CSE881 HW2

Nan Cao, A52871775

Sep 16th, 2016

Problem 1

(a)
x_1, generated from a uniform distribution in the range between 0 and 1, is like to produce the same bins regardless of whether using equal width or equal frequency approaches.
(b)
x_2, generated from a mixture 3 normal distributions, is more suitable for equal frequency tan equal width discretization approaches.
(c)
x_3, generated frrom an exponetial distribution, is appropriate for neither equal frequency nor equal width discretization approaches.
(d)
After discretization, they're all order attributes.

Problem 2

(a)

$$\begin{aligned}
A &= [a_{ij}]_{N*d} \\
\frac{1}{N-1}\boldsymbol{A}^T[\boldsymbol{i}_N - \frac{1}{N}\mathbf{1}_N]\boldsymbol{A} &= \frac{1}{N-1}[\boldsymbol{A} - [\bar{\boldsymbol{a}}_{\cdot j}]_{N*d}]^T[\boldsymbol{A} - [\bar{\boldsymbol{a}}_{i\cdot}]_{N*d}] \\
&= \frac{1}{N-1}\boldsymbol{A}^T[a_{ij} - \bar{\boldsymbol{a}}_{i\cdot}]_{N*d} \\
&= [\frac{1}{N-1}\sum_{N}^{i=0}(a_{ij} - \bar{\boldsymbol{a}}_{i\cdot})]_{N*d} \\
&= {\sigma_{ij}^2}_{N*d} = C
\end{aligned}$$

(b)

$$\begin{aligned}
(N-1)\boldsymbol{X}\boldsymbol{\Lambda}\boldsymbol{X}^T &= (N-1)\boldsymbol{C} \\
&= \boldsymbol{A}^T[\boldsymbol{I}_N - \frac{1}{N}\mathbf{1}_N]\boldsymbol{A} \\
&= \boldsymbol{A}^T\boldsymbol{I}_N\boldsymbol{A} - \frac{1}{N}\boldsymbol{A}^T\mathbf{1}_N\boldsymbol{A} \\
&= \boldsymbol{V}\boldsymbol{\Sigma}^2\boldsymbol{V}^T
\end{aligned}$$

(c)

$$\begin{aligned}
\boldsymbol{A} \textit{ has been centered} \\
\boldsymbol{C} = \frac{1}{N-1}\boldsymbol{A}^T\boldsymbol{A} &= \boldsymbol{X\Lambda X}^T \\
(\boldsymbol{U\Sigma V}^T)^T\boldsymbol{U\Sigma V}^T &= \boldsymbol{X\Lambda X}^T \\
\boldsymbol{V\Sigma}^T(\boldsymbol{U}^T\boldsymbol{U})\boldsymbol{\Sigma V}^T &= \boldsymbol{X\Lambda X}^T \\
\boldsymbol{V}(\boldsymbol{\Sigma}^T\boldsymbol{\Sigma})\boldsymbol{V}^T &= \boldsymbol{X\Lambda X}^T \\
\boldsymbol{V\Lambda V}^T &= \boldsymbol{X\Lambda X}^T \\
\boldsymbol{V} &= \boldsymbol{X}
\end{aligned}$$

Problem 3

d Class of 0 and Class of 1 are easier to be discerned by the first two components and CLasses of 2 and 3 are harder to be discerned?
e Class of 2 and is easier to be discerned.(From my opinion of view, 0 looks similar to 1 and 2 looks similar to 3, if my outcome picture is correct.)
3f Yes, I can visually discern more digit images correctly with the increasing rank of the matrix W.

Problem 4

For digit1, kernel PCA can better discriminate their images than PCA.

Following are the pictures.

