The Data Assimilation Problem

Elias D. Nino, Ph.D.

Universidad del Norte

enino@uninorte.edu.co

August 19, 2020

Outline

1. Outline
2. The Data Assimilation Problem
   - What is Data Assimilation?
   - Componentes in Data Assimilation
3. The Bayes Theorem

What is Data Assimilation? I

- Data Assimilation is the process by which forecasts of imperfect numerical models are adjusted according to real-noisy observations.
- Data assimilation is the process by which observations are incorporated into a computer model of a real system.
- In operational data assimilation the number of components in the model state ranges in $\mathcal{O}(10^8)$.
- What is the problem?
- We want to estimate:

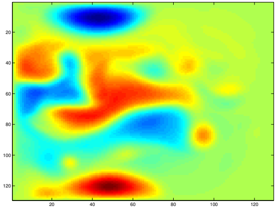
$$\mathbf{x}_k^* = \mathcal{D}_{t_{k-1} \rightarrow t_k} \left(\mathbf{x}_{k-1}^* \right), \text{ for } 1 \leq k \leq T,$$

  where $\mathbf{x}_k^* \in \mathbb{R}^{n \times 1}$.
  - $\mathcal{D}$ is unknown.
  - T number of discrete times.
- The components in data assimilations are,

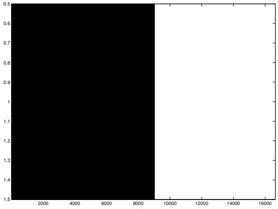
What is Data Assimilation? II

- An imperfect numerical model.
- An observation.
- An observational operator.

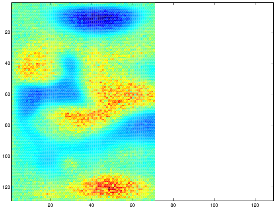
- Noisy observations:



(a) $\mathbf{x}_k^*$



(b) $\mathbf{H}_k \approx \mathcal{H}_k$



(c) $\mathbf{y}_k = \mathbf{H}_k \cdot \mathbf{x}_k^* + \boldsymbol{\theta}_k^o$

with

$$\mathbf{y}_k = \mathcal{H}_k\left(\mathbf{x}_k^*\right) + \boldsymbol{\theta}_k^o \in \mathbb{R}^{m\times 1},\ \text{for } 0 \le k \le T,$$

where $\mathcal{H}_k : \mathbb{R}^n \to \mathbb{R}^m$ and $\boldsymbol{\theta}_k^o \sim \mathcal{N}\left(\mathbf{0}_m,\, \mathbf{R}\right)$.$m \sim \mathcal{O}(10^7)$.

- Imperfect numerical model:

$$\mathbf{x}_k = \mathcal{M}_{t_{k-1}\to t_k}\left(\mathbf{x}_{k-1}\right),\ \text{for } 1 \le k \le T,$$

with $\mathbf{x}_k \in \mathbb{R}^{n\times 1}$.

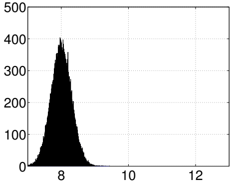
The Bayes Theorem I

- Prior distribution: *probability distribution before the evidence is accounted.*
- Likelihood: *which is the probability of the evidence given the parameters.*
- Posterior distribution: *best estimate after observations are assimilated.*

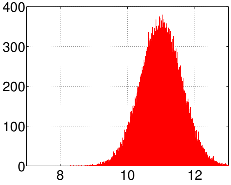
$$\begin{aligned} \mathcal{P}(x|y) &= \frac{\mathcal{P}^b(x) \cdot \mathcal{L}(x|y)}{\mathcal{P}(y)} \\ &\propto \mathcal{P}^b(x) \cdot \mathcal{L}(x|y) \end{aligned}$$

- Prior help us to understand our beliefs...

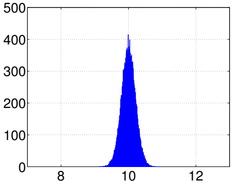
The Bayes Theorem II



(a) Prior 1



(b) Prior 2



(c) Prior 3

Figure: Some prior beliefs...

