The EM filter as a computationally efficient method for data assimilation

David Grypma and Sam Pimentel Department of Mathematical Sciences, Trinity Western University, Langley, BC, Canada

Introduction

The SMART filter (simultaneous multiplicative algebraic reconstruction technique) and the EM filter (expectation maximization) are relatively new filtering algorithms that minimize a cost function of weighted cross entropy (Kullback-Leibler) distances rather than the standard Euclidean distance. Both were originally developed to solve ill-posed inverse problems that arise in reconstructing a time-varying medical image. These algorithms hold potential for data assimilation applications in geophysical fluid problems where we are also interested in time-varying variables of large scale systems. These algorithms have advantages over the frequently used Kalman filter in that they do not involve matrix-matrix multiplication or matrix inversion and therefore are computationally more efficient. Here, we present our research on the EM filter and compare and contrast these findings to the SMART filter and Kalman filter. We implement the EM filter on a simple data assimilation application and compare results with those of the Kalman filter and the SMART filter. We demonstrate the algorithm's potential benefits for geophysical data assimilation applications and show that it is an even better alternative to the Kalman filter than the SMART filter.

Data Assimilation

Data assimilation seeks to determine the optimal state of a system given two types of information. The model:

$$x_{k+1} = Mx_k + \varepsilon_k$$

and the observations:

$$y_k = Hx_k + \delta_k$$

each containing errors. The error statistics for the initial condition, the observations, and the model are characterized as follows:

$$\mathbb{E}[x_o - x_o^b] = 0, \mathbb{E}[(x_o - x_o^b)(x_o - x_o^b)^T] = B_o$$

$$\mathbb{E}[\delta_k] = 0, \mathbb{E}[\delta_k \delta_k^T] = R$$

$$\mathbb{E}[\varepsilon_k] = 0, \mathbb{E}[\varepsilon_k \varepsilon_k^T] = Q$$

Kalman Filter

The data assimilation problem can be formulated as minimizing a cost function of the form

$$J = \frac{1}{2}||x_o - x_o^b||^2_{B_o^{-1}} + \frac{1}{2}\sum_{k=0}^{N}||Hx_k - y_k||^2_{R_k^{-1}} + \frac{1}{2}\sum_{k=0}^{N}||\varepsilon_k||^2_{Q_k^{-1}}$$

subject to the equations above.
The solution is provided by the Kalman Filter (see Evensen (2009)):

$$x_k^a = x_k^b + K_k(y_k - Hx_k^b)$$

with

$$K_k = P_k^f H^T(HP_k^f H^T + R_k)^{-1} \qquad P_k^a = (I - K_k H)P_k^f$$

$$x_{k+1}^b = Mx_k^a \qquad P_{k+1}^f = MP_k^a M^T + Q_K$$

SMART Filter

Alternatively the data assimilation problem could be stated as minimizing a cost function of the form

$$J = KL_{B_o^{-1}}(x_o, x_o^b) + \sum_{k=0}^{N} KL_{R_k^{-1}}(Hx_k, y_k) + \sum_{k=0}^{N} KL_{Q_k^{-1}}(x_{k+1}, Mx_k)$$

Instead of using Euclidean distances we use the cross entropy (Kulback-Leibler) measure, $KL(a,b) = a\log\frac{a}{b} + b - a,\ a,b \geq 0$, where we want to minimize information loss when b is used to approximate a. The solution is provided by the SMART Filter (Qranfal & Byrne, 2011b)

$$r_{ij}^{l+1} = H_{ij}\log\frac{y_i}{(Hx^l)_i} \qquad x_{k+1}^b = Mx_k^a$$

$$x_j^{l+1} = (x_j^l)^\alpha((Mx_k)_{ij})^{1-\alpha}\exp\left(\alpha\sum_{i=1}^{M} r_{ij}^{l+1}\right) \quad \text{with } \alpha = \frac{\sigma - 1}{\sigma} \text{ when } R = \sigma^2 I$$

EM Filter

The EM Filter, like the SMART Filter, uses cross entropy measure rather than Euclidean distances. A different filter is derived from the cost function:

$$J = KL_{B_o^{-1}}(x_o^b, x_o) + \sum_{k=0}^{N} KL_{R_k^{-1}}(y_k, Hx_k) + \sum_{k=0}^{N} KL_{Q_k^{-1}}(Mx_k, x_{k+1})$$

