Q2 $\phi(y, w_k^T x + b_k, \sigma_k^2) = \frac{1}{\sqrt{2\pi}} \cdot \frac{1}{\sqrt{\sigma_k^2}} \times \exp\left(-\left\{\frac{(y_i - w_k^T x_i - b_k)^2}{2\sigma_k^2}\right\}\right)$

where $y \in \mathbb{R}$ & $w_k^T x + b_k \in \mathbb{R}$.

Let $w_k^T x + b_k = \mu_k$.

$\therefore \phi(y, \mu_k, \sigma_k^2) = \frac{1}{\sqrt{2\pi}} \times \frac{1}{\sigma_k^2} \times \exp\left(\frac{-(y-\mu_k)^2}{2\sigma_k^2}\right)$

$\therefore f(y \mid x; \theta) = \sum_{k=1}^{K} \epsilon_k \phi(y, \mu_k, \sigma_k^2)$

& $\epsilon_k \geqslant 0$ & $\sum \epsilon_k = 1$

Let $S \in \{1, \ldots K\}$ be a discrete RV such that

$P_r\{S = k\} = \epsilon_k$

$\therefore P_r(y \in A) = \sum_{k=1}^{K} \left(\int_A \phi(y; \mu_k, \sigma_k^2) \times \epsilon_k \, dy\right.$

Let $\underline{y} = (y_1, \ldots y_n)$ $\therefore$ the likelihood is
& $\underline{x} = (x_1, \ldots x_n)$

~~$L(\theta; x)$~~ $= \prod_{i=1}^{n} f(y_i \mid x_i; \theta)$

$L(\theta, \underline{y} \mid \underline{x}) = \prod_{i=1}^{n} \sum_{k=1}^{K} \epsilon_k \phi(y_i; \mu_k; \sigma_k^2)$

1

$\therefore$ log likelihood is.

$$l(\theta|\underline{x}) = \sum_{i=1}^{n} \log\left(\sum_{k=1}^{K} \epsilon_k \, \phi(y_i; \mu_k, \sigma_k^2)\right)$$

Let $\underline{S} = (S_1 \ldots S_n)$

& $\underline{z} = (\underline{y}, \underline{x}, \underline{s})$. Also, let

$$\Delta_{ik} = \begin{cases} 1 & \text{if } s_i = k \\ 0 & \text{if } s_i \neq k \end{cases}.$$

$\therefore$ The complete data log likelihood is,

$$\log l(\theta; \underline{y} | \underline{x}, \underline{s}) = \sum_{i=1}^{n} \sum_{k=1}^{K} \Delta_{ik} \left[\log \epsilon_k + \log \phi(y_i; \mu_k, \sigma_k^2)\right]$$

Now, looking at E-step, we get.

$$Q(\theta, \theta^{(j)}) = E_{\underline{s}|\underline{y},\underline{x}}\left[l(\theta; \underline{y}|\underline{x}, \underline{s}) \,\middle|\, \underline{y}|\underline{x}; \theta^{(j)}\right]$$

$$\therefore Q(\theta, \theta^{(j)}) = \sum_{i=1}^{n} \sum_{k=1}^{K} \gamma_{i,k}^{(j)} \left[\log \epsilon_k + \log \phi(y_i, \mu_k, \sigma_k^2)\right]$$

$$\& \ \gamma_{i,k}^{(j)} = \frac{\epsilon_k^{(j)} \cdot \phi(y_i, \mu_k^{(j)}, \sigma_k^{(j)})}{\sum_{l=1}^{K} \epsilon_l^{(j)} \, \phi(y_i, \mu_l^{(j)}, \sigma_k^{(j)})}$$

$\therefore$ In terms of $\theta, \theta^{(j)}$ & the data for $Q(\theta, \theta^{(j)})$ is

$$\log l(\theta; \underline{y} \mid \underline{x}, \underline{s}) = \sum_{i=1}^{n} \sum_{k=1}^{K} \Delta_{i,k} \Big[\log \epsilon_K + \log \phi(y_i, \omega_k^T x_i + b_k, \sigma_K^2)\Big]$$

