Study on the Convergence Behavior of Expectation Maximization Algorithm for Laplace Mixture Model

Zakiah Ibrahim KALANTAN (zkalanten@kau.edu.sa)
Faculty of Science, King Abdulaziz University, Saidi Arabia

MSc candidate Faten ALREWELY (fmalrawli@ju.edu.sa)
Faculty of Science, King Abdulaziz University, Saudi Arabia and Faculty of Science, Al Jouf University, Saudi Arabia

Kamarul Ariffin MANSOR[1] (ariff118@uitm.edu.my)
Faculty of Computer Science and Mathematics, Universiti Teknologi MARA, Kedah, Malaysia

ABSTRACT

Laplace mixture model is widely used in lifetime applications. The estimation of model parameters is required to analyze the data. In this paper, the expectation maximization algorithm is used to obtain the estimates of parameters. The algorithm is a widely applicable approach to the iterative computation of the maximum likelihood estimates. However, even though the algorithm is useful for more than two components, we discuss it for a two components Laplace mixture model for simplicity. The behavior of the algorithm is explained with mathematical proofs. We also study the parameter estimates with respect to various sample sizes. The implementation of the expectation maximization algorithm is made via functions written in R script. The performance of the algorithm is guaranteed the convergent to a local maximum of the data log-likelihood model as a function of the model parameters. In addition, the results shown that the estimated parameters are closed to the real parameter values when the sample size is large.

Keywords: *Laplace mixture model, Expectation-Maximization algorithm, Simulation case, R software.*

JEL Classification: *C13, C63*

1. Corresponding Author

1. INTRODUCTION

Mixture models are widely used in many applications in statistical analysis and machine learning such as modelling, dimension reduction, data mining, classification, and survival analysis (Dong et al., 2015). Moreover, mixture distributions are often used as outlier models in some methods as generalizes linear regression (Olive, 2017). Laplace mixture can be applied in many different fields. The Laplace mixture and the Gaussian mixture have been applied to air navigation and geophysics of the well (Ali and Nadarajah, 2007). The Laplace distribution aids in the modelling of symmetric data that is associated with long tails. The distributions further arise directly in those cases that a random variable takes place as the difference of variables that have exponential distributions in the same scale (Huang, 2016). It is used for its ability to model the errors, providing motivation when it comes to the use of regression, whereby parameter estimates are based on the minimizing the sum of absolute values associated with the residuals instead of the least squares. When it comes to the theoretical side, the justification of the Laplace distribution as provided follows that the classical central theorem establishes that the distribution of a large sum S of independent as well as identical distribution with finite variance can be approximated by a normal distribution (Reed, 2006). The results are commonly used in the justification of the use of normal distribution for random variables. An additional significance associated with the Laplace distributions is the fact that a Gaussian can generate a Laplace with variance following an exponential distribution. The Laplace distribution, in this case, is used in the modelling of position errors that are observed in the large navigation systems (Kurtz and Song, 2013). Shenoy and Gorinevsky (2014) estimated parameters by using the EM algorithm for a mixture of asymmetric Laplace and Gaussian distributions. In 2015, Lu et al. (2015) estimated the MAP is constructed using Laplace mixture distribution with zero-mean and Gaussian distribution prior distributions, respectively. After obtaining the parameter estimates using the expectation-maximization algorithm.

The EM algorithm is an extensively applicable approach to the iterative computation of the maximum likelihood (ML) estimates that are essential in an assortment of incomplete data problems (Sheng and Hu, 2005). It is additionally applied in cases when optimizing the probability function is analytically intractable but can be simplified through the assumptions of the existence of as well as valued for the additional but missing or hidden parameters (Shenoy and Gorinevsky, 2014). The EM algorithm is a general-purpose method that offers attractive attributes. One of these is the fact that it is the most commonly used estimation strategy in the frequentist framework in addition to the fact that it is relevant in the Bayesian framework (Enders, 2001). In most cases, Bayesian solutions are similar to the penalized likelihood estimates, with the maximum likelihood estimation technique being a ubiquitous technique used extensively in every aspect involving statistical techniques. Recently, Barazandeh and Razaviyayn (2018) estimated the parameters of a two components of Laplace mixture model in the case of equally weighted mixtures and known scale parameters. They mentioned that for equally weighted Gaussian mixtures with > 2 components there are multiple local optima.

