Weighted Ensemble Square Root Filters for Non-linear, Non-Gaussian, Data Assimilation

OR

When Can You Turn an Ensemble Kalman Filter into a Particle Filter?

David Livings Peter Jan van Leeuwen

Data Assimilation Research Centre
University of Reading

International Conference on Ensemble Methods in Geophysical Sciences
13 November 2012

Motivation

- The ensemble Kalman filter (EnKF) is now widely used.
- The particle filter (PF) offers the possibility of better performance in non-linear, non-Gaussian situations.
- Can an existing EnKF be converted to a PF?
- Papadakis et al (2010) introduced the weighted EnKF (WEnKF) combining best features of EnKF and PF; they concentrated on perturbed-observations EnKF.
- Van Leeuwen (2009) sketched a similar approach for ensemble square roots filters (SRFs), but there was a flaw in the theory.
- This work fixes the flaw and simplifies and generalises the theory.
- But does it work in practice?

Outline

Particle Filter with Proposal Density

Proposal Density from Ensemble Square Root Filter

Test Results

Filtering Problem

- Discrete time Markov process with states $\mathbf{x}_t$ $(t = 0, 1, \ldots)$
- Characterised by $p(\mathbf{x}_0)$, $p(\mathbf{x}_t|\mathbf{x}_{t-1})$ $(t \geq 1)$
- Conditionally independent observations $\mathbf{y}_t$ $(t = 1, 2, \ldots)$
- Characterised by $p(\mathbf{y}_t|\mathbf{x}_t)$ $(t \geq 1)$
- The filtering problem is to determine $p(\mathbf{x}_t|\mathbf{y}_{1:t})$ where $\mathbf{y}_{1:t} = (\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_t)$.

Particle Filter with Proposal Density

- Particle representation $p(\mathbf{x}_t|\mathbf{y}_{1:t}) \approx \sum_i w_t^{(i)} \delta(\mathbf{x}_t - \mathbf{x}_t^{(i)})$
- Algorithm is based on proposal density $q(\mathbf{x}_{0:t}|\mathbf{y}_{1:t})$ satisfying

$$q(\mathbf{x}_{0:t}|\mathbf{y}_{1:t}) = q(\mathbf{x}_t|\mathbf{x}_{t-1}, \mathbf{y}_t) q(\mathbf{x}_{0:t-1}|\mathbf{y}_{1:t-1})$$

Particle Filter with Proposal Density: Algorithm

- Build up particles recursively:

$$\begin{aligned} \mathbf{x}_0^{(i)} &\sim q(\mathbf{x}_0) \\ \mathbf{x}_t^{(i)} &\sim q(\mathbf{x}_t|\mathbf{x}_{t-1}^{(i)}, \mathbf{y}_t) \end{aligned}$$

- Update weights recursively:

$$\begin{aligned} w_0^{(i)} &\propto \frac{p(\mathbf{x}_0^{(i)})}{q(\mathbf{x}_0^{(i)})} \\ w_t^{(i)} &\propto \frac{p(\mathbf{y}_t|\mathbf{x}_t^{(i)})p(\mathbf{x}_t^{(i)}|\mathbf{x}_{t-1}^{(i)})}{q(\mathbf{x}_t^{(i)}|\mathbf{x}_{t-1}^{(i)}, \mathbf{y}_t)} w_{t-1}^{(i)} \end{aligned}$$

with constraint $\sum_i w_t^{(i)} = 1$

Particle Filter with Proposal Density: Requirements for Implementation

To implement the algorithm, we must be able to:

- Sample from $q(\mathbf{x}_0)$ and $q(\mathbf{x}_t|\mathbf{x}_{t-1}, \mathbf{y}_t)$
- Evaluate $q(\mathbf{x}_0)$, $q(\mathbf{x}_t|\mathbf{x}_{t-1}, \mathbf{y}_t)$, $p(\mathbf{x}_0)$, $p(\mathbf{x}_t|\mathbf{x}_{t-1})$, $p(\mathbf{y}_t|\mathbf{x}_t)$

