1.贝叶斯平滑问题对数据的适定性定理

对于贝叶斯平滑问题，$Y=Y_N$, 假设

- $\|Y\|_2, \|Y'\|_2 \le R < \infty$
- 定义 $\varphi(X) := \left(\sum_{n=1}^N \|\mathcal{H}(x_n)\|_2^2\right)^{1/2}$，满足 $\mathbb{E}_{\rho_{\text{prior}}(X)}[\varphi^2(X)] < \infty$.

那么存在 $c>0$，后验分布满足

$$\mathcal{D}_H(\rho_{\text{post}}(X;Y), \rho_{\text{post}}(X;Y')) \le c\|Y-Y'\|_2$$

证明

对于贝叶斯平滑问题，我们有

$$\begin{aligned}\rho_{\text{post}}(X;Y) &= \frac{1}{Z}e^{-\Phi(X,Y)}\rho_{\text{prior}}(X)\\ \rho_{\text{post}}(X;Y') &= \frac{1}{Z'}e^{-\Phi(X,Y')}\rho_{\text{prior}}(X)\end{aligned}$$

其中似然函数满足

$$\begin{aligned}L(X) = e^{-\Phi(X,Y)} &= \exp\Big\{\frac{1}{2}\sum_{n=0}^{N-1}\|\Sigma_\eta^{-1/2}(y_{n+1}-\mathcal{H}(x_{n+1}))\|^2\Big\}\\ L'(X) = e^{-\Phi(X,Y')} &= \exp\Big\{\frac{1}{2}\sum_{n=0}^{N-1}\|\Sigma_\eta^{-1/2}(y'_{n+1}-\mathcal{H}(x_{n+1}))\|^2\Big\}\end{aligned}$$

我们仅需要验证关于贝叶斯反问题的适定性定理的假设。由于$\Phi(X,Y)>0$，我们有

$$\sup_X\Big|\sqrt{L(X)}\Big| + \Big|\sqrt{L'(X)}\Big| \le 2$$

由于$|e^{-w_1}-e^{-w_2}| \le |w_1-w_2|$，我们有

$$\begin{aligned}\Big|\sqrt{L(X)}-\sqrt{L'(X)}\Big| &\le \frac{1}{2}\Big|\sum_{n=0}^{N-1}\frac{1}{2}\Big(\|\Sigma_\eta^{-1/2}(y_{n+1}-\mathcal{H}(x_{n+1}))\|^2 - \|\Sigma_\eta^{-1/2}(y'_{n+1}-\mathcal{H}(x_{n+1}))\|^2\Big)\Big|\\ &= \frac{1}{4}\sum_{n=0}^{N-1}(y_{n+1}-y'_{n+1})\Sigma_\eta^{-1}(y_{n+1}+y'_{n+1}-2\mathcal{H}(x_{n+1}))\\ &\le \frac{1}{4}\|\Sigma_\eta^{-1}\|_2\|Y-Y'\|_2\sum_{n=0}^{N-1}(\|y_{n+1}\|_2+\|y'_{n+1}\|_2+2\|\mathcal{H}(x_{n+1})\|_2)\\ &\le c\|Y-Y'\|_2(R+\varphi(X))\end{aligned}$$

我们记 $\delta = \|Y-Y'\|_2$，我们有

$$\mathbb{E}_{\rho_{\text{prior}}(X)}[c^2(R+\varphi)^2] \le 2c^2(R^2+\mathbb{E}_{\rho_{\text{prior}}(X)}[\varphi(X)^2]) \le \infty$$

2. 贝叶斯滤波问题对数据的适定性定理

- $\|Y\|_2, \|Y'\|_2 \le R < \infty$
- 定义 $\varphi(X) := \left(\sum_{n=1}^N \|\mathcal{H}(x_n)\|_2^2\right)^{1/2}$，满足 $\mathbb{E}_{\rho_{\text{prior}}(X)}[\varphi^2(X)] < \infty$.

那么存在 $c>0$，后验分布满足

$$\mathcal{D}_{TV}(\rho_N(x_N;Y_N), \rho_N(x_N;Y'_N)) \le c\|Y-Y'\|_2$$

证明

根据贝叶斯平滑问题的适定性，我们有

$$\mathcal{D}_H(\rho_{\text{post}}(X_N;Y_N), \rho_{\text{post}}(X_N;Y'_N)) \le c\|Y_N-Y'_N\|_2$$