The aim of this paper is to study simulation for the two component Laplace mixture model and estimate its six parameters by using simulation EM algorithm, the probability density function for the two-component Laplace mixture we want to study on is the following (Kalantan and Alrewely, 2019):

$$f(x|\ \underline{\theta}) = \sum_{k=1}^{K=2} \alpha_k\ Laplace(x;\ \mu_k, \lambda_k) = \frac{\alpha_1}{2\lambda_1}\ exp\left(-\frac{|x-\mu_1|}{\lambda_1}\right) + \frac{\alpha_2}{2\lambda_2}\ exp\left(-\frac{|x-\mu_2|}{\lambda_2}\right) \tag{1}$$

where $-\infty < x < \infty, -\infty < \mu_{1,2} < \infty,\ \lambda_{1,2} > 0$, and $\alpha_1 + \alpha_2 = 1$

And the cumulative distribution function (CDF) of X:

$$F(x) = \begin{cases} \alpha_1\ \left[1 - \frac{1}{2} exp\left(-\frac{(x-\mu_1)}{\lambda_1}\right)\right] + \alpha_2\ \left[1 - \frac{1}{2} exp\left(-\frac{(x-\mu_2)}{\lambda_2}\right)\right], for\ x \geq \mu_{1,2} \\ \frac{1}{2}\left[\alpha_1\ exp\left(\frac{x-\mu_1}{\lambda_1}\right) + \alpha_2\ exp\left(\frac{x-\mu_2}{\lambda_2}\right)\right], \qquad for\ x < \mu_{1,2} \end{cases} \tag{2}$$

The mean and variance are:

$$\mu = E(x) = \alpha_1\mu_1 + \ \alpha_2\mu_2 \tag{3}$$

$$\sigma^2 = \ Var[X] = \alpha_1[\mu_1{}^2 + 2\lambda_1{}^2] + \alpha_2[\mu_2{}^2 + 2\lambda_2{}^2] - [\alpha_1\mu_1 + \alpha_2\mu_2]^2 \tag{4}$$

The paper is organized as follows; Section 2 presents the definition of the EM algorithm and the estimation of mixture Laplace model. Section 3 a simulation study procedure of the EM algorithm to find the estimation parameters numerically. The conclusions in Section 4.

2. EXPECTATION MAXIMIZATION (EM) ALGORITHM

2.1 The definition of EM Algorithm

The mixture density consists of linear K components of densities, and defined as

$$f(x|\ \underline{\theta}) = \sum_{k=1}^{K} \alpha_k\ f(x;\ \underline{\theta}_k)$$

where $\boldsymbol{\alpha_i}$ is known as a mixing parameter or a weighted parameter, it satisfies the two constrains; (1) a positive value, and (2) $\sum_{k=1}^{K} \boldsymbol{\alpha_k} = \mathbf{1}$. In order to estimate the mixture model parameters, its favorable use EM algorithm to obtain the latent variables, when dealing with the observations and do not know their source distribution. The posterior probability on the component is an important quantity in the mixture model, where one can think on which mixture component generated the particular data observations. Now, the likelihood for the mixture Laplace model with i.i.d. samples is

$$L = \prod_{i=1}^{n} f(x_i \mid \Theta) \tag{5}$$

$$= \prod_{i=1}^{n} \left(\sum_{k=1}^{K} \alpha_k\ f(x;\ \underline{\theta}_k) \right) \tag{6}$$