Particle Filter with Proposal Density

Proposal Density from Ensemble Square Root Filter

Test Results

Ensemble Kalman Filter

- Dynamical and observation models:

$$
\begin{aligned}
\mathbf{x}_t &= M(\mathbf{x}_{t-1}) + \boldsymbol{\eta}_{t-1} \\
\mathbf{y}_t &= H(\mathbf{x}_t) + \boldsymbol{\epsilon}_t
\end{aligned}
$$

- $\boldsymbol{\eta}_t$, $\boldsymbol{\epsilon}_t$ random, satisfy certain independence conditions
- Ensemble Kalman filter maintains ensemble of states $\mathbf{x}_t^{(i)}$ through alternating forecast and analysis steps.
- Forecast step: $\mathbf{x}_t^{\mathrm{f}(i)} = M(\mathbf{x}_{t-1}^{(i)}) + \boldsymbol{\eta}_{t-1}^{(i)}$
- Analysis step: uses $\mathbf{y}_t$ to transform forecast ensemble $\mathbf{x}_t^{\mathrm{f}(i)}$ to analysis ensemble $\mathbf{x}_t^{(i)}$
- Plan: use analysis ensemble as particles in particle filter
- Question: what $q(\mathbf{x}_t|\mathbf{x}_{t-1}, \mathbf{y}_t)$ have we sampled from?

Evaluation of $p(\mathbf{y}_t|\mathbf{x}_t)$

- $p(\mathbf{y}_t|\mathbf{x}_t) = p_{\epsilon_t}(\mathbf{y}_t - H(\mathbf{x}_t))$
- Can evaluate as long as can evaluate $H(\mathbf{x}_t)$, $p(\epsilon_t)$
- Don't have to use same $\mathbf{y}_t$, $H(\mathbf{x}_t)$, $p(\epsilon_t)$ in PF as in EnKF
- Opens possibility of using PF to assimilate observations additional to those used in EnKF

Evaluation of $p(\mathbf{x}_t|\mathbf{x}_{t-1})$

- $p(\mathbf{x}_t|\mathbf{x}_{t-1}) = p_{\boldsymbol{\eta}_{t-1}}(\mathbf{x}_t - M(\mathbf{x}_{t-1}))$
- Can evaluate as long as can evaluate $M(\mathbf{x}_{t-1})$, $p(\boldsymbol{\eta}_{t-1})$
- Don't have to evaluate $M(\mathbf{x}_{t-1}^{(i)})$ if EnKF stores it
- Don't have to use same $M(\mathbf{x}_{t-1})$, $p(\boldsymbol{\eta}_{t-1})$ in PF as in EnKF
- Possible problem: what if there are multiple forecast steps between observation times?
- Solution: treat intermediate steps as complete EnKF steps with uninformative observations ($p(\mathbf{y}_t|\mathbf{x}_t)$ constant) and analysis step that leaves forecast ensemble unchanged:

$$w_t^{(i)} \propto \frac{p(\mathbf{x}_t^{(i)}|\mathbf{x}_{t-1}^{(i)})}{q(\mathbf{x}_t^{(i)}|\mathbf{x}_{t-1}^{(i)})} w_{t-1}^{(i)}$$

- With same dynamical model, reduces to $w_t^{(i)} = w_{t-1}^{(i)}$

Ensemble Square Root Filters

Tippett et al (2003)

Essentials of analysis algorithm:

$$\begin{aligned}\overline{\mathbf{x}} &= \overline{\mathbf{x}^{\mathrm{f}}} + \mathbf{K}^{\mathrm{e}}(\mathbf{y} - \overline{\mathbf{y}^{\mathrm{f}}}) \\ \mathbf{X} &= \mathbf{X}^{\mathrm{f}}\mathbf{T}\end{aligned}$$

where

- $\overline{\mathbf{x}}$ is analysis ensemble mean
- $\overline{\mathbf{x}^{\mathrm{f}}}$ is forecast ensemble mean
- $\mathbf{K}^{\mathrm{e}}$ is ensemble approximation of Kalman gain
- $\overline{\mathbf{y}^{\mathrm{f}}}$ is mean of forecast observation ensemble $\mathbf{y}^{\mathrm{f}(i)} = H(\mathbf{x}^{\mathrm{f}(i)})$
- $\mathbf{X}$ is analysis ensemble perturbation matrix
- $\mathbf{X}^{\mathrm{f}}$ is forecast ensemble perturbation matrix
- $\mathbf{T}$ is a function of forecast ensemble members; different formulations use different functions.

