Mathematics and numerics for data assimilation and state estimation – Lecture 9

Summer semester 2020

Overview

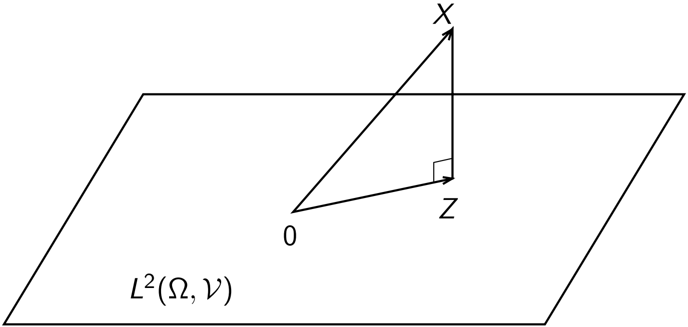
1. Metrics on spaces of probability density functions
2. Approximation result in $Y = G(u) + \eta$ setting
3. Bayesian inversion in different problem setting
4. Linear-Gaussian setting

Summary of lecture 8

Conditional expectations on projections:

For rv $X : \Omega \to \mathbb{R}^d$ and $Y : \Omega \to \mathbb{R}^k$ defined on the same probability space and with $X \in L^2(\Omega, \mathcal{F})$, it holds that

$$\mathbb{E}[X \mid Y] = \mathbb{E}[X \mid \sigma(Y)] = \text{Proj}_{L^2(\Omega,\sigma(Y))} X.$$



Bayesian inversion

Inverse problem

$$Y = G(U) + \eta \tag{1}$$

- observation Y is the observation
- forward model G
- observation noise η
- U is the unknown parameter

Problem assumptions: $\eta \sim \pi_\eta$, $U \sim \pi_U$ and $\eta \perp U$.

Solution:

$$\pi_{U|Y}(u|y) = \frac{\pi_\eta(y - G(u))\pi_U(u)}{\pi_Y(y)}.$$

with $\pi_Y(y)$ often replace by equivalent normalizing constant

$$Z = Z(y) = \int \pi_\eta(y - G(u))\pi_U(u)\,du.$$

Definition 1 (J. Hadamard 1902)

A problem is called **well-posed** if

1. a solution exists,
2. the solution is unique, and
3. the solution is stable with respect to small perturbations in the input.

Objective: For the inverse problem

$$Y = G(u) + \eta,$$

study settings under which condition [3] holds for perturbations in G:

$$\underbrace{|G_\delta - G|}_{(i)} = \mathcal{O}(\delta) \implies \underbrace{d(\pi^\delta(\cdot|y), \pi(\cdot|y))}_{(ii)} = \mathcal{O}(\delta)$$

Namely, give examples where (i) holds and relate this to (ii) for different metrics.

Overview

1. Metrics on spaces of probability density functions
2. Approximation result in $Y = G(u) + \eta$ setting
3. Bayesian inversion in different problem setting
4. Linear-Gaussian setting

Metrics on the space of pdfs

Let us introduce the space of probability density functions on $\mathbb{R}^d$

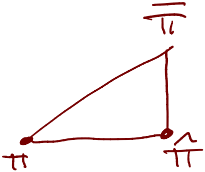
$$\mathcal{M} := \left\{ f \in L^1(\mathbb{R}^d) \mid f \geq 0 \text{ and } \int_{\mathbb{R}^d} f(u)\, du = 1 \right\}$$

and recall that

$$d : \mathcal{M} \times \mathcal{M} \to [0, \infty)$$

is a metric on $\mathcal{M}$ if for all $\pi, \bar{\pi}, \hat{\pi} \in \mathcal{M}$

1. $d(\pi, \bar{\pi}) = 0 \iff \pi \overset{L^1}{=} \bar{\pi}$,
2. $d(\pi, \bar{\pi}) = d(\bar{\pi}, \pi)$,
3. $d(\pi, \bar{\pi}) \leq d(\pi, \hat{\pi}) + d(\hat{\pi}, \bar{\pi})$.