Due to the non-symmetric quality of cross entropy measure a different filter is derived, which is the EM filter (Qranfal & Byrne, 2011a)

$$x_{k+1}^b = Mx_k^a \rightarrow r_{ij}^{l+1} = \frac{P_{ij}s_j^l d_i}{(Ps^l)_i} \leftrightarrow x_j^{l+1} = \alpha\sum_{i=1}^{M} r_{ij}^{l+1} + (1-\alpha)(y_k)_j$$

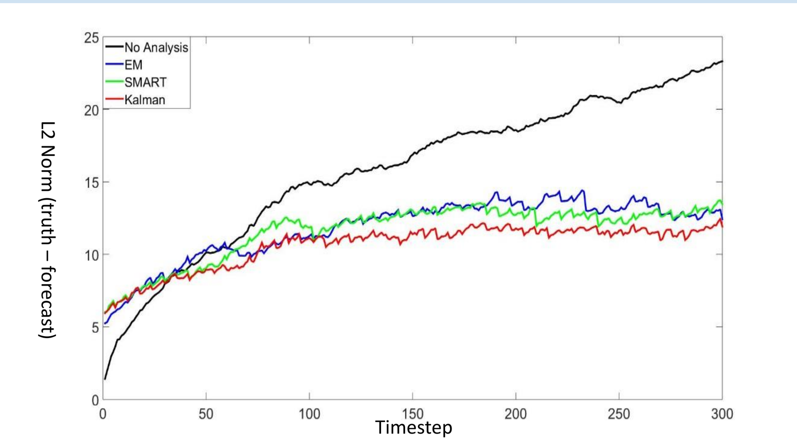
$$\text{with } \alpha = \frac{\sigma - 1}{\sigma} \text{ when } R = \sigma^2 I$$

Experiment and Results

We illustrate EM with 3 experiments. In each experiment we follow a data assimilation problem given in Evensen (2009) whereby a smooth, periodic, pseudorandom wave is advected on a periodic domain. The model has a constant advection speed of 1 m/s, grid spacing of 1 m, a time step of 1 second, and an integration length of 300 seconds. The model is solved analytically.

Experiment 1

Here we show the L2 norm (truth – forecast) vs. time for each filter to show the error propagation of each filter. Along with the general experiment outline this is on a domain size of 200 with 20 equally spaced observations taken every 6 seconds. We take the initial condition to have no error. Model error and observation error is normally distributed with mean of 0 and variance of 0.01.



Experiment 2

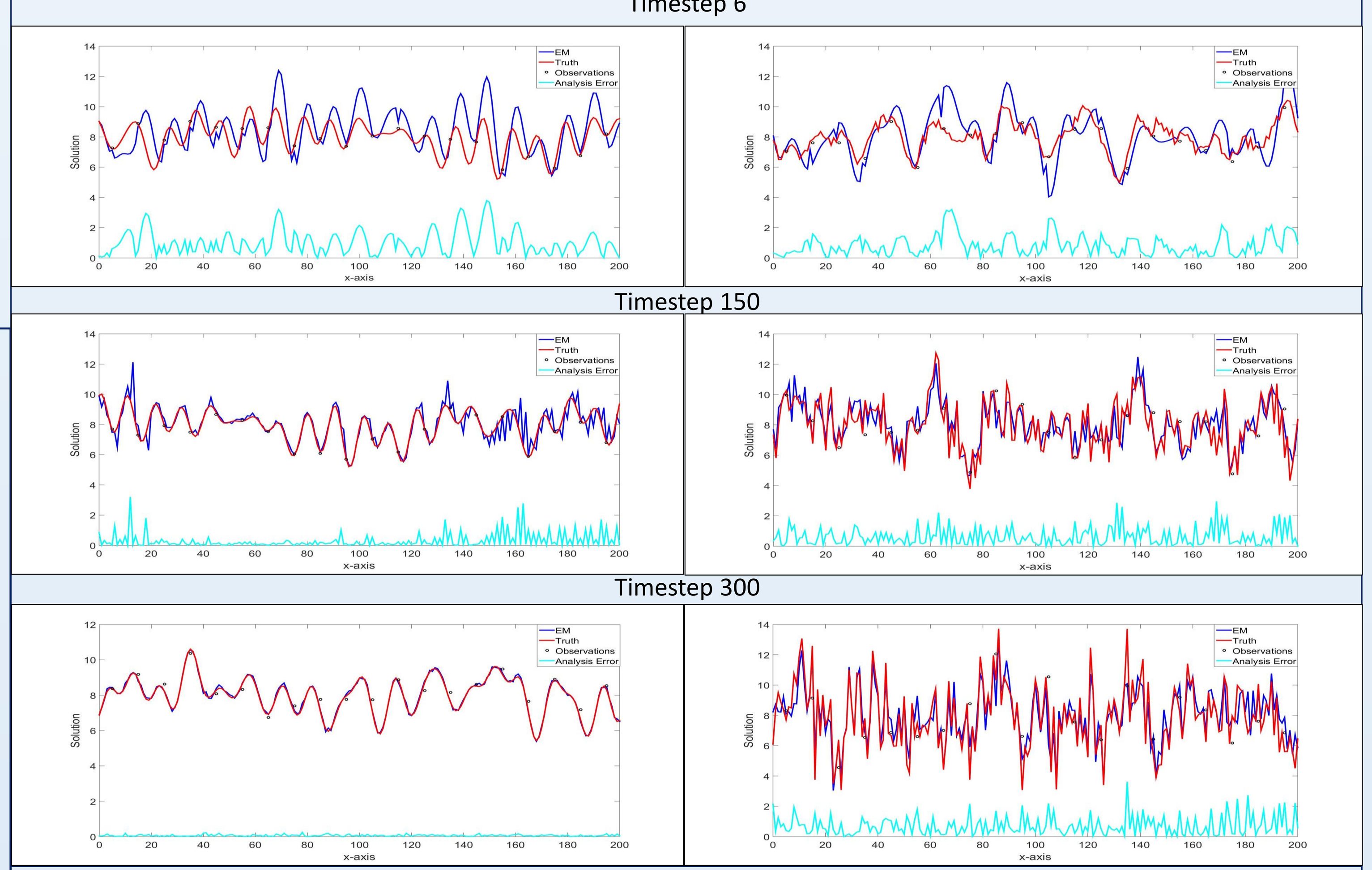
Here we compare the mean runtime and mean absolute error of each filter from a sample of 15 simulations. Along with the general experiment outline this was carried out on a domain size of 1000 with 4 equally spaced observations taken every 5 seconds. We take the initial conditions to have no error. Model error and observation error are normally distributed with mean of 0 and variance of 0.01.