To apply the ML algorithm, we assume that $\boldsymbol{z_{ik}}$ is the indicator and takes the values 1 when the observation belongs to the specific component *k* and 0 for all others. Then, the MLE for model parameters is given as

$$\theta^* = \arg max_\theta\, E\left[\log\left(\sum_{k=1}^{K} \alpha_k\ f(x;\ \underline{\theta}_k)\right)\right] = \sum_{i=1}^{n}\sum_{k=1}^{K} \log[f(x_i;\ \theta_k)\ \alpha_k]\,\omega_{ik} \tag{7}$$

where;

$$\omega_{ik} = P(z_{ik}|x_i;\theta) = \frac{P(x_i|z_{ik};\theta_k)\,P(z_{ik};\theta_k)}{\sum_{j=1}^{K} P(x_i|z_{ij};\theta_j)\,P(z_{ij};\theta_j)} = \frac{P(x_i|z_{ik};\ \theta_k)\,\alpha_k}{\sum_{j=1}^{K} P(x_i|z_{ij};\ \theta_j)\,\alpha_j} \tag{8}$$

where $\boldsymbol{\omega_{ik}}$ is the posterior probability of each unobserved variable, i.e. the weighted membership in a given data model, $\boldsymbol{i = 1, .., n}$ **and** $\boldsymbol{k = 1, 2, \ldots, K}$

When the expectation of the indicator variable is equal to 1 in eqn [8], the component of $\boldsymbol{k^{th}}$ in the mixture is responsible for generating the $\boldsymbol{i\ th}$ observation. Now the two steps of EM are,

- *E-Step:* We calculate the posterior probabilities $\boldsymbol{\omega_{ik}}$.
- *M-Step*: We use these posteriors to update the parameters by maximization from eqn [7]. In this step, the new parameter $\boldsymbol{\theta}$ is computed and repeat the E step again, in order to find the new weight parameter $\boldsymbol{\omega_{ik}}$. Then, update the parameters $\boldsymbol{\theta}$, the step E and M are repeated until the algorithm reach the convergence to obtain the local maximum (Kotz, et al., 2012).

2.2 **The parameter estimations of Mixture Laplace Model using EM Algorithm**

In this section, we try to fit the two components Laplace mixture model to a given data as good as possible. This means searching for suitable estimates for model parameters. The EM algorithm is an iterative computation of the maximum likelihood (ML) estimates, it finds the estimates at the local optima of the data log-likelihood function using given initial values of parameters. For univariate Laplace distribution, the log likelihood is

$$\log L = \prod_{i=1}^{n} \frac{1}{2\lambda}\, exp\left(-\frac{|x-\mu|}{\lambda}\right) = -n\, log(2\lambda) - \frac{1}{\lambda}\sum_{i=1}^{n} |x_i - \mu|$$

Then, the MLE for λ is derived by

$$\frac{\partial \log L}{\partial \lambda} = -\frac{n}{\lambda} + \frac{1}{\lambda^2}\sum_{i=1}^{n} |x_i - \mu|$$

Setting above equation to zero, we have

$$MLE(\lambda) = \hat{\lambda} = \frac{1}{n}\sum_{i=1}^{n} |x_i - \hat{\mu}|$$

To get the MLE of μ, we need to minimize the total absolute deviation $\sum_{i=1}^{n} |x_i - \hat{\mu}|$, regarding to Norton (1984), we assume that $x_1, x_2, \ldots, x_n$ are ordered and the sample median is x_m such that $m = \frac{n+1}{2}$. One can observe that the deviation is continuous and is decreasing for $\mu < x_m$ and increasing when $\mu > x_m$. Therefore, the minimizer is given by $\hat{\mu} = x_m$. Then, the maximum likelihood estimator μ is

$$\hat{\mu} = median\ (x_1, x_2, \ldots \ldots, x_n) = F^{-1}(0.5) \tag{9}$$

and the maximum likelihood estimator for λ is

$$\hat{\lambda} = \frac{1}{n}\sum_{i=1}^{n} |x_i - \hat{\mu}| = average\ (|data\ points - median|) \\ = average(|data\ points - \hat{\mu}|). \tag{10}$$