Definition 2 (Total variation distance)

For any $\pi, \bar{\pi} \in \mathcal{M}$,

$$d_{TV}(\pi, \bar{\pi}) := \frac{1}{2} \int_{\mathbb{R}^d} |\pi(u) - \bar{\pi}(u)|\, du = \frac{1}{2} \|\pi - \bar{\pi}\|_{L^1(\mathbb{R}^d)}$$

Metrics on the space of pdfs

Definition 3 (Hellinger distance) $\pi \in L^1, \; \sqrt{\pi} \in L^2$

For any $\pi, \bar{\pi} \in \mathcal{M}$,

$$d_H(\pi, \bar{\pi}) := \frac{1}{\sqrt{2}} \|\sqrt{\pi} - \sqrt{\bar{\pi}}\|_{L^2(\mathbb{R}^d)} = \frac{1}{\sqrt{2}} \sqrt{\int_{\mathbb{R}^d} \left(\sqrt{\pi} - \sqrt{\bar{\pi}}\right)^2 dx}$$

Lemma 4 (SST Lem 1.8)

For any $\pi, \bar{\pi} \in \mathcal{M}$,

$$0 \leq d_H(\pi, \bar{\pi}) \leq 1 \quad \textit{and} \quad 0 \leq d_{TV}(\pi, \bar{\pi}) \leq 1.$$

Verification for d_{TV}:

$$d_{TV}(\pi, \bar{\pi}) = \frac{1}{2} \|\pi - \bar{\pi}\|_1 \leq \frac{1}{2}\left(\|\pi\|_1 + \|\bar{\pi}\|_1\right) = 1.$$

$$d_H(\pi, \bar{\pi}) = \frac{1}{\sqrt{2}} \sqrt{\int \left(\sqrt{\pi} - \sqrt{\bar{\pi}}\right)^2 dx} = \frac{1}{\sqrt{2}} \sqrt{\int \pi + \bar{\pi} - 2\sqrt{\pi \bar{\pi}}\, dx}$$

$$= \frac{1}{\sqrt{2}} \sqrt{\int \pi + \pi - 2\sqrt{\pi}\sqrt{\pi}\, dx}$$

$$\le \frac{1}{\sqrt{2}} \sqrt{\int \pi + \pi \, dx} \le \frac{\sqrt{2}}{\sqrt{2}} = 1$$

Properties TV and Hellinger distances

Lemma 5

For any $\pi, \bar{\pi} \in \mathcal{M}$,

$$\frac{1}{\sqrt{2}} d_{TV}(\pi, \bar{\pi}) \leq d_H(\pi, \bar{\pi}) \leq \sqrt{d_{TV}(\pi, \bar{\pi})}$$

$$d_{TV}(\pi, \bar{\pi}) = \frac{1}{2} \int_{\mathbb{R}^d} |\pi - \bar{\pi}| \, dx$$

$$= \frac{1}{2} \int_{\mathbb{R}^d} |\sqrt{\pi} - \sqrt{\bar{\pi}}| \, (\sqrt{\pi} + \sqrt{\bar{\pi}}) \, dx$$

$$\leq \frac{1}{2} \sqrt{\int (\sqrt{\pi} + \sqrt{\bar{\pi}})^2 dx} \sqrt{\int (\sqrt{\pi} - \sqrt{\bar{\pi}})^2 dx}$$

Weak errors

$|G_\delta - G| = O(\delta) \Rightarrow d(\pi^\delta, \pi) = O(\delta)$

$|u^\delta_{PM} - u_{PM}| = O(\delta)$

The posterior mean

$$u_{PM}[\pi(\cdot|y)] = \mathbb{E}^{\pi(\cdot|y)}[u] = \int_{\mathbb{R}^d} u\,\pi(u|y)\,du$$

is one possible solution to the inverse problem.

