DIPARTIMENTO DI METODI E MODELLI
PER L'ECONOMIA IL TERRITORIO E LA FINANZA
MEMOTEF

SAPIENZA
UNIVERSITÀ DI ROMA

NEW METHODS FOR PARAMETRIC AND NONPARAMETRIC HIGH ORDER KERNEL DENSITY ESTIMATION

Sergio Pezzulli[1]

ABSTRACT

The Kernel Density Estimation (KDE) method is seen here as the first step of the Expectation Maximization (EM) algorithm for estimating the density of a latent variable when the initial guess is the uniform distribution. The properties of the first EM step are then investigated for different choices of the starting density. When the KDE itself is chosen the asymptotic bias of the EM update has the opposite value of KDE while the variance order is maintained. Thus, the average of the EM update with the KDE reduces the best achievable mean integrated square error from $n^{-4/5}$to $n^{-8/9}$. Another estimator that achieves higher order efficiency (HOE) is directly obtained by the EM update when the initial guess is the square root of the KDE. Moreover, both the average type and the square-root type of HOE estimators have a semiparametric twin estimator, obtained by adopting a parametric estimate in place of the initial guess and, of course, they achieve HOE when the parametric family is exact. These four HOE methods are then tested on simulations and compared to another pair of estimators known in literature which can be seen as a third way to bias-correct the EM update.

Classification JEL: *C10, C13, C14.*
Keywords: *EM, KDE, MISE, Data Analysis.*

Working Paper n° 163
Dicembre 2019

[1] Dipartimento di Metodi e Modelli per l'Economia il Territorio e la Finanza. MEMOTEF. Facoltà di Economia. SAPIENZA – Università di Roma. Via del Castro Laurenziano, 9 – 00161 ROMA. (email: Sergio.Pezzulli@uniroma1.it).