GHILENE RAYANE
LASSO NICOLAS

Ecole Nationale Supérieure de l'Electronique et de ses Applications

Tp1 - Advanced Signal Methods
EM Algorithm

1. Introduction

x = (x1, ..., xN)^t the observation of N independent observations from a mixture of two normal distributions $N(m_1, \sigma_2^2)$ and $N(m_2, \sigma_2^2)$.

• $x_i \in$ Gk if it follows the distribution N (m_k, σ_k^2)

• z_{ik} (avec i = 1, ... ,N et k = 1, 2) is the indicator function that determines the group Gk to which xi belongs (we say that the zik are the hidden data of the problem)

• π_1 et π_2 are the probabilities that x_i is in G_1 or G_2. We note the vector of parameters

$\theta = (\pi_1, \pi_2, m_1, \sigma_1^2, m_2, \sigma_2^2)$ t :

$$p(x_i|\boldsymbol{\theta}, z_{ik}) = \frac{1}{\sqrt{2\pi\sigma_k^2}} \exp\left\{-\frac{1}{2\sigma_k^2}(x_i - m_k)^2\right\}$$

In this Tp, we want to classify the x_i and estimate the unknown parameters. the log of the likelihood is defined as follows:

$$\ln p(\boldsymbol{x}|\boldsymbol{\theta}) = E_z[\ln p(\boldsymbol{x}|\boldsymbol{\theta}, \boldsymbol{z})] = \sum_{i=1}^{N}\sum_{k=1}^{2} z_{ik} \ln[\pi_k p(x_i|\boldsymbol{\theta}, z_{ik})]$$

2. EM for a mixture of Gaussians

2.1 The EM

- Initialisation : $\theta^{(0)}$;
- A l'itération (c) :

— ***Step E :*** We calculate the conditional probabilities $t_{ik}^{(c)}$ that the observation is in the group G_k and x_i the current value of the mixture $\theta^{(c)}$.

We estimate the $Q(\theta,\theta^{(c)})$ function, estimated marginalization of the log-likelihood function, knowing the current value of the mixture $\theta^{(c)}$.

— ***Step M :*** *We estimate the parameter* $\boldsymbol{\theta^{(c+1)}}$ at the iteration (c+1) by maximizing the function $\boldsymbol{Q(\theta, \theta^{(c)})}$ using the parameter $\boldsymbol{\theta}$.

2.2 EM Algorithm

Step E: Estimation

$$\mathrm{p}(\mathrm{x}_i|\theta) = p(\overline{z_{i1}}|\theta)p(x_i|\theta,\overline{z_{i1}}) + p(\underline{z_{i1}}|\theta)$$

$$p(x_i|\theta) = \sum_{k=1}^{2} p(\underline{z_{ik}}|\theta)p(x_i|\theta,\underline{z_{ik}})$$

$$p(x_i|\theta) = \sum_{k=1}^{2} \pi_k p(x_i|\theta,\underline{z_{ik}})$$

$$E_z[z_{ik}|x_i,\theta] = 0 * p(\overline{z_{i1}}|x_i,\theta) + 1 * p(\underline{z_{ik}}|x_i,\theta)$$

$$E_z[z_{ik}|x_i,\theta] = p(\underline{z_{ik}}|x_i,\theta)$$

$$E_z[z_{ik}|x_i,\theta] = \frac{p(x_i|\underline{z_{ik}},\theta)p(\underline{z_{ik}}|\theta)}{p(x_i|\theta)}$$

$$E_z[z_{ik}|x_i,\theta] = \frac{\pi_k p(x_i|\theta,\underline{z_{ik}})}{\pi_1 p(x_i|\theta,\underline{z_{i1}})+\pi_2 p(x_i|\theta,\underline{z_{i2}})}$$

$$Q(\theta|\theta^{(c)}) = E_z[ln(p(x,z|\theta))|x_i,\theta^{(c)}]$$

$$Q(\theta|\theta^{(c)}) = \sum_{i=1}^{N}\sum_{k=1}^{2} E_z[z_{ik}|x_i,\theta^{(c)}]ln(\pi_k p(x_i|\theta,\underline{z_{ik}}))$$

$$Q(\theta|\theta^{(c)}) = \sum_{i=1}^{N}\sum_{k=1}^{2} t_{ik}^{(c)} ln(\pi_k^{(c)} p(x_i|\theta,\underline{z_{ik}}))$$

$$avec : t_{ik}^{(c)} = \frac{\pi_k^{(c)} p(x_i|\theta^{(c)},\underline{z_i k})}{(\pi_1^{(c)} p(x_i|\theta^{(c)},\underline{z_{i1}})+\pi_2^{(c)} p(x_i|\theta^{(c)},\underline{z_{i2}}))}$$

