. Lindsten and T. B. Schön ({johan.dahlin,lindsten,schon}@isy.liu.se)

AUTOMATIC CONTROL
REGLERTEKNIK
LINKÖPINGS UNIVERSITET

Summary

- Parameter estimation when missing data is present is normally carried out using the computationally expensive EM algorithm.
- This work continues earlier work, in which an extension to the CM algorithm was proposed. This method could result in a computationally cheaper alternative to the EM algorithm, solving a similar Maximum Likelihood (ML) problem.

Maximum Likelihood Estimation

The **AR model** of order p has the following structure

$$y_t + \sum_{i=1}^{p} a_i y_{t-i} = e_t, \quad \text{with } e \sim \text{WN}(0, \sigma_e^2).$$

The log-likelihood for $\theta = (a_1, \ldots, a_p)$ follows as

$$\ell_\theta(Y) \propto \frac{1}{2\sigma_e^2} \sum_{t=1}^{N} \left(y_t - \varphi(t)^\top \theta \right)^2.$$

Assuming that only some output values Y are **observed** and the remaining are **missing** Z, results in a joint log-likelihood function on the form

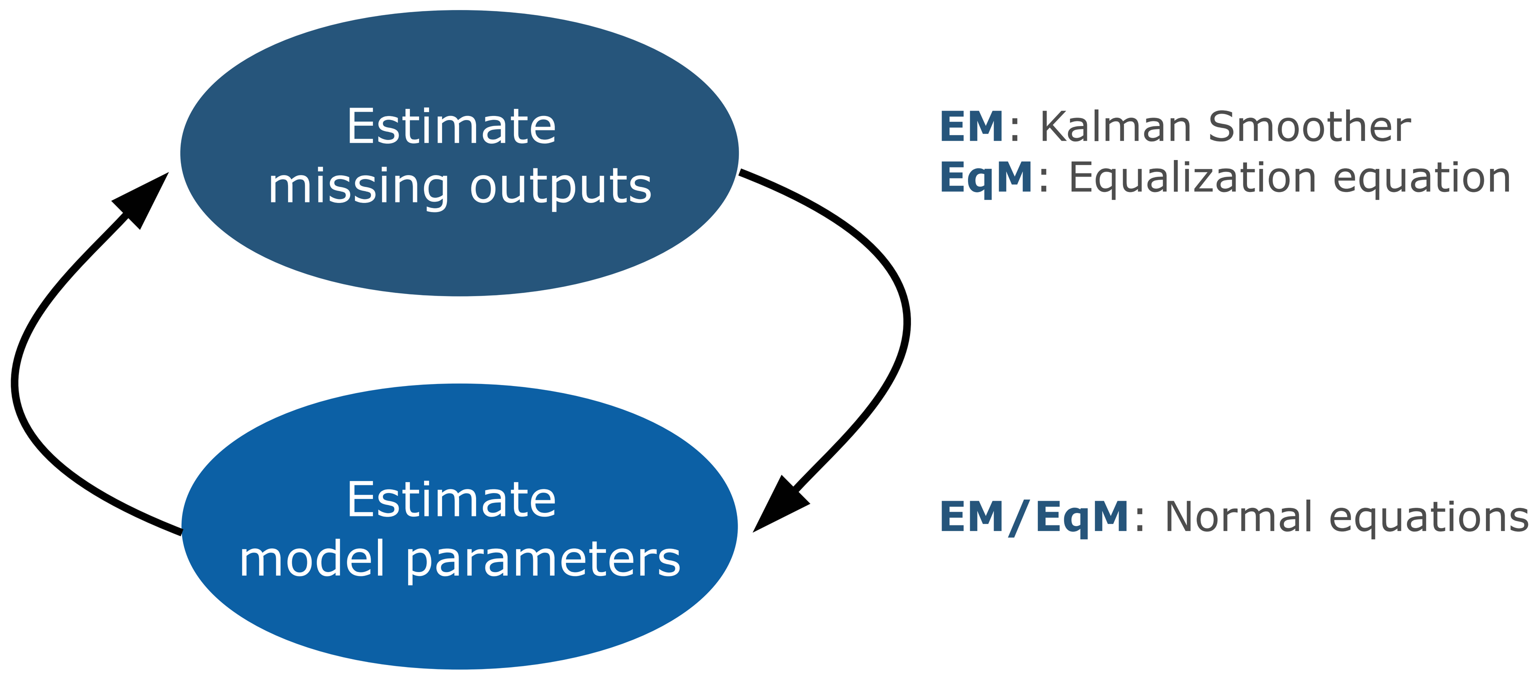
$$\ell_\theta(Z, Y) = \log p_\theta(Z, Y) = \log p_\theta(Z|Y) + \log p_\theta(Y).$$

Expectation Maximization (EM)

The traditional method for solving this ML problem is the use of the EM algorithm with a computationally costly **Kalman smoother**.

EM Algorithm Given an initial parameter estimate, $\widehat{\theta}_0$.

1. **Expectation step** Calculate
$$\mathcal{Q}(\theta, \widehat{\theta}_k) = \mathbb{E}_{\theta_k, Y}\left[\log p_\theta(Z, Y)\right] = \int \log p_\theta(Z, Y) p_{\theta_k}(Z|Y) dZ.$$
2. **Maximization step** Calculate
$$\widehat{\theta}_{k+1} = \arg\max_\theta \mathcal{Q}(\theta, \widehat{\theta}_k).$$
3. End when convergence requirements are met.



