CS 8735: Report for assignment 1

Chanmann Lim

September 17, 2015

Problem 1. In this task, we are given a dataset generated from a mixture density and the job is to implement EM algorithm to learn the parameters of the model. Based on the assumption that the Gaussian Mixture Model has four component Gaussian PDFs with each having a full covariance matrix we will terminate the our EM estimation at the 100^{th} iterations.

The Matlab code for the experiment is in the **Appendix** section.

a) For the first experiment which we named it case **a**, we run EM procedure with the initialization suggested in the assignment.

$$\pi_k^{(0)} = 1/4 \qquad 1 \le k \le 4$$
$$\mu_1^{(0)} = [10\ 2]^T, \mu_2^{(0)} = [5\ 6]^T, \mu_3^{(0)} = [0\ 1]^T, \mu_4^{(0)} = [4\ 3]^T$$
$$\Sigma_k^{(0)} = \mathbf{I}_{2\times 2} \qquad 1 \le k \le 4$$

After the EM procedure terminated, we got

$$\hat{\pi}_1 = 0.3457, \hat{\pi}_2 = 0.1401, \hat{\pi}_3 = 0.1847, \hat{\pi}_4 = 0.3295 \tag{1}$$

$$\hat{\mu}_1 = \begin{bmatrix} 13.0263 \\ 3.0455 \end{bmatrix}, \hat{\mu}_2 = \begin{bmatrix} 4.0619 \\ 7.9674 \end{bmatrix}, \hat{\mu}_3 = \begin{bmatrix} 1.6026 \\ 1.5717 \end{bmatrix}, \hat{\mu}_4 = \begin{bmatrix} 6.9183 \\ 5.9843 \end{bmatrix} \tag{2}$$

$$\hat{\Sigma}_1 = \begin{bmatrix} 6.9183 & 5.9843 \\ -0.7471 & 2.0688 \end{bmatrix}, \hat{\Sigma}_2 = \begin{bmatrix} 0.8788 & 0.2342 \\ 0.2342 & 1.1568 \end{bmatrix}, \tag{3}$$

$$\hat{\Sigma}_3 = \begin{bmatrix} 8.4468 & -0.0635 \\ -0.0635 & 1.0938 \end{bmatrix}, \hat{\Sigma}_4 = \begin{bmatrix} 6.2731 & 2.6295 \\ 2.6295 & 1.9615 \end{bmatrix} \tag{4}$$

Appendix:

problem_1.m

```
% ------------------------------------------------
% CS 8735: Supervised Learning Fall (2015)
%       Unversity of Missouri-Columbia
%             Chanmann Lim
%             September 2015
% ------------------------------------------------
clc;
clear;
close all;

%% Load data
X = load('GMD.dat');

%% EM algorithm
T = 100; % 100 iterations

% Initialization
prior = 1/4 * ones(1, 4);
mu = { [10; 2], [5; 6], [0; 1], [4; 3] };
Sigma = { eye(2), eye(2), eye(2), eye(2) };

[prior, mu, Sigma] = EM(X, T, prior, mu, Sigma);

display(prior);
display(cell2mat(mu));
display(cell2mat(Sigma));
```

EM.m

```
function [ prior, mu, Sigma ] = EM( X, T, prior, mu, Sigma )
%EM - run EM algorithm for T iterations

[~, K] = size(prior);
[N, ~] = size(X);

t = 0;
while t < T
    for k=1:K
        % Expectation step
        g = gamma_nk(X, k, prior, mu, Sigma);
        Nk = sum(g);

        % Maximization step
        mu{k} = 1/Nk * sum(g*ones(1, 2) .* X)';
        X_tilde = X' - mu{k}*ones(1,N);
        Sigma{k} = 1/Nk * (ones(2,1)*g' .* X_tilde * X_tilde');
        prior(k) = Nk / N;
    end

    % Check for convergence
    % We're assuming that EM algorithm will converge in T iteration
    t = t + 1;
end
```

gamma_nk.m

```
function [ g ] = gamma_nk( X, k_i, prior, mu, Sigma )
% GAMMA_NK - gamma n,k in the E-Step of EM algorithm
%            is defined as P(z_n = k|x_n, Theta)
%    where
%        Theta =< prior, mu, Sigma >

    [~, K] = size(prior);
    [N, ~] = size(X);
    denominators = zeros(N, K);
    for k=1:K
        denominators(:, k) = prior(k) * mvnpdf(X, mu{k}, Sigma{k});
    end
    g = denominators(:, k_i) ./ sum(denominators, 2);
end
```

mvnpdf.m

```
function [ y ] = mvnpdf( X, mu, Sigma )
% NORMAL - Multivariate normal density N(x; mu, Sigma)

    [N, d] = size(X);
    y = zeros(N, 1);
    denominator = sqrt((2*pi)^d*det(Sigma));
    for n=1:N
        x = X(n, :)';
        x_tilde = x - mu;
        y(n) = 1/denominator * exp(-0.5 * x_tilde'/Sigma*x_tilde);
    end
end
```