Step M: Maximization

$$\frac{\delta Q(\theta,\theta^{(c)})}{\delta \pi_1} = \sum_{i=1}^{N} \frac{t_{i1}^{(c)}}{\pi_1} - \frac{t_{i2}^{(c)}}{1-\pi_1}$$

$$\frac{\delta Q(\theta,\theta^{(c)})}{\delta \pi_1} = 0 \iff (1-\pi_1)\sum_{i=1}^{N} t_{i1}^{(c)} = \pi_1 \sum_{i=1}^{N} t_{i2}^{(c)}$$

$$\frac{\delta Q(\theta,\theta^{(c)})}{\delta \pi_1} = 0 \iff \pi_1 = \frac{\sum_{i=1}^{N} t_{i1}^{(c)}}{\sum_{i=1}^{N} t_{i1}^{(c)}+\sum_{i=1}^{N} t_{i2}^{(c)}}$$

$$\frac{\delta Q(\theta,\theta^{(c)})}{\delta \pi_1} = 0 \iff \pi_1 = \frac{1}{N}\sum_{i=1}^{N} t_{i1}^{(c)}$$

$$\pi_1^{(c+1)} = argmax_{\pi k}(Q(\theta,\theta^{(c)})) = \frac{1}{N}\sum_{i=1}^{N} t_{ik}^{(c)}$$

$$\frac{\delta Q(\theta,\theta^{(c)})}{\delta m_1} = \sum_{i=1}^{N} \frac{t_{i1}^{(c)} m_1 (x_i - m_1)}{\sigma_1^2}$$

$$\frac{\delta Q(\theta,\theta^{(c)})}{\delta m_1} = 0 \iff \sum_{i=1}^{N} t_{i1}^{(c)} m_1 = \sum_{i=1}^{N} t_{i1}^{(c)} x_1$$

$$\frac{\delta Q(\theta,\theta^{(c)})}{\delta m_1} = 0 \iff m_1 = \frac{\sum_{i=1}^{N} t_{i1}^{(c)} x_1}{\sum_{i=1}^{N} t_{i1}^{(c)}}$$

$$m_k^{(c+1)} = argmax_{mk}(Q(\theta,\theta^{(c)})) = \frac{\sum_{i=1}^{N} t_{ik}^{(c)} x_i}{\sum_{i=1}^{N} t_{ik}^{(c)}}$$

$$\frac{\delta Q(\theta,\theta^{(c)})}{\delta \sigma_1} = \sum_{i=1}^{N} t_{i1}^{(c)} \frac{(x_i - m_1)^2}{\sigma_1^3} - \frac{1}{\sigma_1^2}$$

$$\frac{\delta Q(\theta,\theta^{(c)})}{\delta \sigma_1} = 0 \iff \sum_{i=1}^{N} \frac{t_{i1}^{(c)} (x_i - m_1)^2}{\sigma_1^2} = \sum_{i=1}^{N} t_{i1}^{(c)}$$

$$\frac{\delta Q(\theta,\theta^{(c)})}{\delta m_1} = 0 \iff \sigma_1 = \frac{\sum_{i=1}^{N} t_{i1}^{(c)} (x_i - m_1)^2}{\sum_{i=1}^{N} t_{i1}^{(c)}}$$

$$\sigma_k^{2(c+1)} = argmax_{\sigma k}(Q(\theta,\theta^{(c)}))^2 = \frac{\sum_{i=1}^{N} t_{ik}^{(c)} (x_i - m1)^2}{\sum_{i=1}^{N} t_{ik}^{(c)}}$$

