Applied-Mathematics-for-CS-HW1

Lizhaoyi SA22011035

November 7, 2022

1 Calculation Problems

1.1

Solution 1:
Note that the probability density function can be expressed as:

$$f(x) = \begin{cases} kx, & o \leq x < 3 \\ 2 - \frac{x}{2}, & 3 \leq x < 4 \\ 0, & others \end{cases} \tag{1}$$

With the fact that the definite integration value of $\int_{-\infty}^{\infty} \mathbf{f(x)\,dx}$ equals to 1, we can induce that:

$$\begin{aligned} &\int_{-\infty}^{\infty} \mathbf{f(x)\,dx} = \int_{-\infty}^{0} 0\,dx + \int_{0}^{3} kx\,dx + \int_{3}^{4} (2 - \frac{x}{2})\,dx + \int_{4}^{\infty} 0\,dx \\ &= (\frac{k}{2}x^2)|_{x=0}^{x=3} + (2x - \frac{x^2}{4})|_{x=3}^{x=4} = \frac{9k}{2} + \frac{1}{4} = 1 \\ &\textbf{which solves that } \mathbf{k} = \mathbf{\frac{1}{6}} \end{aligned}$$

So the probability density function for random variable X is actually:

$$f(x) = \begin{cases} \frac{x}{6}, & o \leq x < 3 \\ 2 - \frac{x}{2}, & 3 \leq x < 4 \\ 0, & others \end{cases} \tag{2}$$

With the fact that $\mathbf{E[X]} = \int_{-\infty}^{\infty} \mathbf{x \cdot f(x)\,dx}$ and $\mathbf{Var[X]} = \mathbf{E[X^2]} - \mathbf{E^2[X]} = \int_{-\infty}^{\infty} \mathbf{x \cdot f(x)\,dx}^{\mathbf{2}} - \int_{-\infty}^{\infty} \mathbf{x^2 \cdot f(x)\,dx}$. We can calculate that:

$$\begin{aligned} &\mathbf{E[X]} = \int_{-\infty}^{\infty} x \cdot f(x)\,dx = \int_{-\infty}^{0} 0\,dx + \int_{0}^{3} \frac{x^2}{6}\,dx + \int_{3}^{4} (2x - \frac{x^2}{2})\,dx + \int_{4}^{\infty} 0\,dx = \mathbf{\frac{7}{3}} \\ &\mathbf{E[X^2]} = \int_{-\infty}^{\infty} x^2 \cdot f(x)\,dx = \int_{-\infty}^{0} 0\,dx + \int_{0}^{3} \frac{x^3}{6}\,dx + \int_{3}^{4} (2x^2 - \frac{x^3}{2})\,dx + \int_{4}^{\infty} 0\,dx = \mathbf{\frac{37}{6}} \\ &\mathbf{Var[X]} = E[X^2] - E^2[X] = \frac{37}{6} - (\frac{7}{3})^2 = \mathbf{\frac{13}{18}} \end{aligned}$$

To conclude, $\mathbf{k} = \mathbf{\frac{1}{6}}$, $\mathbf{E[X]} = \mathbf{\frac{7}{3}}$, $\mathbf{Var[X]} = \mathbf{\frac{13}{18}}$.

1.2

Solution 2:
Let the random variable X represent the the weight of this batch of cement from the factory. Since we already

know the population variance, we can induce that the standard error of average :

$$SE_X = \frac{\delta}{\sqrt{n}} = \frac{\sqrt{2.65}}{\sqrt{18}} = 0.38$$

Since that we want to get the 95%-confidence interval and 99%-confidence interval:
For 95%-confidence interval, the significance level $\alpha = 0.05$, $Z_{\frac{\alpha}{2}} = 1.96$.
and we can induce that 95%-confidence interval for random variable X is:

$$[\bar{X} - Z_{\frac{\alpha}{2}} \cdot SE_X, \bar{X} + Z_{\frac{\alpha}{2}} \cdot SE_X] = [24.9 - 1.96 \times 0.38, 24.9 + 1.96 \times 0.38] = [24.2, 25.6]$$

For 99%-confidence interval, the significance level $\alpha = 0.01$, $Z_{\frac{\alpha}{2}} = 2.58$.
and we can induce that 99%-confidence interval for random variable X is:

$$[\bar{X} - Z_{\frac{\alpha}{2}} \cdot SE_X, \bar{X} + Z_{\frac{\alpha}{2}} \cdot SE_X] = [24.9 - 2.58 \times 0.38, 24.9 + 2.58 \times 0.38] = [23.9, 25.9]$$

