1. 练习

考虑反问题：

$$y = \mathcal{G}(\theta) + \eta \qquad \theta = [\theta_1; \theta_2] \qquad \mathcal{G}(\theta) = \theta_1^2 + \theta_2^2$$

已知：

- 观测： $y > 0$
- 噪音： $\eta \sim \mathcal{N}(x; 0, \gamma^2)$
- 先验分布： $0 < \rho_{\text{prior}}(\theta) \leq \rho_{\max} < \infty$

问题：

- 计算最大似然估计
- 计算当$\rho_{\text{prior}}(\theta) = \mathcal{N}(\theta; 0c^2\gamma^2 I)$ 时的最大后验估计
- 探究当$\gamma \to 0$ 时，后验分布的变化

证明

最大似然估计

$$L(\theta; y) \propto \exp\Big(-\frac{(\theta_1^2 + \theta_2^2 - y)^2}{2\gamma^2}\Big)$$

最大似然估计是$\frac{(\theta_1^2+\theta_2^2-y)^2}{2\gamma^2}$的极小值，满足 $\theta_1^2 + \theta_2^2 = y$ 的圆。 不唯一！

最大后验估计

$$\rho_{\text{post}}(\theta; y) \propto \exp\Big(-\frac{(\theta_1^2 + \theta_2^2 - y)^2}{2\gamma^2} - \frac{\theta_1^2 + \theta_2^2}{2c^2\gamma^2}\Big)$$

最大后验估计是$\frac{(\theta_1^2+\theta_2^2-y)^2}{2\gamma^2} + \frac{\theta_1^2+\theta_2^2}{2c^2\gamma^2}$的极小值:

- 当 $c \leq \frac{1}{\sqrt{2y}}, \theta_1 = \theta_2 = 0$
- 当 $c > \frac{1}{\sqrt{2y}}, \theta_1^2 + \theta_2^2 = y - \frac{1}{2c^2}$。不唯一！

后验分布

$$\rho_{\text{post}}(\theta; y) \propto \exp\Big(-\frac{(\theta_1^2 + \theta_2^2 - y)^2}{2\gamma^2}\Big)\rho_{\text{prior}}(\theta)$$

考虑如下两个区域

$$A^+ := \{\theta : |\theta_1^2 + \theta_2^2 - y|^2 \leq \gamma^{2+\delta}\} \qquad A^- := \{\theta : |\theta_1^2 + \theta_2^2 - y|^2 > \gamma^{2-\delta}\}$$

当$\gamma \to 0$， 有对任意$\theta^+ \in A^+$ 和 $\theta^- \in A^-$， 我们有

$$\begin{aligned}\frac{\rho_{\text{post}}(\theta^+; y)}{\rho_{\text{post}}(\theta^-; y)} &= \exp\Big(-\frac{({\theta_1^+}^2 + {\theta_2^+}^2 - y)^2}{2\gamma^2} + \frac{({\theta_1^-}^2 + {\theta_2^-}^2 - y)^2}{2\gamma^2}\Big)\frac{\rho_{\text{prior}}(\theta^+)}{\rho_{\text{prior}}(\theta^-)} \\ &\geq \exp\Big(-\frac{1}{2}\gamma^\delta + \frac{1}{2}\gamma^{-\delta}\Big)\frac{\rho_{\min}}{\rho_{\max}} \to \infty\end{aligned}$$

$\rho_{\min} > 0$ 是 ρ_{prior}在半径为 $2\sqrt{y}$的球里的最小值。

因此，注意到当$\gamma \to 0$时，后验分布$\rho_{\text{post}}(\theta|y)$会集中在半径为 $\sqrt{y}$ 的圆周上。

2. 存在性定理

对于贝叶斯反问题

$$y = \mathcal{G}(\theta) + \eta \qquad \eta \sim \rho_\eta \qquad \theta \sim \rho_{\text{prior}}$$

假设

- θ和η独立
- $Z = \int \rho_\eta(y - \mathcal{G}(\theta))\rho_{\text{prior}} d\theta > 0$

