Journal of Multivariate Analysis 143 (2016) 165–184

Contents lists available at ScienceDirect

Journal of Multivariate Analysis

journal homepage: www.elsevier.com/locate/jmva

On the weak convergence and Central Limit Theorem of blurring and nonblurring processes with application to robust location estimation

Ting-Li Chen [a,*], Hironori Fujisawa [b], Su-Yun Huang [a], Chii-Ruey Hwang [c]

[a] *Institute of Statistical Science, Academia Sinica, Taiwan*
[b] *Institute of Statistical Mathematics, Japan*
[c] *Institute of Mathematics, Academia Sinica, Taiwan*

ARTICLE INFO

Article history:
Received 27 January 2015
Available online 28 September 2015

AMS subject classifications:
60F05
62F35
62H12

Keywords:
Weak convergence
Central Limit Theorem
Blurring process
Robust estimation

ABSTRACT

This article studies the weak convergence and associated Central Limit Theorem for blurring and nonblurring processes. Then, they are applied to the estimation of location parameter. Simulation studies show that the location estimation based on the convergence point of blurring process is more robust and often more efficient than that of nonblurring process.

1. Introduction

In this article we consider two types of processes arisen from mean-shift algorithms [7,15]. We will use notation of triangular arrays [2] for the processes to indicate their dependence on the sample size n. Starting with n points $\{x_{i,n}\}_{i=1}^{n}$ as initials, the nonblurring type process is given by

$$x_{i,n}^{[t+1]} = \frac{\int x\, w(x - x_{i,n}^{[t]})dF_n(x)}{\int w(x - x_{i,n}^{[t]})dF_n(x)} = \sum_{j=1}^{n} \frac{w(x_{j,n} - x_{i,n}^{[t]})}{\sum_{\ell=1}^{n} w(x_{\ell,n} - x_{i,n}^{[t]})} x_{j,n}, \tag{1}$$

where w is a symmetric weight function, $x_{i,n}^{[0]} = x_{i,n}$ and F_n is the empirical distribution function based on the initial points $\{x_{i,n}\}_{i=1}^{n}$. This process (1) consists of n simultaneous updating paths, wherein each path starts from one initial. Another type of updating process, called the blurring type, is considered by replacing F_n with the iteratively updated empirical distribution

* Corresponding author.
E-mail address: tlchen@stat.sinica.edu.tw (T.-L. Chen).

http://dx.doi.org/10.1016/j.jmva.2015.09.009

$F_n^{(t)}$ based on updated points $\{x_{i,n}^{(t)}\}_{i=1}^n$, in addition to the above idea of weighted scores for updating:

$$x_{i,n}^{(t+1)} = \frac{\int x\, w(x - x_{i,n}^{(t)}) dF_n^{(t)}(x)}{\int w(x - x_{i,n}^{(t)}) dF_n^{(t)}(x)} = \sum_{j=1}^n \frac{w(x_{j,n}^{(t)} - x_{i,n}^{(t)})}{\sum_{\ell=1}^n w(x_{\ell,n}^{(t)} - x_{i,n}^{(t)})} x_{j,n}^{(t)}, \tag{2}$$

where $x_{i,n}^{(0)} = x_{i,n}$. Same as the nonblurring process, the blurring process (2) starts with n initials $\{x_{i,n}\}_{i=1}^n$, and then it goes through a simultaneous updating at each iteration. The key difference from the nonblurring process is that, this process (2) takes weighted average according to the updated empirical distribution $F_n^{(t)}$, while the nonblurring process takes weighted average with respect to the initial empirical distribution F_n. That is, at each iteration in the blurring process, not just the weighted centers are updated from $\{x_{i,n}^{(t)}\}_{i=1}^n$ to $\{x_{i,n}^{(t+1)}\}_{i=1}^n$, the empirical distribution is also updated from $F_n^{(t)}$ to $F_n^{(t+1)}$.

