SUPPLEMENTARY MATERIAL OF "VARIATIONAL BAYES' METHOD FOR FUNCTIONS WITH APPLICATIONS TO INVERSE PROBLEMS"

JUNXIONG JIA, QIAN ZHAO, ZONGBEN XU, DEYU MENG, AND YEE LEUNG

Abstract. In this supplementary material, we provide all of the proof details for the lemmas and theorems presented in the main text.

Proof of Lemma 2

Proof. Let $\{\nu_n\}_{n=1}^{\infty} = \{\prod_{i=1}^{M}\nu_n^i\}_{n=1}^{\infty}$ be a sequence of measures in $\mathcal{C}$ that converges weakly to a probability measure ν_*. We want to show that $\nu_* \in \mathcal{C}$. Define

$$\nu_*^i := \int_{\prod_{j\neq i}\mathcal{H}_j} d\nu_*, \quad \text{for} \quad i = 1, 2, \cdots, M. \tag{0.1}$$

Obviously, each ν_*^i is a probability measure. Let f_i be some bounded continuous function defined on $\mathcal{H}_i$ with $i = 1, 2, \cdots, M$. Based on the definition of weak convergence, we obtain

$$\int_{\prod_{j=1}^{M}\mathcal{H}_j} f_i d\nu_n \rightarrow \int_{\mathcal{H}_i} f_i d\nu_*^i, \quad \text{as } n \rightarrow \infty. \tag{0.2}$$

It should be noted that the left hand side of (0.2) is equal to

$$\int_{\mathcal{H}_i} f_i d\nu_n^i, \tag{0.3}$$

and we find that each ν_n^i converges weakly to ν_*^i. Therefore, we find that ν_*^i belongs to $\mathcal{A}_i$. Let f be a bounded continuous function defined on $\prod_{j=1}^{M}\mathcal{H}_j$. Then, it is a bounded continuous function for each variable. Based on the definition of weak convergence, we find that

$$\int_{\prod_{j=1}^{M}\mathcal{H}_j} f d\nu_n \rightarrow \int_{\prod_{j=1}^{M}\mathcal{H}_j} f d\nu_*, \tag{0.4}$$

and

$$\int_{\prod_{j=1}^{M}\mathcal{H}_j} f d\nu_n = \int_{\prod_{j=1}^{M}\mathcal{H}_j} f d\nu_n^1 \cdots d\nu_n^M \rightarrow \int_{\prod_{j=1}^{M}\mathcal{H}_j} f d\nu_*^1 \cdots d\nu_*^M, \tag{0.5}$$

when $n \rightarrow \infty$. Relying on the arbitrariness of f, we conclude that $\nu_* = \prod_{j=1}^{M}\nu_*^j$, which completes the proof. □

Proof of Theorem 5

Proof. From the proof of Lemma 2, we know that ν_n^j converges weakly to ν_*^j for every $j = 1, 2, \cdots, M$. According to $\nu_n^j \ll \nu_*^j$ for $j = 1, 2, \cdots, M$, we have

$$\begin{aligned} D_{\mathrm{KL}}(\nu_n||\nu_*) &= \int \frac{d\nu_n}{d\nu_*} \log\left(\frac{d\nu_n}{d\nu_*}\right) d\nu_* = \sum_{j=1}^{M} \int \log\left(\frac{d\nu_n^j}{d\nu_*^j}\right) d\nu_n^j \\ &= \sum_{j=1}^{M} D_{\mathrm{KL}}(\nu_n^j||\nu_*^j). \end{aligned} \tag{0.6}$$

Using Lemma 2.4 proved in [2] and Lemma 22 shown in [1], we find that ν_n converges to ν_* in the total-variation norm. Combined with the above equality (0.6), the proof is completed. □

Proof of Theorem 9

Proof. For a fixed j, let $B \in \mathcal{M}(\mathcal{H}_j)$, and $\nu_n^j \in \mathcal{A}_j$ be a sequence that converges weakly to ν_*^j and

$$\frac{d\nu_n^j}{d\mu_r^j} = \frac{1}{Z_{nr}^j} \exp(-\Phi_j^{nr}(x_j)). \tag{0.7}$$