& $$Q(\theta, \theta^{(j)}) = \sum_{i=1}^{n} \sum_{k=1}^{K} \gamma_{(i,k)}^{(j)} \Big[\log \epsilon_K + \log \phi(y_i, \omega_k^T x_i + b_K, \sigma_K^2)\Big]$$

& $$\gamma_{(i,k)}^{(j)} = \frac{\epsilon_K^{(j)} \phi(y_i, \omega_k^T x_i + b_k, \sigma_K^2)}{\sum_{l=1}^{K} \epsilon_l^{(j)} \phi(y_i, \omega_l^T x_i + b_l, \sigma_l^2)}$$

M step:

$$\theta^{(j+1)} = \arg\max_{\theta} Q(\theta, \theta^{(j)})$$

$$= \sum_{i=1}^{n} \sum_{k=1}^{K} \gamma_{i,k}^{(j)} \Big[\log \phi(y_i, \omega_k^T x_i + b_k, \sigma_K^2) + \log \epsilon_K\Big]$$

$$\log \phi(y_i, \omega_k^T x_i + b_k, \sigma_K^2) =$$

$$\log\left(\frac{1}{\sqrt{2\pi}} \times \frac{1}{\sqrt{\sigma_k^2}} \times \exp\left(-\left\{\frac{(y_i - \omega_k^T x_i - b_k)^2}{2\sigma_K^2}\right\}\right)\right)$$

$$= \underbrace{\log\left(\frac{1}{\sqrt{2\pi}}\right)}_{C} - \frac{1}{2}\log(\sigma_k^2) - \left\{\frac{(y_i - \omega_k^T x_i - b_k)^2}{2\sigma_k^2}\right\}$$

$$= \frac{-1}{2}\log(\sigma_k^2) - \frac{(y_i - \omega_k^T x_i - b_k)^2}{2\sigma_k^2}$$

$$\therefore Q(\theta, \theta^{(j)}) = \sum_{i=1}^{n}\sum_{k=1}^{K} \gamma_{i,k}^{(j)} \left[\log \epsilon_k - \frac{1}{2}\log(\sigma_k^2) - \frac{(y_i - \omega_k^T x_i - b_k)^2}{2\sigma_k^2}\right]$$

$$\therefore \frac{\partial Q(\theta, \theta^{(j)})}{\partial b_k} = \cancel{\sum_{i=1}^{n} \gamma_{i,k}^{(j)} \times \frac{1}{\sigma_k^2} \times (y_i - \omega_k^T x_i - b_k) = 0}$$

$$\cancel{\text{or } \sum y_i - \omega_k^T \sum x_i = b_k.}$$

$$\cancel{\text{or } b_k = \bar{y} - \omega_k^T \bar{x}}$$

$$\cancel{\frac{\partial Q(\theta, \theta^{(j)})}{\partial \omega_k}} =$$

$$\frac{\partial Q(\theta, \theta^{(j)})}{\partial b_k} = \sum_{i=1}^{n} \gamma_{i,k}^{(j)} \left[\frac{1}{\sigma_k^2} \times (y_i - \omega_k^T x_i - b_k)\right] = 0$$

$$\Rightarrow b_k^{(j+1)} = \frac{\sum \gamma_{i,k}^{(j)} y_i - \omega_k^{T(j+1)} \sum \gamma_{i,k}^{(j)} x_i}{\sum \gamma_{i,k}^{(j)}}$$

$$\bar{x}_\gamma = \frac{\sum \gamma_{i,k}^{(j)} x_i}{\sum \gamma_{i,k}^{(j)}}$$

$$\bar{y}_\gamma = \frac{\sum \gamma_{i,k}^{(j)} y_i}{\sum \gamma_{i,k}^{(j)}}$$

Similarly.

