Polynomial Regression and Serial Ensemble Kalman Filtering

Dan Hodyss\textsuperscript{1}, Bill Campbell\textsuperscript{1}, Jeff Anderson\textsuperscript{2}, Nancy Collins\textsuperscript{2}, Alex Reinecke\textsuperscript{1}

1. Naval Research Laboratory, Monterey, California
2. National Center for Atmospheric Research, Boulder, Colorado
Posterior Mean

Our goal in this talk is to discuss the estimation of the posterior mean

\[ \bar{x}(y) = \int_{-\infty}^{\infty} xp(x \mid y) \, dx \]

where \( p(x \mid y) \) is the posterior density.
The Posterior Mean is a function of $y$

$$\overline{x}(y) = \int_{-\infty}^{\infty} xp(x | y) dx$$
Hodyss (2011; MWR) proved that the curvature was determined by the posterior third moment:

\[
\frac{d^2 \bar{x}}{dy^2} = \frac{T(y)}{R^2}
\]

\(T(y)\) = Posterior Third Moment
\(R\) = Observation error variance

Multivariate generalization can be found in Hodyss (2011)
Hodyss (2011; MWR) shows how to expand the posterior mean as

\[
\bar{x}(y) = \bar{x}_f + G_1 \left[ y - \bar{x}_f \right] + G_2 \left[ \left( y - \bar{x}_f \right)^2 - \alpha_2 \right] + G_3 \left[ \left( y - \bar{x}_f \right)^3 - \alpha_3 \right] + \ldots
\]

The Kalman filter is “optimal” when:

1. The innovation is small
2. Prior is not too skewed
Polynomial Filtering

What’s remarkable about this is it looks just like a Kalman filter!

\[ \bar{x}(y) = \bar{x}_f + Gd \]

Gain:
\[ G = [G_1 \ G_2 \ G_3 \ \ldots] \]

Innovation vector with new pseudo-obs!
\[ d = \begin{bmatrix} y - \bar{x}_f \\ (y - \bar{x}_f)^2 - \alpha_2 \\ (y - \bar{x}_f)^3 - \alpha_3 \\ \vdots \end{bmatrix} \]
All-at-Once vs. Serial-Solve

- Hodyss (2012; MWR) showed how to do the all-at-once solve.
  - Uses a minimization-based technique for the matrix inverse and a post-multiply step
- This talk will introduce the serial formulation.
  - There is a version of DART that now includes quadratic polynomial regression:
    - Ensemble Adjustment Quadratic Filter (EAQF)
    - Ensemble Kalman Quadratic Filter (EnQF)
Define the “squared pseudo-ob” as: \[ y_i^s = \left( y_i - h_i \bar{x}_i^f \right)^2 - r_i \]
Define the “squared state” as: \[ s_j = \left( x_j - \bar{x}_j^f \right)^2 \]

**Step 1: Assimilate regular ob**

\[
\bar{x}_i^a = \bar{x}_i^f + P_2 h_i^T \frac{y_i - h_i \bar{x}_i^f}{P_{2ii} + r_i}
\]

\[
s_i^a = s_i^f + P_3 h_i^T \frac{y_i - h_i \bar{x}_i^f}{P_{2ii} + r_i}
\]

**Step 2: Assimilate the “pseudo-ob”**

\[
\bar{x}_i^a = \bar{x}_i^f + P_3 h_i^T \frac{y_i^s - s_i^f}{P_{4ii} + r_{2i}}
\]

\[
s_i^a = s_i^f + P_4 h_i^T \frac{y_i^s - s_i^f}{P_{4ii} + r_{2i}}
\]
Serial Cubic Nonlinear Regression

**Step 1: Assimilate regular ob**

\[
\bar{x}^a = \bar{x}^f + P_2 h_i^T \frac{y_i - h_i \bar{x}^f}{P_{2ii} + r_i} \quad s^a = s^f + P_3 h_i^T \frac{y_i - h_i \bar{x}^f}{P_{3ii} + r_i} \quad c^a = c^f + P_4 h_i^T \frac{y_i - h_i \bar{x}^f}{P_{4ii} + r_i}
\]

**Step 2: Assimilate the “squared” ob**

\[
\bar{x}^a = \bar{x}^f + P_3 h_i^T \frac{y_i^s - s_i^f}{P_{4ii} + r_{2i}} \quad s^a = s^f + P_4 h_i^T \frac{y_i^s - s_i^f}{P_{4ii} + r_{2i}} \quad c^a = c^f + P_5 h_i^T \frac{y_i^s - s_i^f}{P_{4ii} + r_{2i}}
\]