	Kalman Filter	SMART Filter	EM Filter
Mean Runtime (seconds)	19.04	3.92	1.27
Mean Absolute Error	1.27	1.29	1.30

Experiment 3

Here we compare the convergence of the EM filter when there is no model error and when there is model error applied to the system. Along with the general experiment outline this was carried out on a domain size of 200, with 20 equally spaced observations of the truth every 6 seconds, and observation error with mean 0 and variance 0.01. When model error is applied, it has mean 0 and variance 0.01. The initial condition is offset from the truth by adding a normally distributed error with mean 0, variance 1, and de-correlation length of 20 m.

No Model Error — Model Error



Current and Future Work

The EM filter as implemented here assumes white noise in both observation and model errors. In situations where either the observation error or the model error are not white, then a pre-whitening transform must be applied to the corresponding system of equations. This process is outlined in Qranfal & Byrne (2011b):

"For instance, assume we have an error or noise vector w that is colored. That is, $\mathbb{E}[w] = \mu_w \neq 0$ or $\mathbb{E}[(w - \mu_w)(w - \mu_w)^T] \neq \sigma^2 I$ Let

$$v = \Lambda^{-1/2}E^T(w - \mu_w)$$

where E is the orthonormal matrix of eigenvectors and Λ is the diagonal matrix of positive eigenvalues of the spectral decomposition of the definite positive covariance matrix."

We have successfully implemented a pre-whitening transform for non-white observation errors, and are currently working on an implementation of the transform for non-white model error for both the SMART and EM filter. Implementation of this transform will open up more realistic geophysical applications for the SMART and EM filters.

Sources

J. Qranfal and C. Byrne, 'EM Filter for Time-Varying SPECT Reconstruction', *International Journal of Pure and Applied Mathematics*, vol. 73, no. 4, pp. 379-403, 2011.
J. Qranfal and C. Byrne, 'SMART Filter for Dynamic SPECT Image Reconstruction', *International Journal of Pure and Applied Mathematics*, vol. 73, no. 4, pp. 405-434, 2011.
G. Evensen, Data Assimilation, The Ensemble Kalman Filter, Springer, 2009.

Acknowledgements

This work was gratefully supported by a Trinity Western University Undergraduate Student Research Award to David Grypma with funding from Sam Pimentel's NSERC Discovery Grant, a private donor, and the TWU Mathematics department. We thank Joe Qranfal for valuable discussions