$$\frac{\partial Q(\theta,\theta^{(j)})}{\partial \omega_k} = \frac{\partial}{\partial \omega_k}\left(\sum_{i=1}^{n} \frac{\gamma_{i,k}^{(j)}}{2\sigma_k^2}\left[-\left(y_i - \omega_k^T x_i - \left(\frac{\bar{y}_\gamma - \omega_k^T \bar{x}_\gamma}{\gamma_{i,k}^{(j)}}\right)\right)^2\right.\right.$$

$$= \frac{\partial}{\partial \omega_k}\left(\sum_{i=1}^{n} - \frac{\gamma_{i,k}^{(j)}}{2\sigma_k^2}\left[\frac{\gamma_{i,k}^{(j)} y_i - \bar{y}_\gamma - \omega_k^T(\gamma_{i,k}^{(j)} x_i - \bar{x}_\gamma)}{\gamma_{i,k}^{(j)}}\right]^2\right)$$

$$\cancel{= + \frac{\gamma_{i,k}^{(j)}}{\sigma_k^2}\left(y_i - \omega_k^T x_i - \frac{\bar{y}_\gamma - \omega_k^T \bar{x}_\gamma}{\gamma_{i,k}^{(j)}}\right) \times (\omega}$$

$$\cancel{= \frac{\partial}{\partial \omega_k}\left(\sum_{i=1}^{n} - \frac{\gamma_{i,k}^{(j)}}{2\sigma_k^2 \gamma_{i,k}^{(j)}} \times \left[\gamma_i^{(j)}\right.\right.}$$

Let $(\gamma_{i,k}^{(j)} y_i - \bar{y}_\gamma)/\sqrt{\gamma_{i,k}} = y_{\alpha i}$

& $(\gamma_{i,k}^{(j)} x_i - \bar{x}_\gamma)/\sqrt{\gamma_{i,k}} = x_{\alpha i}$

$$\cancel{\therefore \frac{\partial}{\partial \omega_k}\left(\sum_{i=1}^{n} - \frac{1}{2\sigma_k^2 \gamma_{i,k}^{(j)}}\left[(y_{\alpha i} - \omega_k^T x_{\alpha i})^2\right]\right.}$$

$$= \cancel{(y_{\alpha i} - \omega_k^T x_{\alpha i})}$$

$$= \frac{\partial}{\partial \omega_k} \sum_{i=1}^{n} (y_{\alpha i} - \omega_k^T x_{\alpha i})^2]$$

$$\Rightarrow \omega_k = (X_\alpha^T X_\alpha)^{-1} X_\alpha^T Y_\alpha$$

where $X_\alpha = \begin{bmatrix} x_{\alpha 1} \\ \vdots \\ x_{\alpha n} \end{bmatrix}$ $\quad Y_\alpha = \begin{bmatrix} y_{\alpha 1} \\ \vdots \\ y_{\alpha n} \end{bmatrix}$

$$\frac{\partial Q(\theta, \theta^{(j)})}{\partial \sigma_k^2} = \sum_{i=1}^{n} \left(-\frac{1}{2} \times \frac{1}{\sigma_k^2} + \frac{(y_i - w_k^T x_i - b_k)^2}{2(\sigma_k^2)^2} \right) \gamma_{i,k}^{(j)} = 0$$

$$\text{or } -\frac{n}{\sigma_k^2} + \frac{\sum \gamma_{i,k}^{(j)} (y_i - w_k^T x_i - b_k)^2}{(\sigma_k^2)^2} = 0$$

If $\sigma_k^2 \neq 0$,

$$\Rightarrow \sigma_k^2 \times n = \sum \gamma_{i,k}^{(j)} (y_i - w_k^T x_i - b_k)^2$$

$$\text{or } \sigma_k^2 = \frac{1}{n} \sum \gamma_{i,k}^{(j)} (y_i - w_k^{T(j+1)} x_i - b_k^{(j+1)})^2$$

