ESCOLA BRASILEIRA DE ECONOMIA E FINANÇAS - EPGE

Luciano Fabio Busatto Venturim

Econometrics 1 - Problem Set 2

Rio de Janeiro

4th Quarter - 2021

Question 1: Prove that with the ”intuitive”optimal bandwidth h for the kernel density estimator, $h^* = n^{-0.2}$, the term $\sqrt{nh}$ is of order $O(n^{0.4})$.

Solution: We have that $\frac{\sqrt{nh^*}}{n^{0.4}} = \frac{\sqrt{n^{0.8}}}{n^{0.4}} = \frac{n^{0.4}}{n^{0.4}} = 1$ for all n, hence $\sqrt{nh^*}$ is $O(n^{0.4})$.

Question 2: Consider function f of order $o(h^2)$ and g of order $o(h)$. What is the order of the product $f \cdot g$. Prove your answer.

Solution: See that $\frac{f \cdot g}{h^3} = \frac{f}{h^2} \cdot \frac{g}{h}$. Since $\lim_{h \longrightarrow 0} \frac{f}{h^2} = 0$ and $\lim_{h \longrightarrow 0} \frac{g}{h} = 0$, we have that $\lim_{h \longrightarrow 0} \frac{f}{h^2} \cdot \frac{g}{h} = 0$. Thus, $f \cdot g$ is of order $o(h^3)$.

Question 3: Prove that the globally optimal h^* is of the form $\delta(\int f''(x)^2 dx)^{-0.2} n^{-0.2}$, with δ as defined on the slides, by minimizing $MISE(h)$. (Hint: Disregard the asymptotic terms o and O. In the case of the bias, they are higher order polynomials in h, and their derivatives still go to zero as h goes to zero. In the case of the variance, the asymptotic terms go to zero for $n \longrightarrow \infty$ regardless of h).

Solution: Let $h^* = \underset{h}{\text{argmin}} \int MSE[\hat{f}(x_0)] dx_0$. Disregarding the asymptotic terms, we have that $MSE[\hat{f}(x_0)] = \frac{1}{nh} f(x_0) \int K(z)^2 dz + \left[\frac{1}{2} h^2 f''(x_0) \int z^2 K(z) dz\right]^2$. The FOC is, then,

$$\int \left(\frac{-1}{nh^{*2}} f(x_0) \int K(z)^2 dz + h^{*3} \left[f''(x_0) \int z^2 K(z) dz \right]^2 \right) dx_0 = 0$$

$$\Leftrightarrow \left(\int f(x_0) dx_0 \right) \left(\frac{1}{nh^{*2}} \int K(z)^2 dz \right) = h^{*3} \left(\int f''(x_0)^2 dx_0 \right) \left(\int z^2 K(z) dz \right)^2$$

$$\Leftrightarrow h^{*5} = n^{-1} \frac{\int K(z)^2 dz}{\left(\int z^2 K(z) dz\right)^2} \cdot \left(\int f''(x_0)^2 dx_0 \right)^{-1}$$

$$\Leftrightarrow h^* = n^{-0.2} \delta \cdot \left(\int f''(x_0)^2 dx_0 \right)^{-0.2}$$

where $\delta = \left(\frac{\int K(z)^2 dz}{\left(\int z^2 K(z) dz\right)^2} \right)^{0.2}$.

Question 4: For kernel density estimation and kernel regression alike, the optimal bandwidth choice depends on the unknown f, and replacing with a normal distribution to choose an "optimal" bandwidth makes some people unhappy. A different way of finding the optimal bandwidth is cross-validation. For the regression case, it works in the following way: Delete one observation from the sample, for example (x_1, y_1). Calculate the kernel regression estimator $\overline{m_{-1}}$ using *all other* data points except (x_1, y_1) . Then evaluate $\overline{m_{-1}}$ at x_1 and calculate the squared distance between $\overline{m_{-1}}(x_1)$ and the true y_1. Repeat this for every data point in the sample and sum over the squared errors $\overline{e_i} = y_i - \overline{m_{-i}}(x_i)$. For a given bandwidth h, the value CV is the average squared residual

$$CV(h) = \frac{1}{n}\sum_{i=1}^{n}(y_i - \overline{m_{-i}}(x_i, h))^2$$

Suppose that $y_i = m(x_i) + \epsilon_i$ and x_i i.i.d. with density f and ϵ i.i.d. with mean 0 and finite variance σ^2.Show that minimizing $E_{x,\epsilon}(CV(h))$ is equivalent to minimizing the following version of Integrated Mean Squared Error:

$$IMSE_{n-1}(h) = \int E_{x,\epsilon}[(\overline{m_{-i}}(x, h) - m(x))^2]f(x)dx$$

Solution: We have that $E_{x,\epsilon}(CV(h)) = E_{x,\epsilon}[\frac{1}{n}\sum_{i=1}^{n}(y_i - \overline{m_{-i}}(x_i, h))^2] = E_{x,\epsilon}[\frac{1}{n}\sum_{i=1}^{n}(m(x_i) + \epsilon_i - \overline{m_{-i}}(x_i, h))^2]$. Since the sample is iid, $E_{x,\epsilon}(CV(h)) = E_{x,\epsilon}[(m(x_i) + \epsilon_i - \overline{m_{-i}}(x_i, h))^2] = E_{x,\epsilon}[(m(x_i) - \overline{m_{-i}}(x_i, h))^2] + 2E_{x,\epsilon}[\epsilon_i(m(x_i) - \overline{m_{-i}}(x_i, h))] + E_{x,\epsilon}[\epsilon_i^2] = E_{x,\epsilon}[(m(x_i) - \overline{m_{-i}}(x_i, h))^2] + \sigma^2 = \int E_{x,\epsilon}[(\overline{m_{-i}}(x, h) - m(x))^2]f(x)dx + \sigma^2$, where the last two equalities come from the fact that ϵ_i is independent from $\{x_1, \ldots, x_n\}$ and from the Law of Iterated Expectations. Since σ^2 does not depend on h, the minimization of the two problems are equivalent.

Question 5: Take a break, have some cake.

Solution: Had it!