Machine Learning for Computer Vision, MVA 2014-2015

Instructor: Iasonas Kokkinos, iasonas.kokkinos[at]ecp.fr

Programming Assignment 3 - 3.0 Points/20 Deliverable: 18 December

1 Expectation Maximization for a Mixture-of-Gaussians (2.0/3.0)

The class-specific distributions for the data in the previous programming assignments had been generated by a mixture-of-Gaussians distribution with three components.

Instead of learning a discriminative function for classification, you are now asked to learn a generative model, by performing density estimation for the data belonging to each class.

We assume that we do know the form of the distribution (a mixture of three Gaussians) but do not know its parameters. Your first task is to perform parameter estimation for each class-specific distribution.

.5/2.0 Initialization: use k-means to initialize the centers of the Gaussians. Run k-means 20 times, randomly choosing 3 of the data points as cluster centers each time. Each run will give you, upon convergence, 3 centers.

For each such center evaluate the distortion function minimized by k-means. In the end, pick the set of centers that minimizes this criterion over all runs.

For the best choice of centers, plot the distortion function as a function of k-means iteration.

1/2.0 EM algorithm: Using the clusters found by k-means as initialization, iteratively apply the E- and M-steps of EM. Plot the expected complete observation log-likelihood as a function of EM iteration.

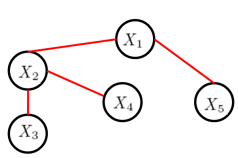
Now you have a method to construct a generative model for both classes.

.5/2.0 Use Bayes' rule to form the class-posteriors. Plot the decision boundaries by adapting the code of the previous exercises.

Optional: Compare them with with those obtained from Adaboost and the SVM classifier from the previous assignments.

2 Sum/Max-Product algorithms (1.0)

Consider the following MRF:



It has discrete variables, each of which can take a distinct number of values. Specifically, $|X_1| = 3, |X_2| = 2, |X_3| = 4, |X_4| = 2, |X_5| = 2$.

The potential functions of the network are as follows:

$$\begin{aligned}\Phi(X_1, X_2) &= \begin{bmatrix} .1 & .9 \\ .5 & .2 \\ \underbrace{.1}_{\Phi(X_1=3, X_2=1)} & .1 \end{bmatrix} \Phi(X_2, X_3) = \begin{bmatrix} .1 & .9 & .2 & .3 \\ .8 & .2 & .3 & .6 \end{bmatrix} \Phi(X_2, X_4) = \begin{bmatrix} .1 & .9 \\ .8 & .2 \end{bmatrix} \Phi(X_1, X_5) = \begin{bmatrix} .1 & .2 \\ .8 & .1 \\ .3 & .7 \end{bmatrix} \\ \Phi(X_j, X_i) &= \Phi(X_i, X_j)^T\end{aligned}$$

while the joint distribution of the state is given by

$$\begin{aligned}P(X) &= \frac{1}{Z}\Phi(X_1, X_2)\Phi(X_2, X_3)\Phi(X_2, X_4)\Phi(X_1, X_5) \\ Z &= \sum_X \Phi(X_1, X_2)\Phi(X_2, X_3)\Phi(X_2, X_4)\Phi(X_1, X_5) \qquad (1)\end{aligned}$$

As we do not know the value Z in advance, we can consider an unnormalized distribution,

$$\tilde{P}(X) = \Phi(X_1, X_2)\Phi(X_2, X_3)\Phi(X_2, X_4)\Phi(X_1, X_5) = P(X)Z, \qquad (2)$$

where, by definition, $Z = \sum_X \tilde{P}(X)$.

- Implement the sum product algorithm; execute it for $\tilde{P}(X)$ in order to compute

$$\begin{aligned}\tilde{P}(X_1) &= \sum_{X_2, X_3, X_4, X_5} \tilde{P}(X) \qquad (3) \\ \tilde{P}(X_2) &= \sum_{X_1, X_3, X_4, X_5} \tilde{P}(X) \qquad (4)\end{aligned}$$

- Use these to compute the Partition function, Z
- Use Z to compute the normalized marginal distributions for $P(X_1), P(X_2)$ and report your results.
- Construct a function that returns P(X). Evaluate it for

$$\begin{aligned}X &= \{1, 2, 4, 2, 1\} \\ X &= \{3, 1, 2, 1, 2\} \\ X &= \{2, 2, 1, 2, 1\}\end{aligned}$$

- Implement the max-product algorithm for this network.

Bonus, 0.5 Consider we know $X_5 = 1, X_4 = 2, X_1 = 3$. Compute $P(X_2), P(X_3)$. You will need to modify the potentials and therefore recompute the partition function.