

Bioinformatics

21L-5294 BCS-8B

Exercise 1

Centres $(2,4)$, $(6,7)$, $(7,3)$

$$DC(8,7) : \sqrt{6^2 + 3^2} = 6.7$$

$$\sqrt{2^2 + 0^2} = 2 \quad \checkmark$$

$$\sqrt{1^2 + 4^2} = 4.123$$

$$DC(1,6) : \sqrt{1^2 + 2^2} = 2.23 \quad \checkmark$$

$$\sqrt{5^2 + 1^2} = 5.099$$

$$\sqrt{6^2 + 3^2} = 6.7$$

$$DC(5,6) : \sqrt{3^2 + 2^2} = 3.6$$

$$\sqrt{1^2 + 1^2} = 1.4 \quad \checkmark$$

$$\sqrt{2^2 + 3^2} = 3.6$$

$$DC(3,4) : \sqrt{1^2 + 0^2} = 1 \quad \checkmark$$

$$\sqrt{3^2 + 3^2} = 4.24$$

$$\sqrt{4^2 + 1^2} = 4.123$$

$$DC(1,3) : \sqrt{1^2 + 1^2} = \cancel{\sqrt{1.414}} \quad \checkmark$$

$$\sqrt{5^2 + 4^2} = 6.403$$

$$\sqrt{8^2 + 0^2} = 8$$

$$DC(10,3) : \sqrt{8^2 + 1^2} = 8.06$$

$$\sqrt{4^2 + 4^2} = 5.65$$

$$\sqrt{3^2 + 0^2} = 3 \quad \checkmark$$

$$DC(5,2) : \sqrt{3^2 + 2^2} = 3.6 \quad \cancel{\text{as}}$$

$$\sqrt{1^2 + 5^2} = 5.099$$

$$\sqrt{2^2 + 1^2} = 2.23 \quad \checkmark$$

$$D(7,1) : \sqrt{5^2+3^2} = 5.83$$

$$\sqrt{1^2+6^2} = 6.08$$

$$\sqrt{0^2+2^2} = 2 \checkmark$$

$$\max(2, 2.23, 1.4, 1, 1.4143, 2.23, 2)$$

$$= 3$$

Exercise 2

Centres $(3,4.5), (6,1.5), (9,5)$

$$D(8,7)$$

$$\sqrt{5^2+25^2} = 5.59 \quad \sqrt{2^2+55^2} = 8.85 \quad \sqrt{1^2+2^2} = 2.23 \checkmark$$

$$D(1,6) \checkmark$$

$$\sqrt{2^2+15^2} = 2.5 \quad \sqrt{5^2+4.5^2} = 6.72 \quad \sqrt{8^2+1^2} = 8.06$$

$$D(5,6)$$

$$\sqrt{2^2+15^2} = 2.5 \quad \sqrt{1^2+4.5^2} = 4.6 \quad \sqrt{4^2+8^2} = 4.123$$

$$D(3,4)$$

$$\sqrt{0^2+0.5^2} = 0.5 \checkmark \quad \sqrt{3^2+2.5^2} = 3.9 \quad \sqrt{6^2+1^2} = 6.08$$

$$D(1,3) \checkmark$$

$$\sqrt{2^2+1.5^2} = 2.5 \quad \sqrt{5^2+1.5^2} = 5.22 \quad \sqrt{8^2+2^2} = 8.24$$

$$D(10,3)$$

$$\sqrt{7^2+1.5^2} = 7.1589 \quad \sqrt{4^2+1.5^2} = 4.27 \quad \sqrt{1^2+2^2} = 2.23$$

$$D(5,2)$$

$$\sqrt{2^2+2.5^2} = 3.2 \quad \sqrt{1^2+0.5^2} = 1.11 \quad \sqrt{4^2+3^2} = 5$$

$$D(7,1)$$

$$\sqrt{4^2+3.5^2} = 5.31 \quad \sqrt{1^2+0.5^2} = 1.11 \quad \sqrt{2^2+4^2} = 4.47$$