The posterior estimate I

- From Bayes' theorem we know that,

$$\mathcal{P}^a(x) = \mathcal{P}(x|y) \propto \mathcal{P}^b(x) \cdot \mathcal{L}(x|y)$$

- The prior is assumed to be Gaussian,

$$\begin{aligned}\mathcal{P}^b(x) &= \frac{1}{\sqrt{2 \cdot \pi}\sigma_b} \cdot \exp\left(-\frac{1}{2} \cdot \frac{(x - x^b)^2}{\sigma_b^2}\right) \\ &\propto \exp\left(-\frac{1}{2} \cdot \frac{(x - x^b)^2}{\sigma_b^2}\right)\end{aligned}$$

- The likelihood is assumed Gaussian as well,

$$\begin{aligned}\mathcal{L}(x|y) &= \frac{1}{\sqrt{2 \cdot \pi}\sigma_o} \cdot \exp\left(-\frac{1}{2} \cdot \frac{(y - x)^2}{\sigma_o^2}\right) \\ &\propto \exp\left(-\frac{1}{2} \cdot \frac{(y - x)^2}{\sigma_o^2}\right)\end{aligned}$$

The posterior estimate II

- Hence, the posterior reads,

$$\begin{aligned} \mathcal{P}^a(x) &\propto \exp\left(-\frac{1}{2}\cdot\frac{(x-x^b)^2}{\sigma_b^2}\right)\cdot\exp\left(-\frac{1}{2}\cdot\frac{(y-x)^2}{\sigma_o^2}\right) \\ &= \exp\left(-\frac{1}{2}\cdot\frac{(x-x^b)^2}{\sigma_b^2}-\frac{1}{2}\cdot\frac{(y-x)^2}{\sigma_o^2}\right) \\ &= \exp\left(-Q(x)\right)\,, \end{aligned}$$

where,

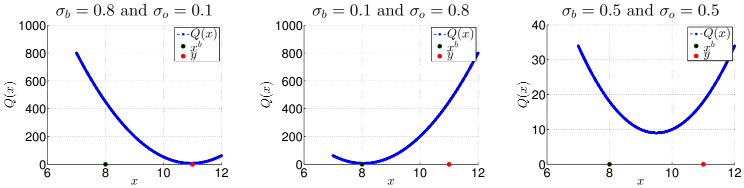
$$Q(x) = \frac{1}{2}\cdot\frac{(x-x^b)^2}{\sigma_b^2}+\frac{1}{2}\cdot\frac{(y-x)^2}{\sigma_o^2}$$

The posterior estimate III

- The Maximum A Posteriori (MAP) , this is, the value of x which maximizes the analysis probability $\mathcal{P}^a(x)$ is the optimal value of,

$$x^a = x^{optimal} = \arg\min_{x} Q(x)\,.$$

In the data assimilation context, $x^{optimal}$ is known as the *analysis state* x^a.



(a) (0.8, 0.1) (b) (0.1, 0.8) (c) (0.5, 0.5)

Figure: $Q(x)$ depending on (σ_b, σ_o).

The posterior estimate IV

- The derivative of $Q(x)$ reads,

$$Q'(x) = \frac{(x - x^b)}{\sigma_b^2} - \frac{(y - x)^2}{\sigma_o^2}\,.$$

Setting this derivative to zero, we obtain,

$$\begin{aligned} 0 &= Q'(x) = \frac{(x^a - x^b)}{\sigma_b^2} - \frac{(y - x^a)}{\sigma_o^2} \\ &= \frac{x^a}{\sigma_b^2} - \frac{x^b}{\sigma_b^2} - \frac{y}{\sigma_o^2} + \frac{x^a}{\sigma_o^2} = \left(\frac{1}{\sigma_b^2} + \frac{1}{\sigma_o^2}\right) \cdot x^a - \left(\frac{x^b}{\sigma_b^2} + \frac{y}{\sigma_o^2}\right)\,, \end{aligned}$$

and therefore, the MAP reads,

$$x^a = \left(\frac{1}{\sigma_b^2} + \frac{1}{\sigma_o^2}\right)^{-1} \cdot \left(\frac{x^b}{\sigma_b^2} + \frac{y}{\sigma_o^2}\right)\,.$$

The posterior estimate V

We can work a little bit on this estimate in order to find out something interesting,

$$
\begin{aligned}
x^a &= \left(\frac{1}{\sigma_b^2} + \frac{1}{\sigma_o^2}\right)^{-1} \cdot \left(\frac{x^b}{\sigma_b^2} + \frac{y}{\sigma_o^2}\right) \\
&= \left(\frac{\sigma_b^2 + \sigma_o^2}{\sigma_b^2 \cdot \sigma_o^2}\right)^{-1} \cdot \left(\frac{x^b}{\sigma_b^2} + \frac{y}{\sigma_o^2}\right) \\
&= \left(\frac{\sigma_b^2 \cdot \sigma_o^2}{\sigma_b^2 + \sigma_o^2}\right) \cdot \left(\frac{x^b}{\sigma_b^2} + \frac{y}{\sigma_o^2}\right) \\
&= \left(\frac{\sigma_o^2}{\sigma_b^2 + \sigma_o^2}\right) \cdot x^b + \left(\frac{\sigma_b^2}{\sigma_b^2 + \sigma_o^2}\right) \cdot y \\
&= \alpha \cdot x^b + (1 - \alpha) \cdot y \,,
\end{aligned}
$$

where $\alpha = \frac{\sigma_o^2}{\sigma_b^2 + \sigma_o^2} \in [0,\, 1]$.