Assuming that $\mu_r^j(B) = 0$ and by assumption (16) in the main text, we have

$$\nu_n^j(B) = \int_B \frac{1}{Z_{nr}^j} \exp(-\Phi_j^{nr}(x_j)) \mu_r^j(dx_j) = 0.$$

Define

$$B_m = \{x \in B \,|\, \mathrm{dist}(x, B^c) \geq 1/m\}, \tag{0.8}$$

and let $f_m > 0$ be a positive continuous function that satisfies

$$f_m(x) = \begin{cases} 1, & x \in B_m, \\ 0, & x \in B^c. \end{cases}$$

Then, we have

$$\nu_*^j(B_m) \leq \int_{\mathcal{H}_j} f_m d\nu_*^j = \lim_{n\to\infty} \int_{\mathcal{H}_j} f_m d\nu_n^j \leq \lim_{n\to\infty} \nu_n^j(B) = 0, \tag{0.9}$$

and

$$\nu_*^j(B) = \sup_m \nu_*^j(B_m) = 0, \tag{0.10}$$

based on the inner regular property of finite Borel measures. Therefore, there exists a constant and a continuous function denoted by Z_r^j and $\Phi_j^r(\cdot)$ such that

$$\frac{d\nu_*^j}{d\mu_r^j}(x_j) = \frac{1}{Z_r^j} \exp\big(-\Phi_j^r(x_j)\big). \tag{0.11}$$

To complete the proof, we should verify the almost surely positiveness of the right-hand side of the above equality. Assume that $\frac{1}{Z_r^j} \exp\big(-\Phi_j^r(x_j)\big) = 0$ on a set $B \subset \mathcal{H}_j$ with $\mu_r^j(B) > 0$. If $B \subset \mathcal{H}_j \backslash \sup_N T_N^j$, then it holds that $\mu_r^j(B) = 0$ by our assumption. Therefore, $B \cap \sup_N T_N^j$ is not empty, and there exists a constant $\tilde{N}$ such that for all $N \geq \tilde{N}$, $B \cap T_N^j$ is not empty. Denote $B_N = B \cap T_N^j$, and then for a sufficiently large N, we have $\mu_r^j(B_N) \geq \frac{1}{2}\mu_r^j(B)$. Let

$$B_N^m = \{x \in B_N \,|\, \mathrm{dist}(x, B_N^c) \geq 1/m\},$$

and define a function g_m similar to f_m with B_m replaced by B_N^m. Given that $\mu_r^j(B_N) = \sup_m \mu_r^j(B_N^m)$, for a large enough m, we find that

$$\mu_r^j(B_N^m) \geq \frac{1}{2}\mu_r^j(B_N) \geq \frac{1}{4}\mu_r^j(B) > 0.$$

By the definition of weak convergence, we have

$$\lim_{n\to\infty}\int_{\mathcal{H}_j} g_m(x)\frac{1}{Z_{nr}^j}\exp\big(-\Phi_j^{nr}(x)\big)d\mu_r^j = \int_{\mathcal{H}_j} g_m(x)\frac{1}{Z_r^j}\exp\big(-\Phi_j^r(x)\big)d\mu_r^j. \tag{0.12}$$

The right hand side of the above equation is equal to 0, but for a large enough m, the left hand side is positive and the lower bound is

$$\frac{1}{4}\exp(-C_N)\mu_r^j(B). \tag{0.13}$$

This is a contradiction, and thereby the closedness of $\mathcal{A}_j (j = 1, \cdots, M)$ have been proved. Combining the obtained results with the statements in Theorem 3, we obviously obtain the existence of a solution which completes the proof. □