我们定义 $F(x_1, x_2, \cdots, x_N) = F(X_N) := f(x_N)$

$$\begin{aligned}\mathcal{D}_{TV}(\rho_N(x_N;Y_N), \rho_N(x_N;Y'_N)) &= \frac{1}{2}\sup_{|f|_\infty\le 1}\Big|\mathbb{E}_{\rho_N(x_N;Y_N)}[f(x_N)] - \mathbb{E}_{\rho_N(x_N;Y'_N)}[f(x_N)]\Big|\\ &\le \frac{1}{2}\sup_{|F|_\infty\le 1}\Big|\mathbb{E}_{\rho_{\text{post}}(X_N;Y_N)}[F(X_N)] - \mathbb{E}_{\rho_{\text{post}}(X_N;Y_N)}[F(X_N)]\Big|\\ &= \mathcal{D}_{TV}(\rho_{\text{post}}(X_N;Y_N), \rho_{\text{post}}(X_N;Y'_N))\\ &\le \sqrt{2}\mathcal{D}_H(\rho_{\text{post}}(X_N;Y_N), \rho_{\text{post}}(X_N;Y'_N))\\ &\le c\|Y_N-Y'_N\|_2\end{aligned}$$

这里我们用了 $\mathcal{D}_{TV} \le \mathcal{D}_H$.

3. 练习

$$\begin{aligned}\frac{1}{\sqrt{2}}\mathcal{D}_{TV}(\rho_A,\rho_B) \le \mathcal{D}_H(\rho_A,\rho_B) &\le \sqrt{\mathcal{D}_{TV}(\rho_A,\rho_B)}\\ \mathcal{D}_{TV}(\rho_A,\rho_B) &= \frac{1}{2}\sup_{|f|_\infty\le 1}|\mathbb{E}_{\rho_A}[f(x)] - \mathbb{E}_{\rho_B}[f(x)]|\end{aligned}$$

证明

使用 Cauchy–Schwarz 不等式

$$\begin{aligned}\mathcal{D}_{TV}(\rho_A,\rho_B) &= \frac{1}{2}\int\Big|\sqrt{\rho_A(x)}-\sqrt{\rho_B(x)}\Big|\Big|\sqrt{\rho_A(x)}+\sqrt{\rho_B(x)}\Big|dx\\ &\le \left(\frac{1}{2}\int\Big|\sqrt{\rho_A(x)}-\sqrt{\rho_B(x)}\Big|^2 d\right)^{1/2}\left(\frac{1}{2}\int\Big|\sqrt{\rho_A(x)}+\sqrt{\rho_B(x)}\Big|^2 dx\right)^{1/2}\\ &\le \mathcal{D}_H(\rho_A,\rho_B)\left(\frac{1}{2}\int\big(2\rho_A(x)+2\rho_B(x)\big)dx\right)^{1/2}\\ &= \sqrt{2}\mathcal{D}_H(\rho_A,\rho_B).\end{aligned}$$

由于我们有$\sqrt{\rho_A(x)}, \sqrt{\rho_B(x)} \ge 0$，因此 $|\sqrt{\rho_A(x)}-\sqrt{\rho_B(x)}| \le |\sqrt{\rho_A(x)}+\sqrt{\rho_B(x)}|$。

因此我们有

$$\begin{aligned}\mathcal{D}_H(\rho_A,\rho_B) &= \left(\frac{1}{2}\int\Big|\sqrt{\rho_A(x)}-\sqrt{\rho_B(x)}\Big|^2 du\right)^{1/2}\\ &\le \left(\frac{1}{2}\int\Big|\sqrt{\rho_A(x)}-\sqrt{\rho_B(x)}\Big|\Big|\sqrt{\rho_A(x)}+\sqrt{\rho_B(x)}\Big|du\right)^{1/2}\\ &\le \left(\frac{1}{2}\int|\rho_A(x)-\rho_B(x)|du\right)^{1/2}\\ &= \sqrt{\mathcal{D}_{TV}(\rho_A,\rho_B)}.\end{aligned}$$

对于全变差距离，我们有

$$\begin{aligned}\Big|\mathbb{E}_{\rho_A}[f(x)] - \mathbb{E}_{\rho_B}[f(x)]\Big| &= \Big|\int f(x)\big(\rho_A(x)-\rho_B(x)\big)dx\Big|\\ &\le 2|f|_\infty\cdot\frac{1}{2}\int|\rho_A(x)-\rho_B(x)|dx\\ &= 2|f|_\infty\mathcal{D}_{TV}(\rho_A,\rho_B).\end{aligned}$$

定义 $f(x) := \text{sign}\Big(\rho_A(x)-\rho_B(x)\Big)$，我们有 $f(x)\Big(\rho_A(x)-\rho_B(x)\Big) = |\rho_A(x)-\rho_B(x)|$，而且 $|f|_\infty = 1$。我们有

$$\begin{aligned}\mathcal{D}_{TV}(\rho_A,\rho_B) &= \frac{1}{2}\int|\rho_A(x)-\rho_B(x)|\,du\\ &= \frac{1}{2}\int f(x)\Big(\rho_A(x)-\rho_B(x)\Big)\,du\\ &= \frac{1}{2}\Big|\mathbb{E}_{\rho_A}[f(x)] - \mathbb{E}_{\rho_B}[f(x)]\Big|.\end{aligned}$$

In [ ]: