The EM algorithm for Poisson data

In this project we study the surprising behavior of the so called expectation maximization (EM) algorithm for Poisson data.

In the context of positron emission tomography (PET), the following model arises. For $i = 1, \ldots, n$, one records independent random variables Y_i that follow a Poisson distribution with intensity $\sum_{j=1}^{m} a_{ij}\lambda_j$. Here, the non-negative real numbers $(a_{ij})_{ij}$ are known in advance and one wants to estimate/reconstruct the λ_j's from the data.

A natural way to do this is via the EM algorithm (in the physics literature sometimes called Richardson-Lucy algorithm). This is an iterative method for estimation, given incomplete data. For the Poisson estimation problem described above, the EM algorithm shows some unexpected behavior. During the first few iterations the method seems to converge rapidly to the truth, but then starts to build in artefacts in the reconstruction. Even worse, in the long-run the EM algorithm yields completely degenerate reconstructions.

The EM algorithm has been applied successfully in practice, by using stopping rules that terminate the iterations (more or less) at the right moment, just before the EM algorithm starts to deteriorate. Up to now, however, there has been only very little work in understanding the behavior of the EM algorithm.

The outline of the project is as follows: After getting familiar with the EM algorithm, we will study it for different distributions. By exploring the described phenomenon by computer simulations, we then hope to derive some heuristics/conjectures concerning the number of iterations until which the EM algorithm shows convergence type behavior. In a final stage we will try to prove a result.