For a perturbation in the forward model $G_\delta = G + \mathcal{O}(\delta)$ that leads to a perturbed posterior density $\pi^\delta(u|y)$, we need to bound the following to verify stability

$$|u_{PM} - u^\delta_{PM}| = |\mathbb{E}^{\pi(\cdot|y)}[u] - \mathbb{E}^{\pi^\delta(\cdot|y)}[u]|$$

More generally, for a mapping $f : \mathbb{R}^d \to \mathbb{R}^k$, we may be interested in bounding

$$|\mathbb{E}^{\pi(\cdot|y)}[f] - \mathbb{E}^{\pi^\delta(\cdot|y)}[f]| = \left|\int f(u)\left(\pi(u|y) - \pi^\delta(u|y)\right)du\right|$$

$f(u) = u \Rightarrow \|f\|_\infty = \infty$

Lemma 6 (SST Lem 1.10)

Let $f : \mathbb{R}^d \to \mathbb{R}^k$ satisfy $\|f\|_{L^\infty(\mathbb{R}^d)} = \operatorname{ess\,sup}_{u \in \mathbb{R}^d} |f(u)| < \infty$. Then for any $\pi, \bar{\pi} \in \mathcal{M}$,

$$|\mathbb{E}^\pi[f] - \mathbb{E}^{\bar{\pi}}[f]| \leq 2\|f\|_\infty d_{TV}(\pi, \bar{\pi})$$

Verification:

$$|\mathbb{E}^\pi[f] - \mathbb{E}^{\bar{\pi}}[f]| = \left| \int_{\mathbb{R}^d} f(u)(\pi(u) - \bar{\pi}(u))\, du \right|$$

$$\leq \int_{\mathbb{R}^d} |f(u)|\, |\pi(u) - \bar{\pi}(u)|\, du$$

$$\leq \|f\|_\infty \int_{\mathbb{R}^d} |\pi(u) - \bar{\pi}(u)|\, du$$

$$= 2\|f\|_\infty d_{TV}(\pi, \bar{\pi}).$$

Lemma 7 (SST Lem 1.11)

Given $\pi, \bar{\pi} \in \mathcal{M}$, *assume that* $f : \mathbb{R}^d \to \mathbb{R}^k$ *satisfies*

$$f_2^2[\pi, \bar{\pi}] := \mathbb{E}^{\pi}[|f(u)|^2] + \mathbb{E}^{\bar{\pi}}[|f(u)|^2] < \infty.$$

Then

$$|\mathbb{E}^{\pi}[f] - \mathbb{E}^{\bar{\pi}}[f]| \leq 2 f_2 \, d_H(\pi, \bar{\pi}).$$

Proof:

$$|\mathbb{E}^{\pi}[f] - \mathbb{E}^{\bar{\pi}}[f]| = \left| \int_{\mathbb{R}^d} f(u)(\pi(u) - \bar{\pi}(u)) \, du \right|$$

$$= \int_{\mathbb{R}^d} |f(u)| \left(\sqrt{\pi(u)} + \sqrt{\bar{\pi}(u)}\right) \left|\sqrt{\pi(u)} - \sqrt{\bar{\pi}(u)}\right| du$$

$$\leq \sqrt{\int_{\mathbb{R}^d} |f(u)|^2 \left(\sqrt{\pi} + \sqrt{\bar{\pi}}\right)^2 du} \; \sqrt{\int_{\mathbb{R}} \left(\sqrt{\pi} - \sqrt{\bar{\pi}}\right)^2 du}$$

$$\leq \dots$$

Application of Lemma 1~~8~~7 to perturbed posterior means.

$$\begin{aligned} |u_{PM}[\pi(\cdot|y)] - u_{PM}[\pi^\delta(\cdot|y)]| &= |\mathbb{E}^{\pi(\cdot|y)}[u] - \mathbb{E}^{\pi^\delta(\cdot|y)}[u]| \\ &\le 2f_2\, d_H(\pi(\cdot|y), \pi^\delta(\cdot|y)). \end{aligned}$$

where $f(u) = u$ for the posterior mean, and thus

$$f_2^2 = \int_{\mathbb{R}^d} |u|^2 (\pi(u|y) + \pi^\delta(u|y))\, du.$$

Example 8 (Extension of MAP estimator example, Lecture 8)

Consider the problem (1) with $\eta \sim N(0, \gamma^2)$, $U \sim U[\cancel{0}\,{-1}, 1]$, $G(u) = u$ and $G_\delta(u) = u + \delta$ for some fixed $gamma > 0$ and $\delta > 0$.

