

Parameter Estimation for Gaussian Mixture Processes based on Expectation-Maximization Method

Xue Xia, Xuebo Zhang^a and Xiaohui Chen

Science & Technology on Underwater Acoustic Antagonizing Laboratory, Zhanjiang 524022, China;

^axuebo_zhang@sina.cn

Keywords: Gaussian mixture processes, expectation-maximization, parameters estimation.

Abstract. Expectation-Maximization (EM) iteration is one of the most efficient algorithms for parameter estimation for Gaussian mixture model, which is a characteristic probability density function model for non-Gaussian processes. In general, EM iteration for multi-dimensional Gaussian mixture is too complicated to realize in practice. Fortunately, for fitting of the background's probability density function in active detection, the single dimensional Gaussian mixture is adequate. Therefore, EM iteration can be simplified efficiently. In view of active detection, followed with descriptions of single-dimensional Gaussian mixture model and its parameter estimation problem, a practicable simplified EM iteration is derived. Initialization and order determination is important in EM iteration. Schemes for initialization and order determining are proposed for high calculating speed, high estimation accuracy, and for the compromise of the two cases. Finally, a numerical simulation is given.

1. Introduction

Gaussian mixture probability density model (GMM) has the advantages, such as structure simplicity, parameter fewness and high-performance of fitting. It's an intensively used typical non-Gaussian probability density model in the signal processing field of image, medical treatment, phone, communication and so on. As non-Gaussian signal processing is getting more and more important in modern signal processing domain, GMM modeling (esp. estimation of model parameters) is gaining more and more extensive attention.

Zabin [1] applies EM iterative algorithm to model parameters estimation of Middleton's Class A model. Seeing the similarity between GMM and Class A model, it offers possibility of applying EM iterative algorithm to GMM model parameters estimation. Wang [2] establishes a GMM model which is fit for signal detection, and applies EM iterative algorithm to parameter estimation of this GMM model, and show the result of stimulation and experiment data to verify it. However, in practical application, the method of setting initial value used in [2] doesn't always work, and the validity of estimation of parameter order using experience still need to be verified. Liu studied on colored non-Gaussian process simulation [3] and compared two methods of mixture Gaussian parameter estimation, giving a brief sketch of advantages and shortages of these two methods [4]. Verbeek proposed an improved Greed-EM algorithm for image processing [5], and Wei simplified it for acoustic reverberation GM stimulation successfully [6].

This article affirms the vilification of using EM algorithm to do GMM parameter estimation, and discusses initial value setting and order estimation problems, and draws a practical conclusion, and show the result of stimulation data to verify it.

2. GMM

In a general way, probability density function (PDF) of GMM can be formulated as equation (1),

$$f(x_n) = \sum_{m=1}^M \varepsilon_m \cdot f_m(x_n) \quad (1)$$

Here, M is the model order. ε_m is the mixture parameter, satisfying Eq. (2). f_m is the probability density function of the m th Gaussian component, shown as Eq. (3). ε_m , σ_m^2 are its mean value and variance respectively.

$$\sum_{m=1}^M \varepsilon_m = 1 \quad (2)$$

$$f_m(x_n) = \frac{1}{\sqrt{2\pi\sigma_m^2}} \exp\left[-\frac{(x_n - \mu_m)^2}{2\sigma_m^2}\right] \quad (3)$$

It can be seen that PDF of GMM is in series form. It matches any probability density no matter in narrow-band or broadband, single-bell-shaped or multi-bell-shaped.

3. EM Method

The most common and direct way of estimating parameters of PDF from sample data is maximum likelihood estimation (MLE), i.e. gaining the estimated value $\theta_{ML} = \arg \max_{\theta} \{f(x|\theta)\}$. However it's difficult to realize in reality [7]. In practice, approximations are used to achieve it. Expectation Maximization (EM) estimation is such kind of variant MLE. It changes maximizing likelihood into maximizing expectation. Thus, it keeps the estimation accuracy and improves the reliability of the estimation algorithm.