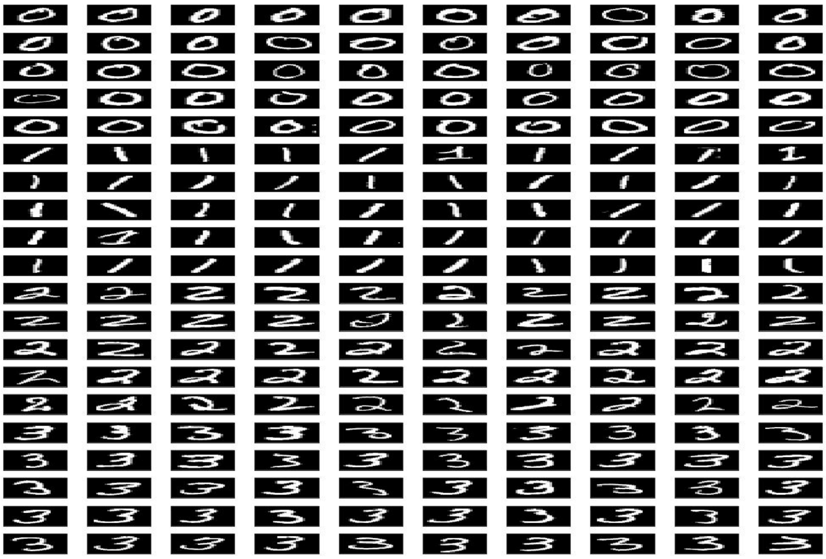

Figure 1: Problem 3b

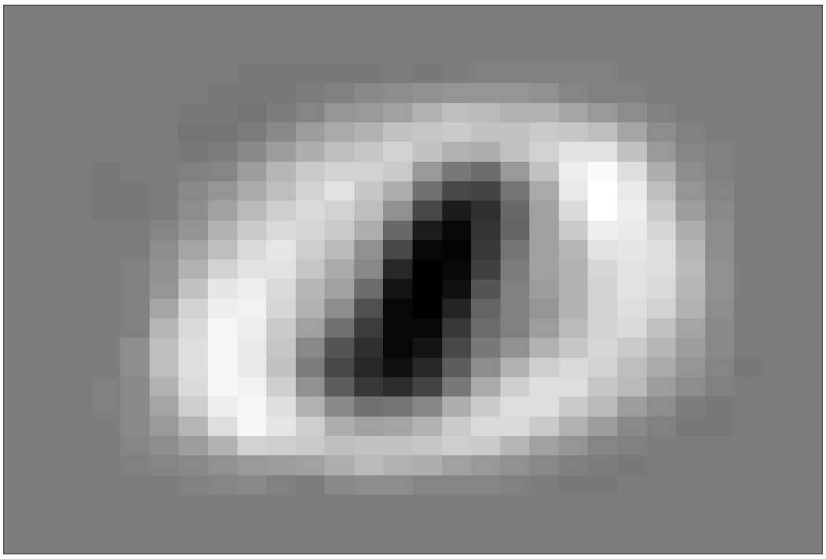

Figure 2: Problem 3c-1

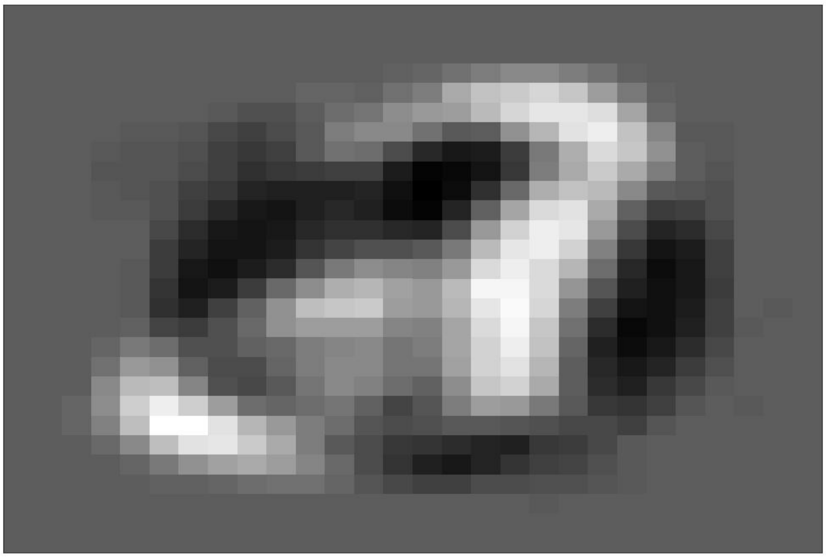

Figure 3: Problem 3c-2

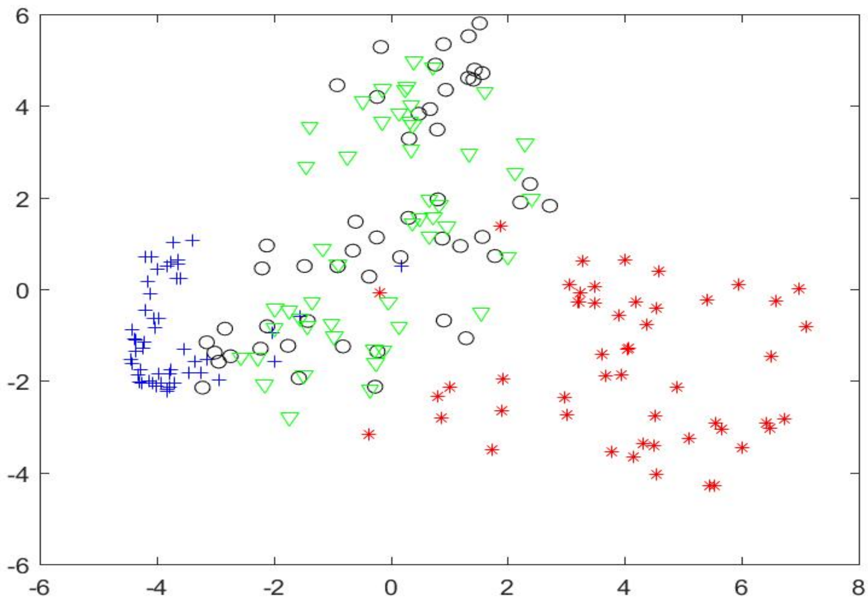

