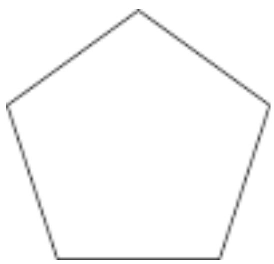
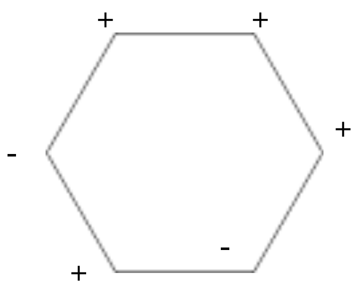
Question 1.

(a).

VC dimension = 5



If we arrange the points like this, we can easily shatter them no matter what label they are assigned.



But for Hexagon, we cannot shatter this situation.

(b).

VC dimension = 2

It cannot shatter segments like + - +

Question 2.

(a).

$$f(x\,|\,\theta) = \frac{1}{\theta}e^{-\frac{x}{\theta}}$$

$$\theta_{MLE} = \mathrm{argmax}_\theta\, L(X\,|\,\theta)$$

$$L(X\,|\,\theta) = \sum \log \frac{1}{\theta}e^{-\frac{x_i}{\theta}} = -N\log\theta - \frac{\sum x_i}{\theta}\log e$$

$$\frac{\partial L(X|\theta)}{\partial \theta} = -\frac{N}{\theta} + \frac{\sum x_i}{\theta^2} \log e = 0$$

$$\theta = \frac{\sum x_i}{N} \log e$$

(b).

$$f(x|\theta) = \theta x^{\theta-1}$$

$$\theta_{MLE} = \mathrm{argmax}_\theta L(X|\theta)$$

$$L(X|\theta) = \sum \log \theta x_i^{\theta-1} = N \log \theta + (\theta - 1) \sum \log x_i$$

$$\frac{\partial L(X|\theta)}{\partial \theta} = \frac{N}{\theta} + \log \prod x_i = 0$$

$$\theta = -\frac{N}{\log \prod x_i} = -\frac{N}{\sum \log x_i}$$

(c).

$$f(x|\theta) = \frac{1}{\theta}$$

$$\theta_{MLE} = \mathrm{argmax}_\theta L(X|\theta)$$

$$L(X|\theta) = \prod f(x_i|\theta) = \frac{1}{\theta^N}$$

The observation here is the less the parameter θ, the larger the likelihood function.

Let $x_{max} = \max(x_1, x_2, \ldots x_N)$, we have $\theta \geq x_{max}$

So $\theta = x_{max}$

Question 3.
Assume:
X1 is the data samples which are labeled 1.
X2 is the data samples which are labeled 2.
(a).
S1 and S2 are learnt independently, to get the MLE of S, we have

$$S_i = \Sigma_i = \frac{\sum (x_i^t - \mu)(x_i^t - \mu)^T}{N}$$

S1 = cov(X1)

S2 = cov(X2)

In my Matlab code,
err_independent = 0.17

(b).
If S1 = S2, we can learn it from the whole training. And the MLE of S is

$$S = \Sigma = \frac{\sum (x^t - \mu)(x^t - \mu)^T}{N}$$

S1 = S2 = cov([X1;X2])
In my Matlab code
err_share = 0.22

(c).
Assume attributes are independent, so in cov(X), all x_{ij} term equals 0 when $i \neq j$
S1 is the diagonal of cov(X1), S2 is the diagonal of cov(X2)
In my Matlab code
err_diagnal = 0.16

(d).
Alpha is the average of the σ of all attribute

$$\frac{\partial L(\chi \,|\, \mu, \Sigma)}{\partial \Sigma^{-1}} = \frac{\partial L}{\partial \alpha} = \frac{\partial \sum -\frac{d}{2}\log 2\pi - \frac{d}{2}\log \alpha - \frac{1}{2}\frac{(x^t-\mu)^T(x^t-\mu)}{\alpha}}{\partial \alpha}$$

Set $\dfrac{\partial(-\frac{d}{2}N\,\log \alpha - \sum_1^N \frac{1}{2\alpha}\sum_i^d (x_i^t - \mu_i)^2)}{\partial \alpha} = 0$

$\alpha = \dfrac{\sum_1^d \sum_1^N (x_i^t - \mu_i)}{dN} = \dfrac{\sum_1^d \sigma_i^2}{d}$, which is the average of variance of each attribute.

