Data Assimilation and Climate
AIMS Rwanda, March, 2020

Lecture Notes

Kalman Filter

Kalman Filter

Apply the Bayesian method to a discrete time stochastic process model and data model, with the following assumptions:

- The prior distribution on the state is normal.
- The process model is linear.
- The conditional distribution on the data given the state is normal, with a mean that depends linearly on the state.

These assumptions result in a normally distributed posterior distribution for the state given the data. We can use matrix manipulations to find its mean and variance. This is the Kalman Filter.

Activities 1.

Let

$$\begin{aligned}\mathbf{X}(i) &= \mathbf{M}_i\mathbf{X}(i-1) + \boldsymbol{\delta}_i \quad i = 1, \ldots, N \\ \mathbf{Y}(i) &= \mathbf{H}_i\mathbf{X}(i) + \boldsymbol{\epsilon}_i, \quad i = 1, \ldots, N\end{aligned}$$

with $\mathbf{X}(0) \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, $\boldsymbol{\epsilon}_i \sim \mathcal{N}(\mathbf{0}, \mathbf{R}_i)$, $\boldsymbol{\delta}_i \sim \mathcal{N}(\mathbf{0}, \mathbf{Q}_i)$.

The Bayesian method applied to the discrete time process and data models, i.e. the Kalman Filter, gives:

$$\begin{aligned}\boldsymbol{\mu}_{i|i-1} &= \mathbf{M}_i\boldsymbol{\mu}_{i-1|i-1} \\ \boldsymbol{\Sigma}_{i|i-1} &= \mathbf{Q}_i + \mathbf{M}_i\boldsymbol{\Sigma}_{i-1|i-1}\mathbf{M}_i^T \\ \mathbf{K}_i &= \boldsymbol{\Sigma}_{i|i-1}\mathbf{H}_i^T\left(\mathbf{H}_i^T\boldsymbol{\Sigma}_{i|i-1}\mathbf{H}_i + \mathbf{R}_i\right)^{-1} \\ \boldsymbol{\mu}_{i|i} &= \boldsymbol{\mu}_{i|i-1} + \mathbf{K}_i\left(\mathbf{y}(i) - \mathbf{H}_i\boldsymbol{\mu}_{i|i-1}\right) \\ \boldsymbol{\Sigma}_{i|i} &= \left(\mathbf{I} - \mathbf{K}_i\mathbf{H}_i\right)\boldsymbol{\Sigma}_{i|i-1}\end{aligned}$$

for $i = 1, \ldots, N$ and with $\mathbf{K}_i$ the Kalman gain matrix.

We see in these formulas how as new data or measurements became available, we can easily update the previous optimal estimates without having to recompute everything.

Forecasting, Filtering and Reanalysis

- Forecasting

$$\begin{aligned}
\mathbf{X}(1) &\sim \mathcal{N}(\boldsymbol{\mu}_{1|0}, \boldsymbol{\Sigma}_{1|0}), \\
\mathbf{X}(2)|\mathbf{y}(1) &\sim \mathcal{N}(\boldsymbol{\mu}_{2|1}, \boldsymbol{\Sigma}_{2|1}) \\
&\vdots \\
\mathbf{X}(i)|\mathbf{y}(1{:}i-1) &\sim \mathcal{N}(\boldsymbol{\mu}_{i|i-1}, \boldsymbol{\Sigma}_{i|i-1}) \\
&\vdots \\
\mathbf{X}(N)|\mathbf{y}(1{:}N-1) &\sim \mathcal{N}(\boldsymbol{\mu}_{N|N-1}, \boldsymbol{\Sigma}_{N|N-1}).
\end{aligned}$$