1.3

Solution 3:
Let the random variable X represent the the weight of this batch of cement from the factory. Since we do not know the population variance, we can induce that the standard error of average by standard error of samples:

$$SE_X = \frac{std}{\sqrt{n-1}} = \frac{10}{\sqrt{19}} = 2.29$$

Since that we want to get the 95%-confidence interval and 99%-confidence interval:
For 95%-confidence interval, the significance level $\alpha = 0.05$, $T_{\frac{\alpha}{2}} = 2.093$.
and we can induce that 95%-confidence interval for random variable X is:

$$[\bar{X} - T_{\frac{\alpha}{2}} \cdot SE_X, \bar{X} + T_{\frac{\alpha}{2}} \cdot SE_X] = [155 - 2.093 \times 2.29, 155 + 2.093 \times 2.29] = [150.21, 159.79]$$

For 99%-confidence interval, the significance level $\alpha = 0.01$, $T_{\frac{\alpha}{2}} = 2.861$.
and we can induce that 99%-confidence interval for random variable X is:

$$[\bar{X} - T_{\frac{\alpha}{2}} \cdot SE_X, \bar{X} + T_{\frac{\alpha}{2}} \cdot SE_X] = [155 - 2.861 \times 2.29, 155 + 2.861 \times 2.29] = [148.45, 161.55]$$

1.4

Solution 4:
Notation: Betweenness centrality of node $v \triangleright \mathbf{c_v}$.
Notice that nodes: I, H, B are not on any shortest path between any 2 nodes (except for nodes I, H, B).
So according to the definition of the Betweenness centrality, we can induce that: $\mathbf{c_I = c_H = c_B = 0}$.
According to Symmetry, we can induce that: $c_E = c_F$, $c_C = c_A$.
(1) For $\mathbf{c_G}$, notice that the any shortest path between node I and node $X \in \{H, E, F, D, C, A, B\}$ contains node G, and any other shortest path between nodes $X, Y \in \{H, E, F, D, C, A, B\}$ does not contain node G, so we can induce that $\mathbf{c_G = 7}$.
(2) For $\mathbf{c_D}$, notice that the any shortest path between node $X \in \{I, H, G, E, F\}$ and node $Y \in \{C, A, B\}$ contains node D, and any other shortest path between nodes $X, Y \in \{I, H, E, F, G, C, A, B\}$ does not contain node D, so we can induce that $\mathbf{c_D = 3 \times 5 = 15}$.
(3) For $\mathbf{c_E}$, $\mathbf{c_F}$, notice that the half shortest paths between node $X \in \{H, I, G\}$ and node $Y \in \{D, C, A, B\}$ contain node E (and the other half contain node F), and any other shortest path between nodes $X, Y \in \{H, I, G, D, C, A, B\}$ does not contain node E, F, so we can induce that $\mathbf{c_E = c_F = 3 \times 4 \times \frac{1}{2} = 6}$.
(4) For $\mathbf{c_A}$, $\mathbf{c_C}$, notice that the half shortest paths between node $X \in \{B\}$ and node $Y \in \{D, E, F, H, I, G\}$ contain node A (and the other half contain node C), and any other shortest path between nodes $X, Y \in \{H, I, G, D, B, E, F\}$ does not contain node A, C, so we can induce that $\mathbf{c_A = c_C = 1 \times 6 \times \frac{1}{2} = 3}$.
To conclude,

Node	A	B	C	D	E	F	G	H	I
c_{Node}	3	0	3	15	6	6	7	0	0

2 Programming Problem

2.1 Gaussian Mixture Model and EM algorithm

Results of programming 1:
Optional: I will show results for different covariance structures
first, to conclude, I will show the performance for these 4 different covariance type:

TYPE:	FULL	TIED	DIAGONAL	SPHERICAL
ACC:	0.967	0.960	0.900	0.893

The means, covariances and weights are listed as follows:
full:
mean:

	feature1	feature2	feature3	feature4
$Gaussian_1$	5.006	3.428	1.462	0.246
$Gaussian_2$	6.54639415	2.94946365	5.48364578	1.98726565
$Gaussian_3$	5.9170732	2.77804839	4.20540364	1.29848217

covariance:

Gaussian1	feature1	feature2	feature3	feature4
feature1	0.12176499999999997	0.09723199999999989	0.016027999999999973	0.010124000000000005
feature2	0.09723199999999989	0.14081699999999964	0.01146399999999994	0.009112000000000025
feature3	0.016027999999999973	0.01146399999999994	0.029556999999999993	0.005948000000000005
feature4	0.010124000000000005	0.009112000000000025	0.005948000000000006	0.010885000000000008
Gaussian2	feature1	feature2	feature3	feature4
feature1	0.38744092783846173	0.09223275515347401	0.3024430207201765	0.06087396923326452
feature2	0.09223275515347401	0.11040913522103273	0.08385111880047248	0.05574334203172931
feature3	0.3024430207201764	0.0838511188004725	0.3258957418180867	0.07276775904883839
feature4	0.06087396923326451	0.05574334203172931	0.07276775904883839	0.08484505047999581
Gaussian3	feature1	feature2	feature3	feature4
feature1	0.27551709999389695	0.09662294713804297	0.1854707207322301	0.05478900528729243
feature2	0.09662294713804297	0.09255152043028306	0.09103430846133617	0.0429989947992187
feature3	0.1854707207322301	0.09103430846133617	0.20235849310079082	0.06171382694220664
feature4	0.05478900528729243	0.0429989947992187	0.06171382694220664	0.03233774636980292

weights:

Gaussian	$Gaussian_1$	$Gaussian_2$	$Gaussian_3$
weights	0.333	0.365	0.301

tied:
mean:

	feature1	feature2	feature3	feature4
$Gaussian_1$	5.006	3.428	1.462	0.246
$Gaussian_2$	6.54639415	2.99716202	5.55284949	2.05143145
$Gaussian_3$	5.95728298	2.75675429	4.31039897	1.33031316

covariance:
For Corvariancetype = tied, 3 Covariance Matrixs are identical for 3 Gaussian distributions

Gaussian	feature1	feature2	feature3	feature4
feature1	0.263331700398019	0.08760866750172987	0.17315729974004476	0.03743309815944638
feature2	0.08760866750172987	0.11040074252195713	0.04836891387035546	0.027011423996912833
feature3	0.17315729974004174	0.04836891387035546	0.20285285254138927	0.04347376256154576
feature4	0.03743309815944638	0.027011423996912833	0.04347376256154576	0.03619052614733872

weights:

Gaussian	$Gaussian_1$	$Gaussian_2$	$Gaussian_3$
weights	0.333	0.320	0.347

diagonal:
mean:

	feature1	feature2	feature3	feature4
$Gaussian_1$	5.006	3.428	1.462	0.246
$Gaussian_2$	6.8060823	3.07023103	5.71889409	2.10305305
$Gaussian_3$	5.92570673	2.74947486	4.40355614	1.41204165

covariance:
For Corvariancetype = diagonal, 3 Covariance Matrixs are diagonal matrixs.

Gaussian	feature1	feature2	feature3	feature4
feature1	0.121765	0	0	0
feature2	0	0.140817	0	0
feature3	0	0	0.029557	0
feature4	0	0	0	0.010885
Gaussian	feature1	feature2	feature3	feature4
feature1	0.28520059	0	0	0
feature2	0	0.08200975	0	0
feature3	0	0	0.25128562	0
feature4	0	0	0	0.06109484
Gaussian	feature1	feature2	feature3	feature4
feature1	0.23145596	0	0	0
feature2	0	0.08738014	0	0
feature3	0	0	0.27563102	0
feature4	0	0	0	0.0688713

weights:

Gaussian	$Gaussian_1$	$Gaussian_2$	$Gaussian_3$
weights	0.333	0.255	0.412

spherical:
mean:

	feature1	feature2	feature3	feature4
$Gaussian_1$	5.006	3.428	1.462	0.246
$Gaussian_2$	6.84191531	3.07023103	3.0717748	2.07049253
$Gaussian_3$	5.90165943	2.74786641	4.39838776	1.43087513

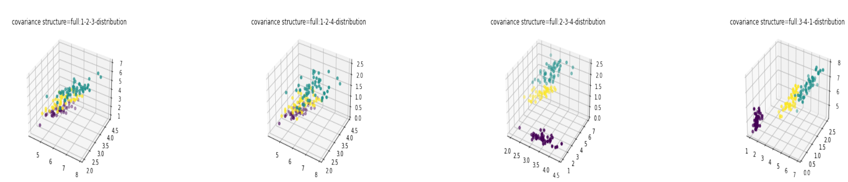
covariance:
For Corvariancetype = spherical, 3 Covariance Matrixs are diagonal matrixs, and the elements in one matrix are entirely indentical.

