线性贝叶斯反问题

黄政宇
北京大学北京国际数学研究中心
北京大学国际机器学习研究中心

本堂课大纲

- 课程内容简介
  - 线性贝叶斯反问题
  - 超定系统
  - 欠定系统

线性贝叶斯反问题

- 贝叶斯反问题

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

- 假设

线性假设：$\mathcal{G}(\theta) = G \cdot \theta \quad G \in R^{N_y \times N_\theta}$

高斯先验分布：$\rho_{\text{prior}}(\theta) = \mathcal{N}(\theta; r_0, \Sigma_0)$

高斯噪声：$\rho_\eta(x) = \mathcal{N}(x; 0, \Sigma_\eta)$

线性贝叶斯反问题

- 似然函数

$$\rho(y|\theta) = \mathcal{N}(y; G\theta, \Sigma_\eta) \propto e^{-\Phi(\theta, y)}$$

$$\Phi(\theta, y) = \frac{1}{2} \| \Sigma_\eta^{-\frac{1}{2}} (y - G\theta) \|^2$$

- 后验分布

$$\rho_{\text{post}}(\theta; y) \propto \rho(y|\theta)\rho_{\text{prior}}(\theta) \propto e^{-\Phi_R(\theta, y)}$$

$$\rho_{\text{post}}(\theta; y) = \frac{1}{Z} e^{-\Phi_R(\theta, y)}$$

$$\Phi_R(\theta, y) = \frac{1}{2} \| \Sigma_\eta^{-\frac{1}{2}} (y - G\theta) \|^2 + \frac{1}{2} \| \Sigma_0^{-\frac{1}{2}} (\theta - r_0) \|^2$$

线性贝叶斯反问题

线性贝叶斯反问题的后验分布

对于线性贝叶斯反问题

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

它的后验分布仍然是高斯

$$\rho_{\text{post}}(\theta; y) = \mathcal{N}(\theta; m_{\text{post}}, C_{\text{post}})$$

后验方差矩阵C_{post}和期望m_{post}满足

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

线性贝叶斯反问题

Sherman-Morrison-Woodbury矩阵等式

对于可逆矩阵 $C \in R^{n\times n}, A \in R^{m\times m}$， 以及矩阵 $U \in R^{n\times m}, V \in R^{n\times m}$, 我们有等式

$$(A + U^T CV)^{-1} = A^{-1} - A^{-1}U^T(VA^{-1}U^T + C^{-1})^{-1}VA^{-1}$$

线性贝叶斯反问题

- 线性贝叶斯反问题的后验分布

$$
\begin{aligned}
C_{\text{post}} &= \left(G^T\Sigma_\eta^{-1}G + \Sigma_0^{-1}\right)^{-1} \\
&= \Sigma_0 - \Sigma_0 G^T\left(G\Sigma_0 G^T + \Sigma_\eta\right)^{-1} G\Sigma_0 \\
m_{\text{post}} &= \left(G^T\Sigma_\eta^{-1}G + \Sigma_0^{-1}\right)^{-1}\left(G^T\Sigma_\eta^{-1}y + \Sigma_0^{-1}r_0\right) \\
&= r_0 - \left(G^T\Sigma_\eta^{-1}G + \Sigma_0^{-1}\right)^{-1} G^T\Sigma_\eta^{-1}(Gr_0 - y) \\
&= r_0 - \Sigma_0 G^T\left(G\Sigma_0 G^T + \Sigma_\eta\right)^{-1}(Gr_0 - y)
\end{aligned}
$$

- 后验方差矩阵不依赖于数据 y
- 方差矩阵变小了 $C_{\text{post}} \preccurlyeq \Sigma_0$
- 后验期望和最大后验估计一致
- 牛顿优化方法

线性贝叶斯反问题

➢ 反问题的优化方法

$$\Phi_R(\theta, y) = \frac{1}{2} \| \Sigma_\eta^{-\frac{1}{2}} (y - G\theta) \|^2 + \frac{1}{2} \| \Sigma_0^{-\frac{1}{2}} (\theta - r_0) \|^2$$

- 梯度： $\nabla_\theta \Phi_R(\theta, y) = G^T \Sigma_\eta (G\theta - y) + \Sigma_0 (\theta - r_0)$
- 海瑟矩阵： $\nabla_\theta^2 \Phi_R(\theta, y) = G^T \Sigma_\eta^{-1} G + \Sigma_0^{-1}$
- 极小值： $m_* = r_0 - \left(G^T \Sigma_\eta^{-1} G + \Sigma_0^{-1}\right)^{-1} G^T \Sigma_\eta^{-1} (G r_0 - y)$

$\nabla_\theta^2 \Phi_R(r_0, y)$ $\nabla_\theta \Phi_R(r_0, y)$

➢ 特殊情况： $\Sigma_\eta = I$, $\Sigma_0 = \frac{1}{\lambda}$

$$\Phi_R(\theta, y) = \frac{1}{2} \| y - G\theta \|^2 + \frac{\lambda}{2} \| \theta - r_0 \|^2$$

