

1. 线性贝叶斯反问题的后验分布

后验分布满足

$$p(\theta; y) = \frac{1}{Z} e^{-\Phi_R(\theta, y)} \quad \Phi_R = \frac{1}{2} \|\Sigma_\eta^{-1/2}(y - G\theta)\|^2 + \frac{1}{2} \|\Sigma_0^{-1/2}(\theta - r_0)\|^2$$

我们能将 Φ_R 重新写为

$$\begin{aligned} \Phi_R &= \frac{1}{2} \left((y - G\theta)^T \Sigma_\eta^{-1} (y - G\theta) + (\theta - r_0)^T \Sigma_0^{-1} (\theta - r_0) \right) \\ &= \frac{1}{2} \left(\theta^T G^T \Sigma_\eta^{-1} G \theta + \theta^T \Sigma_0^{-1} \theta + 2\theta^T G^T \Sigma_\eta^{-1} y + 2\theta^T \Sigma_0^{-1} r_0 \right) + const. \\ &= \frac{1}{2} \left(\theta^T (G^T \Sigma_\eta^{-1} G + \Sigma_0^{-1}) \theta + 2\theta^T (G^T \Sigma_\eta^{-1} y + \Sigma_0^{-1} r_0) \right) + const. \\ &= \frac{1}{2} (\theta - m_{\text{post}})^T C_{\text{post}}^{-1} (\theta - m_{\text{post}}) + const. \end{aligned}$$

因此我们有

$$C_{\text{post}} = (G^T \Sigma_\eta^{-1} G + \Sigma_0^{-1})^{-1} \quad m_{\text{post}} = (G^T \Sigma_\eta^{-1} G + \Sigma_0^{-1})^{-1} (G^T \Sigma_\eta^{-1} y + \Sigma_0^{-1} r_0)$$

2. 超定系统后验一致性

对于超定线性系统 $\text{Null}(G) > 0$, $N_y > N_\theta$, 其中数据满足

$$y = G\theta^\dagger + \gamma\eta^\dagger$$

这里 θ^\dagger 是真实参数 η^\dagger 是未知的真实误差。我们假设噪声满足 $\eta \sim \mathcal{N}(x; 0, \gamma^2 \Sigma_\eta)$, 先验分布满足 $\theta \sim \mathcal{N}(\theta; r_0, \Sigma_0)$ 。

对于任意序列 $M(\gamma) \rightarrow \theta$, $\gamma \rightarrow 0$, 我们有

$$P_{\text{post}}(|\theta - \theta^\dagger|^2 > M(\gamma)\gamma^2) \rightarrow 0 \quad (1)$$

当我们假设 $M(\gamma) = \frac{\epsilon^2}{\gamma^2}$, 我们有

$$P_{\text{post}}(|\theta - \theta^\dagger|^2 > \epsilon^2) \rightarrow 0 \quad (2)$$

证明

我们有后验分布

$$C_{\text{post}} = \gamma^2 (G^T \Sigma_\eta^{-1} G + \gamma^2 \Sigma_0^{-1})^{-1} \quad m_{\text{post}} = (G^T \Sigma_\eta^{-1} G + \gamma^2 \Sigma_0^{-1})^{-1} (G^T \Sigma_\eta^{-1} y + \gamma^2 \Sigma_0^{-1} r_0)$$

为了研究 $\gamma \rightarrow 0$ 的情形, 我们定义

$$C^+ = (G^T \Sigma_\eta^{-1} G)^{-1} \quad m^+ = (G^T \Sigma_\eta^{-1} G)^{-1} (G^T \Sigma_\eta^{-1} y)$$

对于后验分布取期望, 我们有

$$\exists c > 0 \quad \text{s. t.} \quad \mathbb{E}[|\theta - \theta^\dagger|^2] \leq c \left(\mathbb{E}[|\theta - m_{\text{post}}|^2] + |m_{\text{post}} - m^+|^2 + |m^+ - \theta^\dagger|^2 \right). \quad (3)$$

接下来, 我们将放缩每一项。对于第一项

$$\begin{aligned} \mathbb{E}[|\theta - m_{\text{post}}|^2] &= \mathbb{E}[(\theta - m_{\text{post}})^T (\theta - m_{\text{post}})] = \mathbb{E}[\text{Tr}[(\theta - m_{\text{post}}) \otimes (\theta - m_{\text{post}})]] \\ &= \text{Tr} \mathbb{E}[(\theta - m_{\text{post}}) \otimes (\theta - m_{\text{post}})] \\ &= \text{Tr}(C_{\text{post}}) \leq \gamma^2 \text{Tr}[(G^T \Sigma_0^{-1} G)^{-1}]. \end{aligned}$$

对于第二项, 由于

$$\begin{aligned} (G^T \Sigma_\eta^{-1} G) m^+ &= G^T \Sigma_\eta^{-1} y, \\ (G^T \Sigma_\eta^{-1} G + \gamma^2 \Sigma_0^{-1}) m_{\text{post}} &= G^T \Sigma_\eta^{-1} y + \gamma^2 \Sigma_0^{-1} r_0. \end{aligned}$$

我们有

$$m_{\text{post}} - m^+ = \gamma^2 (G^T \Sigma_\eta^{-1} G)^{-1} (\Sigma_0^{-1} r_0 - \Sigma_0^{-1} m_{\text{post}})$$

由于 m_{post} 会收敛到 m^+ , 因此有界。我们有对第二项的估计

$$|m_{\text{post}} - m^+|^2 \leq c\gamma^4$$

对于第三项, 我们有

$$\begin{aligned} m^+ &= (G^T \Sigma_\eta^{-1} G)^{-1} G^T \Sigma_\eta^{-1} G \theta^\dagger + \gamma (G^T \Sigma_\eta^{-1} G)^{-1} G^T \Sigma_\eta^{-1} \eta^\dagger \\ &= \theta^\dagger + \gamma (G^T \Sigma_\eta^{-1} G)^{-1} G^T \Sigma_\eta^{-1} \eta^\dagger \end{aligned}$$

