

Stat 992: Lecture 13

Kernel Smoothing II.

Moo K. Chung mchung@stat.wisc.edu

February 18, 2004

Problem 4. Investigate the property of the kernel estimator $K_H * Y$ when $H \rightarrow \infty$. There are a couple of ways to prove what happens when $H \rightarrow \infty$. It will be shown later that the estimator is a solution to a heat equation and the condition $H \rightarrow \infty$ is equivalent to the steady state when we diffuse heat for infinite amount of time.

Solution. We will show it for 1D isotropic case. General case follows similarly. Note that $K_\sigma \leq \frac{1}{\sqrt{2\pi}\sigma}$. So

$$|K_\sigma * Y(x)| \leq \frac{1}{\sqrt{2\pi}\sigma} \left| \int Y(x) dx \right|.$$

As $\sigma \rightarrow \infty$, it converges to 0. So it seems the kernel smoothing is completely useless when the bandwidth is extremely large. However things change dramatically if we use truncated kernel $\tilde{K}_\sigma$. Assume Y is defined in finite domain Ω . We truncate and normalize kernel such that

$$\tilde{K}_\sigma = \frac{K_\sigma(x) \mathbf{1}_\Omega(x)}{\int_\Omega K_\sigma(x) dx} = \frac{e^{-x^2/2\sigma^2} \mathbf{1}_\Omega(x)}{\int_\Omega e^{-x^2/2\sigma^2} dx}$$

where $\mathbf{1}_\Omega(x)$ is an indicator function. Then trivially

$$\lim_{\sigma \rightarrow \infty} \tilde{K}_\sigma = \frac{\mathbf{1}_\Omega(x)}{\int_\Omega dx}.$$

Hence

$$\lim_{\sigma \rightarrow \infty} \tilde{K}_\sigma * Y(x) = \frac{\int_\Omega Y(x) dx}{\int_\Omega dx}$$

which is the average signal over Ω .

The kernel smoothing corresponds to the solution of a heat equation

$$\frac{\partial f}{\partial t} = \frac{\partial^2 f}{\partial x^2}$$

with initial condition $f(x, 0) = Y(x)$ at time $t = 2\sigma^2$. As $\sigma \rightarrow \infty, t \rightarrow \infty$. Heat diffusion after infinite amount of time corresponds to steady state $\frac{\partial f}{\partial t} = 0$ and the same conclusion can be drawn.

Problem 5.

$$\inf \mu(t) \leq \mathbb{E}\hat{\mu}(t) \leq \sup \mu(t). \quad (1)$$

Inequality (1) implies that the smoothed signal will be smaller than the maximum and larger than the minimum of the signal. Can you come up with sharper bounds than (1)? Hint: assume μ has a finite support and work on the neighborhood of t .

Solution. We will prove it for isotropic case but anisotropic case is similar. We need to assume $|\mu| < \infty$ and μ to have compact support. Let us partition $\mathbb{R}^n = \cup_{j=1}^m \Omega_j$ be a partition. Then

$$\begin{aligned} \mathbb{E}\hat{\mu}(t) &= K_\sigma * \mu(t) \\ &= \sum_{j=1}^m \int_{\Omega_j} K_\sigma(t-s) \mu(s) ds \\ &\leq \sum_{j=1}^m \epsilon_j \sup_{s \in \Omega_j} \mu(s), \end{aligned}$$

where $\epsilon_j = \int_{\Omega_j} K_\sigma(t-s) ds$ and $\sum_{j=1}^m \epsilon_j = 1$. Suppose $\arg \max_{1 \leq j \leq m} \sup_{s \in \Omega_j} \mu(s) = j_0$. We can construct partition such that only single partition Ω_{j_0} gives the maximum. So the global maximum is obtained in partition Ω_{j_0} , i.e.

$$\sup_{s \in \mathbb{R}^n} \mu(s) = \sup_{s \in \Omega_{j_0}} \mu(s).$$

Then the largest value of $\sum_{j=1}^m \epsilon_j \sup_{s \in \Omega_j} \mu(s)$ is obtained if $\epsilon_{j_0} = 1$ and other ϵ_j 's vanish which do not happen since we can always make $0 < \epsilon_{j_0} < 1$. Hence

$$\sum_{j=1}^m \epsilon_j \sup_{s \in \Omega_j} \mu(s) < \sup_{s \in \mathbb{R}^n} \mu(s).$$

Similarly we can bound by below so that

$$\inf_{s \in \mathbb{R}^n} \mu(s) < \sum_{j=1}^m \epsilon_j \sup_{s \in \Omega_j} \mu(s).$$

There is a degenerate situation this proof breaks down. Consider $\mu(t) = c < \infty$. In that case, we should have a single partition $\Omega_{j_0} = \mathbb{R}^n$ and everything collapses to constant $\inf_{s \in \mathbb{R}^n} \mu(s) = K_\sigma * \mu = \sup_{s \in \mathbb{R}^n} \mu(s) = c$.

Problem 24. Can you come up with a sharper bound than this? There is. :) Hint: For fixed t consider a partition Ω_{j_0} that contains t .