那么后验分布存在

$$\rho_{\text{post}}(\theta; y) = \frac{1}{Z}\rho_\eta(y - \mathcal{G}(\theta))\rho_{\text{prior}}(\theta)$$

证明

对于条件概率，我们有

$$\begin{aligned}\rho(\theta, y) = \rho(\theta|y)\rho(y) \qquad &\text{当} \qquad \rho(y) > 0 \\ \rho(\theta, y) = \rho(y|\theta)\rho(\theta) \qquad &\text{当} \qquad \rho(\theta) > 0\end{aligned}$$

根据假设$\rho(y) > 0$， 当$\rho_{\text{prior}}(\theta) > 0$, 我们有

$$\rho(\theta|y) = \frac{\rho(\theta, y)}{\rho(y)} = \frac{1}{Z}\rho_\eta(y - \mathcal{G}(\theta))\rho_{\text{prior}}(\theta)$$

当$\rho_{\text{prior}}(\theta) = 0$, 我们有

$$\rho(\theta|y) = 0 = \frac{1}{Z}\rho_\eta(y - \mathcal{G}(\theta))\rho_{\text{prior}}(\theta)$$

3. 适定性定理

对于贝叶斯反问题

$$y = \mathcal{G}(\theta) + \eta \qquad \eta \sim \rho_\eta \qquad \theta \sim \rho_{\text{prior}}$$

定义

$$L(\theta) = \rho_\eta(y - \mathcal{G}(\theta)) \qquad L_\eta(\theta) = \rho_\eta(y - \mathcal{G}_\eta(\theta))$$

假设存在$K_1, K_2 > 0, \delta^+ > 0$， 对任意 $\delta \in (0, \delta^+)$， 我们有

- $|\sqrt{L(\theta)} - \sqrt{L_\eta(\theta)}| \leq \varphi(\theta)\delta \qquad \mathbb{E}_{\theta \sim \rho_{\text{prior}}}[\varphi(\theta)^2] \leq K_1^2$
- $\sup_\theta |\sqrt{L(\theta)} + \sqrt{L_\eta(\theta)}| \leq K_2$

那么存在$\Delta > 0, c > 0$，后验分布满足

$$\mathcal{D}_H(\rho_{\text{post}}(\theta; \mathcal{G}), \rho_{\text{post}}(\theta; \mathcal{G}_\delta)) < c\delta \qquad \forall \delta \in (0, \Delta)$$

证明

我们将总误差分为两部分，一部分反映了L和L_δ之间的差异, 另一部分反映了$Z := \int L(\theta)\rho_{\text{prior}}(\theta)d\theta$ and $Z_\delta := \int L_\delta(\theta)\rho_{\text{prior}}(\theta)d\theta$之间的差异：

$$\begin{aligned}\mathcal{D}_H(\rho_{\text{post}}(\theta; \mathcal{G}), \rho_{\text{post}}(\theta; \mathcal{G}_\delta)) &= \frac{1}{\sqrt{2}}\Big\|\sqrt{\rho_{\text{post}}(\theta; \mathcal{G})} - \sqrt{\rho_{\text{post}}(\theta; \mathcal{G}_\delta)}\Big\|_{L^2} \\ &= \frac{1}{\sqrt{2}}\Big\|\sqrt{\frac{L\rho_{\text{prior}}}{Z}} - \sqrt{\frac{L\rho_{\text{prior}}}{Z_\delta}} + \sqrt{\frac{L\rho_{\text{prior}}}{Z_\delta}} - \sqrt{\frac{L_\delta\rho_{\text{prior}}}{Z_\delta}}\Big\|_{L^2} \\ &\leq \frac{1}{\sqrt{2}}\Big\|\sqrt{\frac{L\rho_{\text{prior}}}{Z}} - \sqrt{\frac{L\rho_{\text{prior}}}{Z_\delta}}\Big\|_{L^2} + \frac{1}{\sqrt{2}}\Big\|\sqrt{\frac{L\rho_{\text{prior}}}{Z_\delta}} - \sqrt{\frac{L_\delta\rho_{\text{prior}}}{Z_\delta}}\Big\|_{L^2}.\end{aligned}$$

