Study of the Algorithmic Complexity of the Ensemble Kalman Filter and its Efficient Implementations

Jhon E. Hinestroza R.
PhD. Student in Mathematical Engineering

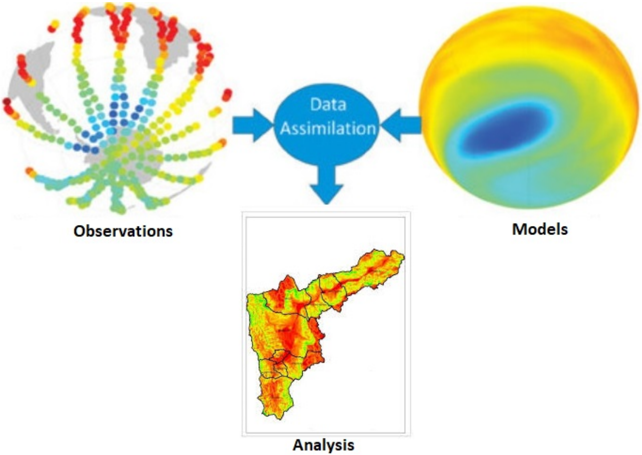
PhD. Olga L. Quintero M. and PhD. Angela M. Rendón P.
Advisor and Co-Advisor

Universidad EAFIT
Doctoral Seminar II

May 24, 2019

Outline

Introduction



Introduction

Linear and Non-Linear Data Assimilation

Gaussian

Ensemble Kalman Filter-EnKF

Variational Data Assimilation

Non-Gaussian

Particle Filters

UNIVERSIDAD EAFIT

Introduction

Linear and Non-Linear Data Assimilation

Gaussian

- Ensemble Kalman Filter-EnKF
- Variational Data Assimilation

Non-Gaussian

- Particle Filters

Current Work

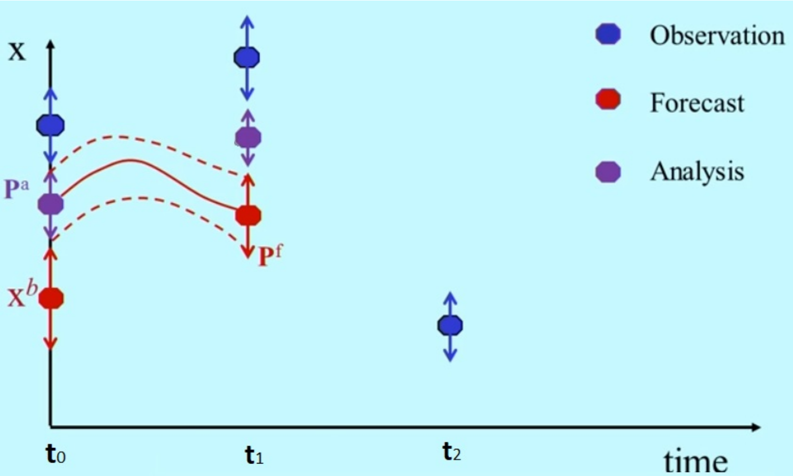
- REnKF
- En 4DVar
- PF 4DVar

UNIVERSIDAD EAFIT

Foundation of the Problem

$$p(\mathbf{x}|\mathbf{y}) = \frac{p(\mathbf{y}|\mathbf{x}) \cdot p(\mathbf{x})}{p(\mathbf{y})}$$

Kalman Filter



Kalman Filter

Assume we seek to estimate the state $\mathbf{x} \in \mathbb{R}^n$

$$\mathbf{x}_{k+1} = \mathbf{M}\left(\mathbf{x}_k, t_k\right) + \mathbf{w}_k,$$

using the measurements $\mathbf{y} \in \mathbb{R}^m$

$$\mathbf{y}_k = \mathbf{H}_k \mathbf{x}_k + \mathbf{v}_k,$$

with

$$\mathbf{w}_k \sim \mathbf{N}(\mathbf{0}, \mathbf{Q}_k),$$
$$\mathbf{v}_k \sim \mathbf{N}(\mathbf{0}, \mathbf{R}_k),$$

$\mathbf{Q}_k \in \mathbb{R}^{n\times n}$, $\mathbf{R}_k \in \mathbb{R}^{m\times m}$.
$\mathbf{M} : \mathbb{R}^n \rightarrow \mathbb{R}^n$, $\mathbf{H} : \mathbb{R}^m \rightarrow \mathbb{R}^n$.