Now, the complete data likelihood for two components Laplace mixture (A2-CLPM) is given from eqn [1] in eqn [6], with i.i.d. samples is

$$\boldsymbol{L = \prod_{i=1}^{n} \sum_{k=1}^{K=2} \alpha_k\, Laplace(x_i;\ \mu_k, \lambda_k)} \tag{11}$$

$$\boldsymbol{log\ L = \prod_{i=1}^{n} \log \sum_{k=1}^{K=2} \alpha_k\, Laplace(x_i;\ \mu_k, \lambda_k)} \tag{12}$$

In order to deal with the EM algorithms, we assume some initial values for the model parameters, for each observation, we will calculate the expectation eqn [12] through the E step.

Then proceed to the M step. we calculate the weighted estimators based on the expectation given in the E step.

Then by using EM algorithm, where the M-step is updating the parameter value, the updated parameter estimates for location, scale and mixing parameter are, respectively;

$$\hat{\mu}_1{}^{new} = \sum_{i=1}^{n} \omega_{i1} \; median\,(x_1, x_2, \ldots \ldots, x_n) \tag{13}$$

$$\hat{\mu}_2{}^{new} = \sum_{i=1}^{n} \omega_{i2} \; median\,(x_1, x_2, \ldots \ldots, x_n) \tag{14}$$

$$\hat{\lambda}_1{}^{new} = average\,(\,\omega_{i1}|data\ points - \hat{\mu}_1{}^{new}|) \tag{15}$$

$$\hat{\lambda}_2{}^{new} = average\,(\,\omega_{i2}|data\ points - \hat{\mu}_2{}^{new}|) \tag{16}$$

$$\hat{\alpha}_1{}^{new} = \frac{1}{n} \sum_{i=1}^{n} \omega_{i1} \tag{17}$$

$$\hat{\alpha}_2{}^{new} = \frac{1}{n} \sum_{i=1}^{n} \omega_{i2} \tag{18}$$

In addition, the value of log-likelihood is recorded for each iteration and studied its steady. The E- and the M-step are iterated until the difference between log-likelihood values of two consecutive iterations reaches a small threshold, at this point the EM algorithm has converged. As final process the weight matrix is obtained with rank (n* K) and the sum of each row in the matrix is equal to 1.

3. SIMULATION STUDY

To fit the Laplace mixture model on a given data, we use the EM algorithm to obtain the parameter estimates numerically with simulation case. In this section, the implementation of EM algorithm is discussed via function written in R using R software (R Core Team, 2019). A generalization data is carried out with different sample sizes of 50, 150, 500, 1000 and 1500. We the study different cases; equal weighted parameters versus unequal one, and equal scale parameters versus unequal parameters. The model parameters are setting to different values, as follows.

The first model, the Laplace mixture model is generated with location parameters equal $\mu_1 = -2$ and $\mu_2 = 10$, the values of scale parameters equal (1,1) while the vector of mixing parameters equals (0.5, 0.5). The second model, we generate a two components Laplace mixture model with the same previous (location and mixing) parameters, while the scale parameters are equal to 3 and 1, respectively.

The estimates of the model parameters are obtained using the EM algorithm for the two cases, and the results of fitted Laplace mixture on the first model, case two, are summarized in Table 1. The results of fitted Laplace mixture on second model are summarized in Table 2. The Histogram of the simulation data for two cases are displayed in Figure 1. It appears from Figure 1 that the scale parameters are influenced on the spread of data points even the both components have the same location and mixing parameters, this is due to the appearance of the outliers in the corresponding data. The results in Table 1 and Table 2 illustrate the bias of scale parameters, in other hand, we observe that for large sample, the estimated model parameters of weighted and location parameters are close to the real values, while the scale parameters are slightly close to the real values.