因此, 我们有对第三项的估计

$$|m^+ - \theta^\dagger|^2 \leq c\gamma^2$$

最后我们使用马尔科夫不等式, 当 $\gamma \rightarrow 0$, 我们有

$$P_{\text{post}}(|\theta - \theta^\dagger|^2 > M(\gamma)\gamma^2) \leq \frac{\mathbb{E}[|\theta - \theta^\dagger|^2]}{M(\gamma)\gamma^2} \leq \frac{c}{M(\gamma)} \rightarrow 0$$

3. 欠定系统无后验一致性

对于欠定线性系统 $\text{Rank}(G) = N_y$, $N_y < N_\theta$

$$y = G\theta + \gamma\eta \quad \eta \sim \mathcal{N}(x; 0, \Sigma_\eta) \quad \theta \sim \mathcal{N}(\theta; r_0, \Sigma_0)$$

对 G 做正交化 $G = G_0 Q_1^T$, 其中 $G_0 \in R^{N_y \times N_y}$ 可逆, $Q_1 \in R^{N_\theta \times N_\theta}$ 列向量正交。我们可以对 Q_1 进行正交补全 $Q = [Q_1; Q_2] \in R^{N_\theta \times N_\theta}$, $Q^T Q = I$

我们有

$$G = [G_0; 0] Q^T = [G_0; 0] [Q_1; Q_2]^T = G_0 Q_1^T$$

其中 $G_0 \in R^{N_\theta \times (N_y - N_\theta)}$ 是核空间 $G Q_2 = 0$ 。

当 $\gamma \rightarrow 0$, 我们有

$$\begin{aligned} m_{\text{post}} &\rightarrow m^+ = \Sigma_0 Q_1 (Q_1^T \Sigma_0 Q_1)^{-1} G_0 y + Q_2 (Q_2^T \Sigma_0^{-1} Q_2)^{-1} Q_2^T \Sigma_0^{-1} r_0 \\ C_{\text{post}} &\rightarrow C^+ = Q_2 (Q_2^T \Sigma_0^{-1} Q_2)^{-1} Q_2^T \end{aligned}$$

证明

对于后验协方差矩阵, 我们有

$$C_{\text{post}} = \Sigma_0 - \Sigma_0 G^T (G \Sigma_0 G^T + \gamma^2 \Sigma_\eta)^{-1} G \Sigma_0$$

由于 $G \Sigma_0 G^T \in R^{N_y \times N_y}$ 可逆, 当 $\gamma \rightarrow 0$

$$\begin{aligned} C_{\text{post}} &= \Sigma_0 - \Sigma_0 G^T (G \Sigma_0 G^T)^{-1} G \Sigma_0 \\ &= \Sigma_0 - \Sigma_0 Q_1 G_0^T (G_0 Q_1^T \Sigma_0 Q_1 G_0^T)^{-1} G_0 Q_1^T \Sigma_0 \\ &= \Sigma_0 - \Sigma_0 Q_1 (Q_1^T \Sigma_0 Q_1)^{-1} Q_1^T \Sigma_0 \\ &= Q_2 (Q_2^T \Sigma_0^{-1} Q_2)^{-1} Q_2^T \end{aligned}$$

对于最后一个等式, 我们使用了如下恒等式

$$I = R := \Sigma_0 Q_1 (Q_1^T \Sigma_0 Q_1)^{-1} Q_1^T + Q_2 (Q_2^T \Sigma_0^{-1} Q_2)^{-1} Q_2^T \Sigma_0^{-1}$$

这个恒等式是由于 $Q_1^T (R - I) = 0$ 和 $Q_2^T \Sigma_0^{-1} (R - I) = 0$ 。

对于后验期望, 我们有

$$m_{\text{post}} = r_0 - \Sigma_0 G^T (\gamma^2 \Sigma_\eta + G \Sigma_0 G^T)^{-1} (G r_0 - y)$$

由于 $G \Sigma_0 G^T \in R^{N_y \times N_y}$ 可逆, 当 $\gamma \rightarrow 0$

$$\begin{aligned} m_{\text{post}} &= r_0 + \Sigma_0 G^T (G \Sigma_0 G^T)^{-1} (y - G r_0) \\ &= r_0 + \Sigma_0 Q_1 G_0^T (G_0 Q_1^T \Sigma_0 Q_1 G_0^T)^{-1} (y - G_0 Q_1^T r_0) \\ &= \Sigma_0 Q_1 (Q_1^T \Sigma_0 Q_1)^{-1} G_0^{-1} y + r_0 - \Sigma_0 Q_1 (Q_1^T \Sigma_0 Q_1)^{-1} Q_1^T r_0 \\ &= \Sigma_0 Q_1 (Q_1^T \Sigma_0 Q_1)^{-1} G_0^{-1} y + \left(\Sigma_0 - \Sigma_0 Q_1 (Q_1^T \Sigma_0 Q_1)^{-1} Q_1^T \Sigma_0 \right) \Sigma_0^{-1} r_0 \\ &= \Sigma_0 Q_1 (Q_1^T \Sigma_0 Q_1)^{-1} G_0^{-1} y + Q_2 (Q_2^T \Sigma_0^{-1} Q_2)^{-1} Q_2^T \Sigma_0^{-1} r_0 \end{aligned}$$

值得注意, 后验协方差矩阵是奇异的, 并且即使 γ 趋于 0, 后验分布依赖于先验分布, 不能近似真实值。