Gaussian	feature1	feature2	feature3	feature4
feature1	0.075756	0	0	0
feature2	0	0.075756	0	0
feature3	0	0	0.075756	0
feature4	0	0	0	0.075756
Gaussian	feature1	feature2	feature3	feature4
feature1	0.1644286	0	0	0
feature2	0	0.1644286	0	0
feature3	0	0	0.1644286	0
feature4	0	0	0	0.16442866
Gaussian	feature1	feature2	feature3	feature4
feature1	0.16244304	0	0	0
feature2	0	0.16244304	0	0
feature3	0	0	0.16244304	0
feature4	0	0	0	0.16244304

weights:

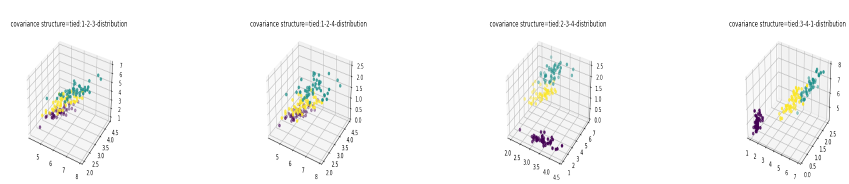
Gaussian	$Gaussian_1$	$Gaussian_2$	$Gaussian_3$
weights	0.333	0.256	0.411

the distribution graphs are listed as follows:(for there is 4 different features for each datapoint, I will choose 3 out of them to draw distribution figure.)



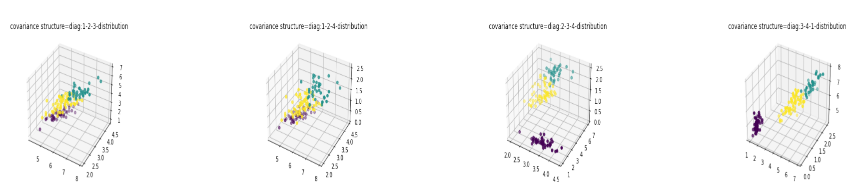
(a) (b) (c) (d)

Figure 1: FULL



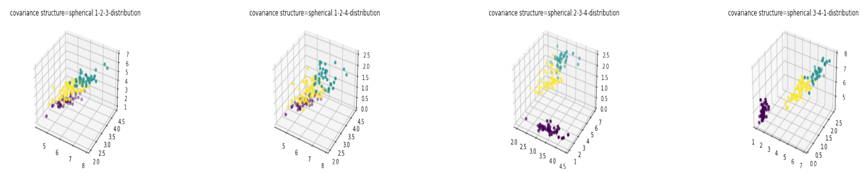
(a) (b) (c) (d)

Figure 2: TIED



(a) (b) (c) (d)

Figure 3: DIAG



(a) (b) (c) (d)

Figure 4: SPHERICAL

2.2 PageRank

Results of programming 2:
Refering Google's PageRank Algorithm, initally I set hyerparameter β as 0.85.
Please find my 'PageRankresult.csv' file in the attachment.
Top20 nodes and their pagerank values:
top20-id:
'7237', '7498', '7339', '7162', '7224', '7435',
'6519', '7100', '7199', '7595', '4811', '6101',
'7536', '4785', '7489', '7587', '7488', '6617',
'7226', '7416'.
top20-pagerank:
0.007385263137179232, 0.00681009860765043,
0.006406924018043255, 0.003162036786812069,
0.0029170626276129818, 0.002763272800815785,
0.0027062061687505874, 0.0025896020131115152,
0.0025699388989470213, 0.002563931924916656,
0.0024410816119636507, 0.0023056856477816105,
0.00224877983433263, 0.0022168779223014656,
0.001975366361434149, 0.0018521968597531642,
0.0018327286451309875, 0.0017622294437587298,
0.001729532478683448, 0.0017161698520741196.
Optional:
I increase the hyper-parameter β from $0.1 \rightarrow 0.2 \rightarrow 0.3 \rightarrow ... \rightarrow 1.0$
the mean of top20 PageRank value varies as follows:

β	0.1	0.2	0.3	0.4	0.5
mean of top20 pagerank	0.00040	0.00066	0.00095	0.00124	0.00157
β	0.6	0.7	0.8	0.9	1.0
mean of top20 pagerank	0.00192	0.00232	0.00276	0.00326	0.00379

To conclude, the larger (β), the less chance for one node randomly jumping to other nodes, the larger (top page-rank values), the lower(for variance of the page-rank values for all the nodes)