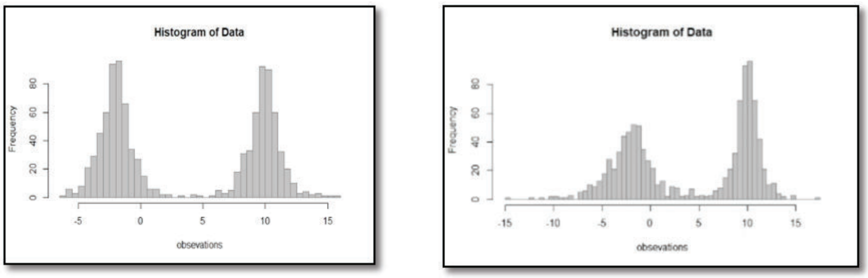
For illustration, in case one, the mixing and location parameters are equal to $\alpha_1 = \alpha_2 = 0.5 \quad \mu_1 = -2,\ \mu_2 = 10.$ In Table 1, when the sample size equals 1500, we can see that $\hat{\alpha}_1 = 0.517941, \hat{\alpha}_2 = 0.482059, \hat{\mu}_1 = -1.921763$ and $\hat{\mu}_2 = 9.981839$ are close to the given parameters $\alpha_1 = \alpha_2 = 0.5\ \mu_1 = -2,\ \mu_2 = 10$. While as we mentioned, the scale parameters $\hat{\lambda}_1 = 0.5200006$ and $\hat{\lambda}_2 = 0.5025787$ are slightly close to $\lambda_1 = \lambda_2 = 1$. To be more specifics, the errors calculated between the estimated parameters and the real parameter values decreases with the increase of the sample size (note that the real values of parameters are given within the caption of table).

Hence, to discuss the impact of mixing parameters, the third model is handled with mixing parameters equal 0.3 and 0.6, while the values of location and scale parameters, respectively, equal -2, 10, 1, and 1. The histogram of this data is presented in Figure 2 (a) and the results of fitting Laplace mixture model are summarized in Table 3. One can deduce that the parameter estimates of the density with large value of mixing parameters are closed to the true value compared with second density, this conclusion is consistent with fourth model. A two component Laplace mixture model is drawn with the same previous (mixing and location) parameters, while $\lambda_1 = 3$ and $\lambda_2 = 1$, the results are presented in Table 4.

Simulation data from Laplace mixture model;
(a) **where** $\boldsymbol{\alpha_1 = 0.5,\ \mu_1 = -2,\ \mu_2 = 10, \lambda_1 = \lambda_2 = 1}$.

(b)where $\boldsymbol{\alpha_1 = 0.5,\ \mu_1 = -2,\ \mu_2 = 10, \lambda_1 = 3,\ \lambda_2 = 1}$.

Figure 1

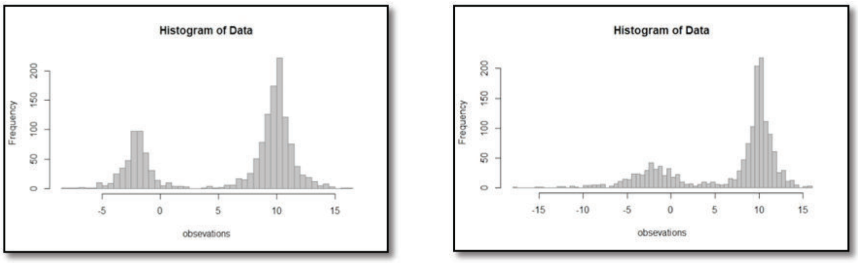


(a) (b)

Simulation data from Laplace mixture model;
(a) where $\boldsymbol{\alpha_1 = 0.3,\ \alpha_2 = 0.7,\ \mu_1 = -2,\ \mu_2 = 10,\ \lambda_1 = \lambda_2 = 1}$.