Now, for ϵ_k

Let

$$J(\theta, \theta^{(j)}) = \sum_{i=1}^{n} \sum_{k=1}^{K} \gamma_{i,k}^{(j)} \log(\epsilon_k) + \beta \left(\sum_{i=1}^{k} \epsilon_k - 1 \right)$$

$$\frac{\partial J}{\partial \epsilon_k} = \sum_{i=1}^{n} \frac{\gamma_{i,k}}{\epsilon_k} + \beta = 0$$

$$\therefore \epsilon_k = \frac{\sum \gamma_{i,k}^{(j)}}{-\beta}$$

Now, using $\sum_n \epsilon_k = 1$, we get

$$-\beta = \sum_{i=1}^{n} \sum_{k=1}^{K} \gamma_{i,k}^{(j)}$$

$$-\beta = k$$

$$\therefore \epsilon_k = \frac{\sum_{i=1}^{n} \gamma_{i,k}^{(j)}}{n}$$

Scanned by CamScanner

Solving for b_k & ω_k

Let $x_i = [1 \quad x_i]$

$\beta_k = [b_k \quad \omega_k]$

$\therefore$

$$\therefore \frac{\partial Q(\theta, \theta^{(j)})}{\partial \beta_k} = \frac{\partial}{\partial \beta_k}\left(\sum_{i=1}^{k} \frac{\gamma_{i,k}^{(j)}}{\sigma_k^2}(y_i - \beta_k^T x_i)^2\right)$$

$$= \frac{\partial}{\partial \beta_k} \sum_{i=1}^{n} \gamma_{i,k}^{(j)} (y_i - \beta_k^T x_i)^2$$

$$\therefore \beta_k = (X^T W_k^* X)^{-1} X^T W_k^* y$$

where $W_k^* = \begin{bmatrix} \gamma_{1,k}^{(j)} & & & \\ & \gamma_{2,k}^{(j)} & & \\ & & \ddots & \\ & & & \gamma_{n,k}^{(j)} \end{bmatrix}$

d) The # iteration $= 10$

$b = [-0.0192 \quad 0.3160]$

$\omega = [0.5328 \quad -1.3234]$

$V = [3.7 \times 10^{-11} \quad 0.3468]$

$e = [3.2 \times 10^{-5} \quad 1]$

ω^* |

Q1 a) # iterations = 225

The portions which are present at a lot of pixels in the ~~of~~ image & ~~there~~ where there is not sudden change in intensity.

~~Comp~~ Compression ratio = 0.0707

R M Abs. error = 0.0232

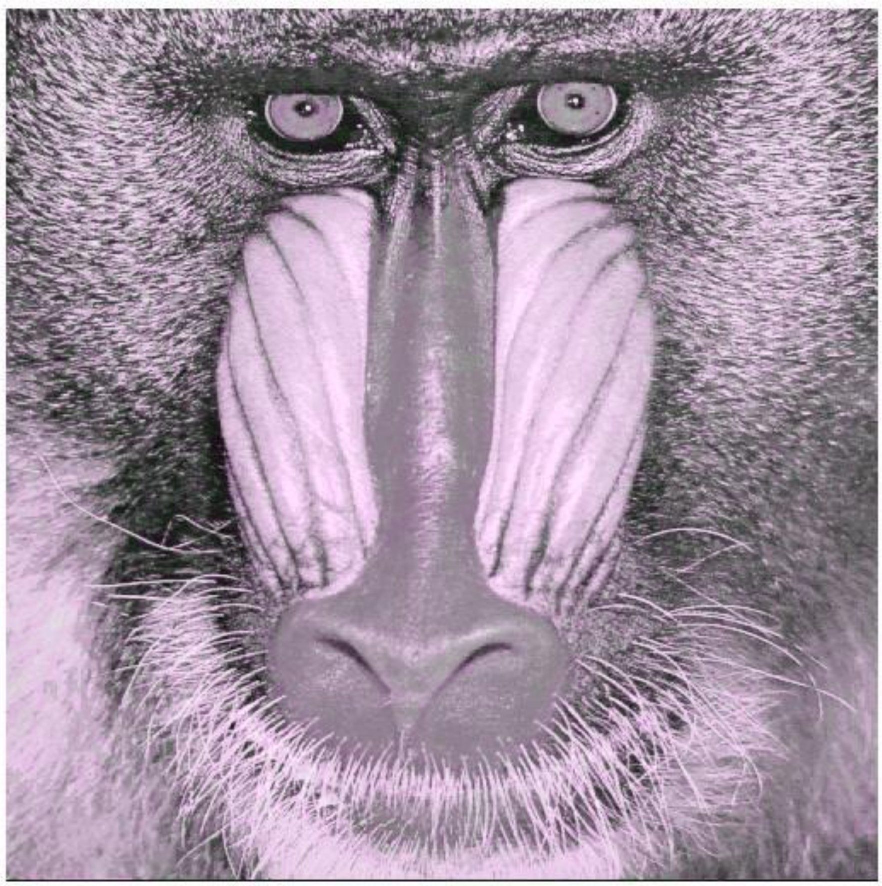
b) bpp = $\frac{k \times m^2 \times 24 + \log_2 k \times N^2/m^2}{N^2}$