Proof of Theorem 10

Proof. Here, we focus on the deduction of formula (21) presented in the main text. By inserting the prior probability measure into the Kullback-Leibler divergence between ν and μ, for each i $(i = 1, 2, \cdots, M)$ we find that

$$\begin{aligned}
D_{\mathrm{KL}}(\nu||\mu) &= \int_{\mathcal{H}} \log\left(\frac{d\nu}{d\mu_r}\right) - \log\left(\frac{d\mu_0}{d\mu_r}\right) - \log\left(\frac{d\mu}{d\mu_0}\right)d\nu \\
&= \int_{\mathcal{H}}\left(-\sum_{j=1}^{M}\Phi_j^r(x_j) + \Phi^0(x) + \Phi(x)\right)d\nu + \mathrm{Const} \\
&= \int_{\mathcal{H}_i}\left[\int_{\prod_{j\neq i}\mathcal{H}_j}\Big(\Phi^0(x) + \Phi(x)\Big)\prod_{j\neq i}\nu^j(dx_j)\right]\nu^i(dx_i) \\
&\quad - \int_{\mathcal{H}_i}\Phi_i^r(x_i)\nu^i(dx_i) + \text{terms not related to } \Phi_i(x_i).
\end{aligned}$$

For $i = 1, 2, \cdots, M$, let $\tilde{\nu}^i$ be a probability measure defined as follows:

$$\frac{d\tilde{\nu}^i}{d\mu_r^i} \propto \exp\left(-\int_{\prod_{j\neq i}\mathcal{H}_j}\Big(\Phi^0(x) + \Phi(x)\Big)\prod_{j\neq i}\nu^j(dx_j)\right). \tag{0.14}$$

By assumption (19) and (20) shown in the main text, we know that the right-hand side of (0.14) is positive almost surely. Then, we easily know that the measures $\tilde{\nu}^i$ and μ_r^i are equivalent with each other. Therefore, we obtain

$$\begin{aligned}
D_{\mathrm{KL}}(\nu||\mu) &= -\int_{\mathcal{H}_i}\log\left(\frac{d\tilde{\nu}^i}{d\mu_r^i}\right)d\nu^i + \int_{\mathcal{H}_i}\log\left(\frac{d\nu^i}{d\mu_r^i}\right)d\nu^i + \mathrm{Const} \\
&= D_{\mathrm{KL}}(\nu^i||\tilde{\nu}^i) + \text{terms not related to } \nu^i.
\end{aligned} \tag{0.15}$$

Obviously, in order to attain the infimum of the Kullback-Leibler divergence, we should take $\nu^i = \tilde{\nu}^i$. Comparing formula (0.14) with definition (14) in the main text, we notice that the condition $\nu^i = \tilde{\nu}^i$ implies the following equality:

$$\Phi_i^r(x_i) = \int_{\prod_{j\neq i}\mathcal{H}_j} \Big(\Phi^0(x) + \Phi(x)\Big) \prod_{j\neq i} \nu^j(dx_j) + \text{Const},$$

which completes the proof. □

Verify conditions in Theorem 10 for the linear inverse problem introduced in Subsection 3.1

At last, we provide a detailed verification of the conditions in Theorem 10 for the example employed in Subsection 3.1. As stated in Remark 14, we consider $\lambda' = \log\lambda$ and $\tau' = \log\tau$ as hyper-parameters. For a sufficiently small $\epsilon > 0$, taking $a_u(\epsilon, u) := \|u\|^2_{\mathcal{H}_u}$, $a_{\lambda'}(\epsilon, \lambda') := \max\big\{-\lambda', \exp(\epsilon\exp(\lambda'))\big\}$ and $a_{\tau'}(\epsilon, \tau') := \max\big\{-\tau', \exp(\epsilon\exp(\tau'))\big\}$, then we try to verify conditions (19) and (20). In the following, the notation C is a constant that may be different from line to line. In this example, we take $x_1 = u$, $x_2 = \lambda'$, and $x_3 = \tau'$. As shown in the main text, we have

$$\Phi^0(u, \lambda', \tau') = \frac{1}{2}\sum_{j=1}^{K}(u_j - u_{0j})^2(e^{\lambda'} - 1)\alpha_j^{-1} - \frac{K}{2}\lambda',$$
$$\Phi(u, \lambda', \tau') = \frac{e^{\tau'}}{2}\|Hu - d\|^2 - \frac{N_d}{2}\tau'.$$

With these preparations, we firstly verify

$$T^1 := \sup_{u\in T_N^u} \sup_{\substack{\nu^{\lambda'}\in\mathcal{A}_{\lambda'}\\ \nu^{\tau'}\in\mathcal{A}_{\tau'}}} \int_{\mathbb{R}}\int_{\mathbb{R}} (\Phi^0 + \Phi)1_A(u, \lambda', \tau')\nu^{\lambda'}(d\lambda')\nu^{\tau'}(d\tau') < \infty. \tag{0.16}$$