- Filtering

$$\begin{aligned}
\mathbf{X}(1)|\mathbf{y}(1) &\sim \mathcal{N}(\boldsymbol{\mu}_{1|1}, \boldsymbol{\Sigma}_{1|1}),\\
\mathbf{X}(2)|\mathbf{y}(1{:}2) &\sim \mathcal{N}(\boldsymbol{\mu}_{2|2}, \boldsymbol{\Sigma}_{2|2})\\
&\vdots\\
\mathbf{X}(i)|\mathbf{y}(1{:}i) &\sim \mathcal{N}(\boldsymbol{\mu}_{i|i}, \boldsymbol{\Sigma}_{i|i})\\
&\vdots\\
\mathbf{X}(N)|\mathbf{y}(1{:}N) &\sim \mathcal{N}(\boldsymbol{\mu}_{N|N}, \boldsymbol{\Sigma}_{N|N}).
\end{aligned}$$

- Reanalysis

$$\begin{aligned}
\mathbf{X}(N)|\mathbf{y}(1{:}N) &\sim \mathcal{N}(\boldsymbol{\mu}_{N|N}, \boldsymbol{\Sigma}_{N|N}), \text{(same as filtering)} \\
\mathbf{X}(N-1)|\mathbf{y}(1{:}N) &\sim \mathcal{N}(\boldsymbol{\mu}_{N-1|N}, \boldsymbol{\Sigma}_{N-1|N}) \\
&\vdots \\
\mathbf{X}(i)|\mathbf{y}(1{:}N) &\sim \mathcal{N}(\boldsymbol{\mu}_{i|N}, \boldsymbol{\Sigma}_{i|N}) \\
&\vdots \\
\mathbf{X}(0)|\mathbf{y}(1{:}N) &\sim \mathcal{N}(\boldsymbol{\mu}_{0|N}, \boldsymbol{\Sigma}_{0|N}).
\end{aligned}$$

$$\begin{aligned}
\boldsymbol{\mu}_{i|N} &= \boldsymbol{\mu}_{i|i} + \mathbf{J}_i \left(\boldsymbol{\mu}_{i+1|N} - \boldsymbol{\mu}_{i+1|i}\right) \\
\boldsymbol{\Sigma}_{i|N} &= \boldsymbol{\Sigma}_{i|i} + \mathbf{J}_i \left(\boldsymbol{\Sigma}_{i+1|N} - \boldsymbol{\Sigma}_{i+1|i}\right) \mathbf{J}_i^T
\end{aligned}$$

where

$$\mathbf{J}_i = \boldsymbol{\Sigma}_{i|i}\mathbf{M}_i^T\boldsymbol{\Sigma}_{i+1|i}^{-1}$$

Extended Kalman Filter

Nonlinear process and data models

$$
\begin{aligned}
\mathbf{X}(i) &= \mathcal{M}_i(\mathbf{X}(i-1)) + \boldsymbol{\delta}_i \quad i = 1, \ldots, N \\
\mathbf{Y}(i) &= \mathcal{H}_i(\mathbf{X}(i)) + \boldsymbol{\epsilon}_i, \quad i = 1, \ldots, N
\end{aligned}
$$

$\mathbf{X}(0) \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, $\boldsymbol{\epsilon}_i \sim \mathcal{N}(\mathbf{0}, \mathbf{R}_i)$, $\boldsymbol{\delta}_i \sim \mathcal{N}(\mathbf{0}, \mathbf{Q}_i)$.

Approach: Linearize about the current mean and covariance. Use the linear Kalman Filter algorithm with $\mathbf{M}_i$ and $\mathbf{H}_i$ the tangent linear operator evaluated at the current time step i.e.

$$
(\mathbf{M}_i)_{jk} = \frac{\partial\left(\mathcal{M}_i\right)_j}{\partial x_k}, \quad (\mathbf{H}_i)_{jk} = \frac{\partial\left(\mathcal{H}_i\right)_j}{\partial x_k}
$$

