# CS 5525

Biplav Choudhury (906089826)

19 October 2020

## 1 Solutions to Assignment 5

### 1.1 Problem 1

#### 1.1.1 Part 1

After 3 iterations, the the centroids are 0.1, 0.3 and 0.7 where the assignments are -

- 0.1 - 0.1
- 0.3 - 0.2, 0.4
- 0.7 - 0.5, 0.6, 0.8, 0.9

#### 1.1.2 Part 2

The SSE is 0.12

#### 1.1.3 Part 3

Using bisecting k-means, the new centroids are 0.267, 0.0.55 and 0.85. The clusters are -

- 0.267 - 0.1, 0.2, 0.4
- 0.55 - 0.5, 0.6
- 0.85 - 0.8, 0.9

The SSE here is 0.060067

#### 1.1.4 Part 3

As the SSE in bisecting k means is larger, bisecting k means is better.

### 1.2 Problem 2

The dendogram in case of single linkage is

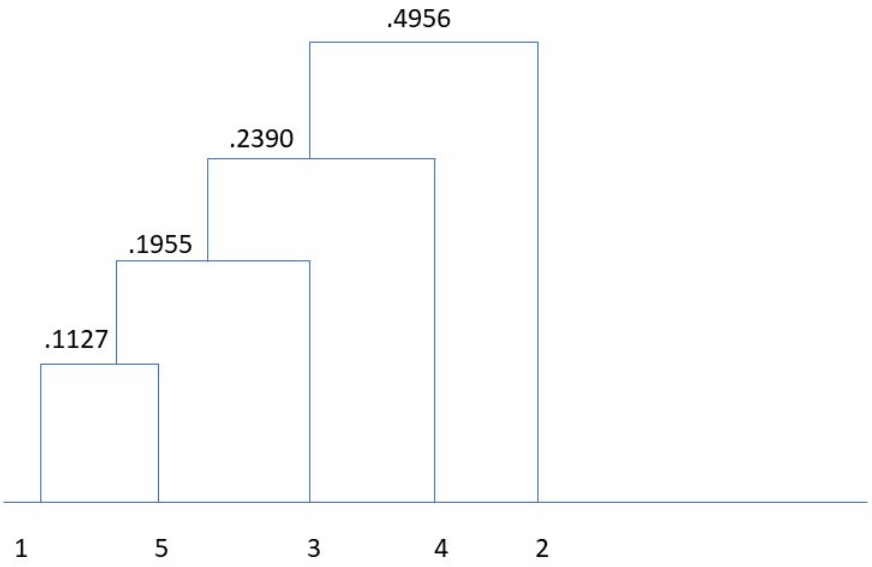


Figure 1: Single Linkage

- Merge 1 and 5
- Merge (1,5) and 3
- Merge (1,5,3) and and 4
- Merge (1,5,3,4) and 2.