Figure 4: Problem 3d

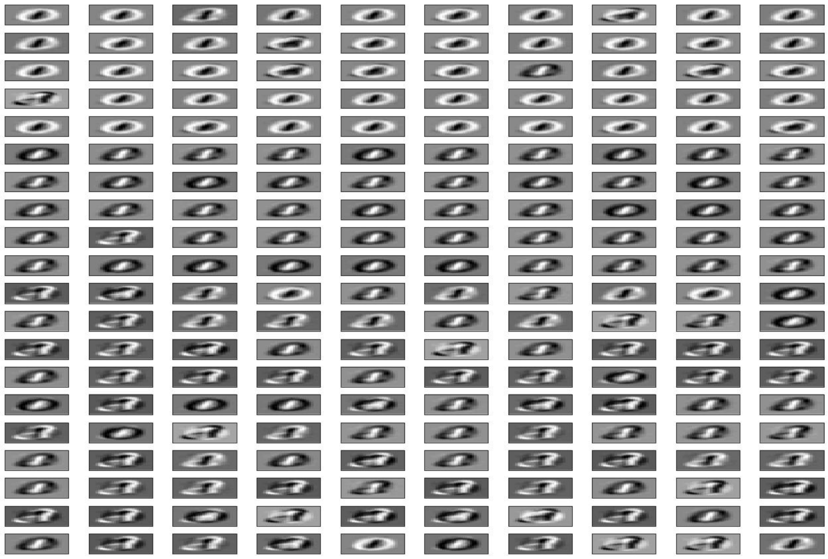

Figure 5: Problem 3e

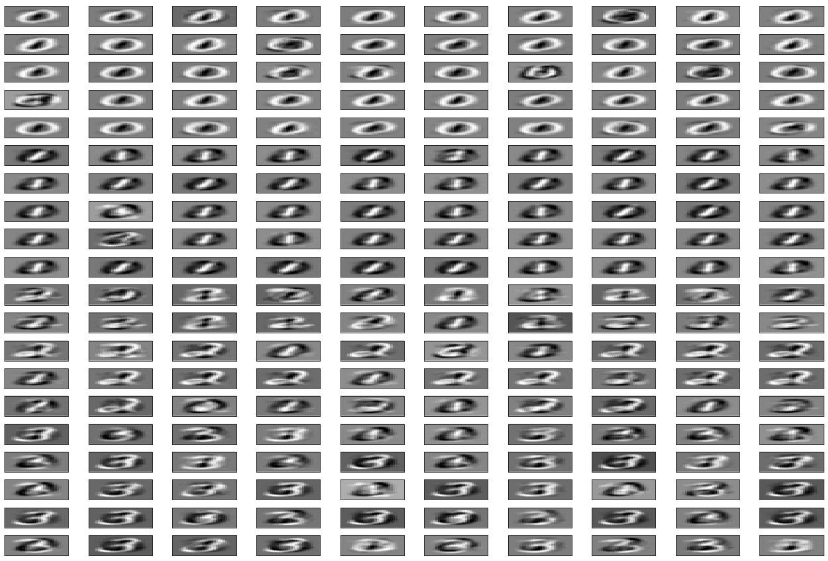

Figure 6: Problem 3f

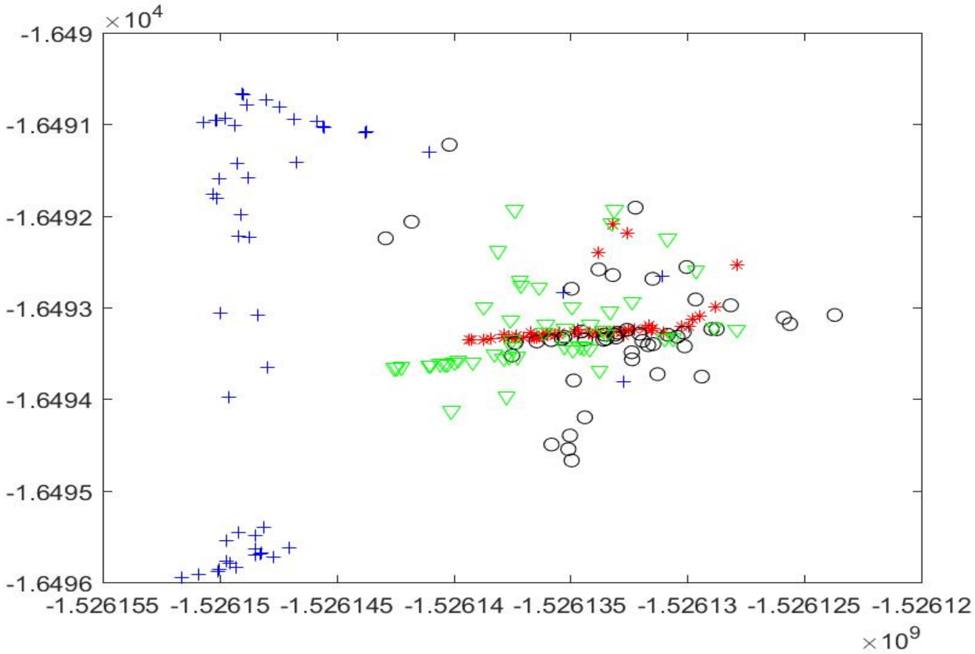

Figure 7: Problem 4 scatter plot