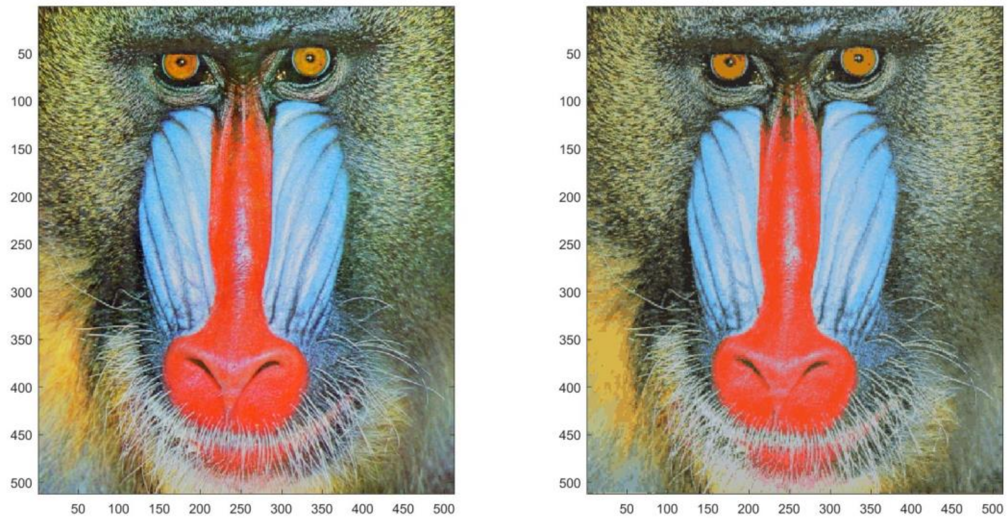
Compression ratio = $\frac{k \times m^2 \times 24 + \log_2 k \times N^2/m^2}{N^2 \times 24}$