Solutions:

$$\pi(u|y) = \frac{\overbrace{\pi(Y|u)}^{\pi_\eta(Y - G(u))}\,\pi_U(u)}{\pi_\Sigma(Y)} = \frac{e^{-(y-u)^2/2\gamma^2}\mathbb{1}_{(-1,1)}(u)}{2Z(y)}$$

$Y = G(U) + \eta$

$\Sigma = Y - \delta$

and

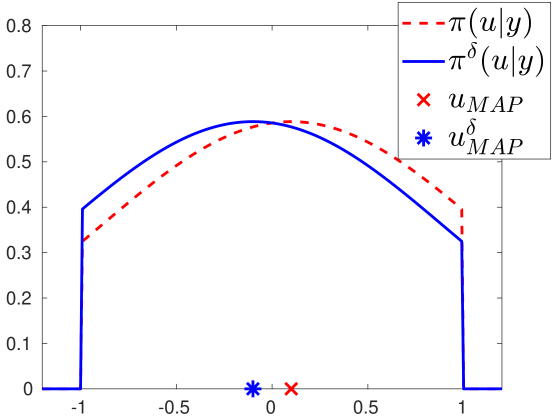
$$\pi^\delta(u|y) = \frac{e^{-(y-(u+\delta))^2/2\gamma^2}\mathbb{1}_{(-1,1)}(u)}{2Z(y-\delta)} = \pi(u|y-\delta)$$

Recalling that

$$u_{MAP}[\pi(\cdot|y)] = \arg\max_{u\in\mathbb{R}} \pi(u|y) = \begin{cases} y & \text{if } y \in (-1,1) \\ -1 & \text{if } y \leq -1 \\ 1 & \text{if } y \geq 1 \end{cases}$$

implies that $\quad |u_{MAP}[\pi(\cdot|y)] - u_{MAP}[\pi^\delta(\cdot|y)]| \leq \delta.$

Distance between u_{MAP} and u^{δ}_{MAP} when $\gamma = 1$, $y = 0.1$ and $\delta = 0.2$.



Exercise

Prove that also

$$|u_{PM}[\pi(\cdot|y)] - u_{PM}[\pi^{\delta}(\cdot|y)]| = \mathcal{O}(\delta).$$

Overview

1. Metrics on spaces of probability density functions
2. Approximation result in $Y = G(u) + \eta$ setting
3. Bayesian inversion in different problem setting
4. Linear-Gaussian setting

Approximation assumptions

By introducing the notation

$$g(u) := \pi_{\eta}(y - G(u)) \quad \text{and} \quad g_\delta(u) := \pi_{\eta}(y - G_\delta(u)),$$

we have

$$\pi(u|y) = \frac{g(u)\pi_U(u)}{Z} \quad \text{and} \quad \pi^\delta(u|y) = \frac{g_\delta(u)\pi_U(u)}{Z^\delta}.$$

Assumption 1

Assume there exists constant $K_1, K_2 > 0$ such that for sufficiently small $\delta > 0$,

(i) $\cancel{\|\sqrt{g} - \sqrt{g_\delta}\|_{L^2(\mathbb{R}^d)} \leq K_1\delta}$ $\sqrt{\left(E^{\pi_U}\left[|\sqrt{g} - \sqrt{g_\delta}|^2\right]\right)} \leq K_1 \delta$

(ii) $\|\sqrt{g}\|_{L^\infty(\mathbb{R}^d)} + \|\sqrt{g_\delta}\|_{L^\infty(\mathbb{R}^d)} \leq K_2$

Approximation results

Theorem 9

If Assumption 1 holds, then there exists $c_1, c_2, c_3 > 0$ *such that for sufficiently small* $\delta > 0$

$$|Z - Z^\delta| \leq c_1\delta \quad \textit{and} \quad Z, Z^\delta > c_2 \quad \textit{[SST Lemma 1.15]}$$

and

$$d_H(\pi(\cdot|y), \pi^\delta(\cdot|y)) \leq c_3\delta \quad \textit{[SST Theorem 1.14]}$$

where we recall that

$$d_H(\pi, \bar{\pi}) = \frac{1}{\sqrt{2}}\|\sqrt{\pi} - \sqrt{\bar{\pi}}\|_{L^2}.$$