The logarithmic likelihood function of complete data u can be expressed as:

$$l_c(\theta) = \ln \prod_{n=1}^N f(u_n | \theta) = \ln \prod_{n=1}^N \prod_{m=1}^M [f(x_n | m, \theta) \varepsilon_m]^{z_{nm}} = \sum_{n=1}^N \sum_{m=1}^M [z_{nm} \ln f(x_n | m, \theta) + z_{nm} \ln \varepsilon_m] \quad (4)$$

In Eq. (4), $f(m) = \varepsilon_m$ is the PDF of data coming from the m th Gaussian source. z_{nm} is the m th element of the n th sample data's indicating array. If the sample data is from the m th Gaussian source, $z_{nm} = 1$, otherwise $z_{nm} = 0$. Likelihood function $f(x_n | m, \theta)$ is Gaussian PDF, expressed as Eq. (3)

For the likelihood function shown in Eq. (4), z_{nm} is not known, so it's not easy to solve the extreme value. However, the mean value $\langle z_{nm} \rangle$ can be found out, it's practicable to count the mean value of the posterior probability of z in both sides of Eq. (4), as shown in Eq. (5).

$$\langle l_c(\theta) \rangle = \sum_{n=1}^N \sum_{m=1}^M [\langle z_{nm} \rangle \ln f(x_n | m, \theta) + \langle z_{nm} \rangle \ln \varepsilon_m] \quad (5)$$

3.1 Expectation Step

The typical EM algorithm contains two steps. One is Expectation-Step (E step), which does the expecting work. The other is Maximization-step (M step), which does the maximizing work. E step is to gain $\langle z_{nm} \rangle$, the expectation of z_{nm} . While the expectation of z_{nm} in regard to $f(x|m, \theta)$ is the posterior probability density of sample data coming from the m th source, as follow:

$$\langle z_{nm} \rangle = f(x_n | m, \theta) = \frac{f(x_n | m, \theta) \varepsilon_m}{\sum_{j=1}^M f(x_n | j, \theta) \varepsilon_j} = \frac{f(x_n | m, \theta) \varepsilon_m}{\sum_{j=1}^M f(x_n | j, \theta) \varepsilon_j} \quad (6)$$

3.2 Maximization Step

Substitute (3), (6) into (5), and solve Eq. (5) of θ , i.e. solve Eq. (7), then we can get the estimation value, shown as (8).

$$\partial \langle l_c(\theta) \rangle / \partial \mu_m = 0, \quad \partial \langle l_c(\theta) \rangle / \partial \sigma_m^2 = 0, \quad \partial \langle l_c(\theta) \rangle / \partial \varepsilon_m = 0 \quad (7)$$

$$\hat{\mu}_m = \frac{\sum_{n=1}^N (\langle z_{nm} \rangle x_n)}{\sum_{n=1}^N (\langle z_{nm} \rangle)}, \quad \sigma_m^2 = \frac{\sum_{n=1}^N [\langle z_{nm} \rangle (x_n - \mu_m)^2]}{\sum_{n=1}^N (\langle z_{nm} \rangle)}, \quad \varepsilon_m = \frac{\sum_{n=1}^N (\langle z_{nm} \rangle)}{N} \quad (8)$$

4. Simulated Experiment

4.1 Random Noise

Table 1 GMM Parameters

μ_1	μ_2	μ_3	σ_1^2	σ_2^2	σ_3^2	ε_1	ε_2	ε_3
2.8	0.8	1.3	0.4	0.1	0.2	0.01	0.19	0.8

Using the setting values in Table 1, we generate waveform of the third order GMM series as shown in Fig.1 (a). Fig. 1(b) shows the histogram of the generated noise, which indicates that the noise model order $2 \leq M \leq 4$. Based on the presented method, the parameters are estimated in the next section.