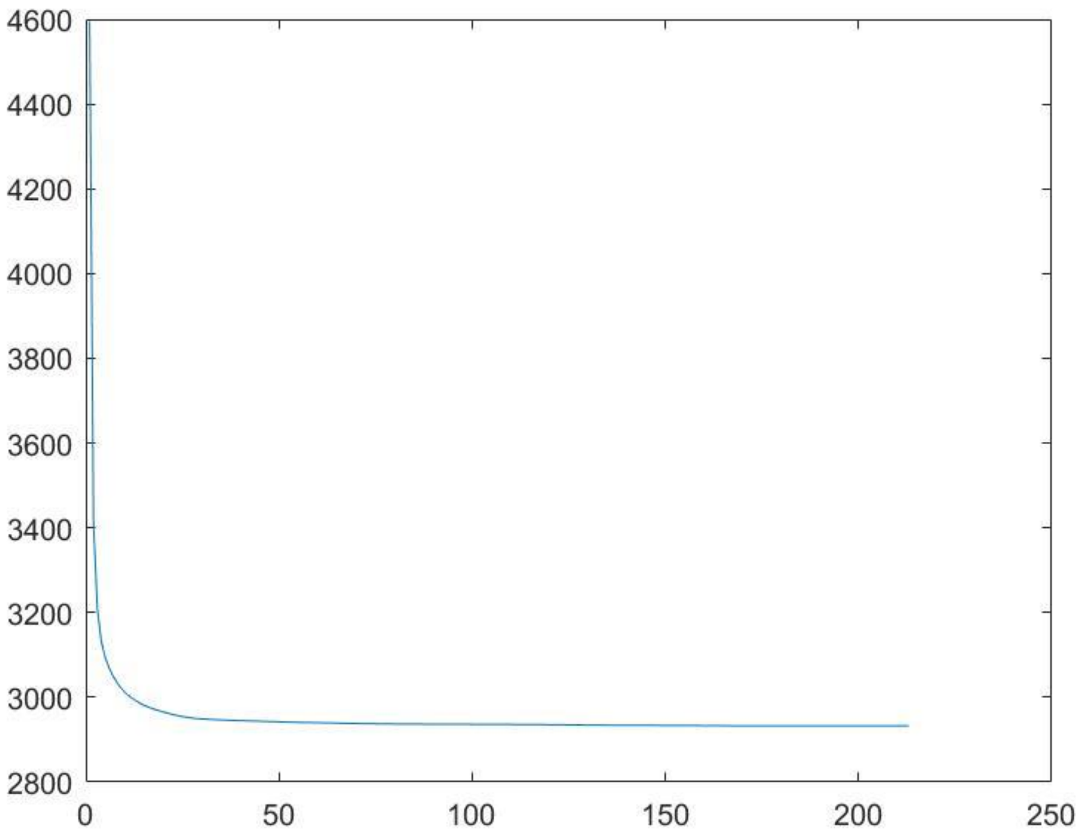
=

```matlab
clc;
clear;

% load image
y = imread('mandrill.tiff');
% extract blocks
M = 2; % block side-length
k = 100;
n = numel(y)/(3*M*M); % number of blocks
d = size(y,1); % image length/width

c=0; % counter
x = zeros(n,3*M*M);
for i=1:M:d % loop through blocks
    for j=1:M:d
        c = c+1;
        x(c,:) = reshape(y(i:i+M-1,j:j+M-1,:),[1,M*M*3]);
    end
end

% K means algorithm
rng(0);
perm = randperm(n);
centroids = x(perm(1:k), :);
obj_vec = zeros(500, 1);
for iter = 1:500
    dist_frm_centroids = dist2(x, centroids);
    [min_dist, cluster_map] = min(dist_frm_centroids,[], 2);
    obj = mean(min_dist);
    obj_vec(iter) = obj;
    centroids_old = centroids;
    centroids = zeros(100, 12);
    for i = 1:k
        centroids(i, :) = mean(x(find(cluster_map == i),:));
    end
    if norm(centroids - centroids_old) < 0.01
        break
    end
end

obj_vec = obj_vec(1:iter);
plot(1:iter, obj_vec)

x_new = centroids(cluster_map,:);

y_new = zeros(512, 512, 3);
c = 0;
for i=1:M:d % loop through blocks
    for j=1:M:d
        c = c+1;
        y_new(i:i+M-1,j:j+M-1,:) = reshape(x_new(c, :), M, M, 3);
    end
end

y_new = uint8(y_new);

subplot(1,2, 1)
```

```
imagesc(y)
subplot(1,2,2)
imagesc(y_new)
drawnow

subplot(1, 1, 1)
imshowpair(y,y_new);

rel_mean_abs_error = sum(abs(y_new(:) - y(:))) / (3 * 512 * 512 * 256);
% 0.0232

N = 512;
compression_ratio = (k * M*M*24 + log2(k) * (N*N)/(M*M))/ (N*N*24)
```

