1. 线性贝叶斯反问题的后验分布

后验分布满足

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

我们能把 Φ_R 重新写为

$$\begin{aligned}\Phi_R&=\frac{1}{2}\Big((y-G\theta)^T\Sigma_\eta^{-1}(y-G\theta)+(\theta-r_0)^T\Sigma_0^{-1}(\theta-r_0)\Big)\\&=\frac{1}{2}\Big(\theta^TG^T\Sigma_\eta^{-1}G\theta+\theta^T\Sigma_0^{-1}\theta+2\theta^TG^T\Sigma_\eta^{-1}y+2\theta^T\Sigma_0^{-1}r_0\Big)+const.\\&=\frac{1}{2}\Big(\theta^T\big(G^T\Sigma_\eta^{-1}G+\Sigma_0^{-1}\big)\theta+2\theta^T\big(G^T\Sigma_\eta^{-1}y+\Sigma_0^{-1}r_0\big)\Big)+const.\\&=\frac{1}{2}(\theta-m_{\rm post})^TC_{\rm post}^{-1}(\theta-m_{\rm post})+const.\end{aligned}$$

因此我们有

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

2. 超定系统后验一致性

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

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

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

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

$$P_{\rm post}\big(|\theta-\theta^\dagger|^2>M(\gamma)\gamma^2\big)\to 0 \tag{1}$$

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

$$P_{\rm post}\big(|\theta-\theta^\dagger|^2>\epsilon^2\big)\to 0 \tag{2}$$

证明

我们有后验分布

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

为了研究 $\gamma\to 0$ 的情形， 我们定义

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

对于后验分布取期望，我们有

$$\exists c>0\quad \mathrm{s.t.}\quad \mathbb{E}\big[|\theta-\theta^\dagger|^2\big]\le c\Big(\mathbb{E}\big[|\theta-m_{\rm post}|^2\big]+|m_{\rm post}-m^+|^2+|m^+-\theta^\dagger|^2\Big). \tag{3}$$

接下来，我们将放缩每一项。对于第一项

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

对于第二项， 由于

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

我们有

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

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

$$|m_{\rm post}-m^+|^2\le c\gamma^4$$

对于第三项， 我们有

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

因此，我们有对第三项的估计

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

最后我们使用马尔科夫不等式， 当$\gamma\to 0$， 我们有

$$P_{\rm post}\big(|\theta-\theta^\dagger|^2>M(\gamma)\gamma^2\big)\le\frac{\mathbb{E}[|\theta-\theta^\dagger|^2]}{M(\gamma)\gamma^2}\le\frac{c}{M(\gamma)}\to 0$$

3. 欠定系统无后验一致性

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

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

对G做正交化$G=G_0Q_1^T$，其中 $G_0\in R^{N_y\times N_y}$ 可逆，$Q_1\in R^{N_y\times N_\theta}$列向量正交。我们可以对$Q_1$进行正交补全 $Q=[Q_1;Q_2]\in R^{N_\theta\times N_\theta}$，$Q^TQ=I$

我们有

$$G=[G_0;0]Q^T=[G_0;0][Q_1;Q_2]^T=G_0Q_1^T$$

其中$G_0\in R^{N_\theta\times(N_y-N_\theta)}$是核空间$GQ_2=0$。

当$\gamma\to 0$，我们有

$$\begin{aligned}&m_{\rm post}\to m^+=\Sigma_0Q_1(Q_1^T\Sigma_0Q_1)^{-1}G_0y+Q_2(Q_2^T\Sigma_0^{-1}Q_2)^{-1}Q_2^T\Sigma_0^{-1}r_0\\&C_{\rm post}\to C^+=Q_2(Q_2^T\Sigma_0^{-1}Q_2)^{-1}Q_2^T\end{aligned}$$

证明

对于后验协方差矩阵，我们有

$$C_{\rm post}=\Sigma_0-\Sigma_0G^T(G\Sigma_0G^T+\gamma^2\Sigma_\eta)^{-1}G\Sigma_0$$

由于$G\Sigma_0G^T\in R^{N_y\times N_y}$ 可逆，当$\gamma\to 0$

$$\begin{aligned}C_{\rm post}&=\Sigma_0-\Sigma_0G^T(G\Sigma_0G^T)^{-1}G\Sigma_0\\&=\Sigma_0-\Sigma_0Q_1G_0^T(G_0Q_1^T\Sigma_0Q_1G_0^T)^{-1}G_0Q_1^T\Sigma_0\\&=\Sigma_0-\Sigma_0Q_1(Q_1^T\Sigma_0Q_1)^{-1}Q_1^T\Sigma_0\\&=Q_2(Q_2^T\Sigma_0^{-1}Q_2)^{-1}Q_2^T\end{aligned}$$

对于最后一个等式，我们使用了如下恒等式

$$I=R:=\Sigma_0Q_1(Q_1^T\Sigma_0Q_1)^{-1}Q_1^T+Q_2(Q_2^T\Sigma_0^{-1}Q_2)^{-1}Q_2^T\Sigma_0^{-1}$$

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

$$m_{\rm post}=r_0-\Sigma_0G^T(\gamma^2\Sigma_\eta+G\Sigma_0G^T)^{-1}(Gr_0-y)$$

由于$G\Sigma_0G^T\in R^{N_y\times N_y}$ 可逆，当$\gamma\to 0$

$$\begin{aligned}m_{\rm post}&=r_0+\Sigma_0G^T(G\Sigma_0G^T)^{-1}(y-Gr_0)\\&=r_0+\Sigma_0Q_1G_0^T(G_0Q_1^T\Sigma_0Q_1G_0^T)^{-1}(y-G_0Q_1^Tr_0)\\&=\Sigma_0Q_1(Q_1^T\Sigma_0Q_1)^{-1}G_0^{-1}y+r_0-\Sigma_0Q_1(Q_1^T\Sigma_0Q_1)^{-1}Q_1^Tr_0\\&=\Sigma_0Q_1(Q_1^T\Sigma_0Q_1)^{-1}G_0^{-1}y+\Big(\Sigma_0-\Sigma_0Q_1(Q_1^T\Sigma_0Q_1)^{-1}Q_1^T\Sigma_0\Big)\Sigma_0^{-1}r_0\\&=\Sigma_0Q_1(Q_1^T\Sigma_0Q_1)^{-1}G_0^{-1}y+Q_2(Q_2^T\Sigma_0^{-1}Q_2)^{-1}Q_2^T\Sigma_0^{-1}r_0\end{aligned}$$

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。