Data Assimilation and Climate
AIMS Rwanda, April, 2020

Lecture Notes

Variational Data Assimilation

Variational assimilation is based on optimal control theory. The state is estimated by minimizing a cost function. Maximum Likelihood Estimation showed us that minimizing this cost function can be viewed as maximizing a probability density function. We will now consider **discrete** variational data assimilation.

Stationary case: 3DVar

Define the cost function

$$\mathcal{J}(\mathbf{x}) = \frac{1}{2}(\mathbf{x}-\mathbf{x}^b)^T\mathbf{B}^{-1}(\mathbf{x}-\mathbf{x}^b) + \frac{1}{2}(\mathbf{H}\mathbf{x}-\mathbf{y})^T\mathbf{R}^{-1}(\mathbf{H}\mathbf{x}-\mathbf{y}),$$

where $\mathbf{x}$ is the state in three dimensions, $\mathbf{x}^b$ is the background state, and $\mathbf{y}$ contains the observations. Assume that $\mathbf{x}, \mathbf{x}^b \in \mathbb{R}^n$ and $\mathbf{y} \in \mathbb{R}^m$.

Activity: Identify the dimensions of $\mathbf{B}$, $\mathbf{R}$ and $\mathbf{H}$.

If we consider

$$\begin{aligned}\mathbf{x} &= \mathbf{x}^b + \boldsymbol{\eta} \\ \mathbf{y} &= \mathbf{H}\mathbf{x} + \boldsymbol{\epsilon}\end{aligned}$$

Minimizing this cost function can be viewed as

- Minimizing the errors in the background state ($\boldsymbol{\eta}$) and data ($\boldsymbol{\epsilon}$) in a weighted least squares sense, with weights $\mathbf{B}$ and $\mathbf{R}$.
- Maximizing the pdf of the errors in the state and data, with $\boldsymbol{\eta} \sim \mathcal{N}(0, \mathbf{B})$ and $\boldsymbol{\epsilon} \sim \mathcal{N}(0, \mathbf{R})$

Activity: Show that the $\boldsymbol{\nabla}\mathcal{J}(\mathbf{x}) = \mathbf{B}^{-1}(\mathbf{x}-\mathbf{x}^b) - \mathbf{H}^T\mathbf{R}^{-1}(\mathbf{y}-\mathbf{H}\mathbf{x})$. *Hint:* $\boldsymbol{\nabla}\left((\mathbf{x}-\mathbf{x}^b)^T\mathbf{B}^{-1}(\mathbf{x}-\mathbf{x}^b)\right) = 2\mathbf{B}^{-1}(\mathbf{x}-\mathbf{x}^b)$.

Solution: Use the hint and take the gradient of the cost function term by term

Activity: Solve $\boldsymbol{\nabla}\mathcal{J}(\hat{\mathbf{x}}) = \mathbf{0}$ for $\hat{\mathbf{x}}$, and show that

$$\hat{\mathbf{x}} = \left(\mathbf{B}^{-1} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}\right)^{-1}\left(\mathbf{B}^{-1}\mathbf{x}^b + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{y}\right)$$

Solution:

$$\begin{aligned}
\mathbf{B}^{-1}\left(\hat{\mathbf{x}}-\mathbf{x}^b\right) &= \mathbf{H}^T\mathbf{R}^{-1}\left(\mathbf{y}-\mathbf{H}\hat{\mathbf{x}}\right) \\
\left(\mathbf{B}^{-1} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}\right)\hat{\mathbf{x}} &= \mathbf{B}^{-1}\mathbf{x}^b + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{y}
\end{aligned}$$

Re-writing $\hat{\mathbf{x}}$

$$\begin{aligned}
\hat{\mathbf{x}} &= \left(\mathbf{B}^{-1} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}\right)^{-1}\left(\left(\mathbf{B}^{-1} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}\right)\mathbf{x}^b - \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}\mathbf{x}^b + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{y}\right) \\
&= \mathbf{x}^b + \left(\mathbf{B}^{-1} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}\right)^{-1}\mathbf{H}^T\mathbf{R}^{-1}\left(\mathbf{y} - \mathbf{H}\mathbf{x}^b\right) \\
&= \mathbf{x}^b + \mathbf{K}\left(\mathbf{y} - \mathbf{H}\mathbf{x}^b\right)
\end{aligned}$$

Note that

$$\begin{aligned}
\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{H}^T &= \mathbf{H}^T\mathbf{R}^{-1}\left(\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R}\right) \\
&= \left(\mathbf{B}^{-1} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}\right)\mathbf{B}\mathbf{H}^T
\end{aligned}$$

so we have $\mathbf{K} = \left(\mathbf{B}^{-1} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}\right)^{-1}\mathbf{H}^T\mathbf{R}^{-1} = \mathbf{B}\mathbf{H}^T\left(\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R}\right)^{-1}$.

