Fast Expectation Maximization Algorithm for Parameter Estimation of Finite Mixtures

Objective: To develop a fast method for the expectation maximization (EM) algorithm within the frame of parameter estimation of Gaussian and Poisson finite mixtures to be used in signal segmentation or classification of tomographic count data.

Methodology: Let $\boldsymbol{f}=\{f_{ij}\}$ be the intensity values (pixels) of the reconstructed image and let us assume that $\boldsymbol{f}$ is a finite mixture over K_s disjoint, spatially connected image classes. Then we can define the likelihood for the image distribution

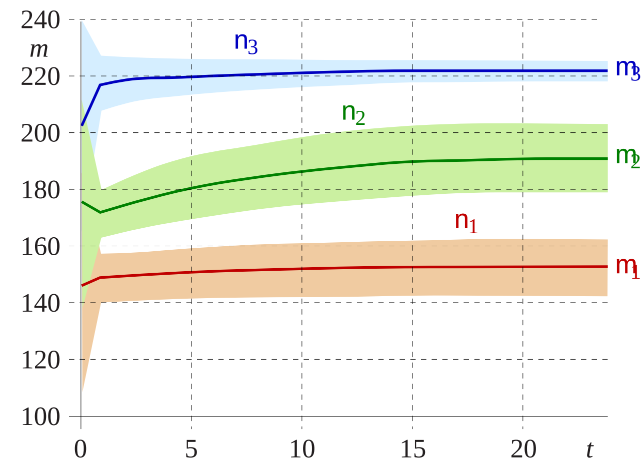
$$G(\boldsymbol{f}|\Omega,\Theta)=\Pi_{i,j}\Sigma_{k_s}\omega_{k_s}\, p(f_{i,j}|\theta_{k_s})$$

with the model parameters $\Theta=\{\theta_{k_s}\}$ and the weights $\Omega=\{\omega_{k_s}\}$, $\Sigma_{k_s}\omega_{k_s}=1$, which represents the ratio of the number of pixels in class k_s and the number of image pixels. Commonly used statistical moments admitted in the parameter vector θ_{k_s} are the intensity mean μ_{k_s} and the standard deviation σ_{k_s}. The estimation of the model parameters of each class is carried out using a maximum likelihood expectation maximization (MLEM) approach. In order to specify the relation between f_{ij} and k_s, Liang and colleagues [1] introduced an unobserved indicator vector $z_{ij}=\{z_{k_s}\}$ which can be interpreted as conditional probability that f_{ij} belongs to class k_s, by a given realization $\boldsymbol{f}$ as well as parameters Ω and Θ. Note that $Z=\{z_{ij}\}$ spans a vector space with (N^2K_s) elements. Since the solution trial is described completely with methods of statistics and a description of the image as a function of local indexes i and j within the EM algorithm is unnecessary, the dimension of Z can be reduced drastically. Let $\boldsymbol{H}(\boldsymbol{f})$ be the image histogram defined by $\boldsymbol{H}(\boldsymbol{f})=\text{cardial}\{i,j\text{: } f_{ij}=m\}$, then the EM algorithm can be applied to the histogram vector instead of the image matrix. Under this consideration, the dimension of the indicator vector $z_m=\{z_{k_s}\}$ is reduced to (MK_s) elements with $M=\max\{f_{ij}\}-\min\{f_{ij}\}$. The value of z_{k_s} can now be interpreted as conditional probability that the intensity value m (and so each signal value which satisfies $f_{ij}=m$) belongs to class k_s by a given realization $\boldsymbol{f}$ as well as the parameters Ω and Θ. Consequently, the likelihood function to be maximized is now defined by

$$G(\{m,z_{k_s}\}|\Omega,\Theta)=\Pi_m\Sigma_{k_s}\omega_{k_s}\, p(m,z_{k_s}|\theta_{k_s})$$

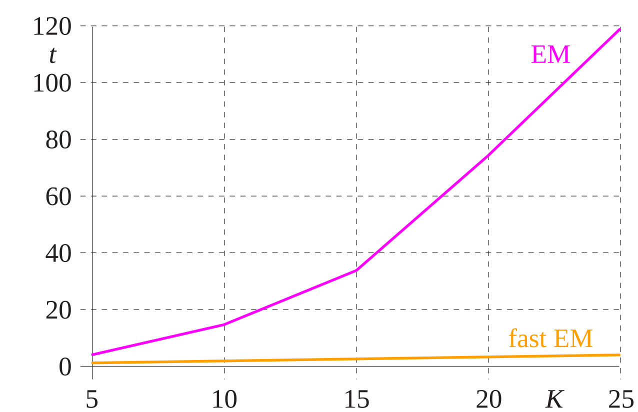
and the classical EM algorithm introduced by Dempster *et al.* [2] can be applied.

[1] Liang *et al.*, IEEE TNS vol. 39, no. 4, pp. 1126-1133, 1997.
[2] Dempster *et al.*, JRSSB vol. 38, pp. 1-38, 1977.

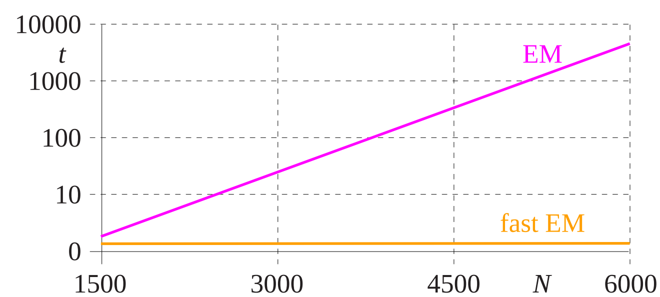


Dynamics of class mean and variance values during the EM process shown for K=3

The formulation of this fast EM algorithm yields an significant improvement in computational time, although the total number of analyzed random variables has not changed. Estimated model parameters as well as the convergence dynamics are identical to [2].



Time performance versus number of mixture sub-population (signal classes)



Time performance versus number of elements

IEEE TNS vol. 44, no. 4, pp. 1583-1590 (Appendix), 1997.