Kalman Filter

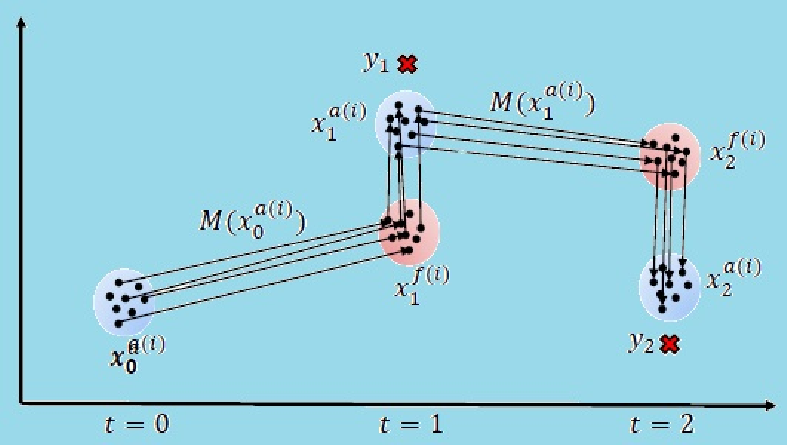
1. Forecast Step:

$$\mathbf{x}^f_{k+1} = \mathbf{M}_{k+1}\mathbf{x}^a_k,$$
$$\mathbf{P}^f_{k+1} = \mathbf{M}_{k+1}\mathbf{P}^a_k\mathbf{M}^T_{k+1} + \mathbf{Q}_{k+1}.$$

2. Analysis Step:

$$\mathbf{K}_{k+1} = \mathbf{P}^f_{k+1}\mathbf{H}^T\left(\mathbf{H}\mathbf{P}^f_{k+1}\mathbf{H}^T + \mathbf{R}_{k+1}\right)^{-1},$$
$$\mathbf{x}^a_{k+1} = \mathbf{x}^f_{k+1} + \mathbf{K}_{k+1}\left(\mathbf{y}_{k+1} - \mathbf{H}\mathbf{x}^f_{k+1}\right),$$
$$\mathbf{P}^a_{k+1} = \left(\mathbf{I} - \mathbf{K}_{k+1}\mathbf{H}\right)\mathbf{P}^f_{k+1}.$$

Ensemble Kalman Filter EnKF



UNIVERSIDAD EAFIT

Ensemble Kalman Filter EnKF

1. Forecast Step:

$$\mathbf{x}_{k+1}^{f} = \mathbf{M}_{k+1}\left(\mathbf{x}_{k}^{a}\right),$$
$$\mathbf{P}_{k+1}^{f} = \frac{1}{N-1}\sum_{i=1}^{N}\left(\mathbf{x}_{i}^{f} - \overline{\mathbf{x}}^{f}\right)\left(\mathbf{x}_{i}^{f} - \overline{\mathbf{x}}^{f}\right)^{T},$$

with N, number of ensemble members and

$$\overline{\mathbf{x}}^{f} = \frac{1}{N}\sum_{i=1}^{N}\mathbf{x}_{i}^{f}.$$

2. Analysis Step:

$$\mathbf{K}_{k+1} = \mathbf{P}_{k+1}^{f}\mathbf{H}^{T}\left(\mathbf{H}\mathbf{P}_{k+1}^{f}\mathbf{H}^{T} + \mathbf{R}_{k+1}\right)^{-1},$$
$$\mathbf{x}_{k+1}^{a} = \mathbf{x}_{k+1}^{f} + \mathbf{K}_{k+1}\left(\mathbf{y}_{k+1} - \mathbf{H}\mathbf{x}_{k+1}^{f}\right).$$

UNIVERSIDAD EAFIT®

Efficient Implementation

- **Advantage**
  1. To reduce the computational cost, in terms of the number of operations, of assimilating large data sets.
  2. The resulting algorithms scale linearly with respect to the number of observations.

