

Agglomerative Hierarchical Clustering

James Loomis & Paul Romanczyk

October 27, 2010

Outline

Introduction

Distance

Example

S.L. Dist. Matrices

Cluster Plots

Dendrograms

Summary

References

Questions

Extra Stuff

T.L. Dist. Matrices

A.L. Dist. Matrices

Outline

Outline

Introduction

Distance

Example

Single Linkage Distance Matrices

Cluster Plots

Dendrograms

Summary

References

Questions

Extra Stuff

Total Linkage Distance Matrices

Average Linkage Distance Matrices

Outline

Introduction

Distance

Example

S.L. Dist. Matrices

Cluster Plots

Dendrograms

Summary

References

Questions

Extra Stuff

T.L. Dist. Matrices

A.L. Dist. Matrices

Hierarchical Clustering

Hierarchical clustering:

- ▶ Clustering using a hierarchy of clusters
- ▶ May be represented in a tree structure (*dendrogram*)
- ▶ Root - a single cluster containing all observations
- ▶ Leaves - individual observations.

Outline

Introduction

Distance

Example

S.L. Dist. Matrices

Cluster Plots

Dendrograms

Summary

References

Questions

Extra Stuff

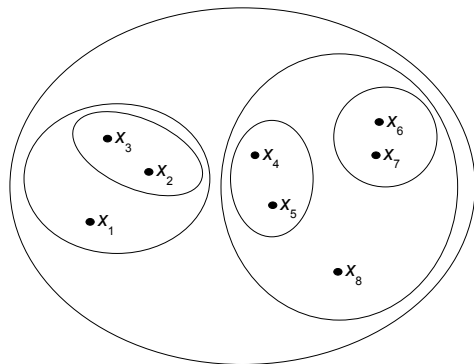
T.L. Dist. Matrices

A.L. Dist. Matrices

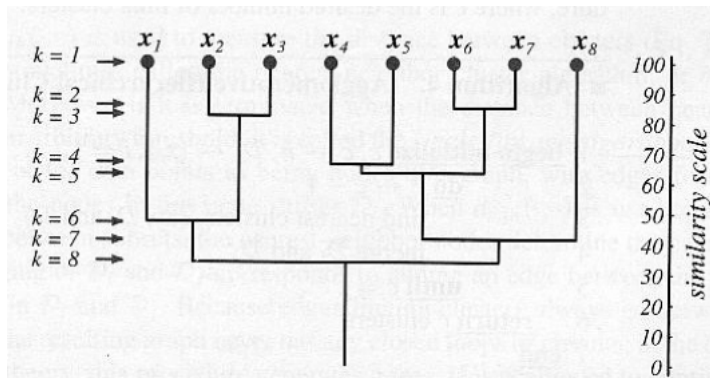
Hierarchical Clustering

Hierarchical clustering:

- ▶ Clustering using a hierarchy of clusters
- ▶ May be represented in a tree structure (*dendrogram*)
- ▶ Root - a single cluster containing all observations
- ▶ Leaves - individual observations.



Dendrogram



[Duda et al., 2001] Figure 10.11

Outline

Introduction

Distance

Example

S.L. Dist. Matrices

Cluster Plots

Dendrograms

Summary

References

Questions

Extra Stuff

T.L. Dist. Matrices

A.L. Dist. Matrices

Two Distinct Approaches:

Outline

Introduction

Distance

Example

S.L. Dist. Matrices

Cluster Plots

Dendrograms

Summary

References

Questions

Extra Stuff

T.L. Dist. Matrices

A.L. Dist. Matrices

Two Distinct Approaches:

- ▶ Agglomerative (*bottom up, clumping*)
 - ▶ Start with n *singleton* clusters
 - ▶ Successively merge ("*clump*") clusters
 - ▶ Computation from one level to another generally simpler
 - ▶ For small number of clusters, takes many iterations

Outline

Introduction

Distance

Example

S.L. Dist. Matrices

Cluster Plots

Dendrograms

Summary

References

Questions

Extra Stuff

T.L. Dist. Matrices

A.L. Dist. Matrices

Two Distinct Approaches:

- ▶ Agglomerative (*bottom up, clumping*)
 - ▶ Start with n *singleton* clusters
 - ▶ Successively merge ("*clump*") clusters
 - ▶ Computation from one level to another generally simpler
 - ▶ For small number of clusters, takes many iterations
- ▶ Divisive (*top down, splitting*)
 - ▶ Start with one cluster
 - ▶ Successively split clusters
 - ▶ Single iteration is more expensive
 - ▶ With fewer clusters, fewer iterations