Proof idea Lemma 1.15

$$|Z - Z^\delta| = \left| \int (g(u) - g_\delta(u))\pi_U(u)du \right|$$

$$\le \int |\sqrt{g} + \sqrt{g_\delta}|\,|\sqrt{g} - \sqrt{g_\delta}|\,\pi_U(u)\,du$$

$$\le \sqrt{\int |\sqrt{g} + \sqrt{g_\delta}|^2 \pi_U(u)du}\sqrt{\int |\sqrt{g} - \sqrt{g_\delta}|^2 \pi_U(u)\,du}$$

$$\lesssim K_2 \sqrt{\left(E^{\pi_U}\left[|\sqrt{g} - \sqrt{g_\delta}|^2\right]\right)} \le K_2 K_1 \delta$$

Positivity: $Z = \pi_Y(y) > 0$ by assumption, so by . . .

$$Z_\delta \geq Z_\delta - Z + Z \to Z > 0 \text{ as } \delta \downarrow 0$$

$$\Rightarrow Z_\delta > 0 \quad \text{for } \delta \text{ sufficiently small}$$

Proof idea Thm 1.14

$$\pm\sqrt{\frac{g_\delta \pi_U}{2}}$$

$$
\begin{aligned}
d_H(\pi(\cdot|y), \pi^\delta(\cdot|y)) &= \frac{1}{\sqrt{2}}\|\sqrt{\pi} - \sqrt{\pi^\delta}\|_2 \\
&= \frac{1}{\sqrt{2}}\left\|\sqrt{\frac{g\pi_U}{Z}} - \sqrt{\frac{g_\delta \pi_U}{Z^\delta}}\right\|_2 \\
&\leq \frac{1}{\sqrt{2}}\left\|\sqrt{\frac{g\pi_U}{Z}} - \sqrt{\frac{g_\delta \pi_U}{Z}}\right\|_2 + \frac{1}{\sqrt{2}}\left\|\sqrt{\frac{g_\delta \pi_U}{Z}} - \sqrt{\frac{g_\delta \pi_U}{Z^\delta}}\right\|_2 \\
&\leq \frac{1}{\sqrt{2}\,2}\|(\sqrt{g} - \sqrt{g_\delta})\sqrt{\pi_U}\|_2 \\
&\quad + \frac{1}{\sqrt{2}}\|\sqrt{g_\delta}\sqrt{\pi_U}\|_2 \left|\frac{1}{2} - \frac{1}{2^\delta}\right| \\
&\leq C\delta
\end{aligned}
$$

Summary of well-posedness result

Recall that

$$g(u) := \pi_{\eta}(y - G(u)) \quad \text{and} \quad g_\delta(u) := \pi_{\eta}(y - G_\delta(u)),$$

which yields

$$\pi(u|y) = \frac{g(u)\pi_U(u)}{Z} \quad \text{and} \quad \pi^\delta(u|y) = \frac{g_\delta(u)\pi_U(u)}{Z^\delta}.$$

Summary results: If for sufficiently small $\delta > 0$

(i) $\cancel{\|\sqrt{g} - \sqrt{g_\delta}\|_{L^2(\mathbb{R}^d)} = \mathcal{O}(\delta)}$ $\sqrt{\mathbb{E}^{\pi_U}\left[|\sqrt{g} - \sqrt{g_\delta}|^2\right]} = \mathcal{O}(\delta)$

(ii) $\|\sqrt{g}\|_{L^\infty(\mathbb{R}^d)} + \|\sqrt{g_\delta}\|_{L^\infty(\mathbb{R}^d)} < \infty$

Then the well-posedness condition [3] holds in the following sense:

$$d_H(\pi(\cdot|y), \pi^\delta(\cdot|y)) = \mathcal{O}(\delta).$$

Example with unspecified model where (i) and (ii) hold

Consider setting where $\|G_\delta - G\|_\infty = \mathcal{O}(\delta)$

$$\|G\|_\infty + \|G_\delta\|_\infty < \infty \quad \text{and} \quad \eta \sim N(0,1).$$