The blurring process was developed and named SUP (self-updating process, [9]) and was recently applied to cryo-em image clustering [8]. It is also known as the blurring type mean-shift algorithm [4,5,7,10]. Blurring mean-shift can be viewed as a homogeneous self-updating process. Algorithm convergence and location estimation consistency of the blurring and nonblurring processes were discussed in Cheng [7], Comaniciu and Meer [10], Li et al. [20], Chen [6], and Ghassabeh [16]. In this article, we study their weak convergence and associated Central Limit Theorem. Due to the complicated dependent structure of random variables in $\{x_{i,n}^{[t]}\}_{i=1}^n$ and $\{x_{i,n}^{(t)}\}_{i=1}^n$, the study of their asymptotic behavior becomes challenging.

The convergence point of the blurring and nonblurring process can be used for location estimation, which is one of the most basic and commonly used tasks in statistical analysis as well as in computer vision. It is well-known that the sample mean is not a robust location estimator and it is sensitive to outliers and data contamination. To reduce the influence from deviant data, there is a wide class of robust M-estimators in statistics literature using weighted scores [17,18,24]. Consider a weighted score equation for the mean μ:

$$\sum_{i=1}^n w(x_{i,n} - \mu)(x_{i,n} - \mu) = 0. \tag{3}$$

The weighted mean that satisfies the estimating Eq. (3) can be shown to take the following form

$$\mu = \frac{\sum_{i=1}^n x_{i,n} w(x_{i,n} - \mu)}{\sum_{i=1}^n w(x_{i,n} - \mu)} = \frac{\int x w(x-\mu) dF_n(x)}{\int w(x-\mu) dF_n(x)}. \tag{4}$$

This estimator (4) can be obtained by the fixed-point iteration algorithm at convergence [3,22], where the iterative update is given by

$$\mu^{[t+1]} = \frac{\int x w(x - \mu^{[t]}) dF_n(x)}{\int w(x - \mu^{[t]}) dF_n(x)}, \quad t = 0, 1, 2, \dots. \tag{5}$$

As the sample size $n \to \infty$, the empirical distribution converges to the underlying distribution, which is assumed to be symmetric and unimodal. Then, it is easy to check that $\mu^{[t]}$ moves toward μ in each update, and hence converges. However, $\mu^{[t]}$ may not converge, as $t \to \infty$, to the true mean μ for any initial $\mu^{[0]}$, unless the initial is close enough to the target μ. Considering the updating process (5) with simultaneous starting initials $\{x_{i,n}\}_{i=1}^n$, it then leads to the nonblurring process given in (1). By replacing F_n with $F_n^{(t)}$, we have the blurring process given in (2). The iterative updating process based on either (1) or (5) has been adopted for robust mean estimation ([13,14,21,25], among others), and robust clustering [23]. It is also known as the nonblurring type mean-shift algorithm. On the other hand, robust estimation based on blurring approach is rather rare in the literature. Here we strongly recommend it as an alternative choice. From our simulation studies in Section 3, the blurring type algorithm is often more robust with smaller mean square error. Thus, the blurring type algorithm deserves more attention and further exploration.

The contribution of this article is twofold. First, we derive theoretical properties of the blurring and nonblurring processes including their weak convergence to a Brownian bridge-like process and associated Central Limit Theorem. These theoretical results are presented in Section 2, with all technical proofs being placed in the Appendix. Second, we apply the derived Central Limit Theorem to location estimation. Simulation studies comparing location estimation based on using blurring and nonblurring processes are presented in Section 3. Our simulation results suggest that the blurring type algorithm is often more robust than the existent nonblurring type algorithm for robust M-estimation.

2. Main results

Let $\{x_{i,n} \in \mathbb{R} : i = 1, \dots, n\}$, $n \in \mathbb{N}$, be a triangular array of random variables. In this section, we establish some theoretical properties of blurring and nonblurring processes induced by these random variables. All the Lemmas, Theorems and a Corollary in this section are under the assumption of the following conditions C1–C3.