	Innovation	Gain
3DVar	$\mathbf{y} - \mathbf{H}\mathbf{x}^b$	$\mathbf{B}\mathbf{H}^T\left(\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R}\right)^{-1}$
Kalman Filter	$\mathbf{y}(i) - \mathbf{H}_i\boldsymbol{\mu}_{i\|i-1}$	$\boldsymbol{\Sigma}_{i\|i-1}\mathbf{H}_i^T\left(\mathbf{H}_i^T\boldsymbol{\Sigma}_{i\|i-1}\mathbf{H}_i + \mathbf{R}_i\right)^{-1}$

Non-stationary case: 4DVar

$$\mathbf{x}(0) = \mathbf{x}^b + \boldsymbol{\eta}$$
$$\mathbf{y}(i) = \mathbf{H}_i\mathbf{x}(i) + \boldsymbol{\epsilon}_i$$

Strong contraint 4DVar

$$\mathbf{x}(i) = \mathbf{M}_i\mathbf{x}(i-1)$$

with cost function

$$\mathcal{J}(\mathbf{x}(0)) = \frac{1}{2}(\mathbf{x}(0)-\mathbf{x}^b)^T\mathbf{B}^{-1}(\mathbf{x}(0)-\mathbf{x}^b)+\frac{1}{2}\sum_{i=1}^{N}(\mathbf{H}_i\mathbf{x}(i)-\mathbf{y}(i))^T\mathbf{R}_i^{-1}(\mathbf{H}_i\mathbf{x}(i)-\mathbf{y}(i)).$$

Note that the given initial condition, $\mathbf{x}(0)$, defines a unique state, $\mathbf{x}(i)$, so both terms in the cost function depend on $\mathbf{x}(0)$.

Activity: Use the process model $\mathbf{x}(i) = \mathbf{M}_i\mathbf{x}(i-1)$ to find a formula for the state $\mathbf{x}(i)$ as a function of the initial condition $\mathbf{x}(0)$, for any (i).

Solution: $\mathbf{x}(1) = \mathbf{M}_1\mathbf{x}(0), \mathbf{x}(2) = \mathbf{M}_2\mathbf{M}_1\mathbf{x}(0), \ldots \rightarrow \mathbf{x}(i) = \mathbf{M}_i \ldots \mathbf{M}_2\mathbf{M}_1\mathbf{x}(0)$.

Activity: Write the cost function explicitly as a function of $\mathbf{x}(0)$.

Solution:

$$\begin{aligned}\mathcal{J}(\mathbf{x}(0)) &= \frac{1}{2}(\mathbf{x}(0) - \mathbf{x}^b)^T\mathbf{B}^{-1}(\mathbf{x}(0) - \mathbf{x}^b) \\ &+ \frac{1}{2}\sum_{i=1}^{N}(\mathbf{H}_i\mathbf{M}_i \ldots \mathbf{M}_2\mathbf{M}_1\mathbf{x}(0) - \mathbf{y}(i))^T\mathbf{R}_i^{-1}(\mathbf{H}_i\mathbf{M}_i \ldots \mathbf{M}_2\mathbf{M}_1\mathbf{x}(0) - \mathbf{y}(i)).\end{aligned}$$

Weak contraint 4DVar

$$
\begin{aligned}
\mathbf{x}(0) &= \mathbf{x}^b + \boldsymbol{\eta} \\
\mathbf{x}(i) &= \mathbf{M}_i \mathbf{x}(i-1) + \boldsymbol{\delta}_i \\
\mathbf{y}(i) &= \mathbf{H}_i \mathbf{x}(i) + \boldsymbol{\epsilon}_i
\end{aligned}
$$

Define the cost function

$$
\begin{aligned}
\mathcal{J}(\mathbf{x}(0), \mathbf{x}(1), \ldots, \mathbf{x}(N)) &= \frac{1}{2}(\mathbf{x}(0) - \mathbf{x}^b)^T \mathbf{B}^{-1} (\mathbf{x}(0) - \mathbf{x}^b) \\
&+ \frac{1}{2} \sum_{i=1}^{N} (\mathbf{H}_i \mathbf{x}(i) - \mathbf{y}(i))^T \mathbf{R}_i^{-1} (\mathbf{H}_i \mathbf{x}(i) - \mathbf{y}(i)) \\
&+ \frac{1}{2} \sum_{i=1}^{N} (\mathbf{x}(i) - \mathbf{M}_i \mathbf{x}(i-1))^T \mathbf{Q}_i^{-1} (\mathbf{x}(i) - \mathbf{M}_i \mathbf{x}(i-1))
\end{aligned}
$$