对于Z和Z_δ， 我们有

$$\begin{aligned}|Z - Z_\delta| &= \Big|\int \big(L(\theta) - L_\delta(\theta)\big)\rho_{\text{prior}}(\theta)d\theta\Big| \\ &\leq \Big(\int\Big|\sqrt{L(\theta)} - \sqrt{L_\delta(\theta)}\Big|^2\rho_{\text{prior}}(\theta)d\theta\Big)^{1/2}\Big(\int\Big|\sqrt{L(\theta)} + \sqrt{L_\delta(\theta)}\Big|^2\rho_{\text{prior}}(\theta)d\theta\Big)^{1/2} \\ &\leq \Big(\int \delta^2\phi(\theta)^2\rho_{\text{prior}}(\theta)d\theta\Big)^{1/2}\Big(\int K_2^2\rho_{\text{prior}}(\theta)d\theta\Big)^{1/2} \\ &\leq K_1K_2\delta, \quad \delta \in (0, \delta^+).\end{aligned}$$

由于$Z > 0$， 取 $\Delta = \frac{Z}{2K_1K_2}$， 当 $\delta < \Delta < \delta^+$, 我们有

$$Z_\delta \geq Z - |Z - Z_\delta| \geq \frac{1}{2}Z$$

对于第一项， 当$\delta < \Delta$，我们有

$$\begin{aligned}\Big\|\sqrt{\frac{L\rho_{\text{prior}}}{Z}} - \sqrt{\frac{L\rho_{\text{prior}}}{Z_\delta}}\Big\|_{L^2} &= \Big|\frac{1}{\sqrt{Z}} - \frac{1}{\sqrt{Z_\delta}}\Big|\Big(\int L(\theta)\rho_{\text{prior}}(\theta)d\theta\Big)^{1/2} \\ &= \frac{|Z - Z_\delta|}{(\sqrt{Z} + \sqrt{Z_\delta})\sqrt{Z_\delta}} \\ &\leq \frac{K_1K_2}{Z}\delta,\end{aligned}$$

对于第二项， 当$\delta < \Delta$，我们有

$$\begin{aligned}&\Big\|\sqrt{\frac{L\rho_{\text{prior}}}{Z_\delta}} - \sqrt{\frac{L_\delta\rho_{\text{prior}}}{Z_\delta}}\Big\|_{L^2} \\ =&\frac{1}{\sqrt{Z_\delta}}\Big(\int\Big|\sqrt{L(\theta)} - \sqrt{L_\delta(\theta)}\Big|^2\rho_{\text{prior}}(\theta)d\theta\Big)^{1/2} \\ \leq&\sqrt{\frac{2K_1^2}{Z}}\delta.\end{aligned}$$

因此

$$\mathcal{D}_H(\rho_{\text{post}}(\theta; \mathcal{G}), \rho_{\text{post}}(\theta; \mathcal{G}_\delta)) \leq \frac{K_1K_2}{Z} + \sqrt{\frac{2K_1^2}{Z}}\delta = c\delta,$$

其中 $c = \frac{K_1K_2}{Z} + \sqrt{\frac{2K_1^2}{Z}}$, 不依赖于 δ.

4. 练习

考虑如下常微分方程描述的系统

$$\frac{dx}{dt} = f(x, \theta) \qquad x(0) = x_0$$

贝叶斯反问题

$$y = \mathcal{G}(\theta) + \eta \qquad y = x(1) \qquad \eta \sim \mathcal{N}(x; 0, \delta_\eta^2 I)$$

假设一致有界和Lipschitz连续，即存在常数$f_{\max}$：

$$\begin{aligned}&|f(x, \theta)|, |\partial_x f(x, \theta)| < f_{\max} \qquad \forall x, \theta \\ &|f(x, \theta) - f(x', \theta)| < f_{\max}|x - x'| \qquad \forall x, x', \theta\end{aligned}$$