Efficient Implementation

- **Advantage**
  1. To reduce the computational cost, in terms of the number of operations, of assimilating large data sets.
  2. The resulting algorithms scale linearly with respect to the number of observations.
- **Issues**
  1. The computational cost of the subsequent matrix operations can become expensive.
  2. The additional operations may contribute significantly to the total computational cost of the implementation.

UNIVERSIDAD EAFIT

SVD Implementation

procedure (SVD-EnKF)$\left(\mathbf{X}, \mathbf{X}', \mathbf{HX}', \mathbf{D}, \mathbf{E}\right)$

  $[\Sigma,\ \mathbf{U},\ \mathbf{V}] \leftarrow \text{SVD}(\mathbf{HX}' + \mathbf{E})\ (mN^2)$

  $\Lambda \leftarrow \Sigma\Sigma^T\ (m)$

  $s \leftarrow \sum_i \lambda_{i,i}\ (m)$

  $p \leftarrow max\left\{k \mid \sum_k \lambda_{k,k|s<0.99}\right\}$

  $\mathbf{X}_1 \leftarrow \Lambda^{-1}\mathbf{U}^T\ (mp)$

  $\mathbf{X}_2 \leftarrow \mathbf{X}_1\mathbf{D}\ (mnp)$

  $\mathbf{X}_3 \leftarrow \mathbf{U}\mathbf{X}_2\ (mNp)$

  $\mathbf{X}_4 \leftarrow \left(\mathbf{HX}'\right)^T\mathbf{X}_3\ (mN^2)$

  $\mathbf{X}^a \leftarrow \mathbf{X} + \mathbf{X}'\mathbf{X}_4\ (nN^2)$

  return $\mathbf{X}^a$

end procedure

Computational cost: $O\left(nN^2 + mN^2 + mNp + mN + m\right)$

Cholesky Decomposition Implementation

procedure (CHOL-EnKF)$\left(\mathbf{X}, \mathbf{X}', \mathbf{HX}', \mathbf{D}, \mathbf{E}\right)$

$\mathbf{R} \leftarrow \frac{1}{N-1}\text{diag}\left(\mathbf{E}\mathbf{E}^T\right)$

$\mathbf{Q} \leftarrow (N-1)\,\mathbf{I} + \left(HX'\right)^T \mathbf{R}^{-1}\left(\mathbf{HX}'\right)\ \left(mN^2\right)$

$\mathbf{L}\mathbf{L}^T \leftarrow \textbf{CHOLESKYM}\left(\mathbf{Q}\right)\ \left(N^3\right)$

$\mathbf{Z} \leftarrow \left(\mathbf{HX}'\right)^T \mathbf{R}^{-1}\mathbf{D}\ \left(mN^2\right)$

$\mathbf{W} \leftarrow \mathbf{Q}^{-1}\mathbf{Z}\ \left(N^3\right)$

$\mathbf{M} \leftarrow \mathbf{R}^{-1}\left[\mathbf{I} - \left(\mathbf{HX}'\mathbf{W}\right)\right]\ \left(mN^2\right)$

$\mathbf{Z} \leftarrow \left(\mathbf{HX}'\right)^T \mathbf{M}\ \left(mN^2\right)$

$\mathbf{X}^a \leftarrow \mathbf{X} + \frac{1}{N-1}\mathbf{X}'\mathbf{Z}\ \left(nN^2\right)$