$$\max(\underline{x}) =$$

$$\max(2.23, 2.5, 2.5, 0.5, 2.5, 2.23, 1.11, 1.11)$$

$$= 2.5$$

$$\text{distortion} = \frac{1}{8} (2.23^2 + 2.5^2 + 2.5^2 + 0.5^2 + 2.5^2 + 2.23^2 + 1.11^2 + 1.11^2) \\ = 3.92$$

$$\text{Centres } (5/3, 13/3), (6.5, 6.5), (22/3, 2)$$

	(1.66, 4.33)	(6.5, 6.5)	(7.33, 2)
(8, 7)	6.87	1.58 ✓	5.044
(1, 6)	1.795 ✓	5.52	7.48
(5, 6)	3.73	1.58 ✓	4.629
(3, 4)	1.38 ✓	4.3	4.769
(1, 3)	1.48 ✓	6.519	6.408
(5, 2)	4.07	4.74	2.33 ✓
(10, 3)	8.44	4.94	2.85 ~
(7, 1)	6.29	5.52	105 ✓

$$\max(1.58, 1.795, 1.58, 1.38, 1.48, 2.33, 2.85, 1.05)$$

$$= 2.85$$

$$\text{distortion} = \frac{1}{8} (1.58^2 + 1.795^2 + 1.58^2 + 1.38^2 + 1.48^2 + 2.33^2 + 2.85^2 + 1.05^2)$$

$$= 3.13$$

Conclusion: even though the max distance increased we saw the ^{second} centres have lower distortion. showing better centres.

Exercise 3

$(0, 1, 1.9, 3)$

0	1	1.9	3	
✓	✓	✓	✓	
1	1	0	0.9	2

$$\text{Distortion} = \frac{1}{4}(1^2 + 0^2 + 0.9^2 + 0) = 0.4525$$

cluster 1 : $(1, 0, 0.9)$

$$\text{centre} = \frac{1 + 0 + 0.9}{3} = 0.633$$

cluster 2 = $\frac{3}{1} = 3$

0	1	1.9	3	
✓	✓	✓	✓	
0.633	0.633	0.367	1.267	2.367

3	2	1.1	0	
✓	✓	✓	✓	

$$\text{Distortion} = \frac{1}{4}(0.633^2 + 0.367^2 + 1.1^2 + 0) = 0.4363$$

cluster 1 : $\frac{0+1}{2} = 0.5$

cluster 2 : $\frac{1.9+3}{2} = 2.45$

0	1	1.9	3	
✓	✓	✓	✓	

0.5	0.5	0.5	1.4	2.5	
✓	✓	✓	✓	✓	

2.45	2.45	1.45	0.55	0.55	
✓	✓	✓	✓	✓	

$$\text{Distortion} = \frac{1}{4}(0.5^2 + 0.5^2 + 0.55^2 + 0.55^2) = 0.27025$$

cluster 1 : $\frac{1+0}{2} = 0.5$

cluster 2 : $\frac{1.9+3}{2} = 2.45$

Exercise 4

	g_1	g_2	g_3	g_4	g_5	g_6	g_7	g_8	g_9	g_{10}
g_1	0									
g_2	8.1	0								
g_3	9.2	12	0	.	.					
g_4	7.7	0.9	11.2	0	.					
g_5	9.3	12	0.7	11.2	0	.				
g_6	2.3	9.5	11.1	9.2	11.2	0				
g_7	5.1	10.1	8.1	9.5	8.5	5.6	0			
g_8	10.2	12.8	1.1	12	1	12.1	9.1	0		
g_9	6.1	2	10.5	1.6	10.6	7.7	8.3	11.4	0	
g_{10}	7.0	1	11.5	1.1	11.6	8.5	9.3	12.4	11	0

