浅谈最小二乘法的及其应用

December 9, 2019

摘要

本文讨论了在不同背景问题下的最小二乘法的应用。

1 导言——Gauss–Markov 定理

定义 1. 若有概率空间 $(H, \mathcal{F}, \nu)$ 和其上的随机变量 $X : (H, \mathcal{F}, \nu) \to (H', \mathcal{F}')$，其中 H 和 H' 为 Hilbert 空间、$\mathcal{F}$ 和 $\mathcal{F}'$ 分别为 H 和 H' 上内积生成的拓扑结构的 Borel 集。定义如下范畴

$$\mathrm{Ob}(\mathcal{C}) = H'$$
$$a \longrightarrow b \quad \text{iff.} \quad \langle b, a\rangle = \int_{H'} \langle X(y), b\rangle \, d\nu(y)$$

若在 $\mathcal{C}$ 中存在终对象，则称其为随机变量 X 相对于概率测度 ν 的期望值，记为 $\mathrm{E}_\nu[X]$。另定义 $\mathrm{Var}_\nu(X) = \mathrm{E}_\nu[(X - \mathrm{E}_\nu(X))^2]$，若存在，为随机变量 X 相对于概率测度 ν 的方差。当概率测度可从上下文明确时，我们也会省略下标 ν。

Hilbert 空间 H' 的仿射子空间[1]M 被称为一个线性模型。我们假设 $\mathrm{E}(X) \in M$，并称 M 上的线性泛函为参数函数。注意所有的参数函数都是可估计的，即若 ϕ 为一参数函数、Q 为任意一线性投影 $H' \to M$，则有对于所有概率测度 ν

$$\int_{H'} \phi(Q(y)) \, d\nu(y) = \phi(Q(\mathrm{E}_\nu[X])) = \phi(\mathrm{E}_\nu[X])$$

而这意味着所有线性投影 ϕ 都是参数函数的无偏差估计量，并且可以看出，不存在其他的无偏差估计量。计算估计量 $\phi(Q)$ 的方差：

$$\mathrm{Var}(\phi(Q)) = (m \mapsto \phi(m) \cdot Q^*(m))(\mathrm{Var}(m \mapsto \phi(m) \cdot Q^*(m)))$$

其中 Q^* 是 Q 的 Hermitian 伴随。

定义 2. 若线性投影 $P : H' \to M$ 满足

$$P(\mathrm{Var}(f)(m)) = \mathrm{Var}(P^*)(m) \quad \forall f : M \to H', \, m \in M$$

则称 P 为相对于概率密度 ν、随机变量 X 和线性模型 M 的一般最小二乘估计量。

定理 3（Gauss-Markov）. 如上定义的 P 是所有无偏差估计量中方差最小的，且仿射子空间 $P(H)$ 和 $(\mathrm{Id} - P)(H)$ 不相关，即

$$E[\langle P(X(h)), (\mathrm{Id} - P)(X(h))\rangle] = \langle E[P(X(h))], E[(\mathrm{Id} - P)(X(h))]\rangle \quad \forall h \in H$$

在进入关于 Riemannian 流形的讨论前，我们将会描述一定条件下的随机过程中最小二乘法的使用。

2 随机过程

参考文献

[1] 又称线性流形