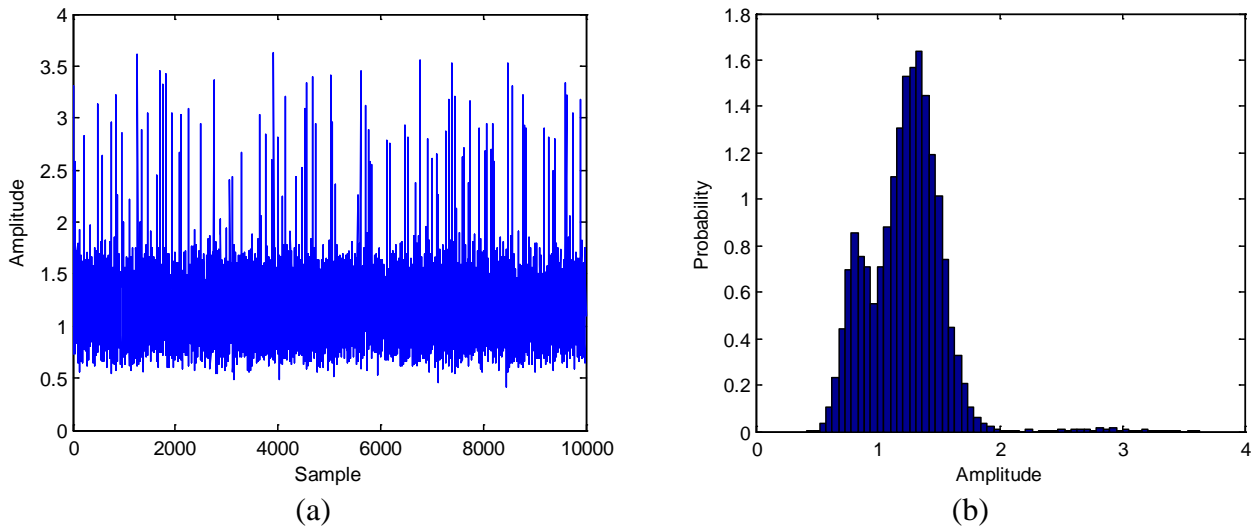


Fig. 1 The 3rd order GMM noise waveform. (a) The generated noise. (b) The histogram of the noise

4.2 Parameter Estimation

Here, we use $M=3$, and random initial value setting to start the EM algorithm.

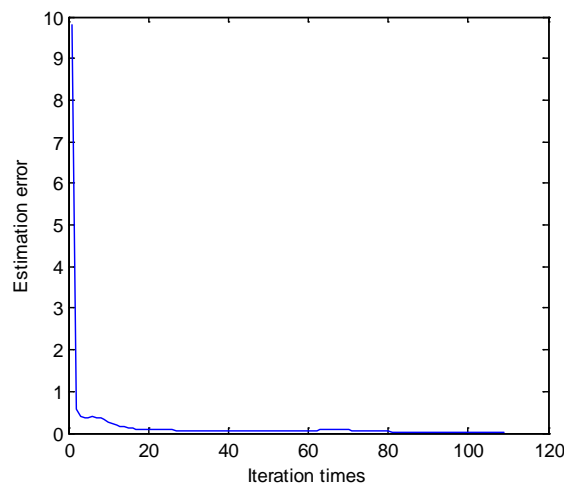


Fig.2 Estimated error

As shown in Fig.2, Estimated Error sharply decreased at the beginning 2nd or 3rd times, and then slowly getting near to the threshold.

Fig.3 (a) - Fig.3 (c) show the iteration processes of estimated parameters respectively. It can be seen clearly that parameters get close to the setting values as iteration goes on. Also have situations that the estimated value goes away from the setting value but later comes back.