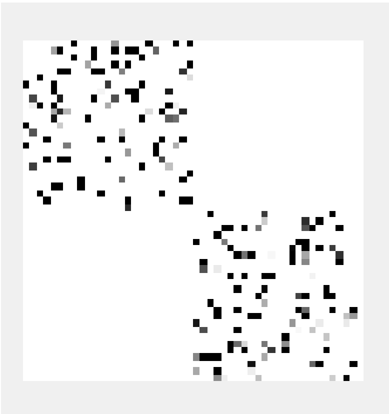
3. Implementation

AlgoEM Function

```
function [theta_out, iteration] = algoEM(vect_x, theta_param_0)

epsilon = 0.1;
max_iter = 1000;

theta_cell = num2cell(theta_param_0);
[pi_1, pi_2, m_1, sig_1, m_2, sig_2] = theta_cell{:};

for iteration = 1:max_iter

    t_1 = pi_1 * exp((-1/2) * ((vect_x - m_1) / abs(sig_1)).^2) / (abs(sig_1) * sqrt(2*pi));
    t_2 = pi_2 * exp((-1/2) * ((vect_x - m_2) / abs(sig_2)).^2) / (abs(sig_2) * sqrt(2*pi));
    t_sum = t_2 + t_1;
    t_1 = t_1 ./ t_sum;
    t_2 = t_2 ./ t_sum;

    prior_theta = [pi_1, pi_2, m_1, sig_1, m_2, sig_2];

    pi_1 = sum(t_1) / length(vect_x);
    pi_2 = sum(t_2) / length(vect_x);

    m_1 = sum(t_1 .* vect_x) / sum(t_1);
    m_2 = sum(t_2 .* vect_x) / sum(t_2);

    sig_1 = sqrt(sum(t_1 .* (vect_x - m_1).^2) / sum(t_1));
    sig_2 = sqrt(sum(t_2 .* (vect_x - m_2).^2) / sum(t_2));

    theta_out = [pi_1, pi_2, m_1, sig_1, m_2, sig_2];

    if norm(theta_out - prior_theta) < epsilon
        break;
    end
end

if iteration == max_iter
    disp("Convergence not reached within maximum iterations.");
end

end
```

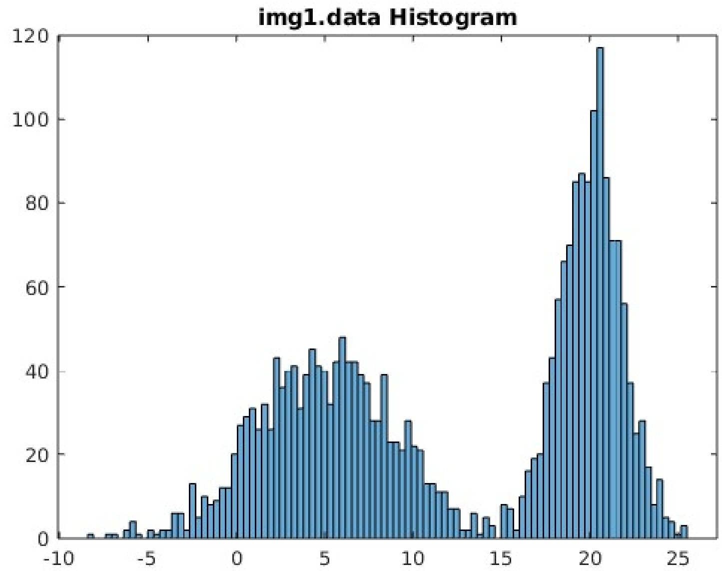
Plot 1: algoEM

We started by loading the image (50*50) from the file img.dat and stored its pixel values in a vector called vect_x which is of size (2500*1)



Plot2: img1

We then calculated the histogram of this image:

