Mathematical Statistics
Stockholm University
Bachelor Thesis **2019:9**
http://www.math.su.se

Applying the EM algorithm to pixel-based approximations of bimodal data

Magnus Pierrau*

June 2019

Abstract

In this thesis we develop and study an approximative version of the Expectation-Maximization algorithm, modified to analyze quantized data. The quantized data consists of a grid of pixels and emulates low resolution images of simplified protein structures. We perform several experiments designed to capture the performance of the approximative and regular EM algorithms under different circumstances. We find that the approximative variant consistently performs worse than the regular algorithm but provides acceptable results. However, as the size of the pixels decrease, the approximative variant approaches the results of the regular algorithm. We find that the approximative variant has a highly unstable convergence rate and that the assumption of homoscedasticity is not suited well for this variant.

*Postal address: Mathematical Statistics, Stockholm University, SE-106 91, Sweden. E-mail: magnus.pierrau@gmail.com. Supervisor: Ola Hössjer, Christian Blau.