Bayesian Variance-Stabilizing Bandwidth Selection for a Kernel Density Estimator Using a Conjugate Prior

Kiheiji NISHIDA

Lecturer, General Education Center, Hyogo University of Health Sciences.
1-3-6, Minatojima, Chuo-ku, Kobe, Hyogo, JAPAN.
kiheiji.nishida(at)gmail.com

A Kernel-type density or regression estimator does not produce constant estimator variance over domains. To correct this problem, Nishida and Kanazawa (2015) and Nishida (2017) proposed a variance-stabilizing (VS) bandwidth for a local linear regression estimator. In this presentation, we show that variance-stabilization is possible using the bayesian approach with a conjugate prior in the case of the kernel density estimator.

Let $X_1, ..., X_n$ be the i.i.d. observations from an unknown density $f_X(\cdot)$ to be estimated. The kernel density estimator is defined by

$$\widehat{f_h}(x)=\frac{1}{nh}\sum_{i=1}^{n} K\left(\frac{X_i - x}{h}\right), \tag{1}$$

where $K(\cdot)$ is a nonnegative symmetric kernel function and h is a bandwidth satisfying $h \to 0$ and $nh \to \infty$ as $n \to 0$. To introduce the bayesian approach, we denote $\pi(h)$ to be the pdf over all possible values of h and the posterior distribution of h given $X_1, ..., X_n$ is defined as

$$\widehat{\pi}(h|X_1, ..., X_n)=\frac{\widehat{f_h}(x)\pi(h)}{\int \widehat{f_h}(x)\pi(h)dh}. \tag{2}$$

The Bayes estimator for the local variable bandwidth $h(x)$ in (1) is the posterior mean of (2) at x given by $h^*(x) = \int h\widehat{\pi}(h|X_1, ..., X_n, x)dh$ if the squared loss function is employed. One possible approach to calculate the posterior mean of (2) is by assuming conjugacy in the distribution of h. Gangopadhyay and Cheung (2002) supposed that the kernel function is Gaussian and the prior of h is given by the inverse gamma distribution written as

$$\tau(h)=\frac{2}{\Gamma(\alpha)\beta^{\alpha}}\frac{1}{h^{2\alpha+1}}\exp\left(-\frac{1}{\beta h^2}\right), \quad \alpha > 0,\ \beta > 0,\ h > 0. \tag{3}$$

They reported the closed form of the posterior mean of $h(x)$ as,

$$h^*(x) = \frac{\Gamma(\alpha)}{\sqrt{2\beta}\Gamma(\alpha+\frac{1}{2})}\frac{\sum_{i=1}^{n}\left[\frac{1}{\beta(X_i-x)^2+2}\right]^{\alpha}}{\sum_{i=1}^{n}\left[\frac{1}{\beta(X_i-x)^2+2}\right]^{\alpha+\frac{1}{2}}}. \tag{4}$$

To estimate the VS bandwidth for (1), the hyper parameters α and β in (3) should be determined to satisfy $h^*(x) = f_X(x)C$ at every value of x, where C is a positive constant, simultaneously satisfying the condition $\alpha > 0$ and $\beta \to \infty$, $n\beta^{-\frac{1}{2}} \to \infty$ as $n \to \infty$ in Kulasekera and Gallagher (2010). In the upcoming presentation, we will show results about the hyperparameters to be determined and also provide simulation. Moreover, we will present ideas of other possible bayesian VS bandwidth estimation procedures.

References

[1] Gangopadhyay,A., and Cheung,K.N.(2002). "Bayesian Approach to the Choice of Smoothing Parameter in Kernel Density Estimation", *J. Nonparametric Statistics*, **14**, 6, pp.655-664.

[2] Kulasekera,Q.K.B., and Gallagher,C.(2010). "Local Adaptive Smoothing in Kernel Regression Estimation", *Statistics and Probability Letters*, **80**, pp.540-547.

[3] Nishida,K.. and Kanazawa,Y. (2015). "On Variance-Stabilizing Multivariate Nonparametric Regression Estimation", *Communications in Statistics–Theory and Methods*, **44**, pp.2151-2175.

[4] Nishida,K.(2017). "Skewing Methods for Variance-Stabilizing Local Linear Regression Estimation", arXiv:1704.04356.