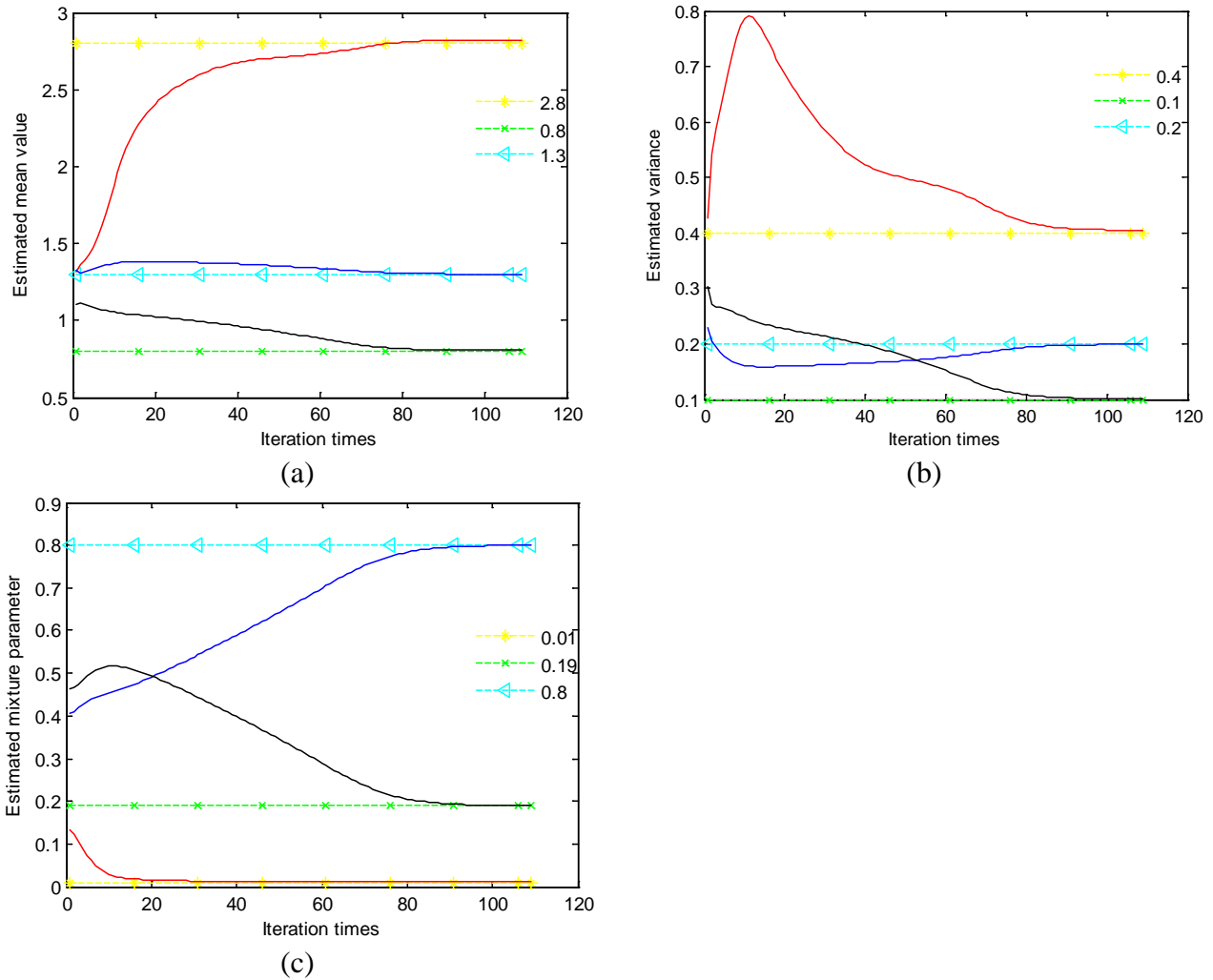


Fig.3 Iteration of parameter estimation with the presented method. (a) Mean values. (b) Variance. (c) Mixture parameter

In Fig.4, theoretical PDF means the curve is generated using the setting parameters, and statistical PDF means the curve is generated using the statistical information of data, and the GMM PDF means the curve is generated using the estimated parameters. We can see that, these three curves matching each other well.

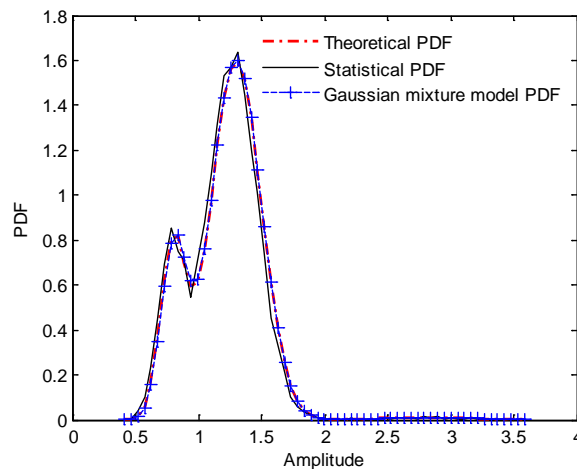


Fig. 4 PDF curves ($M=3$)

4.3 the Influence of Model Order

Fig. 5 shows the PDF when the model order $M=2$, and $M=4$.

As shown in Fig.5 (a) and Fig.5 (b), we can see that, if M is smaller than setting value, it would poorly matching the theoretical PDF curve. However, bigger than setting value doesn't make it any

better. In general, M should be equal or bigger than 2, but not necessary to be bigger than twice the PDF line peak numbers.