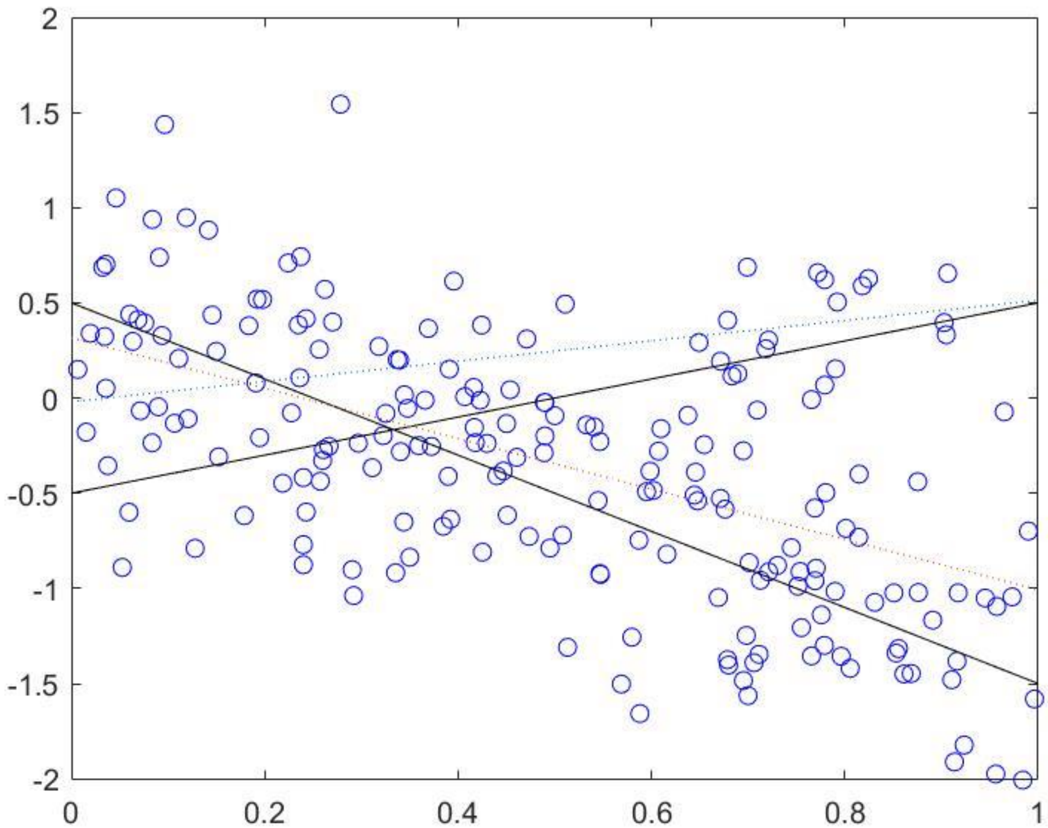


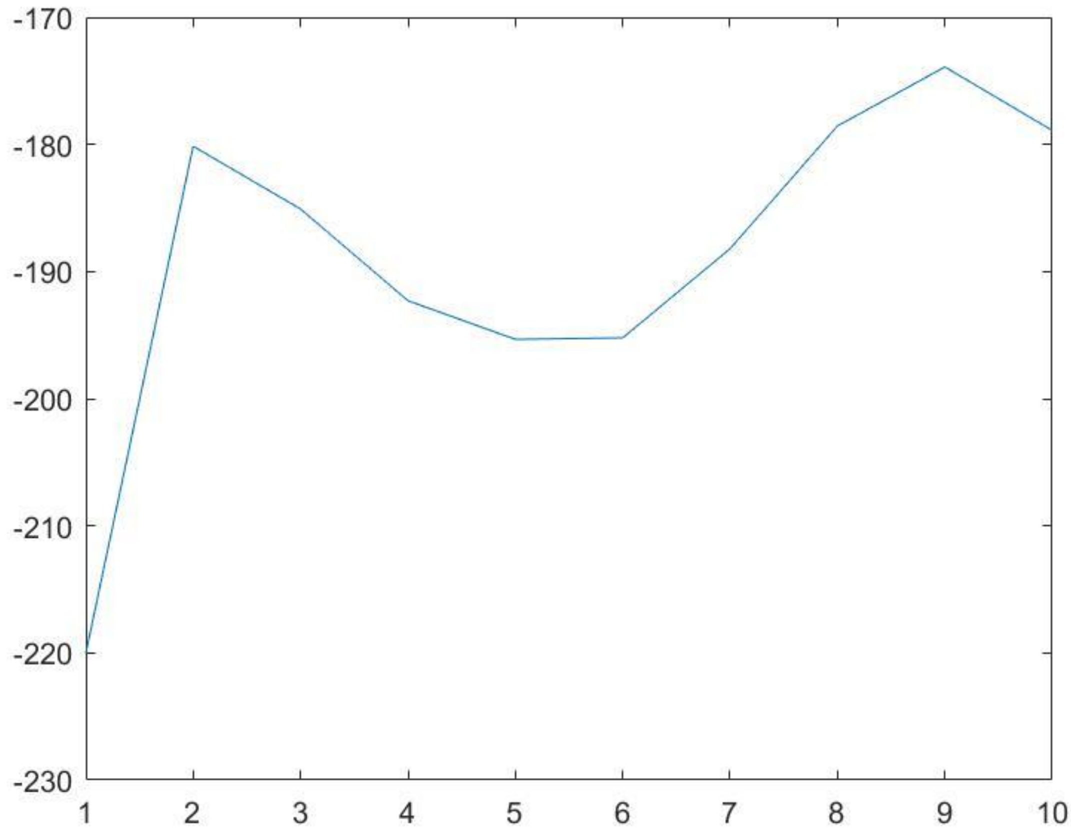

50
100
150
200
250
300
350
400
450
500
50 100 150 200 250 300 350 400 450 500
50
100
150
200
250
300
350
400
450
500
50 100 150 200 250 300 350 400 450 500


4600
4400
4200
4000
3800
3600
3400
3200
3000
2800
0 50 100 150 200 250

```
clear all
close all
rng(0);
n = 200; % sample size
K = 2; % number of lines
e = [.7 .3]; % mixing weights
w = [-2 1]; % slopes of lines
b = [.5 -.5]; % offsets of lines
v = [.2 .1]; % variances
for i=1:n
x(i) = rand;
if rand < e(1);
y(i) = w(1)*x(i) + b(1) + randn*sqrt(v(1));
else
y(i) = w(2)*x(i) + b(2) + randn*sqrt(v(2));
end
end
plot(x,y,'bo')
hold on
t=0:0.01:1;
plot(t,w(1)*t+b(1),'k')
plot(t,w(2)*t+b(2),'k')

% EM algorithm
e = [.5 .5];
w = [1 -1]; % slopes of lines
b = [0 0]; % offsets of lines
v = repmat(var(y), 1, 2); % variances
log_likelihood = zeros(500, 1);
for iter = 1:10
    u = w' * x + b' * ones(n, 1)';
    w_mat = repmat(w', 1, 200);
    x_mat = vertcat(x, x);
    y_mat = vertcat(y, y);
    v_mat = v' * ones(1, n);