return $\mathbf{X}^a$

end procedure

Computational cost: $O\left(N^3 + nN^2 + mN^2\right)$

Sherman Morrison Implementation

procedure (MF-EnKF)$\left(\mathbf{X}, \mathbf{X}', \mathbf{HX}', \mathbf{D}, \mathbf{E}\right)$
    $\mathbf{R} \leftarrow \text{diag}\left(\mathbf{E}\mathbf{E}^T\right)$
    call **SM** $\left(\mathbf{R}, \mathbf{HX}', \mathbf{HX}', \mathbf{d}_1, \mathbf{z}_1\right)$ $\left(mN^2\right)$
    $\mathbf{w} \leftarrow \mathbb{X}' \left(\mathbf{HX}'\right)^T \mathbf{z}_1$ $\left(nN\right)$
    $\mathbf{x}_1^a \leftarrow \mathbf{x}_1 + \mathbf{w}$ $\left(n\right)$
    for do $i \leftarrow 2, \ldots N$ do
        call **SIMPLIFIED** $\left(\mathbf{R}, \mathbf{HX}', \mathbf{d}_i, \mathbf{z}_i\right)$ $\left(mN\right)$
        $\mathbf{w} \leftarrow \mathbf{X}' \left(\mathbf{HX}'\right)^T \mathbf{z}_i$ $\left(nN\right)$
        $\mathbf{x}_1^a \leftarrow \mathbf{x}_i + \mathbf{w}$ $\left(n\right)$
    end for
    return $\mathbf{X}^a$
end procedure

Computational cost: $O\left(mN^2 + nN + mN + n\right)$

Thanks!

References

- Godinez, H. C. and Moulton, J. D. (2012). An efficient matrix-free algorithm for the ensemble Kalman filter. *Computational Geosciences*, 16(3):565–575.
- Kalman, R. (1960). A New Ap proach to Linear Filtering and Prediction Problems. *Journal of Basic Engineering*, 21(2):125–147.
- Kalman, R. E. and Bucy, R. S. (1961). New Results in Linear Filtering and Prediction Theory. *Journal of Basic Engineering*, 83(1):95.
- Majda, A. J. and Tong, X. T. (2018). Performance of Ensemble Kalman Filters in Large Dimensions. *Communications on Pure and Applied Mathematics*, 71(5):892–937.
- Mandel, J. (2006). Efficient Implementation of the Ensemble Kalman Filter. (231):CCM Report 231.
- Segers, A. (2002). *Data assimilation in atmospheric chemistry models using Kalman filtering*. PhD thesis, Delft University, Netherlands.

Sherman-Morrison Solver

Sherman-Morrison solver as described in Evensen 1994 y Maponi 2007.

procedure (SM)$(\mathbf{A}_0, \mathbf{U}, \mathbf{V}, \mathbf{b}, \mathbf{x})$
    Solve $\mathbf{A}_0\mathbf{x}_0 \leftarrow \mathbf{b}$
    Solve $\mathbf{A}_0\mathbf{y}_{0,k} \leftarrow \mathbf{u}_k$ for $k \leftarrow 1, \ldots, N$
    for do $i \leftarrow 1 \ldots, N-1$
        $$\mathbf{x}_i \leftarrow \mathbf{x}_{i-1} - \frac{\mathbf{v}_i^T \mathbf{x}_{i-1}}{1+\mathbf{v}_i^T \mathbf{y}_{i-1,i}} \mathbf{y}_{i-1,i}$$
        for do $k \leftarrow i+1, \ldots, N$
            $$y_{i,k} \leftarrow y_{i-1,k} - \frac{\mathbf{v}_i^T \mathbf{y}_{i-1,k}}{1+\mathbf{v}_i^T \mathbf{y}_{i-1,i}} \mathbf{y}_{i-1,i}$$
        end for
    end for
    $$\mathbf{x}_N \leftarrow \mathbf{x}_{N-1} - \frac{\mathbf{v}_N^T \mathbf{x}_{N-1}}{1+\mathbf{v}_N^T \mathbf{y}_{N-1,N}} \mathbf{y}_{i-1,i}$$
    return $\mathbf{x}$
end procedure

Simplified Sherman-Morrison Solver Subsequent Right-hand Sides

procedure (SIMPLIFIED)($\mathbf{A}_0$,$\mathbf{V}$,$\mathbf{b}$,$\mathbf{x}$)
    Solve $\mathbf{A}_0\mathbf{x}_0 \leftarrow \mathbf{b}$
    for **do** $i \leftarrow 1, \ldots, N$
        $\mathbf{x}_i \leftarrow \mathbf{x}_{i-1} - \frac{\mathbf{v}_i^T \mathbf{x}_{i-1}}{1+\mathbf{v}_i^T \mathbf{y}_{i-1,i}} \mathbf{y}_{i-1,i}$
    end for
    return $\mathbf{x}$
end procedure