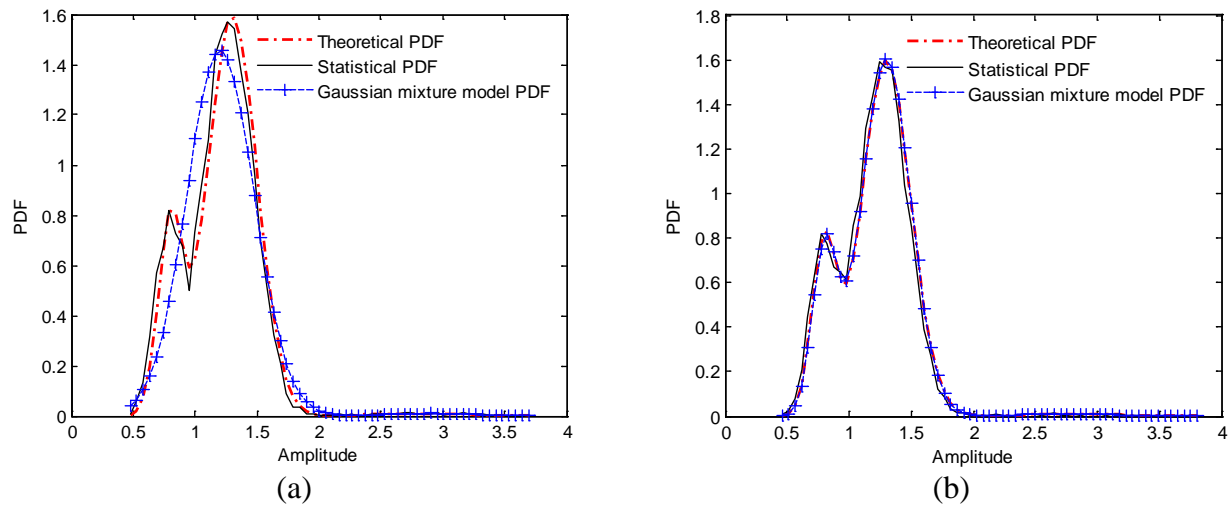


Fig. 5 PDF curves. (a) $M=2$. (b) $M=4$.

5. Conclusions

As we can see, the processing results of the presented method show the consistency between the setting values and the estimated values. It means that EM algorithm has a good performance in estimating GMM parameters. And repeated experiments on this segment of data get the approximate results. Value of M has a high impact on estimating result. We can use such strategy to determine M , that observe the shape of histogram to count the main peak numbers, M should be equal or bigger than it, but not necessary to be bigger than twice the peak numbers.

Acknowledgment

This work was financially supported by Key Laboratory Foundation under grant 9140C290401150C29132, and the National Nature Science Foundation of China under grant 61601473.

References

- [1] S. M. Zabin, H. V. Poor, Efficient estimation of Class A noise parameters via the EM algorithm, *IEEE Transactions on Information Theory*, 37(1991)60-72.
- [2] P. B. Wang, Z. M. Cai, W. S. Liu, EM estimation of PDF parameters for Gaussian mixture processes, *Technical Acoustics*, 126(2007)499-502.
- [3] W. S. Liu, P. B. Wang, X. F. Gu, et al, A simulation approach to colored non-Gaussian process, *Technical Acoustics*, 32(2013)228-232.
- [4] W. S. Liu, P. B. Wang, X. F. Gu, Comparison of two EM algorithms for Gaussian mixture parameter estimation, *Technical Acoustics*, 33(2014)539-543.
- [5] J. J. Verbeek, N. Valassis, B. Krose, Efficient greedy learning of gaussian mixture models [R], The Netherlands: Computer Science Institute, University of Amsterdam, 2001.
- [6] H. K. Wei, P. B. Wang, Z. M. Cai, Study of reverberation for gaussian mixture model, *Technical acoustics*, 26(2007)514-518.
- [7] M. V. Shawn, M. O. James, T. L. Jeffrey, V. O. Alan, Parameters estimation for auto-regressive Gaussian-mixture processes: the EMAX algorithm, *IEEE Transactions on Signal Processing*, 46(1998)2744-2756.