Bayes' theorem and data assimilation

Ivan Kasanický[1,2], Jan Mandel[1,3]

[1]Institute of Computer Sciences, The Czech Academy of Sciences,
[2] Faculty of Mathematics and Physics, Charles University
[3]Department of Mathematical and Statistical Sciences, CU Denver

The research has been partially supported by
National Science Foundation grant DMS-1216481,
Czech Science Foundation grant GA13-34856S.
and grant SVV 2016 No. 260334.

ROBUST 2016

State space model

$\mathcal{H}$: separable Hilbert space.

- Underlying dynamical system:

$$X^{(0)} \sim \mathcal{N}\left(m^{(0)}, \mathsf{P}^{(0)}\right),\ X^{(0)} \in \mathcal{L}(\mathcal{H})$$
$$X^{(t)} = \Psi\left(X^{(t-1)}\right) + V^{(t)},\ V^{(t)} \sim \mathcal{N}\left(0, \mathsf{Q}^{(t)}\right)$$
$$\Psi : \mathcal{H} \to \mathcal{H} \text{ measurable}$$

- Noisy observations:

$$Y^{(t)} = \mathsf{H}X^{(t)} + W^{(t)},\ W^{(t)} \sim \mathcal{N}_w\left(0, \mathsf{R}^{(t)}\right)$$
$$\mathsf{H} : \mathcal{H} \to \mathcal{G} \text{ bunded, linear}$$

- In real world applications: $\mathcal{H} = \mathbb{R}^n$, $\mathcal{G} = \mathbb{R}^m$.
- Theory: $\mathcal{H} = \ell^2$

Data assimilation

- Sequential process.
- Using a observation $Y^{(t)} = y^{(t)}$ update a forecast (prior) estimate $X^{(t),f}$ to produce an analysis (posterior) estimate $X^{(t),a}$.
- Data assimilation $\approx$ filtration problem.
- $X^{(t),f} \approx X^{(t)} | Y^{(t-1)} = y^{(t-1)}, \ldots, Y^{(1)} = y^{(1)}$
- $X^{(t),a} \approx X^{(t)} | Y^{(t)} = y^{(t)}, \ldots, Y^{(1)} = y^{(1)}$

Simplifications

For $t \in \mathbb{N}$ fixed:

- $\mu^f \sim \mathcal{N}\left(m^f, \mathsf{P}^f\right)$ measure induced by $X^f = X^{(t),f}$,
- $X^a = X^{(t),a}$,
- $\mathsf{H} = \mathsf{I}$, so $Y = Y^{(t)} = X + W$.

Bayes' update

Analysis (posterior) distribution

$$\mu^a(B) = \frac{1}{c(y)} \int_B \exp\left(-\frac{1}{2}|y-x|^2_{\mathsf{R}^{-1}}\right) d\mu^f(x) \Leftrightarrow c(y) > 0,$$

where

$$\begin{aligned}
c(y) &= \int_{\mathcal{H}} \exp\left(-\frac{1}{2}|y-x|^2_{\mathsf{R}^{-1}}\right) d\mu^f(x), \\
|z|^2_{\mathsf{R}^{-1}} &= z^*\mathsf{R}^{-1}z \qquad\qquad \text{if } \dim(\mathcal{H}) < \infty, \\
|z|^2_{\mathsf{R}^{-1}} &= \begin{cases} \left\langle \mathsf{R}^{-1/2}z, \mathsf{R}^{-1/2}z \right\rangle & \text{if } z \in \mathsf{R}^{1/2}(\mathcal{H}), \\ \infty & \text{if } z \notin \mathsf{R}^{1/2}(\mathcal{H}), \end{cases} \\
\mathsf{R}^{1/2}(\mathcal{H}) &= \mathrm{range}(\mathsf{R}^{1/2}).
\end{aligned}$$

Problem

When

$$\mathsf{Trace}(\mathsf{R}) < \infty,$$

$\mathsf{R}^{1/2}(\mathcal{H})$ is Cameron-Martin space of a measure $\mu_{\mathsf{R}} \sim \mathcal{N}(0, \mathsf{R})$.

$$\dim(\mathcal{H}) < \infty \Rightarrow \mu_{\mathsf{R}}\left(\mathsf{R}^{1/2}(\mathcal{H})\right) = 1$$
$$\dim(\mathcal{H}) = \infty \Rightarrow \mu_{\mathsf{R}}\left(\mathsf{R}^{1/2}(\mathcal{H})\right) = 0$$

Therefore, if measures μ^f and μ_{R} are equivalent,

$$c\,(y) = \int_{\mathcal{H}} \exp\left(-\frac{1}{2}\,|y - x|^2_{\mathsf{R}^{-1}}\right) d\mu^f\,(x) \leq \mu^f\left(\mathsf{R}^{1/2}(\mathcal{H})\right) = 0$$

$$\forall y \in \mathcal{H}.$$

Theorem

If $\dim(\mathcal{H}) = \infty$, operators P^f and R commute, i.e., $\mathsf{P}^f\mathsf{R} - \mathsf{R}\mathsf{P}^f = 0$, then

$$c\,(y) = \int_{\mathcal{H}} \exp\left(-\frac{1}{2}\,|y-x|^2_{\mathsf{R}^{-1}}\right) d\mu^f\,(x) > 0 \quad \forall y \in \mathcal{H}$$

$$\Leftrightarrow \quad r = \inf_{i\in\mathbb{N}} \{r_i\} > 0$$

where r_i, $i \in \mathbb{N}$, are eigenvalues of R.

Additionally,
- $\mathrm{Trace}(\mathsf{R}) = \sum_{i=1}^{\infty} r_i < \infty \;\Rightarrow\; c(y) = 0\ \forall y \in \mathcal{H}$,
- $r = 0 \;\Rightarrow\; c(y) = 0\ \forall y \in A \subset \mathcal{H}$, and A is dense.

When $\mathrm{Trace}(\mathsf{R}) = \sum_{i=1}^{\infty} r_i = \infty$:
- $\mathcal{N}_w(0, \mathsf{R})$ is not a σ-additive,
- $\mathcal{N}_w(0, \mathsf{R})$ is defined only on cylindrical sets,
- data noise W is only a week random variable (white noise, etc.).

Conclusions

When a forecast (prior) distribution is Gaussian and an observation operator is linear:

- $\dim(\mathcal{H}) < \infty \Rightarrow$ no problem with the Bayes' update,
- $\dim(\mathcal{H}) = \infty$ and $\mathrm{Trace(R)} < \infty \Rightarrow$ the Bayes' formula is useless,
- $\dim(\mathcal{H}) = \infty$ and R bounded from below $\Rightarrow$ the Bayes' update is well defined,
- white noise is good.

Conclusions

When a forecast (prior) distribution is Gaussian and an observation operator is linear:

• $\dim(\mathcal{H}) < \infty \Rightarrow$ no problem with the Bayes' update,

• $\dim(\mathcal{H}) = \infty$ and Trace(R) $< \infty \Rightarrow$ the Bayes' formula is useless,

• $\dim(\mathcal{H}) = \infty$ and R bounded from below $\Rightarrow$ the Bayes' update is well defined,

• white noise is good.

Thank you!