State estimates from Variational Data Assimilation

- 3DVar

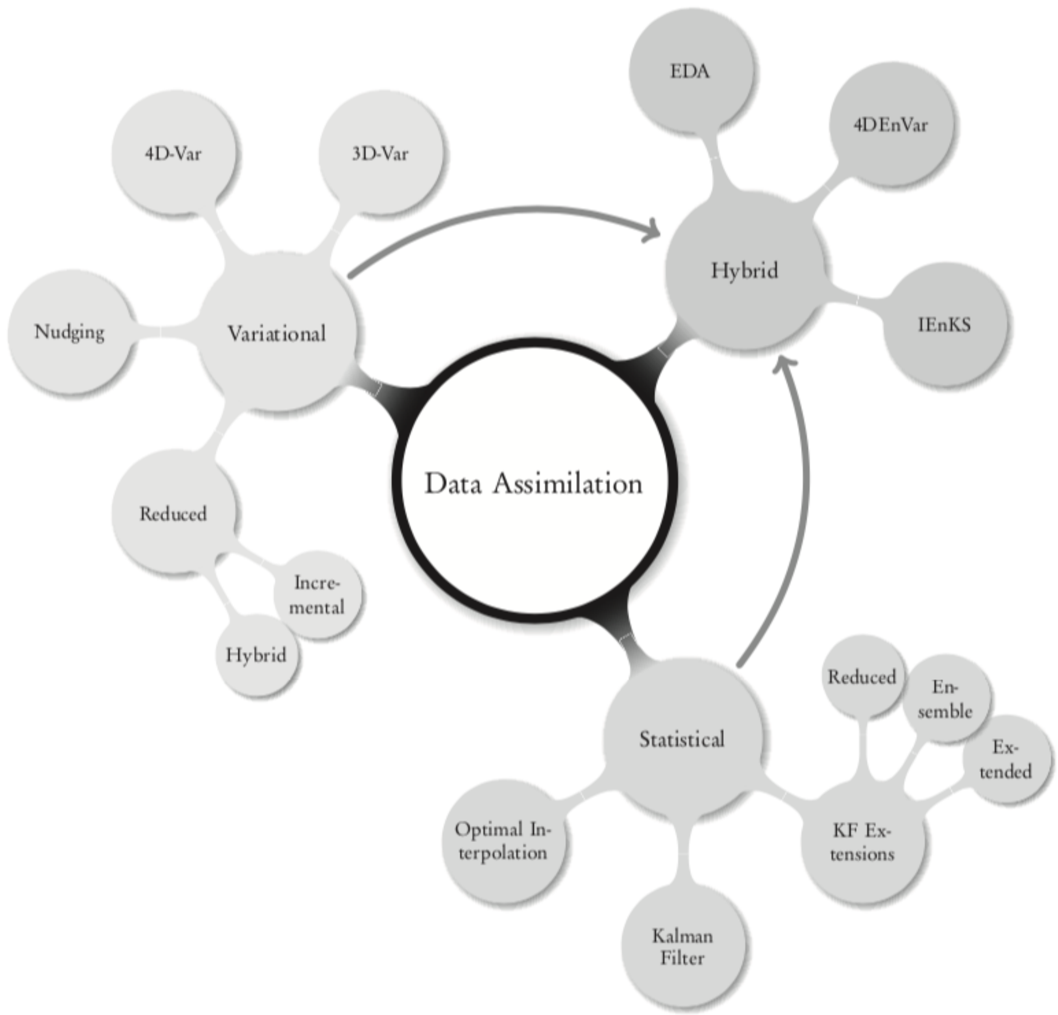
  Solving $\nabla_{\mathbf{x}}\mathcal{J}(\mathbf{x}) = \mathbf{0}$ gives background state estimate $\hat{\mathbf{x}}$ found with formula.

- Strong constraint 4DVar

  Solving $\nabla_{\mathbf{x}(0)}\mathcal{J}(\mathbf{x}(0)) = \mathbf{0}$ gives background state estimate $\hat{\mathbf{x}}(0)$ from which remaining states are estimated by $\hat{\mathbf{x}}(i) = \mathbf{M}_i\hat{\mathbf{x}}(i-1)$. The estimate $\hat{\mathbf{x}}(0)$ is typically found with *adjoint methods*.

- Weak constraint 4DVar

  $\nabla_{\mathbf{x}(0),\mathbf{x}(1),\ldots,\mathbf{x}(N)}\mathcal{J}\left(\mathbf{x}(0),\mathbf{x}(1),\ldots,\mathbf{x}(N)\right) = \mathbf{0}$ gives background state estimates $\hat{\mathbf{x}}(0),\hat{\mathbf{x}}(1),\ldots,\hat{\mathbf{x}}(N)$. These estimates are typically found with *adjoint methods*.



Courtesy of *Data assimilation: Methods, algorithms, and applications, fundamentals of algorithms.*, SIAM 2016