Potential Problem for Particle Filter

$$
\begin{aligned}
\overline{\mathbf{x}} &= \overline{\mathbf{x}^{\mathrm{f}}} + \mathbf{K}^{\mathrm{e}}(\mathbf{y} - \overline{\mathbf{y}^{\mathrm{f}}}) \\
\mathbf{X} &= \mathbf{X}^{\mathrm{f}}\mathbf{T}
\end{aligned}
$$

- Analysis step destroys independence of particles.
- For one thing, $\overline{\mathbf{x}}$ depends on all $\mathbf{x}^{\mathrm{f}(i)}$ via $\overline{\mathbf{x}^{\mathrm{f}}}$.
- Fortunately, $\overline{\mathbf{x}^{\mathrm{f}}}$ tends to property of system as whole as ensemble size $N \to \infty$.
- Similarly for $\mathbf{K}^{\mathrm{e}}$ and $\overline{\mathbf{y}^{\mathrm{f}}}$
- But there are two problems with $\mathbf{T}$.
- $\mathbf{T}$ mixes up ensemble members.
- $\mathbf{T}$ is $N \times N$, so cannot tend to limit as $N \to \infty$.

Solution

- Most ensemble square root filters can be written in pre-multiplier form:

$$\mathbf{X} = \mathbf{A}\mathbf{X}^{\mathrm{f}}$$

- This applies to:
  - EAKF (Anderson, 2001)
  - Filter of Whitaker and Hamill (2002)
  - Symmetric version of ETKF (Wang et al, 2004)
  - Any ensemble square root filter that is unbiased and nondegenerate in a certain sense (Livings et al, 2008)
- $\mathbf{A}$ has fixed size, so there is hope it tends to limit as $N \to \infty$.
- For specific filters above, $\mathbf{A}$ does tend to property of system as whole as $N \to \infty$.
- If $\mathbf{A}$ is property of system as whole, independence of particles is preserved.

Evaluation of $q(\mathbf{x}_t|\mathbf{x}_{t-1}, \mathbf{y}_t)$ and Weights

- Generalise to any analysis step of the form

$$\mathbf{x}_t^{(i)} = \mathbf{A}(\mathbf{y}_t)\mathbf{x}_t^{\mathrm{f}(i)} + \mathbf{b}(\mathbf{y}_t)$$

  where $\mathbf{A}(\mathbf{y}_t)$, $\mathbf{b}(\mathbf{y}_t)$ are properties of system as whole.
- Then $q(\mathbf{x}_t^{(i)}|\mathbf{x}_{t-1}^{(i)}, \mathbf{y}_t) = |\mathbf{A}(\mathbf{y}_t)|^{-1} q(\mathbf{x}_t^{\mathrm{f}(i)}|\mathbf{x}_{t-1}^{(i)})$.
- $|\mathbf{A}(\mathbf{y}_t)|^{-1}$ is the same for all particles, so can be dropped from the weights:

$$w_t^{(i)} \propto \frac{p(\mathbf{y}_t|\mathbf{x}_t^{(i)})p(\mathbf{x}_t^{(i)}|\mathbf{x}_{t-1}^{(i)})}{q(\mathbf{x}_t^{\mathrm{f}(i)}|\mathbf{x}_{t-1}^{(i)})} w_{t-1}^{(i)}$$

- $q(\mathbf{x}_t^{\mathrm{f}(i)}|\mathbf{x}_{t-1}^{(i)})$ is just $q(\boldsymbol{\eta}_{t-1}^{(i)})$.