**Step 3: Assimilate the “cubed” ob**

\[
\bar{x}^a = \bar{x}^f + P_4 h_i^T \frac{y_i^c - c_i^f}{P_{6ii} + r_{3i}} \quad s^a = s^f + P_5 h_i^T \frac{y_i^c - c_i^f}{P_{6ii} + r_{3i}} \quad c^a = c^f + P_6 h_i^T \frac{y_i^c - c_i^f}{P_{6ii} + r_{3i}}
\]
Scalar Test Problem: Large Ensemble

DA Experiment:
Gaussian Ob Likelihood (R=1)
Gamma Prior with (k,θ) chosen such that

\[ P = 1 \text{ but varying skewness} \]

Kalman and Quadratic formulas evaluated
with known analytic moments
(i.e. infinite ensemble)

Particle filter uses \(10^6\) particles

Experiment is repeated \(10^6\) times for
different truths and different obs

10% reduction in MSE when prior
skewness is about 1.5
Scalar Test Problem: Small Ensemble

Is truncating the expansion a useful method for reducing sampling error?
DA in Lorenz-63

Obs of x and z
Ensemble size: 1000 members
Ran for 1000 cycles

Adaptive prior inflation (inf_flavor = 2)

Posterior RMSE of X-Variable

R = 0.1

EAKF – Square Root
EnKF – Perturbed Obs
EAQF – Square Root
EnQF – Perturbed Obs

R = 0.5

Number of Time Steps Between Obs

Number of Time Steps Between Obs
Prior Skewness in Lorenz-63

Average absolute value of skewness over last 800 cycles.

Solid:  \( R = 0.1 \)
Dashed:  \( R = 0.5 \)

For \( R = 0.1 \) skewness is substantial after 6 time steps between observations

For \( R = 0.5 \) skewness is substantial after 3 time steps between observations

This appears to explain the posterior separation between KPO and QPO on previous slide.
DA in the Bgrid GCM

Resolution: 60x30x5 (approximately 6 degrees)
Obs: 10 soundings of U, V, T, Ps with $R = 1$
Ensemble Size: 1000 members

Adaptive prior inflation ($\text{inf\_flavor} = 2$)
Ran DA for 365 cycles

- **Solid** = Prior: EAKF-EAQF
- **Solid** = Prior: EnKF-EnQF
- **Dashed** = Posterior: EAKF-EAQF
- **Dashed** = Posterior: EnKF-EnQF

RMSE Difference between Methods
**Bgrid Prior Skewness: Obs 1 Day Apart**

Average absolute value of skewness over last 200 cycles.

**Zonal Wind: Lowest Layer**

Red Circles = Ob locations
Bgrid Prior Skewness: Obs 10 Days Apart

Average absolute value of skewness over last 200 cycles.

Zonal Wind: Lowest Layer

Red Circles = Ob locations
Summary and Future Work

• Polynomial filtering is easily implementable within an already constructed Ensemble-Based Kalman filter.

• There’s little point to non-Gaussian methods if the skewness is small.
  – QC: Only create “pseudo-obs” for obs in regions of high skewness ... ?

• Future Work: Testing these results with nonlinear ob operators, higher resolution, and trying this out in the Navy’s aerosol model (NAAPS).
NAAPS Prior Skewness:
Total Aerosol Optical Depth
Importance of the Third Moment

If you believe that using a nonlinear observation operator in the numerator of the Kalman gain is useful, viz.

\[ \langle (x - \bar{x})(h(x) - \bar{h}(x)) \rangle \]

Then note the Taylor-series about \( x = \bar{x} \):

\[ h(x) = h(\bar{x}) + \frac{dh}{dx}(x - \bar{x}) + \frac{1}{2} \frac{d^2h}{dx^2}(x - \bar{x})^2 + \ldots \]

Use this in the covariance to obtain

\[ \langle (x - \bar{x})(h(x) - \bar{h}(x)) \rangle = \frac{dh}{dx} \langle (x - \bar{x})^2 \rangle + \frac{1}{2} \frac{d^2h}{dx^2} \langle (x - \bar{x})^3 \rangle + \ldots \]

The lowest-order impact is from the third moment!