(b)where $\boldsymbol{\alpha_1 = 0.3,\ \alpha_2 = 0.7,\ \mu_1 = -2,\ \mu_2 = 10,\ \lambda_1 = 3,\ \lambda_2 = 1}$

Figure 2



(a) (b)

Simulation study (case two): Estimated parameters for Laplace mixture model using EM method where
$\boldsymbol{\alpha_1 = 0.5,\ \mu_1 = -2,\ \mu_2 = 10, \lambda_1 = \lambda_2 = 1}$.

Table 1

Size	Estimated Model Parameters					
	$\hat{\alpha}_1$	$\hat{\alpha}_2$	$\hat{\mu}_1$	$\hat{\mu}_2$	$\hat{\lambda}_1$	$\hat{\lambda}_2$
200	0.5150003	0.4849997	-1.891684	9.983323	0.4587286	0.4513023
500	0.5137642	0.4862358	-1.973138	9.920682	0.4816675	0.4943917
1000	0.5219601	0.4780399	-2.070659	9.943676	0.5078332	0.4603390
1500	0.517941	0.482059	-1.921763	9.981839	0.5200006	0.5025787

Simulation study (case two): Estimated parameters for Laplace mixture model using EM method where

$\boldsymbol{\alpha_1 = 0.5,\ \mu_1 = -2,\ \mu_2 = 10,\ \lambda_1 = 3,\ \lambda_2 = 1}$.

Table 2

Size	Estimated Model Parameters					
	$\hat{\alpha}_1$	$\hat{\alpha}_2$	$\hat{\mu}_1$	$\hat{\mu}_2$	$\hat{\lambda}_1$	$\hat{\lambda}_2$
200	0.44304	0.55696	-1.807426	9.910258	1.2852007	0.6160709
500	0.486334	0.513666	-2.158989	9.905762	1.1883873	0.4701929
1000	0.489318	0.510681	-2.217895	9.990367	1.3075566	0.4854025
1500	0.479044	0.520955	-2.004944	10.040728	1.3393325	0.5817867

Simulation study (second case): Estimated parameters for Laplace mixture model using EM method where

$\boldsymbol{\alpha_1 = 0.3,\ \alpha_2 = 0.7,\ \mu_1 = -2,\ \mu_2 = 10,\ \lambda_1 = \lambda_2 = 1.}$

Table 3

Size	Estimated Model Parameters					
	$\hat{\alpha}_1$	$\hat{\alpha}_2$	$\hat{\mu}_1$	$\hat{\mu}_2$	$\hat{\lambda}_1$	$\hat{\lambda}_2$
200	0.2943424	0.7056576	-1.976900	9.955522	0.2582667	0.7336697
500	0.3003374	0.6996626	-2.149507	9.949568	0.2516404	0.8845739
1000	0.3067515	0.6932485	-1.995943	9.883252	0.3016676	0.7183937
1500	0.3174850	0.6825150	-1.996097	10.034181	0.3007596	0.7248341

Simulation study (second case): Estimated parameters for Laplace mixture model using EM method where
$\alpha_1 = 0.3,\ \alpha_2 = 0.7,\ \mu_1 = -2,\ \mu_2 = 10,\ \lambda_1 = 3,\ \lambda_2 = 1.$

Table 4

Size	Estimated Model Parameters					
	$\hat{\alpha}_1$	$\hat{\alpha}_2$	$\hat{\mu}_1$	$\hat{\mu}_2$	$\hat{\lambda}_1$	$\hat{\lambda}_2$
200	0.2416361	0.7583639	0.800327	9.408884	3.883403	1.161198
500	0.2266359	0.7733641	1.987543	9.520705	3.058183	0.9345327
1000	0.2353731	0.7646269	1.342410	9.519425	3.507642	1.089535
1500	0.2302366	0.7697634	2.059981	9.476455	3.162097	0.9876496

