Prof. Dr. Frank Deinzer
Technical University of Applied Sciences
Würzburg-Schweinfurt

Work Sheet 5

EM-Algorithm

Exercise 1

Realize an implementation of the EM algorithm in a programming language of your choice. Do not use a pre-built implementation. A description of the algorithm is part of the lecture, but can also be found in [1].

Test your implementation on the following scenarios, especially examining the accuracy and stability of the estimate. Vary the initialization of the parameters.

a) Generate a synthetic sample set from a mixture distribution of your choice: n mixture components and correspondingly many multidimensional normal distribution components. Estimate the mixture parameters from this sample set using the correct number of mixture components

b) Change the mean vectors and the covariance matrices for the mixture components. Does this affect the EM algorithm?

c) Vary the size of the sample set and repeat the scenarios from 1 a) and 1 b). What do you observe?

d) Vary the number of mixture components using more or less then used to create the sample set. What do you observe?

e) What happens to the EM estimation process if you significantly increase the dimensionality of the distribution?

References

[1] Marc Peter Deisenroth, A. Aldo Faisal, and Cheng Soon Ong. *Mathematics for Machine Learning*. Cambridge University Press, 2020.