基于向前欧拉方法的贝叶斯反问题

$$y = \mathcal{G}_\delta(\theta) + \eta$$

能近似出θ 的后验分布。

证明

对于向前欧拉方法，我们有

$$x_{n+1} = x_n + \delta f(x_n, \theta)$$

其中时间步长是 $\delta = \frac{1}{N}$。对于近似的贝叶斯反问题，我们有$\mathcal{G}_\delta(\theta) = x_N$。 我们需要验证，假设存在$K_1, K_2 > 0, \delta^+ > 0$， 对任意 $\delta \in (0, \delta^+)$， 我们有

- $|\sqrt{L(\theta)} - \sqrt{L_\eta(\theta)}| \leq \varphi(\theta)\delta \qquad \mathbb{E}_{\theta \sim \rho_{\text{prior}}}[\varphi(\theta)^2] \leq K_1^2$
- $\sup_\theta |\sqrt{L(\theta)} + \sqrt{L_\eta(\theta)}| \leq K_2$

由于噪声 $\eta \sim \mathcal{N}(0, \delta_\eta^2 I)$ 是高斯，我们有

$$\begin{aligned}\sqrt{L(\theta)} &= \sqrt{\rho_\eta\big(y - \mathcal{G}(\theta)\big)} = \frac{1}{(2\pi)^{N_x/4}\delta_\eta^{N_x/2}}\exp\Big(-\frac{1}{4\delta_\eta^2}|y - \mathcal{G}(\theta)|^2\Big) \\ \sqrt{L_\delta(\theta)} &= \sqrt{\rho_\eta\big(y - \mathcal{G}_\delta(\theta)\big)} = \frac{1}{(2\pi)^{N_x/4}\delta_\eta^{N_x/2}}\exp\Big(-\frac{1}{4\delta_\eta^2}|y - \mathcal{G}_\delta(\theta)|^2\Big).\end{aligned}$$

对于假设一， 由于对任意$w_1, w_2 > 0$, 我们有$|e^{-w_1} - e^{-w_2}| \leq |w_1 - w_2|$。因此，我们有

$$\begin{aligned}\Big|\sqrt{L(\theta)} - \sqrt{L_\delta(\theta)}\Big| &\leq \frac{1}{(2\pi)^{N_x/4}\delta_\eta^{N_x/2}} \cdot \frac{1}{4\delta_\eta^2} \cdot \big||y - \mathcal{G}(\theta)|^2 - |y - \mathcal{G}_\delta(\theta)|^2\big| \\ &= \frac{1}{(2\pi)^{N_x/4}\delta_\eta^{N_x/2}} \cdot \frac{1}{4\delta_\eta^2} \cdot |2y - \mathcal{G}(\theta) - \mathcal{G}_\delta(\theta)||\mathcal{G}(\theta) - \mathcal{G}_\delta(\theta)| \\ &\leq c\delta.\end{aligned}$$

这里我们用了关于f的假设，以及向前欧拉方法的一阶全局误差性质。

对于假设二，我们有

$$\begin{aligned}\sqrt{L(\theta)} &= \frac{1}{(2\pi)^{N_x/4}\delta_\eta^{N_x/2}}\exp\Big(-\frac{1}{4\delta_\eta^2}|y - \mathcal{G}(\theta)|^2\Big) \leq \frac{1}{(2\pi)^{N_x/4}\delta_\eta^{N_x/2}} \\ \sqrt{L_\delta(\theta)} &= \frac{1}{(2\pi)^{N_x/4}\delta_\eta^{N_x/2}}\exp\Big(-\frac{1}{4\delta_\eta^2}|y - \mathcal{G}_\delta(\theta)|^2\Big) \leq \frac{1}{(2\pi)^{N_x/4}\delta_\eta^{N_x/2}}.\end{aligned}$$

因此我们可以取 $K_2 = \frac{2}{(2\pi)^{N_x/4}\delta_\eta^{N_x/2}}$.