The dendogram in case of complete linkage is

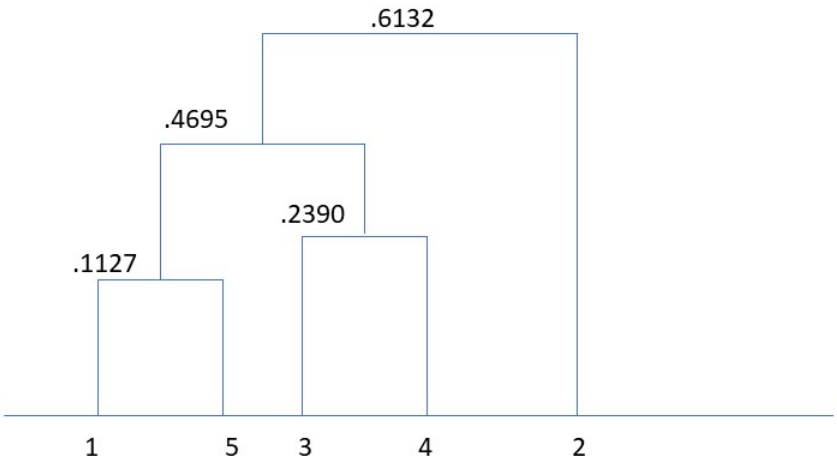


Figure 2: Complete Linkage

- Merge 1 and 5
- Merge 3 and 4
- Merge (1,5) and (3,4)
- Merge (1,5,3,4) and 2

*the calculation were done by hand and pictures have been added here*

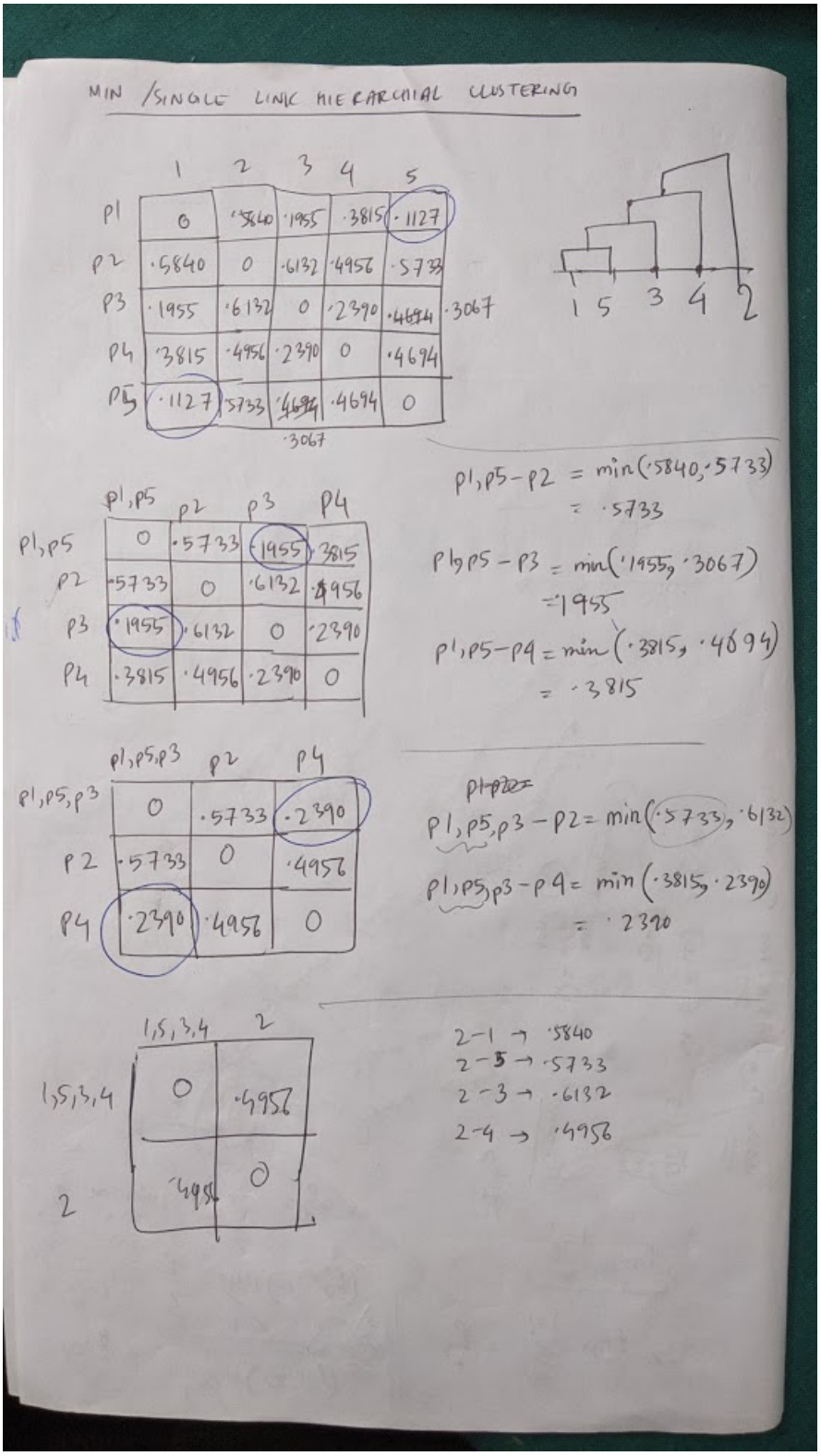


Figure 3: Single Linkage

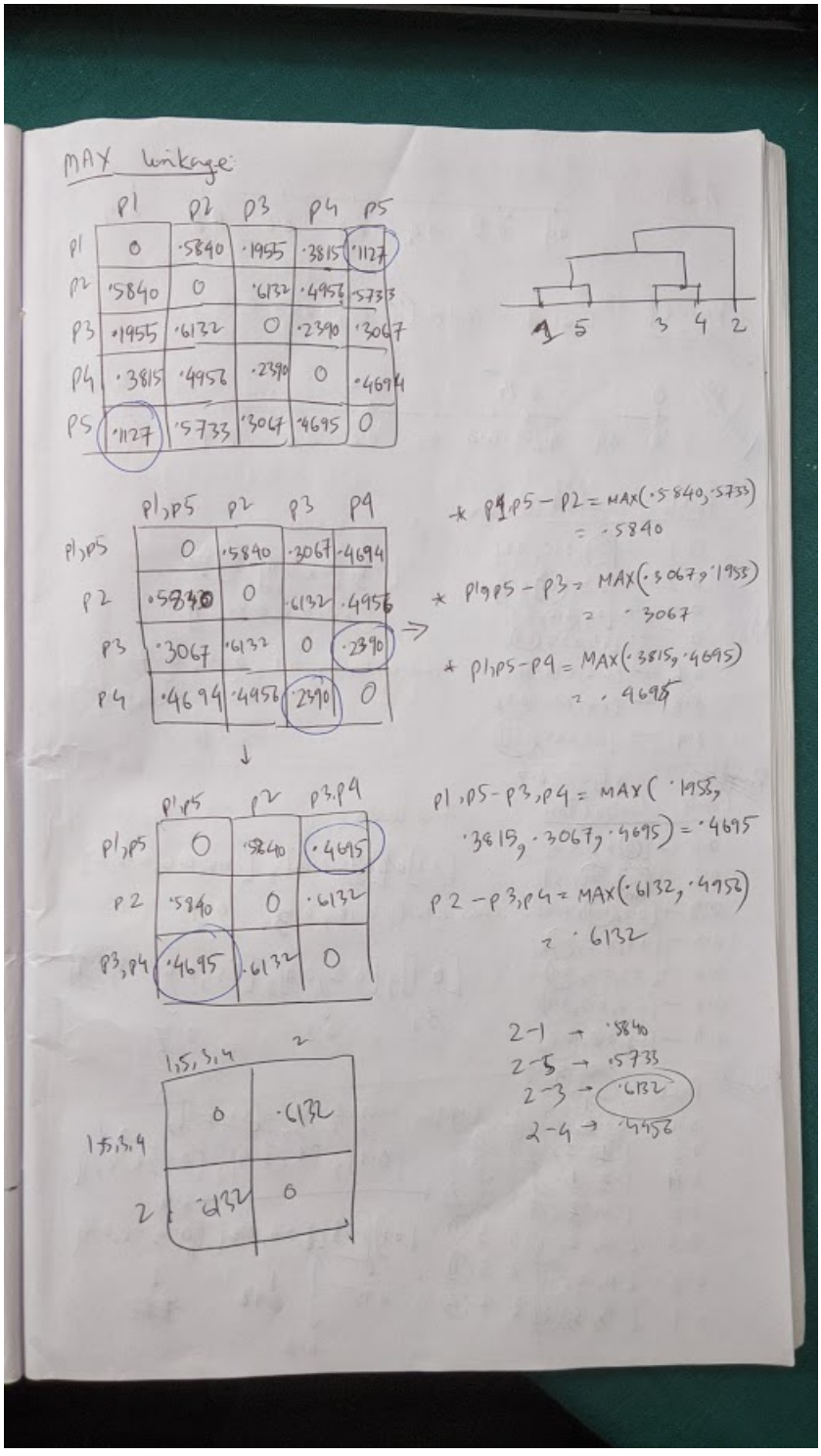


Figure 4: Complete Linkage

## 1.3 Problem 3

### 1.3.1 Part 1

Core points are a,b,c,d,e,f,g,h,i,j,k,l,q,r,s,t,x (17 points)

### 1.3.2 Part 2

Boundary points are p,u,z,y,m,w,v (7 points)

### 1.3.3 Part 3

Noise points are n,o (2 points)

### 1.3.4 Part 4

3 clusters

## 1.4 Problem 4

### 1.4.1 Part 1

For pure, entropy = 0, purity = 1, NMI = 1

### 1.4.2 Part 2

Entropy for both are 0.6068 and 0.5989

### 1.4.3 Part 3

Purity for both are 0.7

### 1.4.4 Part 4

NMI for both are