EM algorithm for Gaussian mixture

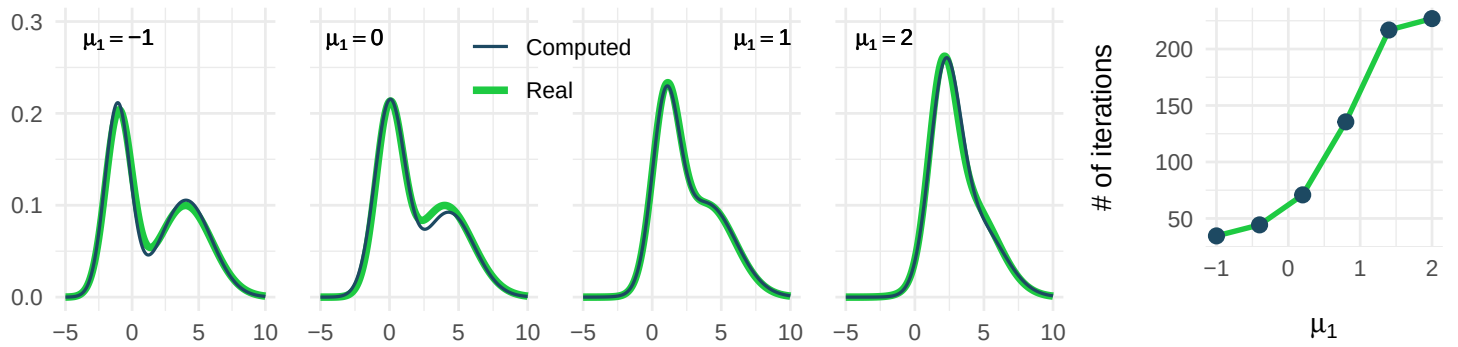
Alessandro Ferrera

The Expectation-Maximization (EM) algorithm is an iterative method for estimating the parameters of a statistical model, especially useful for models with latent variables.
In this report, we apply the EM algorithm to estimate the parameters of a Gaussian mixture distribution and investigate how different parameter configurations impact the convergence rate of the algorithm. Specifically, while analyzing the effect of a given parameter, all others are held constant with reference values: $\mu_1 = -0.5$, $\sigma_1^2 = 1$, $\mu_2 = 4$, $\sigma_2^2 = 4$, and $\tau = 0.5$.

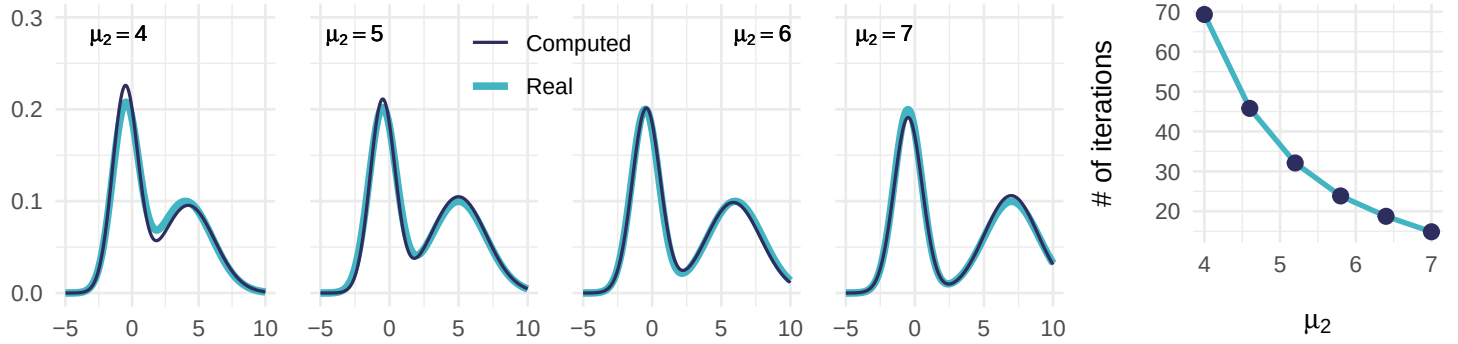
Overlapping

Effect of the Mean

When we increase the mean of the first Gaussian distribution, we observe a substantial increase in the number of iterations required for convergence. This is because, with means close to each other, the distributions overlap significantly, making it challenging for the algorithm to distinguish between them. As shown in the distribution plots, when $\mu_1 = 2$, the two distributions are almost overlapping, and the algorithm requires an enormous number of iterations.

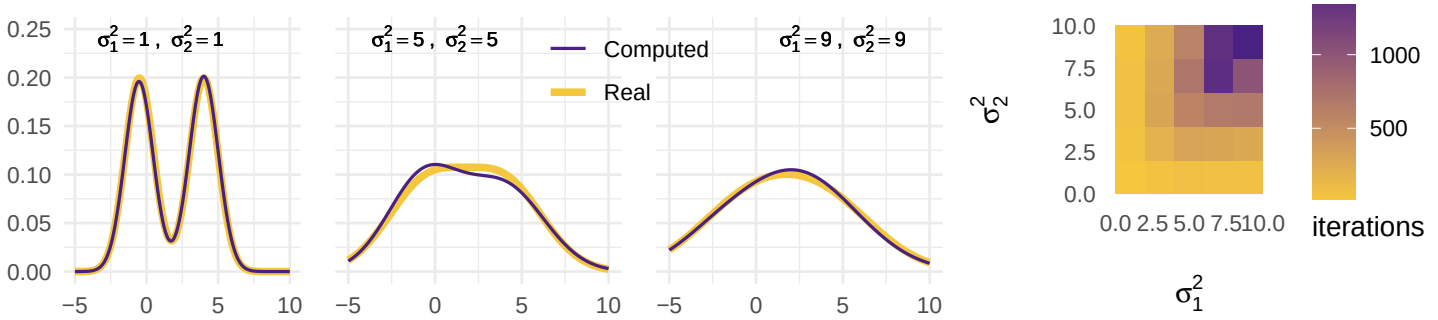


Conversely, when we increase the mean of the second Gaussian distribution, the number of iterations to convergence decreases. This can be attributed to the fact that the two distributions are farther apart, making them easier to separate. As a result, the algorithm converges more quickly due to the clearer separation between the components.



