

CO553 - Introduction to Machine Learning: Unsupervised Learning

Prepared by Josiah Wang

Autumn 2020/2021

Sample answers

The following are the final answers to the questions on unsupervised learning. We are deliberately not providing you with the full workings; these are meant for you to use as a sanity check after you have tried out the exercises yourself. If you are struggling to reproduce the given answers, please post your workings on Piazza and we can help you figure out where you went wrong!

1. (a) Cluster assignments:

i	1	2	3	4	5	6	7	8	9	10
$D(\mu_1, x^{(i)})$	1.1985	1.2307	1.6985	1.7501	1.1611	0.8504	0.1844	2.0436	1.1693	2.2583
$D(\mu_2, x^{(i)})$	1.3446	2.1471	1.6776	0.2400	2.9560	2.2755	2.1352	0.5459	2.8804	0.6935
$c^{(i)}$	1	1	2	2	1	1	1	2	1	2

(b) $\mu_1 = \begin{bmatrix} 0.0317 \\ 0.0900 \end{bmatrix}; \mu_2 = \begin{bmatrix} 1.3725 \\ 1.5300 \end{bmatrix}$

2. 0.1622

3. $\mu = 4.414; \sigma^2 = 5.8423; p(x = 3.13 | \mu, \sigma^2) = 0.1433$

4. GMM-EM:

- (a) r_{ik} :

r_{ik}	$k = 1$	$k = 2$
$i = 1$	0.0353	0.9647
$i = 2$	0.9989	0.0011
$i = 3$	0.9203	0.0797
$i = 4$	0.2274	0.7726
$i = 5$	0.9997	0.0003

(b) $\mu_1 = 2.8145; \sigma_1^2 = 1.2594; \pi_1 = 0.6363$
 $\mu_2 = 5.5078; \sigma_2^2 = 0.2674; \pi_2 = 0.3637$

5. 0.2305