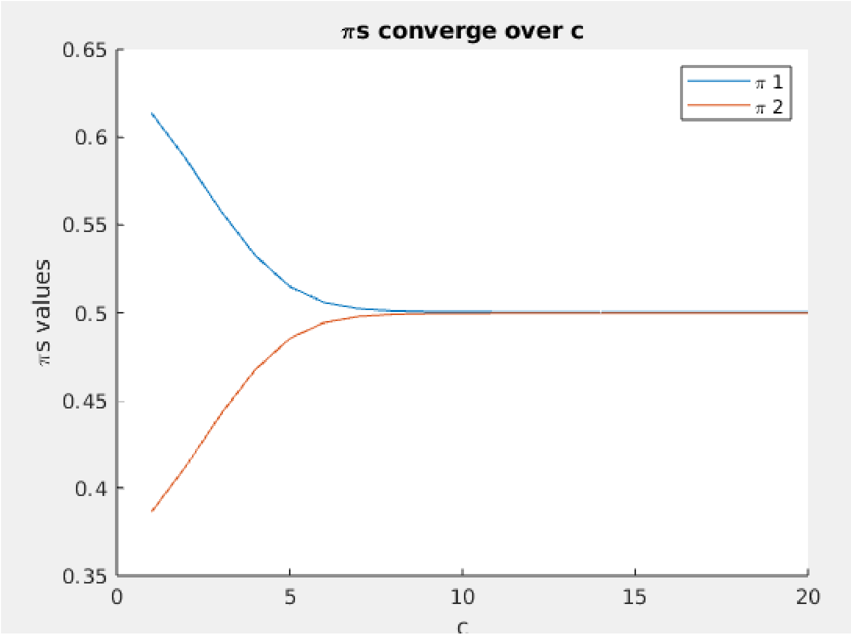


Plot3: Histogram of the image

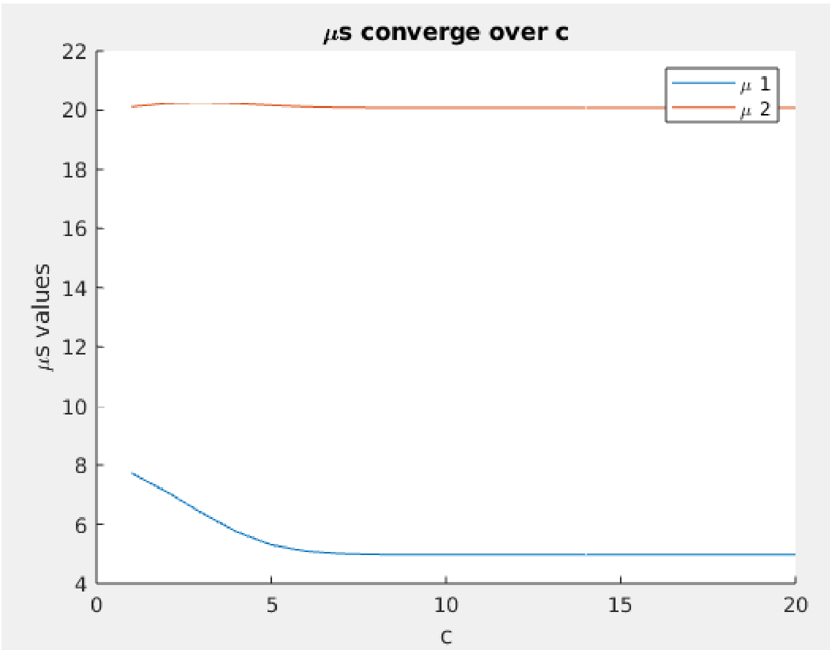
As we can see from the plot, we have two Gaussian figures, one centered at 5 and the other at 20. We are going to apply the algoEM using the following values

$$\theta^{(0)} = (\pi_1, \pi_2, \mu_1, \sigma_1, \mu_2, \sigma_2) = (0.6,\ 0.4,\ 5,\ 5,\ 20,\ 15)$$

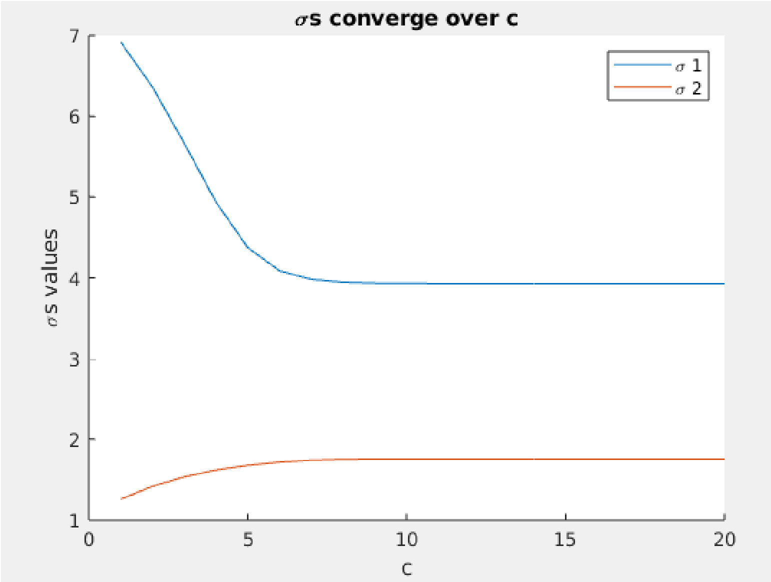
We tested different values of theta to prove that it converges, we started by testing the convergence of π



