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

UQ Mathematics for Uncertainty Quantification | RWTH AACHEN UNIVERSITY

Summer semester 2020

Overview

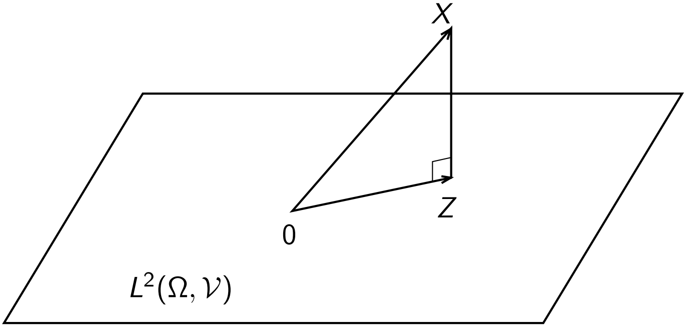
1. Metrics on spaces of probability density functions
2. Approximation result in $Y = G(u) + \eta$ setting
3. Bayesian inversion in different problem setting
4. Linear-Gaussian setting

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

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



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. $d(\pi, \bar{\pi}) = d(\pi, \hat{\pi}) + d(\hat{\pi}, \bar{\pi})$.

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)

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

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

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}:

$$d_{TV}(\pi, \bar{\pi}) =$$

Properties TV and Hellinger distances

Lemma 5

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

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

Weak errors

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

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

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

$$|\mathbb{E}^{\pi(\cdot|y)}[f] - \mathbb{E}^{\pi^\delta(\cdot|y)}[f]|$$

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* $\pi, \bar\pi \in \mathcal{M}$,

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

Verification:

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

$$=$$

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

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

Proof:

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

$$=$$

Application of Lemma 18 to perturbed posterior means.

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

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 $\eta \sim N(0, \gamma^2)$, $U \sim U[0,1]$, $G(u) = u$ and $G_\delta(u) = u + \delta$ for some fixed *gamma* > 0 and $\delta > 0$.
Solutions:

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

and

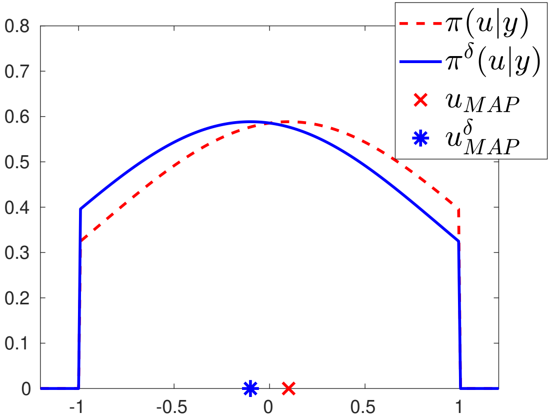
$$\pi^\delta(u|y) = \frac{e^{-(y-(u+\delta))^2/2\gamma^2}\mathbb{1}_{(-1,1)}(u)}{2Z(y-\delta)} = \pi(u|y-\delta)$$

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

Approximation assumptions

By introducing the notation

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

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) $\|\sqrt{g} - \sqrt{g_\delta}\|_{L^2(\mathbb{R}^d)} \leq K_1 \delta$

(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

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

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

Proof idea Thm 1.14

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

Summary of well-posedness result

Recall that

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

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) $\|\sqrt{g} - \sqrt{g_\delta}\|_{L^2(\mathbb{R}^d)} = \mathcal{O}(\delta)$

(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

$$\begin{aligned}
\sqrt{g(u)} - \sqrt{g_\delta(u)} &= \sqrt{\pi_\eta(y - G(u))} - \sqrt{\pi_\eta(y - G_\delta(u))} \\
&= \frac{1}{(2\pi)^{1/4}} \left(\exp(\frac{-(y - G(u))^2}{4}) - \exp(\frac{-(y - G_\delta(u))^2}{4}) \right) \\
&\leq \\
&= \mathcal{O}(\delta).
\end{aligned}$$

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

And a specified model which may lead to stability

Consider the ordinary differential equation

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

and the associated explicit-Euler numerical solution

$$X^{\delta}_{n+1} = X^{\delta}_{n}(1+\delta), \quad X^{\delta}_{0} = u.$$

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

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

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

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

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

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

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

Verification:

Consequently,

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

$$\le$$

$$\le c\delta^2 \sum_{k=0}^{N-1} (1+\delta)^{N-1-k} |X(t_k)| + (1+\delta)^N E_0 \le c\delta e^1 |u| \le c\delta.$$

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

Inverse problem with random model and exact observations

Let us consider a different type of inverse problem

$$Y = G(U)$$

with prior $U \sim U(0,1)$ and, for any $u \in (0,1)$, $G(u) \sim \mathit{Bernoulli}(u)$.