Then

$$\begin{aligned}
\sqrt{g(u)} - \sqrt{g_\delta(u)} &= \sqrt{\pi_\eta(y - G(u))} - \sqrt{\pi_\eta(y - G_\delta(u))} \\
&= \frac{1}{(2\pi)^{1/4}} \left(\exp(\frac{-(y - G(u))^2}{4}) - \exp(\frac{-(y - G_\delta(u))^2}{4}) \right) \\
&\le C \left| (y - G(u))^2 - (y - G_\delta(u))^2 \right| \\
&\le C \left| \big((y - G(u)) + y - G_\delta(u) \big) \big(G(u) - G_\delta(u) \big) \right| \\
&\le C(y, \|G\|_\infty, \|G_\delta\|_\infty)\, \delta \\
&= \mathcal{O}(\delta).
\end{aligned}$$

$$\text{and} \quad \|\sqrt{g}\|_\infty = \|\sqrt{g_\delta}\|_\infty = \frac{1}{(2\pi)^{1/4}}.$$

And a specified model which may lead to stability

Consider the ordinary differential equation

$$\dot{x}(t;u) = \overbrace{x(t;u)}^{f(x)} \quad t > 0 \quad \text{and} \quad x(0;u) = u \quad \text{for } u \in [-1,1],$$

and the associated explicit-Euler numerical solution

$$X_{n+1}^{\delta} = X_n^{\delta}(1+\delta), \quad X_0^{\delta} = u.$$

The forward model is the solution flow map from $t = 0$ to $t = 1$:

$$G(u) = x(1;u) = ue^1 \quad \text{and} \quad G_\delta(u) = X_{\lfloor \delta^{-1} \rfloor}^{\delta}(1 + (1 - \delta\lfloor \delta^{-1} \rfloor)).$$

For simplicity, we assume that $\delta^{-1} = N \in \mathbb{N}$. Then $G_\delta(u) = X_N^{\delta}$.

For $t_k = k\delta$, and note that

$$X(t_{k+1}) = e^{\delta} X(t_k).$$

$X(t) = e^t u$ $\quad X(t_k+\delta) = e^{\delta} X(t_k)$

For $E_k := |X(t_k) - X_k^{\delta}|$ it then holds that

$$E_{k+1} \leq (e^{\delta} - (1+\delta))|X(t_k)| + (1+\delta)E_k$$

Verification:

$$E_{k+1} = |X(t_{k+1}) - X_{k+1}^{\delta}| = |e^{\delta} X(t_k) \underbrace{\pm (1+\delta)X(t_k)} - (1+\delta) X_k^{\delta}|$$

$$\leq (e^{\delta} - (1+\delta))|X(t_k)| + (1+\delta)|X(t_k) - X_k^{\delta}|$$

Consequently,

$$E_N = |G(u) - G_{\delta}(u)| \leq \underbrace{(e^{\delta} - (1+\delta))}_{\leq c\delta^2} |X(t_{N-1})| + (1+\delta)E_{N-1}$$

$$\leq c\delta^2 |X(t_{N-1})| + (1+\delta)\left(c\delta^2 |X(t_{N-2})| + (1+\delta)E_{N-2}\right)$$

$$\leq c\delta^2 \sum_{k=0}^{N-1} \overbrace{(1+\delta)^{N-1-k}|X(t_k)|}^{\leq |X(T)|} + (1+\delta)^N \underbrace{E_0}_{=0} \leq c\delta e^1 |u| \leq \hat{c}\delta.$$

For the **relevant** $u \in [-1, 1]$, we have shown that

$$\|G - G_\delta\|_{L^\infty([-1,1])} \leq c\delta,$$

where $c > 0$ satisfies

$$|e^\delta - (1 + \delta)| \leq c\delta^2 \quad \forall \delta \in (0, \delta^+) \tag{2}$$

Note also that

$$\|G\|_{L^\infty([-1,1])} + \|G_\delta\|_{L^\infty([-1,1])} \leq e^1 + (1 + \delta)^{1/\delta} \leq 2e^1.$$

Exercise: For any $\delta \in (0, \delta^+ = 1)$, show that $c = e^1/2$ satisfies (2).

Comments:

- Relevant u values not being the whole of $\mathbb{R}^d$ may be motivated for instance by π_U having compact support.
- See also [SST 1.1.3] for a more general example of forward models stable under perturbations.

Overview