Why Using Truncated Estimator Technique

Roy

November 21, 2019

In this report, I briefly present the intuition behind using the truncated estimator technique in Silverman's paper [1]. Their objective is to prove that, given the data $\{(X_i, Y_i)\}_{i=1}^n$,

$$\sup_x |h_n(x) - Eh_n(x)| \to 0, \text{ a.s.} \tag{1}$$

where h_n is a kernel estimator of the underlying true function $h : \mathbb{R} \to \mathbb{R}$ using the n data points. Notice that (X_i, Y_i) are i.i.d. random vectors of a relatively general distribution: They only require that Y have bounded moments up to certain order. And it is the **truncated estimator technique** that allows this generalization, rather than just Gaussian.

The intuition behind this genius technique follows from three facts:

1. For a rather general random variable Y, it's empirical distribution should approach to it's distribution function when the size of the data increases. And the empirical distribution is a summation of i.i.d. random variables.

2. Certain summation of i.i.d. random variables can be well approximated by summation of i.i.d. Gaussian random variables [2].

3. The truncated estimator h_n^B should be close to h_n when $B \to \infty$.

Notice that both 1 and 3 can be guaranteed with very high accuracy whenever the data size increases. In 2, the probability of the estimating error greater than a threshold does not depend on the data size; however, this threshold decreases as n increases. *Therefore, all the three items above depend on the sample size, and as it increases, the approximation error is controlled with high probability.*

In Namjoon's proof: The key idea is *Markov inequality*, and the probability of the error bound does not improve when the sample size is increased. The fundamental reason is that, the *Markov inequality does not include the sample size information.* This explains why the upper-bound does not improve as the data size increases.

References

[1] Yue-pok Mack and Bernard W Silverman. Weak and strong uniform consistency of kernel regression estimates. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 61(3):405–415, 1982.

[2] G Tusnády. A remark on the approximation of the sample df in the multidimensional case. *Periodica Mathematica Hungarica*, 8(1):53–55, 1977.