In other words U is a continuous rv, while $Y|(U = u) \sim \mathit{Bernoulli}(u)$ is discrete.

Given a measurement $Y = y$, we may formally proceed as before

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

Problem: $Y|(U = u)$ is a discrete rv!

Alternative measures-based approach:

By the properties, for $y \in \{0, 1\}$,

$$\mathbb{P}(Y = y, U \in du) = \mathbb{P}(Y = y | U \in du)\mathbb{P}(U \in du)$$
$$\mathbb{P}(Y = y, U \in du) = \mathbb{P}(U \in du | Y = y)\mathbb{P}(Y = y)$$

we derive by Bayes' rule the posterior measure

$$\mathbb{P}(U \in du | Y = y) = \frac{\mathbb{P}(Y = y | U \in du)\mathbb{P}(U \in du)}{\mathbb{P}(Y = y)}$$

By $Y = y \mid U = u$, it follows that

$$\mathbb{P}(Y = y \mid U \in du) = (1 - u)^{1-y} u^y$$

and thus

$$\mathbb{P}(U \in du | Y = y) = \frac{(1 - u)^y u^y du}{Z}.$$

With density form

$$\pi_{U|Y}(u|y) = \qquad = \frac{(1 - u)^{1-y} u^y}{Z}. \tag{3}$$

Is the coin fair?

Consider an inverse problem with a sequence of **exact** observations of coin tosses

$$Y_k = G_k(U), \quad \text{for} \quad k = 1, 2, \ldots$$

with $G_k(U)|U = u \sim \mathit{Bernoulli}(u)$, where for any fixed $\tilde{u} \in (0,1)$ $(G_1(\tilde{u}), G_2(\tilde{u}), \ldots)$ is an iid sequence. Hence

$$(Y_1, Y_2, \ldots)|(U = u) = (G_1(u), G_2(u), \ldots)$$

is a (conditionally $U = u$) iid sequence.
Input: Coin-bias prior $U \in U(0,1)$ and flipping coin results $Y = (Y_1, \ldots, Y_n) = (y_1, \ldots, y_n)$.

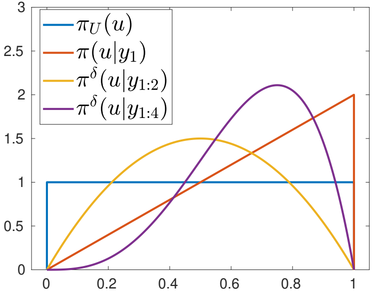
Direct extension of (3) yields

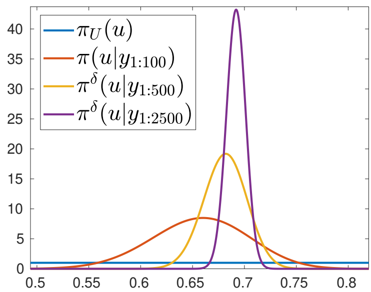
$$\pi_{U|Y}(u|y) = \frac{\prod_{k=1}^{n}(1-u)^{1-y_k}u^{y_k}}{Z} = \frac{(1-u)^{n-\bar{y}_n}u^{\bar{y}_n}}{Z}$$

where $\bar{y}_n = \sum_{k=1}^{n} y_k$.

Computational result given

$$y = (1, 0, 1, 1, \ldots) \quad \text{with} \quad \bar{y}_{100} = 66,\ \bar{y}_{500} = 341,\ \bar{y}_{2500} = 1730$$





Numerical integration gives

$$\mathbb{P}(|U - 0.7| < 0.05 | Y_{1:500} = Y_{1:500}) = 0.9320$$

See ["Data analysis" by D.S. Sivia section 2.1] for more on this example.

