DM3 - Probabilistic Graphical Models

Baptiste Doyen

5th December 2018

1 EM for HMM

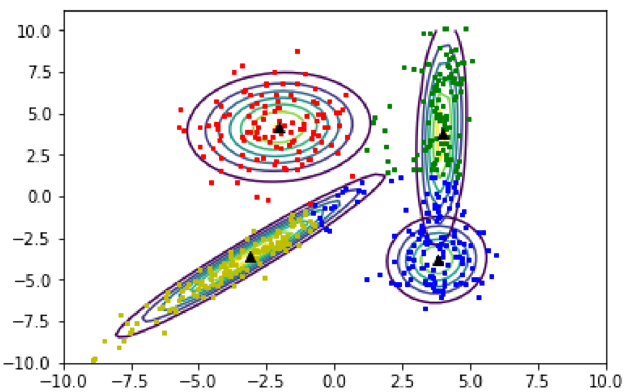
We use the same notations as in lecture notes.
E-step formula isgiven by :

$$E[l(\theta)] = \sum_{i=1}^{K} p(q_0|u_1,...,u_T)log(\pi_i) + \sum_{i=1}^{T}\sum_{i,j=1}^{K} p(q_{t+1}=i, q_t=j|u_1,...,u_T)log(A_{i,j})$$

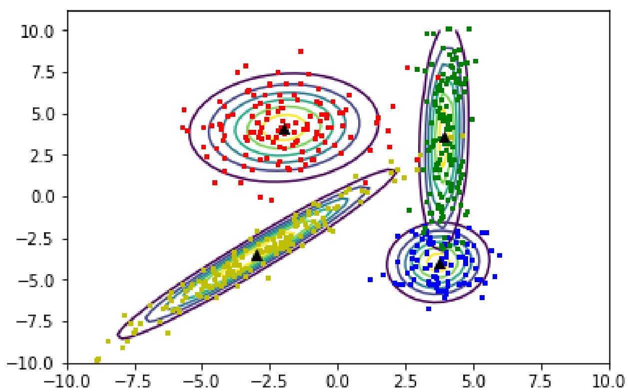
$$+\sum_{i=1}^{T}\sum_{i}^{K} p(q_t=i|u_1,...,u_T)log(f(u_t,q_t))$$

Then, M-step consists in solving : max $E[l(\theta)]$ st. $\sum_{l=1}^{K} A_{k,l}=1$ and $\sum_{i=1}^{K}\pi_i=1$
By computing the Lagrangian of this problem and annuling its gradient, we obtain the following expressions :
$\hat{\pi}_i = p(q_0=i|u_1,...,u_T)$, $\hat{A}_{i,j} = \frac{\sum_{i,j=1}^{K} p(q_{t+1}=i,q_t=j|u_1,...,u_T)}{\sum_{t=1}^{K} p(q_t=j|u_1,...,u_T)}$, $\hat{\mu}_i = \frac{\sum_{t=1}^{T} p(q_t=i|u_1,...,u_T)u_t}{\sum_{t=1}^{T} p(q_t=i|u_1,...,u_T)}$
and $\hat{\Sigma}_i = \frac{\sum_{t=1}^{T} p(q_t=i|u_1,...,u_T)(u_t-\hat{\mu}_i)^T(u_t-\hat{\mu}_i)}{\sum_{t=1}^{T} p(q_t=i|u_1,...,u_T)}$

2 GMM vs. HMM



GMM output plot



HMM output plot

3 Log-likelihoods and comments

After computing the normalized log-likelihoods, we obtain the following results :

Method	HMM	GMM
Train	-1.40	-10.05
Test	-3.66	-33.30

The values obtained for GMM are lower than the ones expected (cf solution of $DM2$).
HMM optimizes better the log-likelihood. This is visible with the above figures, the HMM ellipses captures better the data points and the main difference between the two methods lies in the Σ_i estimation. One has also to consider that the HMM methods has been initialized with the GMM, which is a better initialization compared to only k-means we used to initialize GMM.