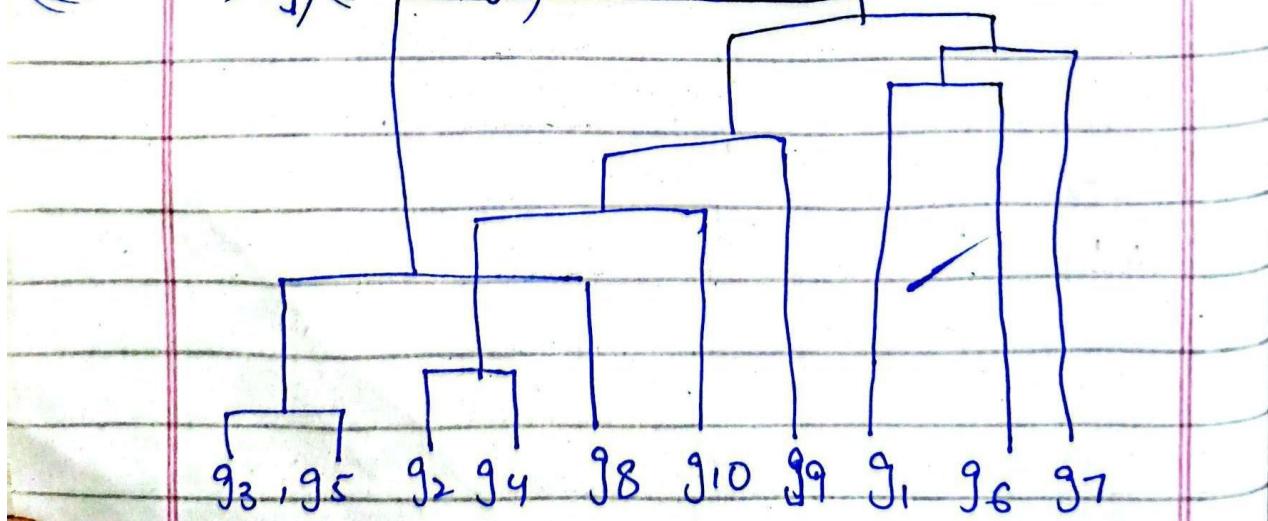
	g_1	g_6	g_7	$\cancel{g_8}$	g_9	g_{10}	$\cancel{g_{11}}$	g_{12}, g_{14}
g_1	0							
g_6	2.3	0						
g_7	5.1	5.6	0					
$\cancel{g_8}$	10.2	12.1	9.1	0				
g_9	6.1	7.7	8.3	11.4	0			
g_{10}	7	8.5	9.3	12.4	1.1	0		
$\cancel{g_{11}}$	9.2	11.1	8.1	(1)	10.5	11.5	0	
g_{12}, g_{14}	7.7	9.2	9.5	12	1.6	1	11.2	0

	g_1	g_6	g_7	g_9	$\cancel{g_8}$	$\cancel{g_{10}}$	$(g_3, g_5)g_8$
g_1	0						
g_6	2.3	0					
g_7	5.1	5.6	0				
$\cancel{g_8}$	6.1	7.7	8.3	0			
$\cancel{g_{10}}$	7	8.5	9.3	1.1	0		
$\cancel{g_{11}}$	7.7	9.2	9.5	1.6	(1)	0	
g_{12}, g_{14}	9.2	11.1	8.1	10.5	11.5	11.7	(1)

	g_1	g_6	g_7	$\cancel{g_8}$	$(g_3, g_5)g_8$	$\cancel{(g_3, g_5)g_8}$
g_1	0					
g_6	2.3	0				
g_7	5.1	5.6	0			
$\cancel{g_8}$	6.1	7.7	8.3	0		
$(g_3, g_5)g_8$	9.2	11.1	8.1	10.5	0	
$\cancel{(g_3, g_5)g_8}$	7	8.5	9.3	(1)	11.2	0

	g_3	g_{35}	g_7	$(\text{g}_3, \text{g}_5)\text{g}_8$	$((\text{g}_2, \text{g}_4)\text{g}_{10})\text{g}_9$
g_3	○				
g_{35}	2.3	○			
g_7	5.1	5.6	○		
$(\text{g}_3, \text{g}_5)\text{g}_8$	9.2	11.1	8.1	○	
$((\text{g}_2, \text{g}_4)\text{g}_{10})\text{g}_9$	9.1	7.7	8.3	10.5	○