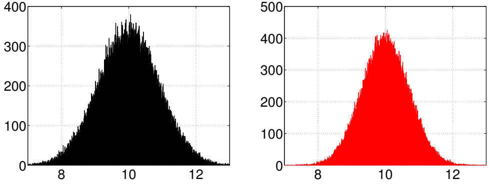
The posterior estimate VI

- Clearly x^a is the posterior mode with variance, $\frac{\sigma_b^2 \cdot \sigma_o^2}{\sigma_b^2 + \sigma_o^2}$, (homework) show that, the moments of the posterior distribution are,

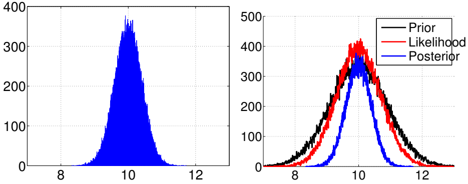
$$x|y \sim \mathcal{N}\left(x^a, \frac{\sigma_b^2 \cdot \sigma_o^2}{\sigma_b^2 + \sigma_o^2}\right)$$

- Let's check some cases,

The posterior estimate VII



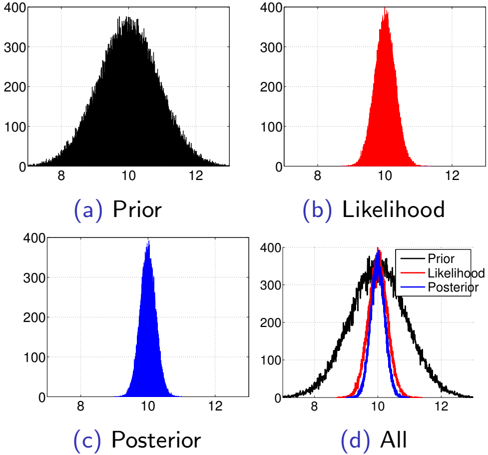
(a) Prior (b) Likelihood



(c) Posterior (d) All

Figure: $(\sigma_b, \sigma_o) = (0.9, 0.7)$.

The posterior estimate VIII



(a) Prior (b) Likelihood (c) Posterior (d) All

Figure: $(\sigma_b, \sigma_o) = (0.9, 0.3)$.

The posterior estimate IX

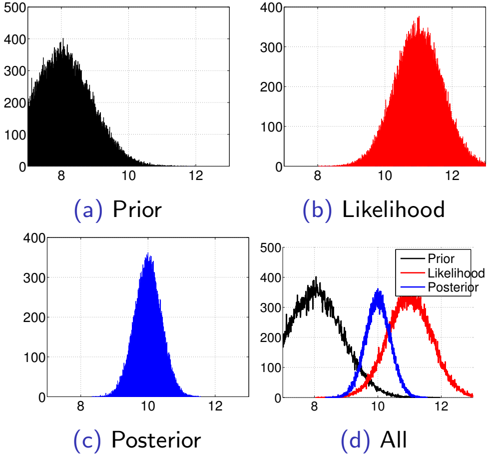
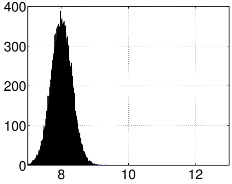
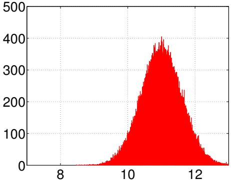


Figure: $(\sigma_b, \sigma_o) = (0.9, 0.7)$.

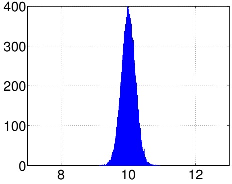
The posterior estimate X



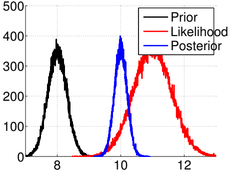
(a) Prior



(b) Likelihood



(c) Posterior



(d) All

Figure: $(\sigma_b, \sigma_o) = (0.9, 0.3)$.