

【题目】 Gaussian 过程与信号估计

近几年来，一类全新的信号估计问题—Super-Level Set Estimation，逐渐吸引了人们的兴趣。寻找随机信号 $\mathbf{x}(\mathbf{t})$ 在某个空间 H 上的最大值，这一问题已经被考虑很长时间了。由于观测到的信号样本中含有噪声，因此人们往往转而考虑估计信号取值 $\mathbf{x}(\mathbf{t})$ 以一定概率大于某一给定门限值 T 的空间点 \mathbf{t} 的集合，即所谓的 Super-Level Set: $\{\mathbf{t} \in H \mid P(\mathbf{x}(\mathbf{t}) > T) > c\}$ 。很显然，这一类问题在科学和工程领域十分常见。

我们关心的估计问题有一点特殊性，即所研究的随机信号是难以观测的。换句话说，想在某个点 \mathbf{t} 上得到 $\mathbf{x}(\mathbf{t})$ ，需要付出很大的代价。这就意味着，我们不能采用简单粗暴的“遍历”方法，对空间 H 中的所有点进行观测。观测的点需要被高效利用。这样一来， $\mathbf{x}(\mathbf{t})$ 的统计结构，就自然地进入了我们的视野。其中，Gaussian 过程由于其良好的解析特性，受到了广泛的重视[1]。[2][3]对[1]进行了改进和延伸，建立了水平集估计的理论框架。[4]对空间进行了划分，然后在小空间上得到了更进一步的结果。[5]作为最近完成的工作，给出了上述估计算法的改进及其理论分析。

我们的目标非常明确，以我们在课程中学习到的 Gaussian 过程知识为基础，通过必要的背景学习，掌握基本的理论分析工具和数值算法，对上述 Super-Level Set 的估计问题进行研究。由于文献中使用到的 Gaussian 过程性质，我们基本都在课程中学习过，并不陌生，所以同学们完全有条件提出具有一定创新性的估计方法，并对其做比较充分的理论分析。我们欢迎各种改进的估计。希望同学们可以像[5]那样，对所提出的估计算法进行深入的理论分析。当然，充分的数值仿真实验也是必要的。最终，希望同学们可以通过解析和数值两种途径，对所提算法的性能进行验证，并形成报告。

【参考文献】

- [1] Active Learning for Level Set Estimation
- [2] Active Learning for Identifying Function Threshold boundaries
- [3] Truncated variance reduction
- [4] Active Area Search via Bayesian Quadrature
- [5] Robust Super-Level Set Estimation using Gaussian Processes