4. CONCLUSIONS

This paper proposed parameter estimates for a two components Laplace mixture model using the expectation maximization algorithm. The behavior of algorithm is discussed using simulation data and the implementation is made using R software. The effectiveness of algorithm is presented with different values of model parameters and sample size. As expected, the parameter estimates are close to the real parameter values when the sample size is increased. The performance of the algorithm is guaranteed the convergent to a local maximum of the data log-likelihood model as a function of the model parameters. In addition, the value of weighted component is also affected on the parameter estimates, and the equal number of weighted parameters lead to good estimates for model parameters.

Acknowledgement

This paper is a component of a Master thesis undertaken by the second author under the supervision of the first author. The authors are also grateful to King Abdul Aziz University and Al Jouf University for the beneficial cooperation between them. The authors would like to thank the reviewers for their helpful comments.

References

1. **Ali, M., Nadarajah, S.**, 2007, *Information matrices for normal and Laplace mixtures: Information Sciences*, 177(3), 947-955.
2. **Barazandeh, B., Razaviyayn, M.**, 2018, *On The Behavior of the Expectation-Maximization Algorithm For Mixture Models*. 2018 IEEE Global Conference on Signal and Information Processing (GlobalSIP), Anaheim, CA, USA, 61-65.
3. **Dong, N., Huang, H., Zheng, L.**, 2015, *Support Vector Machine in Crash Prediction at The Level of Traffic Analysis Zones: Assessing the Spatial Proximity Effect*. Accident Analysis and Prevention, 82(2015) 192-198.
4. **Enders, C.K.**, 2001, *A primer on maximum likelihood algorithms available for use with missing data*. Journal of structural equation modeling, 1070-5511.
5. **Huang, Y.**, 2016, *Quantile regression-based Bayesian semiparametric mixed-effects models for longitudinal data with non-normal, missing and mismeasured covariate*. Journal of Statistical Computation and Simulation, 86(6), 1183-1202.
6. **Kalantan, Z., Alrewely, F.**, 2019, *A 2-Components Laplace Mixture Model*, The 4 International Conference on Computing, Mathematics and Statistics 2019 (iCMS2019). At: Langkawi Island, Malaysia.
7. **Kotz, et al.**, 2012, *The Laplace distribution and generalizations: a revisit with applications to communications, economics, engineering, and finance*, Springer Science & Business Media. 146120173X.
8. **Kurtz, N., Song, J.**, 2013, *Cross-entropy-based adaptive importance sampling using Gaussian mixture*. Structural Safety, 42, 35-44.
9. **Lu, Y., Gao, Q., Sun, D., Xia, Y., Zhang, D.**, 2015, *SAR speckle reduction using Laplace mixture model and spatial mutual information in the directionlet domain*. Elsevier, Neurocomputing 173 (2016) 633–644.
10. **Norton, R.M.**, 1984, *The Double Exponential Distribution: Using Calculus to Find a Maximum Likelihood Estimator*. The American Statistician*, 38*(2), 135-136. doi:10.2307/2683252
11. **Olive. D. J.**, 2017, *'Liner Regression'*. Department of Mathematics Southern Illinois University Carbondale, IL, USA.
12. **Reed, W. J.**, 2006, *The normal-Laplace distribution and its relatives*. In Advances in Distribution Theory, Order Statistics, and Inference (pp. 61-74). Birkhäuser Boston.
13. **R Core Team, 2019**, *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.
14. **Sheng, X., Hu, Y. H.**, 2005, *Maximum likelihood multiple-source localization using acoustic energy measurements with wireless sensor networks*. IEEE Transactions on Signal Processing, 53(1), 44-53.
15. **Shenoy, S., Gorinevsky, D.**, 2014, *Gaussian-Laplace mixture model for electricity market*. Paper presented at the Decision and Control (CDC), 2014 IEEE 53rd Annual Conference on.