    gaussian_pdf = normpdf(y_mat, u, sqrt(v_mat));
    e_mat = repmat(e', 1, 200);
```

```
    phi_num = e_mat .* gaussian_pdf;
    phi_dem = sum(phi_num, 1);
    phi = phi_num ./ repmat(phi_dem, 2, 1);
    log_likelihood(iter,:) =
sum(log(sum(gaussian_pdf .* e_mat,1)));
    if iter ~= 1
        if abs(log_likelihood(iter-1,:) -
log_likelihood(iter,:)) <= (10 ^ (-4))
            break
        end
     end
    x_new = [ones(1,n); x];
    phi_diag_1 = diag(phi(1,:));
    a1 = (inv(x_new * phi_diag_1 * x_new') * x_new
* phi_diag_1 * y')';

    phi_diag_2 = diag(phi(2,:));
    a2 = (inv(x_new * phi_diag_2 * x_new') * x_new
* phi_diag_2 * y')';

    b = [a1(1) a2(1)];
    w = [a1(2) a2(2)];
    u = w' * x + b' * ones(n, 1)';
    v = sum(phi .* ((y_mat - u) .^ 2), 2)' / n;
    e = sum(phi, 2)' / n;
end

hold on
plot(t,w(1)*t+b(1),':')
plot(t,w(2)*t+b(2),':')

hold off
subplot(1, 1, 1);
plot(1:iter, log_likelihood(1:iter,:))
```


2
1.5
1
0.5
0
-0.5
-1
-1.5
-2
0
0.2
0.4
0.6
0.8
1


-170
-180
-190
-200
-210
-220
-230
1
2
3
4
5
6
7
8
9
10