Equalization Maximization (EqM)

The joint distribution $p_\theta(Y, Z)$ can easily be maximized using e.g. the **cyclic maximization (CM) algorithm**. However, we would like to find the maximizing argument of $p_\theta(Y)$. If a function $h(Y, \widehat{\theta}_k)$ satisfies

$$\log p_\theta(Z|Y) = \log p_\theta(h(Y, \widehat{\theta}_k)|Y) = \text{const.},$$

then the log-likelihood can be rewritten as

$$\log p_\theta(Y) = \log p_\theta(Y, \widehat{Z})|_{\widehat{Z}=h(Y,\widehat{\theta}_k)} + \text{const.}$$

and then maximizing $p_\theta(Y, Z)$ is equivalent with maximizing $p_\theta(Y)$.

EqM Algorithm [1] Given an initial parameter estimate, $\widehat{\theta}_0$.

1. **Equalization step** Calculate
$$\widehat{Z}_k = h(Y, \widehat{\theta}_k).$$
2. **Maximization step** Calculate
$$\widehat{\theta}_{k+1} = \arg\max_\theta \log p_\theta\left(Y, \widehat{Z}_k\right).$$
3. End when convergence requirements are met.

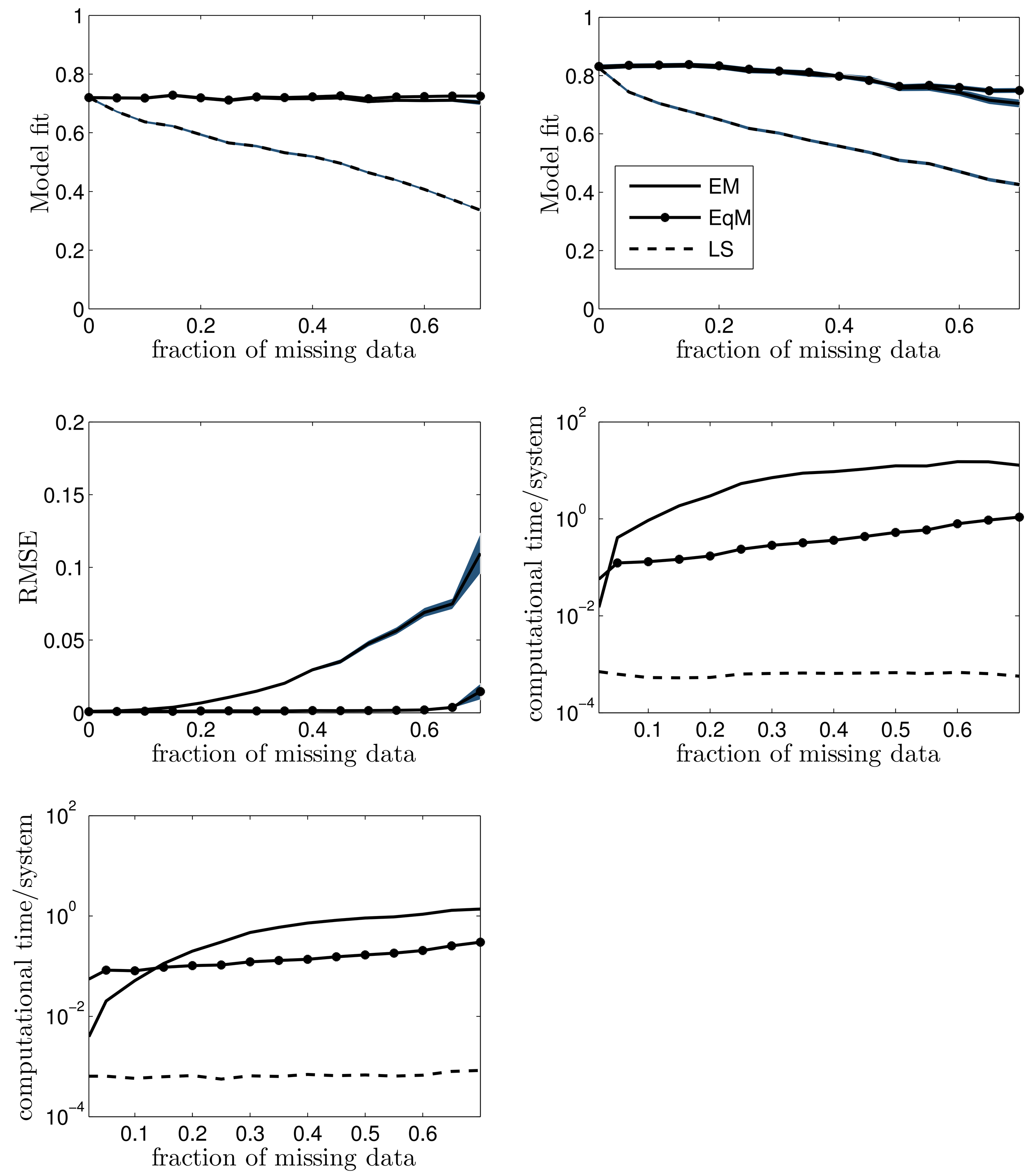
A suitable choice for the **equalization function** for Gaussian processes, can be shown to be

$$h(Y, \widehat{\theta}_k) = \mu(Y, \widehat{\theta}_k) + C_{1\cdot}(\widehat{\theta}_k) \left[\frac{\log(|C(\widehat{\theta}_0)|/|C(\widehat{\theta}_k)|)}{C_{11}(\widehat{\theta}_k)} \right]^{1/2},$$

where μ and C denote the conditional mean and covariance.

Simulation study

A comparison of model fit and RMSE using LS, EM, and EqM on 1000 randomly generated AR systems with model order 2 (left column) and random model orders, $n \in [1, 20]$ (right column).



References

[1] P. Stoica, L. Xu, and J. Li. A new type of parameter estimation algorithm for missing data problems. *Statistics & Probability Letters*, 75(3):219 – 229, 2005.

More information and source code

Available from: `http://www.control.isy.liu.se/~johda87/`.