Taking the specific expressions of Φ^0 and Φ into (0.16), we have

$$T^1 \le C \sup_{u\in T_N^u} \sup_{\substack{\nu^{\lambda'}\in\mathcal{A}_{\lambda'}\\ \nu^{\tau'}\in\mathcal{A}_{\tau'}}} \left(T^{11} + T^{12} + T^{13} + T^{14}\right), \tag{0.17}$$

where

$$T^{11} = \int_{\mathbb{R}^+}\int_{\mathbb{R}} \frac{1}{2}\sum_{j=1}^{K}(u_j - u_{0j})^2(e^{\lambda'} - 1)\alpha_j^{-1} e^{-\Phi^r_{\tau'}(\tau')}e^{-\Phi^r_{\lambda'}(\lambda')}\mu_r^{\tau'}(d\tau')\mu_r^{\lambda'}(d\lambda'),$$
$$T^{12} = \int_{\mathbb{R}^-} -\frac{K}{2}\lambda' e^{-\Phi^r_{\lambda'}(\lambda')}\mu_r^{\lambda'}(d\lambda'),$$
$$T^{13} = \int_{\mathbb{R}} \frac{e^{\tau'}}{2}\|Hu - d\|^2 e^{-\Phi^r_{\tau'}(\tau')}\mu_r^{\tau'}(d\tau'),$$
$$T^{14} = \int_{\mathbb{R}^-} -\frac{N_d}{2}\tau' e^{-\Phi^r_{\tau'}(\tau')}\mu_r^{\tau'}(d\tau').$$

Because the techniques used for estimating these terms are similar, we provide the estimates of T^{13} as an example and omit the details for other terms. Because H is assumed to be a linear bounded operator, we have

$$\begin{aligned} T^{13} \leq & C\int_{\mathbb{R}}(e^{\epsilon e^{\tau'}}+1)e^{-\Phi^r_{\tau'}(\tau')}\mu_r^{\tau'}(d\tau') \\ \leq & C\int_{\mathbb{R}}\max(1, a_{\tau'}(\epsilon,\tau'))e^{-\Phi^r_{\tau'}(\tau')}\mu_r^{\tau'}(d\tau') < \infty. \end{aligned} \tag{0.18}$$

Next, we need to estimate

$$T^2 := \sup_{\lambda'\in T_N^{\lambda'}} \sup_{\substack{\nu^u\in\mathcal{A}_u \\ \nu^{\tau'}\in\mathcal{A}_{\tau'}}} \int_{\mathcal{H}_u}\int_{\mathbb{R}}(\Phi^0+\Phi)1_A(u,\lambda',\tau')\nu^{\tau'}(d\tau')\nu^u(du) < \infty. \tag{0.19}$$

Taking the specific expressions of Φ^0 and Φ into (0.19), we have

$$T^2 \leq C \sup_{\lambda'\in T_N^{\lambda'}} \sup_{\substack{\nu^u\in\mathcal{A}_u \\ \nu^{\tau'}\in\mathcal{A}_{\tau'}}} \left(T^{21}+T^{22}+T^{23}+T^{24}\right), \tag{0.20}$$

where

$$\begin{aligned} T^{21} &= e^{\lambda'}\int_{\mathcal{H}_u}\frac{1}{2}\sum_{j=1}^K(u_j-u_{0j})^2\alpha_j^{-1}e^{-\Phi^r_u(u)}\mu_r^u(du), \\ T^{22} &= \frac{K}{2}|\lambda'|, \\ T^{23} &= \int_{\mathcal{H}_u}\|Hu-d\|^2e^{-\Phi^r_u(u)}\mu_r^u(du)\int_{\mathbb{R}}\frac{1}{2}e^{\tau'}e^{-\Phi^r_{\tau'}(\tau')}\mu_r^{\tau'}(d\tau'), \\ T^{24} &= \frac{N_d}{2}\int_R|\tau'|e^{-\Phi^r_{\tau'}(\tau')}\mu_r^{\tau'}(d\tau'). \end{aligned}$$