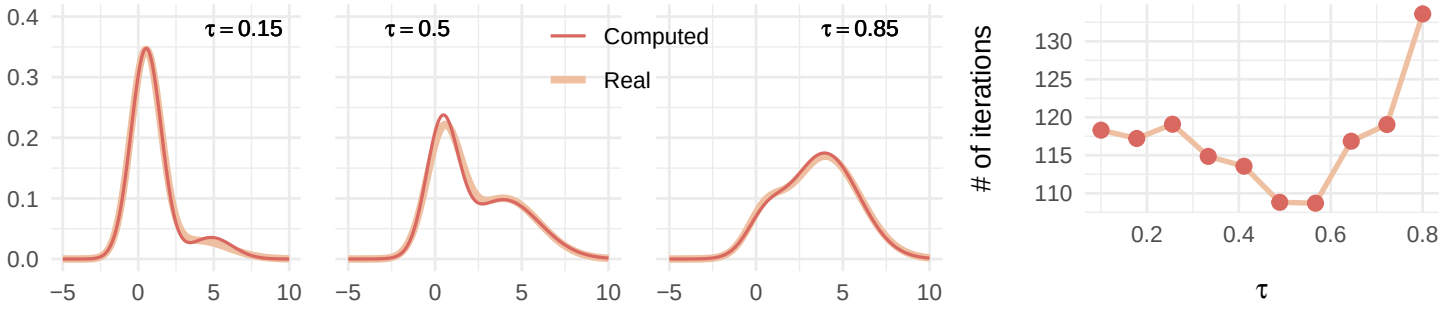
Effect of the Variance

Similarly to the mean, changes in the variance of the Gaussian distributions follow a similar trend. With big variances, distributions overlap and convergence takes longer. Instead, when the variances are small, the distributions are more separated, and the algorithm converges faster.



Effect of Mixing Probability

The behavior of the algorithm also varies with the mixing probability. The EM algorithm converges very slowly when τ is close to the extreme values (0 or 1) and converges fastest when $\tau = 0.6$. This suggests that the algorithm is most efficient when the mixture components are more balanced, slightly favoring the component with higher variance.

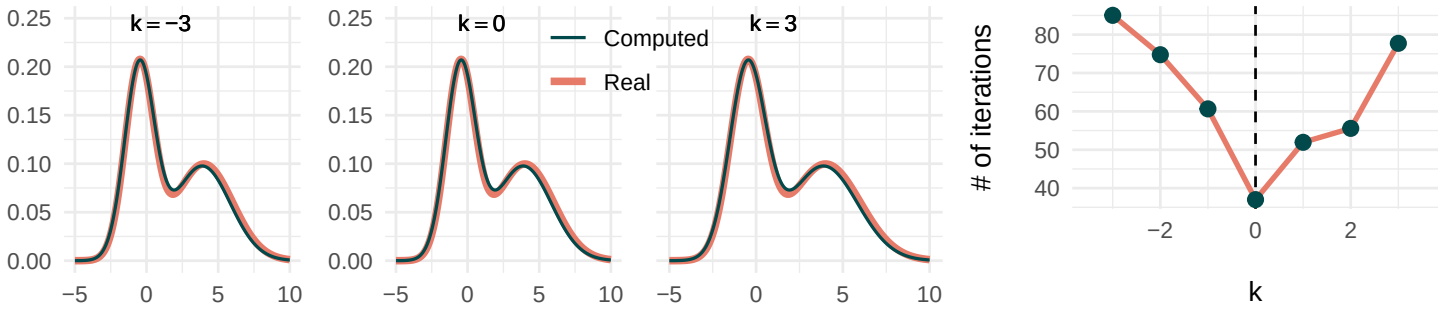


Effect of Initial Parameter Estimates

The initial guesses for the parameters are crucial for the algorithm's convergence rate. We analyzed convergence under various initialisations, using the rule:

$$\mu_0(k) = \mu + k, \quad \tau_0(k) = \tau + k \cdot 0.15, \quad \sigma_0^2(k) = \begin{cases} \sigma^2 \cdot |k|^{\text{sgn}(k)} & \text{if } k \neq 0 \\ \sigma^2 & \text{if } k = 0 \end{cases}$$

where k is a random number between -3 and 3, and μ, τ and σ^2 are the real parameters. With $k < 0$ we are underestimating the real parameters, and with $k > 0$ we are overestimating them. In the following plot, we can see that the algorithm converges faster when the initial guess is close to the real parameters ($k = 0$).



Conclusion

In conclusion, the EM algorithm is a powerful tool for estimating the density of a Gaussian mixture, since we established that in every cases there is only a slight difference between the real and the computed densities. Though, convergence speed is highly sensitive to the parameter settings. Proper initial estimates are essential for faster convergence, and configurations with distinct component distributions or balanced mixing probabilities lead to more efficient convergence.