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

Linear-Gaussian setting

We consider the inverse problem

$$Y = G(U) + \eta \tag{4}$$

with

Assumption 2

- linear forward model $G(u) = Au$ where $A \in \mathbb{R}^{k \times d}$
- and $\eta \sim N(0, \Gamma)$, $U \sim N(\hat{m}, \hat{C})$ where both Γ and $\hat{C}$ are positive definite and $\eta \perp U$.

Given the observation $Y = y$, Bayesian inversion yields

$$\pi(u|y) = \qquad = \frac{\pi_\eta(y - Au)\pi_U(u)}{Z}$$

where we have used that

$$Y|(U = u) = G(u) + \eta \sim N(G(u), \Gamma).$$

Recall that for $X \sim N(\mu, \Sigma)$,

$$\pi_X(x) = \frac{\exp\Big(-\frac{1}{2}|x-\mu|_\Sigma^2\Big)}{Z}$$

with

$$|x-\mu|_\Sigma := |\Sigma^{-1/2}(x-\mu)|.$$

So we may write (for a different normalizing constant Z),

$$\begin{aligned}\pi(u|y) &= \frac{\pi_\eta(y-Au)\pi_U(u)}{Z}\\ &= \frac{\exp\Big(-\frac{1}{2}|y-Au|_\Gamma^2 - \frac{1}{2}|u-\hat{m}|_{\hat{C}}^2\Big)}{Z}\\ &= \frac{\exp(-\mathsf{J}(u))}{Z}\end{aligned}$$

with

$$\mathsf{J}(u) := \frac{|y-Au|_\Gamma^2 + \frac{1}{2}|u-\hat{m}|_{\hat{C}}^2}{2}.$$

Objective: Verify that $U|Y=y$ is Gaussian, and find its density.

On the one hand:

$$\pi(u|y) = \frac{\exp(-\mathrm{J}(u))}{Z}$$

on the other, let us make the ansatz that for some $m \in \mathbb{R}^d$ and pos. def. C,

$$\pi(u|y) = \frac{\exp\left(-\frac{1}{2}|u-m|_C^2\right)}{Z}$$

For this to hold, we must find m and C s.t.,

$$|u-m|_C^2 = 2\mathrm{J}(u).$$

Written in terms of polynomial parts

$$|u-m|_C^2 = (u-m)^T C^{-1}(u-m) = u^T {\color{red}C^{-1}} u - 2u^T {\color{blue}C^{-1} m} + q$$

and

$$\begin{aligned} 2\mathrm{J}(u) &= |y-Au|_\Gamma^2 + |u-\hat{m}|_{\hat{C}}^2 \\ &= (y-Au)^T\Gamma^{-1}(y-Au) + (u-\hat{m})^T\hat{C}^{-1}(u-\hat{m}) \\ &= u^T {\color{red}(A^T\Gamma^{-1}A + \hat{C}^{-1})} u - 2u^T {\color{blue}(A^T\Gamma^{-1}y + \hat{C}^{-1}\hat{m})} + \hat{q} \end{aligned}$$

Enforcing equality for same-order-term coefficients yields

$$u^T C^{-1} u = u^T (A^T \Gamma^{-1} A + \hat{C}^{-1}) u \quad \forall u \in \mathbb{R}^d \implies C = (A^T \Gamma^{-1} A + \hat{C}^{-1})^{-1}$$

and

$$u^T C^{-1} m = u^T (A^T \Gamma^{-1} y + \hat{C}^{-1} \hat{m}) \quad \forall u \in \mathbb{R}^d \implies m = C A^T \Gamma^{-1} y.$$

Theorem 10

If Assumption 2 holds, then

$$\pi(u|y) = \frac{\exp\left(-\frac{1}{2} |u - m|_C^2 \right)}{2} \tag{5}$$

with

$$C = (A^T \Gamma^{-1} A + \hat{C}^{-1})^{-1} \quad \textit{and} \quad m = C A^T \Gamma^{-1} y.$$

Next time

- For the linear-Gaussian setting, study the posterior density in the small noise limit $\eta \sim N(0, \Gamma)$ when $|\Gamma| \to 0$.

- How informative is the MAP estimator?