Remembering that the operator H is bounded and the specific forms of $a_{\tau'}(\epsilon,\tau')$ and $a_u(\epsilon,u)$, we can obtain that the above four terms are all bounded. The following inequality

$$T^3 := \sup_{\tau'\in T_N^{\tau'}} \sup_{\substack{\nu^u\in\mathcal{A}_u \\ \nu^{\lambda'}\in\mathcal{A}_{\lambda'}}} \int_{\mathcal{H}_u}\int_{\mathbb{R}}(\Phi^0+\Phi)1_A(u,\lambda',\tau')\nu^{\lambda'}(d\lambda')\nu^u(du) < \infty \tag{0.21}$$

can be proved similarly, we omit the details. With the above calculations, we verified conditions (19) with $i=1,2,3$. Now, we turn to verify conditions (20). For conditions (20) with $i=2,3$, the inequalities could be verified similarly as for the case of $i=1$. Hence, we only provide details when $i=1$ that is to prove

$$T^4 := \sup_{\substack{\nu^{\lambda'}\in\mathcal{A}_{\lambda'} \\ \nu^{\tau'}\in\mathcal{A}_{\tau'}}} \int_{\mathcal{H}_u}\exp\left(-\int_{\mathbb{R}^2}(\Phi^0+\Phi)1_{A^c}\nu^{\lambda'}(d\lambda')\nu^{\tau'}(d\tau')\right)\max(1,\|u\|^2_{\mathcal{H}_u})\mu_r^u(du) < \infty.$$

Through a direct calculation, we find that

$$\begin{aligned} -\int_{\mathbb{R}^2}(\Phi^0+\Phi)1_{A^c}\nu^{\lambda'}(d\lambda')\nu^{\tau'}(d\tau') \leq & \frac{1}{2}\sum_{j=1}^K\alpha_j^{-1}(u_j-u_{0j})^2\int_{\mathbb{R}-}(1-e^{\lambda'})\nu^{\lambda'}(d\lambda') \\ & + \frac{K}{2}\int_{\mathbb{R}}|\lambda'|e^{-\Phi^r_{\lambda'}(\lambda')}\mu_r^{\lambda'}(d\lambda') \\ & + \frac{N_d}{2}\int_{\mathbb{R}}|\tau'|e^{-\Phi^r_{\tau'}(\tau')}\mu_r^{\tau'}(d\tau'). \end{aligned}$$

Then we have

$$T^4 \leq C \int_{\mathcal{H}_u} \exp\left(\frac{1}{2}\sum_{j=1}^{K}\alpha_j^{-1}(u_j - u_{0j})^2 \int_{\mathbb{R}^-}(1-e^{\lambda'})\nu^{\lambda'}(d\lambda')\right)\max(1, \|u\|_{\mathcal{H}_u}^2)\mu_r^u(du).$$

Considering $\int_{\mathbb{R}^-}(1-e^{\lambda'})\nu^{\lambda'}(d\lambda') < 1$ and the definition of μ_r^u, we know that the right hand side of the above inequality is bounded which completes the proof.

References

[1] M. Dashti and A. M. Stuart. The Bayesian approach to inverse problems. *Handbook of Uncertainty Quantification*, pages 311–428, 2017.

[2] F. J. Pinski, G. Simpson, A. M. Stuart, and H. Weber. Kullback-Leibler approximation for probability measures on infinite dimensional space. *SIAM J. Math. Anal.*, 47(6):4091–4122, 2015.

School of Mathematics and Statistics, Xi'an Jiaotong University, Xi'an, 710049, China

E-mail address: jjx323@xjtu.edu.cn

School of Mathematics and Statistics, Xi'an Jiaotong University, Xi'an, 710049, China

E-mail address: timmy.zhaoqian@xjtu.edu.cn

School of Mathematics and Statistics, Xi'an Jiaotong University, Xi'an, 710049, China

E-mail address: zbxu@mail.xjtu.edu.cn

School of Mathematics and Statistics, Xi'an Jiaotong University, Xi'an, 710049, China

E-mail address: dymeng@mail.xjtu.edu.cn

The Chinese University of Hong Kong, Shatin, Hong Kong, SAR

E-mail address: yeeleung@cuhk.edu.hk