S1_dia2 = sum(diag(cov(X1)))/(m-1) * eye(m-1);
S2_dia2 = sum(diag(cov(X2)))/(m-1) * eye(m-1);

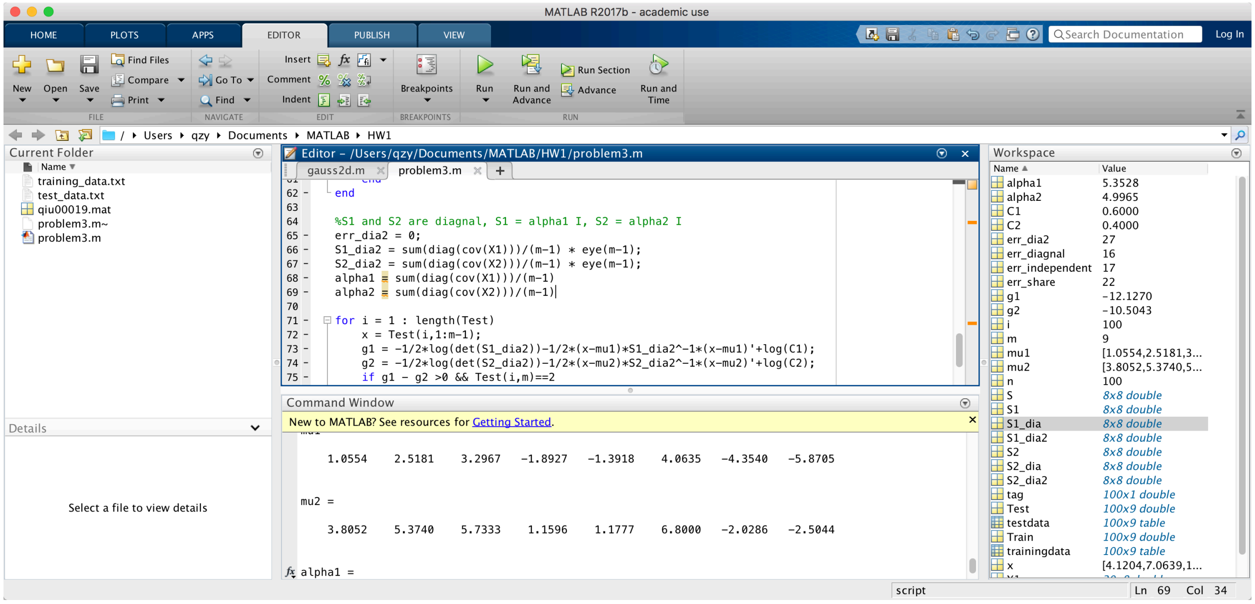
T=In my Matlab code
err_dia2 = 0.27

Clarification:

mu1, mu2, alpha1, alpha2 is shown in command window.

Other values can be seen in work space.

qiu00019.mat contains the raw .txt data


MATLAB R2017b - academic use
HOME
PLOTS
APPS
EDITOR
PUBLISH
VIEW
Search Documentation
Log In
New
Open
Save
Find Files
Compare
Print
FILE
Go To
Find
NAVIGATE
Insert
Comment
Indent
EDIT
Breakpoints
BREAKPOINTS
Run
Run and Advance
Run Section
Advance
Run and Time
RUN
/ Users qzy Documents MATLAB HW1
Current Folder
Name
training_data.txt
test_data.txt
qiu00019.mat
problem3.m~
problem3.m
Details
Select a file to view details
Editor - /Users/qzy/Documents/MATLAB/HW1/problem3.m
gauss2d.m
problem3.m
end
%S1 and S2 are diagnal, S1 = alpha1 I, S2 = alpha2 I
err_dia2 = 0;
S1_dia2 = sum(diag(cov(X1)))/(m-1) * eye(m-1);
S2_dia2 = sum(diag(cov(X2)))/(m-1) * eye(m-1);
alpha1 = sum(diag(cov(X1)))/(m-1)
alpha2 = sum(diag(cov(X2)))/(m-1)
for i = 1 : length(Test)
x = Test(i,1:m-1);
g1 = -1/2*log(det(S1_dia2))-1/2*(x-mu1)*S1_dia2^-1*(x-mu1)'+log(C1);
g2 = -1/2*log(det(S2_dia2))-1/2*(x-mu2)*S2_dia2^-1*(x-mu2)'+log(C2);
if g1 - g2 >0 && Test(i,m)==2
Command Window
New to MATLAB? See resources for Getting Started.
1.0554 2.5181 3.2967 -1.8927 -1.3918 4.0635 -4.3540 -5.8705
mu2 =
3.8052 5.3740 5.7333 1.1596 1.1777 6.8000 -2.0286 -2.5044
alpha1 =
Workspace
Name
Value
alpha1 5.3528
alpha2 4.9965
C1 0.6000
C2 0.4000
err_dia2 27
err_diagnal 16
err_independent 17
err_share 22
g1 -12.1270
g2 -10.5043
i 100
m 9
mu1 [1.0554,2.5181,3...
mu2 [3.8052,5.3740,5...
n 100
S 8x8 double
S1 8x8 double
S1_dia 8x8 double
S1_dia2 8x8 double
S2 8x8 double
S2_dia 8x8 double
S2_dia2 8x8 double
tag 100x1 double
Test 100x9 double
testdata 100x9 table
Train 100x9 double
trainingdata 100x9 table
x [4.1204,7.0639,1...
script
Ln 69 Col 34