Then

$$\begin{aligned}
\boldsymbol{\mu}_{i|i-1} &= \mathcal{M}_i(\boldsymbol{\mu}_{i-1|i-1}) \\
\boldsymbol{\Sigma}_{i|i-1} &= \mathbf{Q}_i + \mathbf{M}_i \boldsymbol{\Sigma}_{i-1|i-1} \mathbf{M}_i^T \\
\mathbf{K}_i &= \boldsymbol{\Sigma}_{i|i-1} \mathbf{H}_i^T \left(\mathbf{H}_i^T \boldsymbol{\Sigma}_{i|i-1} \mathbf{H}_i + \mathbf{R}_i \right)^{-1} \\
\boldsymbol{\mu}_{i|i} &= \boldsymbol{\mu}_{i|i-1} + \mathbf{K}_i \left(\mathbf{y}(i) - \mathcal{H}_i(\boldsymbol{\mu}_{i|i-1}) \right) \\
\boldsymbol{\Sigma}_{i|i} &= \left(\mathbf{I} - \mathbf{K}_i \mathbf{H}_i \right) \boldsymbol{\Sigma}_{i|i-1}
\end{aligned}$$

for $i = 1, \ldots, N$

Ensemble Kalman Filter

Nonlinear process and data models

$$\begin{aligned}\mathbf{X}(i) &= \mathcal{M}_i(\mathbf{X}(i-1)) + \boldsymbol{\delta}_i \quad i = 1, \ldots, N \\ \mathbf{Y}(i) &= \mathbf{H}_i\mathbf{X}(i) + \boldsymbol{\epsilon}_i, \quad i = 1, \ldots, N\end{aligned}$$

$\mathbf{X}(0) \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, $\boldsymbol{\epsilon}_i \sim \mathcal{N}(\mathbf{0}, \mathbf{R}_i)$, $\boldsymbol{\delta}_i \sim \mathcal{N}(\mathbf{0}, \mathbf{Q}_i)$.

- Assume we are given estimates $\hat{\mathbf{x}}_{i-1|i-1}$ and $\hat{\boldsymbol{\Sigma}}_{i-1|i-1}$ at $i-1$.
- Assume $\mathbf{X}(i-1)|\mathbf{y}(1:i-1) \sim \mathcal{N}(\hat{\mathbf{x}}_{i-1|i-1}, \hat{\boldsymbol{\Sigma}}_{i-1|i-1})$ and draw independent samples $\mathbf{x}^j_{i-1|i-1}$ for $j = 1, \ldots, m$.
- Create an ensemble of forecasts $\mathbf{x}^j_{i|i-1} = \mathcal{M}_i(\mathbf{x}^j_{i-1|i-1}) + \boldsymbol{\delta}_i$

- Forecast estimate

$$\hat{\mathbf{x}}_{i|i-1} = \frac{1}{m}\sum_{j=1}^{m} \mathbf{x}^j_{i|i-1},$$

- Find $\hat{\boldsymbol{\Sigma}}_{i|i-1} = \mathrm{Cov}(\hat{\mathbf{x}}_{i|i-1})$

- Compute gain in same manner $\hat{\mathbf{K}}_i = \hat{\boldsymbol{\Sigma}}_{i|i-1}\mathbf{H}_i^T \left(\mathbf{H}_i^T \hat{\boldsymbol{\Sigma}}_{i|i-1}\mathbf{H}_i + \mathbf{R}_i\right)^{-1}$

- Compute an ensemble of simulated filtering states

$$\mathbf{x}^j_{i|i} = \mathbf{x}^j_{i|i-1} + \hat{\mathbf{K}}_i \left(\mathbf{y}(i) + \boldsymbol{\epsilon}_j - \mathbf{H}_i(\mathbf{x}^j_{i|i-1})\right)$$

- Filter estimates $\hat{\mathbf{x}}_{i|i} = \frac{1}{m}\sum_{j=1}^{m} \mathbf{x}^j_{i|i}$, and $\hat{\boldsymbol{\Sigma}}_{i|i} = \mathrm{Cov}(\hat{\mathbf{x}}_{i|i})$.