	g_{35}	$(\text{g}_3, \text{g}_5)\text{g}_8$	$((\text{g}_2, \text{g}_4)\text{g}_{10})\text{g}_9$	$\cancel{\text{g}_1, \text{g}_6}$
g_{35}	○			
$(\text{g}_3, \text{g}_5)\text{g}_8$	8.1	○		
$((\text{g}_2, \text{g}_4)\text{g}_{10})\text{g}_9$	8.3	10.5	○	
$\cancel{\text{g}_1, \text{g}_6}$	5.1	9.2	6.1	○
		$(\text{g}_3, \text{g}_5)\text{g}_8$	$((\text{g}_2, \text{g}_4)\text{g}_{10})\text{g}_9$	$(\text{g}_1, \text{g}_6)\text{g}_7$
$(\text{g}_3, \text{g}_5)\text{g}_8$	○			
$\cancel{(\text{g}_2, \text{g}_4)\text{g}_{10}\text{g}_9}$	10.5			
$\cancel{(\text{g}_1, \text{g}_6)\text{g}_7}$	8.1	6.1	○	
$(\text{g}_3, \text{g}_5, \text{g}_8)$		$(\text{g}_3, \text{g}_5)\text{g}_8$	$((\text{g}_2, \text{g}_4)\text{g}_{10})\text{g}_9$	$((\text{g}_1, \text{g}_6)\text{g}_7)$
$((\text{g}_2, \text{g}_4)\text{g}_{10})\text{g}_9$			8.1	○
$((\text{g}_2, \text{g}_4)\text{g}_{10})\text{g}_9$				



	g ₁	g ₂	g ₃	g ₄	g ₅	g ₆	g ₇	g ₈	g ₉	g ₁₀
g ₁	0									
g ₂	8.1	0								
g ₃	9.2	12	0							
g ₄	7.7	0.9	11/2	0						
g ₅	9.3	12	0.7	11.2	0					
g ₆	2.3	9.5	11.1	9.2	11.2	0				
g ₇	5.1	10.1	8.1	9.5	8.5	5.0	0			
g ₈	0.2	12.8	1.1	12	1	12.1	9.1	0		
g ₉	6.1	2	10.5	1.6	10.6	7.7	8.3	11.4	0	
g ₁₀	7	1	11.5	1.1	11.6	8.5	9.3	12.4	1.1	0

	g_1	g_2	g_3	g_6	g_7	g_8	g_9	g_{10}	g_{395}
g_1	0								
g_2	8.1	0							
g_3	7.7	0.9	0						
g_6	2.3	9.5	9.2	0					
g_7	5.1	10.1	9.5	5.6	0				
g_8	10.2	12.8	12	12.1	9.1	0			
g_9	6.1	2	1.6	12.6 7.7 8.5	8.3 9.3 8.3	11.4 12.4 8.3	11.1 11.1 12.5	0	
g_{10}	7	1	1.1	4.6 4.6	8.3 8.3	8.3 8.3	12.5 12.5	0	
g_{395}	9.25	12	11.4	11.15	8.3	1.05	10.55	11.55	0

	g_1	g_6	g_7	g_{10}	g_9	g_{10}	$\cancel{g_{10}}$	$g_2 g_4$
g_1	0							
g_6	2.3	0						
g_7	5.1	5.6	0					
g_{10}	10.2	12.1	9.4	0				
g_9	6.1	7.7	8.3	11.4	0			
g_{10}	7	8.5	9.3	12.4	1.1	0		
g_{10}	9.25	11.15	8.3	1.05	10.55	11.55	1.4	0
$g_2 g_4$	7.9	9.35	9.8	12.4	2.8	1.05	11.6	0

	g_1	g_6	g_7	g_{10}	g_9	g_{10}	$g_2 g_4$	$g_3 g_5 g_8$
g_1	0							1
g_6	2.3	0						
g_7	5.1	5.6	0					
g_{10}	6.1	7.7	8.3	0				
g_{10}	7	8.5	9.3	1.1	0			
$g_2 g_4$	7.9	9.35	9.8	1.4	1.05	0		
$g_3 g_5 g_8$	9.25	11.15	8.3	10.75	11.75	12.5	0	0

	g_1	g_6	g_7	g_{10}	g_9	g_{10}	$g_2 g_4$	$g_3 g_5 g_8$
g_1	0							
g_6	2.3	0						
g_7	5.1	5.6	0					
g_{10}	7.9	9.35	9.8	0				
$g_3 g_5 g_8$	9.25	11.15	8.3	1.2	0			
$g_2 g_4$	6.55	8.1	8.8	1.225	11.975	0		