Plot4: Convergence of π



Plot5: Convergence of the mean



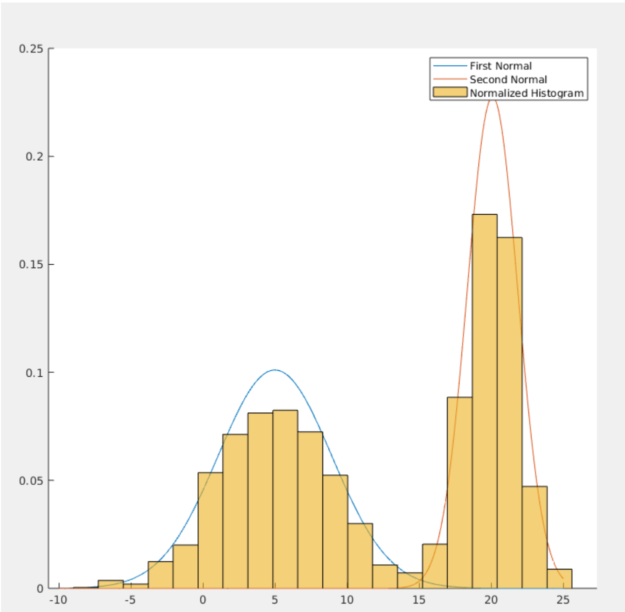
Plot6: Convergence of the standard deviation

1	2	3	4	5	6
0.5003	0.4997	4.9691	3.9300	20.0760	1.7583

As we can see the values of theta converge to

$$\theta^{(0)} = (\pi_1, \pi_2, \mu_1, \sigma_1, \mu_2, \sigma_2) = (0.5003, 0.4997\ , 4.9691\ , 3.93\ , 20.076,\ 1.7583)$$

We can plot the probability distributions deduced from the previous histogram, here are the results:



Plot7:Probability distributions & histogram

The algorithm seems to work very well! The probability distributions perfectly match the shape of the histogram.

Test on the second image

For the second image, we used the following theta values:

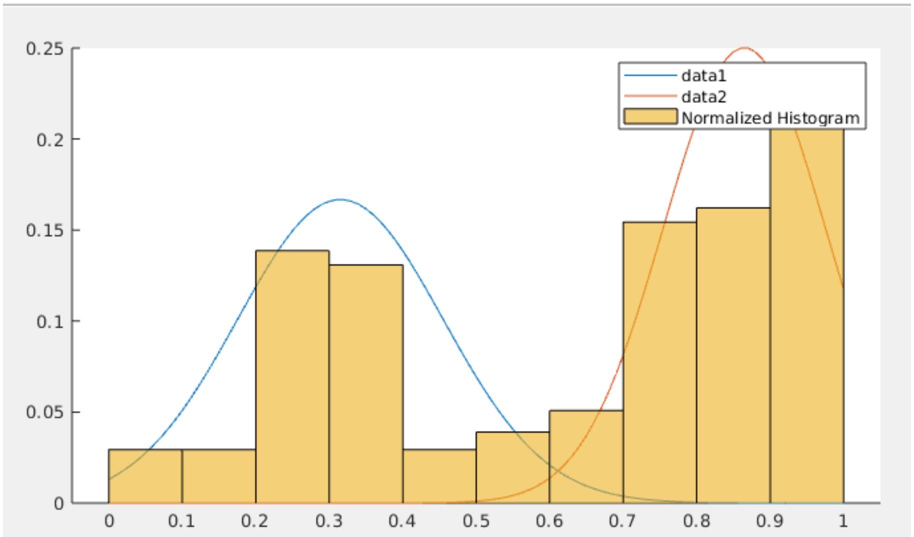
$$\theta^{(0)} = (\pi_1, \pi_2, \mu_1, \sigma_1, \mu_2, \sigma_2) = (0.5, 0.5, 0.3, 0.2, 0.9, 0.1)$$

and after convergence we got the following values:

1	2	3	4	5	6	7
0.4050	0.5950	0.3151	0.1396	0.8648	0.1099	

$$\theta^{(0)} = (\pi_1, \pi_2, \mu_1, \sigma_1, \mu_2, \sigma_2) = (0.4050, 0.5950, 0.3151, 0.1396, 0.8648, 0.1099)$$

here's the plot of the histogram and estimated distributions of the second image:



Here we had to perform manual scaling of our distributions in order to save time because we had a bug in our function, but the estimations were correct.

End.