正则项

线性贝叶斯反问题（超定系统）

- 超定线性系统 $G \in R^{N_y \times N_\theta}$

$$y = G\theta \qquad \text{其中 } \mathrm{Null}(G) = 0，N_y > N_\theta$$

- $0 = G\theta \qquad \Longrightarrow \qquad \theta = 0$
- $G^T \Sigma_\eta G \succ 0$
- $\exists \alpha > 0, \quad \theta^T G^T \Sigma_\eta G\theta \geq \alpha \theta^T \theta \qquad \forall \theta \in R^{N_\theta}$
- $\exists y, \qquad y = G\theta$ 无解

线性贝叶斯反问题（超定系统）

- 超定线性系统

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

其中 $\mathrm{Null}(G) = 0$，$N_y > N_\theta$

- 无先验时的后验分布 $(\Sigma_0 \to \infty)$

$$C_{\text{post}} = \left(G^T \Sigma_\eta^{-1} G\right)^{-1}$$

$$m_{\text{post}} = \left(G^T \Sigma_\eta^{-1} G\right)^{-1} G^T \Sigma_\eta^{-1} y$$

最大似然估计：

$$\hat{\theta}_{\text{MLE}} = \operatorname{argmin}_\theta \frac{1}{2} \| \Sigma_\eta^{-\frac{1}{2}} (y - G\theta) \|^2$$

$$m_{\text{post}} = \hat{\theta}_{\text{MLE}}$$

线性贝叶斯反问题（超定系统）

- 超定线性系统

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

其中 $\mathrm{Null}(G) = 0$，$N_y > N_\theta$

- 小观测噪声时的后验分布 $(\Sigma_\eta \to \gamma^2\Sigma_\eta$，$\gamma \to 0)$

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

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

$$m^+ = \left(G^T\Sigma_\eta^{-1}G\right)^{-1}G^T\Sigma_\eta^{-1}y = \mathrm{argmin}_\theta \frac{1}{2}\|\Sigma_\eta^{-\frac{1}{2}}(y - G\theta)\|^2$$

$$m_{\text{post}} \to m^+ \qquad\qquad \rho_{\text{post}}(\theta; y) \to \delta(\theta - m^+)$$

线性贝叶斯反问题（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

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

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

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

$$P_{\mathrm{post}}(|\theta-\theta^\dagger|^2>M(\gamma)\gamma^2)\to 0$$

线性贝叶斯反问题（欠定系统）

- 欠定线性系统 $G \in R^{N_y \times N_\theta}$

$$y = G\theta \qquad N_y < N_\theta$$

假设 $\mathrm{Rank}(G) = N_y$，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，$Q_1 \in R^{N_\theta \times N_y}$列向量正交。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

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

$Q_2 \in R^{N_\theta \times (N_y - N_\theta)}$是核空间（$GQ_2 = 0$）。

线性贝叶斯反问题（欠定系统）

- 欠定线性系统

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

其中$\mathrm{Rank}(G) = N_y$，$N_y < N_\theta$

- 无先验时的后验分布 $(\Sigma_0 \to \infty)$

$$C_{\text{post}} = \left(G^T \Sigma_\eta^{-1} G\right)^{-1}$$

$$m_{\text{post}} = \left(G^T \Sigma_\eta^{-1} G\right)^{-1} G^T \Sigma_\eta^{-1} y$$

$G^T \Sigma_\eta^{-1} G$ 不可逆

欠定系统的反问题一定需要先验信息

线性贝叶斯反问题（欠定系统）

无后验一致性

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

$$y = G\theta + \eta \qquad \eta \sim \mathcal{N}\left(x; 0, \gamma^2 \Sigma_\eta\right) \qquad \theta \sim \mathcal{N}(\theta; r_0, \Sigma_0)$$

我们做分解$G = [G_0\ 0]Q^T = [G_0\ 0][Q_1\ Q_2]^T$，其中$G_0 \in R^{N_y \times N_y}$可逆, $Q \in R^{N_\theta \times N_\theta}$是正交矩阵, $Q_1 \in R^{N_\theta \times N_y}$, $Q_2 \in R^{N_\theta \times (N_y - N_\theta)}$。

当$\gamma \to 0$，m_{post}和C_{post}会趋于

$$m^+ = \Sigma_0 Q_1 (Q_1^T \Sigma_0 Q_1)^{-1} G_0^{-1} y + Q_2 (Q_2^T \Sigma_0^{-1} Q_2)^{-1} Q_2^T \Sigma_0^{-1} r_0$$

$$C^+ = Q_2 (Q_2^T \Sigma_0^{-1} Q_2)^{-1} Q_2^T$$

依赖于先验信息、C^+奇异