Theoretical Conclusions

- It appears to be easy to convert an ensemble SRF into a PF:

$$w_t^{(i)} \propto \frac{p(\mathbf{y}_t|\mathbf{x}_t^{(i)})p(\mathbf{x}_t^{(i)}|\mathbf{x}_{t-1}^{(i)})}{q(\mathbf{x}_t^{\mathrm{f}(i)}|\mathbf{x}_{t-1}^{(i)})} w_{t-1}^{(i)}$$

- Call this a weighted ensemble SRF.
- No linear or Gaussian assumptions are necessary.
- The filter doesn't even have to be an ensemble SRF; it just has to have an analysis step of the form

$$\mathbf{x}_t^{(i)} = \mathbf{A}(\mathbf{y}_t)\mathbf{x}_t^{\mathrm{f}(i)} + \mathbf{b}(\mathbf{y}_t)$$

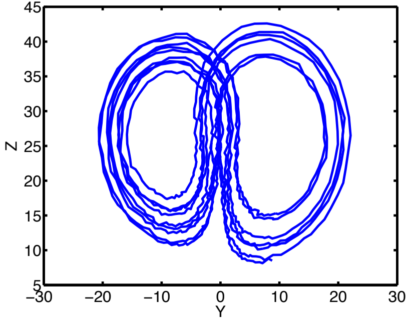
- As long as refinements such as inflation or localisation preserve this form, the particle filter is still correct.

Particle Filter with Proposal Density

Proposal Density from Ensemble Square Root Filter

Test Results

Test System



First 1200 timesteps of truth

- Lorenz-63 as used by Ades and Van Leeuwen (2012) to test equivalent weights particle filter
- Gaussian model noise each timestep, $\mathrm{std} = \sqrt{2\Delta t}$, $\Delta t = 0.01$
- Observations of all coordinates every 40 timesteps, Gaussian errors, $\mathrm{std} = \sqrt{2}$
- Gaussian initial distribution, $\mathrm{std} = \sqrt{2}$
- Weighted ensemble SRF from EAKF with 10 particles

Results

- After assimilating the first observation, the weights of all but one of the particles have collapsed to zero.
- The problem is that the displacement of a particle during the analysis is typically much larger than the single-step model noise, making very small values of $p(\mathbf{x}_t^{(i)}|\mathbf{x}_{t-1}^{(i)})$ likely.

$$w_t^{(i)} \propto \frac{p(\mathbf{y}_t|\mathbf{x}_t^{(i)})\color{red}{p(\mathbf{x}_t^{(i)}|\mathbf{x}_{t-1}^{(i)})}}{q(\mathbf{x}_t^{\mathrm{f}(i)}|\mathbf{x}_{t-1}^{(i)})} w_{t-1}^{(i)}$$

- Replacing the EAKF with an ETKF doesn't help.
- Using 100 particles doesn't help.

Practical Conclusions

- The weighted ensemble SRF is only viable if the single-step model noise is at least comparable in size with the particle displacements during analysis.
- This is not so for the Lorenz-63 test system.
- Is it so for an NWP system?

References I

M. Ades and P. J. van Leeuwen.
An exploration of the equivalent weights particle filter.
Q. J. R. Meteorol. Soc., 2012.
In press.

Jeffrey L. Anderson.
An ensemble adjustment Kalman filter for data assimilation.
Mon. Wea. Rev., 129(12):2884–2903, 2001.

David M. Livings, Sarah L. Dance, and Nancy K. Nichols.
Unbiased ensemble square root filters.
Physica D, 237(8):1021–1028, 2008.

Nicolas Papadakis, Etienne Mémin, Anne Cuzol, and Nicolas Gengembre.
Data assimilation with the weighted ensemble Kalman filter.
Tellus, 62A(5):673–697, 2010.

References II

- Michael K. Tippett, Jeffrey L. Anderson, Craig H. Bishop, Thomas M. Hamill, and Jeffrey S. Whitaker.
  Ensemble square root filters.
  Mon. Wea. Rev., 131(7):1485–1490, 2003.

- Peter Jan van Leeuwen.
  Particle filtering in geophysical systems.
  Mon. Wea. Rev., 137(12):4089–4114, 2009.

- Xuguang Wang, Craig H. Bishop, and Simon J. Julier.
  Which is better, an ensemble of positive-negative pairs or a centered spherical simplex ensemble?
  Mon. Wea. Rev., 132(7):1590–1605, 2004.

- Jeffrey S. Whitaker and Thomas M. Hamill.
  Ensemble data assimilation without perturbed observations.
  Mon. Wea. Rev., 130(7):1913–1924, 2002.