

时间序列分析作业三

10161511309 栗嘉逸

2019 年 12 月 17 日

题目一：

$$a_t p_t \varphi_t = O(1) \quad a.s$$

$$a_t \|p_t \varphi_t\| = O(1) \quad a.s$$

证明

题目二：考虑线性动态系统

$$y(t) = \theta^T x(t) + e(t)$$

$$y(t) \in \mathbb{R}, \theta \in \mathbb{R}^d, x(t) \in \mathbb{R}$$

$$e(t) \sim N(0, \lambda(t)) \quad \text{相互独立}$$

试给出 θ 的极大似然估计

解答

我们假设在 n 个时间 t_1, t_2, \dots, t_n 处进行观测，设这些时刻的输入为 $x(t_1), x(t_2), \dots, x(t_n)$ ，观测到的输出为 $y(t_1), y(t_2), \dots, y(t_n)$

则对应这些观测值的概率密度为：

$$p\{e(t_i) = y(t_i) - x(t_i)\} = \frac{1}{\sqrt{2\pi\lambda(t_i)}} \exp\left\{-\frac{1}{2\lambda(t_i)}[y(t_i) - \theta^T x(t_i)]^2\right\}$$

我们设 $\theta = (\theta_1, \theta_2, \dots, \theta_d) \in \mathbb{R}^d$ ，则似然函数：

$$L(t_1, t_2, \dots, t_n; \theta) = \frac{1}{(2\pi)^{n/2} \prod_{i=1}^n \lambda(t_i)} \exp\left\{-\frac{1}{2} \sum_{i=1}^n [y(t_i) - \sum_{k=1}^d \theta_k x_k(t_i)]^2\right\}$$

取对数得：

$$\ln L = C - 1/2 \sum_{i=1}^n \frac{[y(t_i) - \sum_{k=1}^d \theta_k x_k(t_i)]^2}{\lambda(t_i)}$$

其中 C 为与 θ 无关的已知数。

又由：

$$\frac{\partial \ln L}{\partial \theta_l} = \sum_{i=1}^n \frac{(y(t_i) - \sum_{k=1}^d \theta_k x_k(t_i))(x_l(t_i))}{\lambda(t_i)} = 0$$

可以得到：

$$\sum_{i=1}^n \frac{y(t_i)}{\lambda(t_i)} x_l(t_i) = \sum_{k=1}^d [\sum_{i=1}^n x_k(t_i) x_l(t_i)] \theta_k$$

其中 $l = 1, 2, \dots, d$

这是一个 d 元的线性方程组

我们设 $Y = [\sum_{i=1}^n \frac{y(t_i)}{\lambda(t_i)} x_1(t_i), \sum_{i=1}^n \frac{y(t_i)}{\lambda(t_i)} x_2(t_i), \dots, \sum_{i=1}^n \frac{y(t_i)}{\lambda(t_i)} x_d(t_i)]^T$

并且 $A = [\sum_{k=1}^n x_j(t_k) x_i(t_k)]_{i,j}$

则 A 是一个 $d \times d$ 的矩阵

从而方程组可以表示为

$$Y = A\theta$$

并且一定有解（因为 Y 为 d 维向量而 $\text{rank} A \leq n$ ，且 A 是方阵），

这个解可以表达为：

$$\hat{\theta}_{MLE} = A^+ Y + (I - A^+ A) T$$

（其中 T 为任意的 $d \times d$ 矩阵）

解对应的点为对数似然函数可能的极值点，如果我们能够说明这个点是对数似然函数的最大值点，那么这个点也即是 θ 的极大似然估计。为满足这一条件，我们只需要探讨何时在该点处对数似然函数的黑塞矩阵是负定的即可。（如果黑塞矩阵恒负定，则该点必然为全局的最大值点）

考虑二阶偏导数：

$$\frac{\partial \ln L}{\partial \theta_l \partial \theta_m} = - \sum_{i=1}^n \frac{x_l(t_i) x_m(t_i)}{\lambda(t_i)}$$

于是 $g := \ln L$ 的 Hessian 矩阵为:

$$H := H(g) = - \left(\sum_{k=1}^n \frac{x_i(t_k) x_j(t_k)}{\lambda(t_k)} \right)_{i,j}$$

对于 $\forall \alpha := (\alpha_1, \dots, \alpha_d) \in \mathbb{R}^d$

$$\alpha^T H \alpha = - \sum_{i=1, j=1}^d \left(\sum_{k=1}^n \frac{x_i(t_k) x_j(t_k)}{\lambda(t_k)} \right) \alpha_i \alpha_j = - \sum_{k=1}^n \left(\sum_{i=1}^d x_i(t_k) \alpha_i / \lambda(t_k) \right)^2 \leq 0$$

当等号成立时, 对于任意的 $1 \leq k \leq n$, 我们有:

$$\sum_{i=1}^d x_i(t_k) \alpha_i / \lambda(t_k) = 0$$

我们把上面这的式子看作一个关于 α 的 d 元线性方程组, 并记

$$B = (x_i(t_j) / \lambda(t_j))$$

则当 $\text{rank} B \geq d$ 的时候, 原方程有且仅有零解, 这即说明了上面的不等式中等号成立的条件为 α 为零向量, 于是 H 是负定矩阵, 从而可以保证对数似然函数取到最大值。

综上所述, 当我们在上面定义的矩阵 B 满足条件 $\text{rank} B \geq d$ 的时候, 我们可以给出 θ 的极大似然估计 $\hat{\theta}_{MLE}$ (如前面给出)

Remark. 这即是说明了, 至少有 d 组线性不相关的输入、输出样本, θ 的极大似然估计才是一定存在的。