Outline

Introduction

Distance

Example

S.L. Dist. Matrices

Cluster Plots

Dendrograms

Summary

References

Questions

Extra Stuff

T.L. Dist. Matrices

A.L. Dist. Matrices

Agglomerative Clustering Algorithm

Agglomerative
Hierarchical
Clustering
10/14/2010
6

Loomis &
Romanczyk

```
1  $c, \hat{c} \leftarrow n$ 
2  $D_i \leftarrow \{\mathbf{x}_i\}$  where  $i = 1, \dots, n$ 
3   do  $\hat{c} \leftarrow \hat{c} - 1$ 
4     find nearest clusters  $D_i, D_j$ 
5     merge  $D_i$  and  $D_j$ 
6   until  $c = \hat{c}$ 
7 return  $c$  clusters
```

Outline

Introduction

Distance

Example

S.L. Dist. Matrices

Cluster Plots

Dendrograms

Summary

References

Questions

Extra Stuff

T.L. Dist. Matrices

A.L. Dist. Matrices

Agglomerative Clustering Algorithm

Agglomerative
Hierarchical
Clustering
10/14/2010
6

Loomis &
Romanczyk

```
1   $c, \hat{c} \leftarrow n$ 
2   $D_i \leftarrow \{\mathbf{x}_i\}$  where  $i = 1, \dots, n$ 
3      do  $\hat{c} \leftarrow \hat{c} - 1$ 
4          find nearest clusters  $D_i, D_j$ 
5          merge  $D_i$  and  $D_j$ 
6      until  $c = \hat{c}$ 
7  return  $c$  clusters
```

How do we determine which two clusters are *nearest*?

Outline

Introduction

Distance

Example

S.L. Dist. Matrices

Cluster Plots

Dendrograms

Summary

References

Questions

Extra Stuff

T.L. Dist. Matrices

A.L. Dist. Matrices

Properties of Distance

- ▶ Distance is non-negative.
 - ▶ $D(x, y) \geq 0$
- ▶ $D(x, y) = 0$ if and only if $x = y$.
- ▶ Distance is symmetric.
 - ▶ $D(x, y) = D(y, x)$
- ▶ Distance satisfies the triangle inequality
 - ▶ $D(x, z) \leq D(x, y) + D(y, z)$

Outline

Introduction

Distance

Example

S.L. Dist. Matrices

Cluster Plots

Dendrograms

Summary

References

Questions

Extra Stuff

T.L. Dist. Matrices

A.L. Dist. Matrices

Distance Measures—Between Points

Let $\vec{x}_1 = [x_{1,1} \ x_{1,2} \ \cdots \ x_{1,n}]^T$ and
 $\vec{x}_2 = [x_{2,1} \ x_{2,2} \ \cdots \ x_{2,n}]^T$

| Name | Formula |
|-------------|---|
| Manhattan | $d_1(\vec{x}_1, \vec{x}_2) = \sum_{i=1}^n x_{1,i} - x_{2,i} $ |
| Euclidian | $d_2(\vec{x}_1, \vec{x}_2) = \sqrt{\sum_{i=1}^n x_{1,i} - x_{2,i} ^2}$ |
| P-norm | $d_p(\vec{x}_1, \vec{x}_2) = \sqrt[p]{\sum_{i=1}^n x_{1,i} - x_{2,i} ^p}$ |
| Statistical | $d_s(\vec{x}_1, \vec{x}_2) = \sqrt{\sum_{i=1}^n \left(\frac{x_{1,i} - x_{2,i}}{\sigma_i}\right)^2}$ |
| Mahalanobis | $d_m(\vec{x}_1, \vec{x}_2) = \sqrt{(\vec{x}_1 - \vec{\mu})\Sigma^{-1}(\vec{x}_2 - \vec{\mu})^T}$ |
| Cosine | $d_c(\vec{x}_1, \vec{x}_2) = \frac{\vec{x}_1^T \vec{x}_2}{\ \vec{x}_1\ \cdot \ \vec{x}_2\ }$ |
| Chebyshev | $d_C(\vec{x}_1, \vec{x}_2) = \max(x_{1,i} - x_{2,i})$ |

Outline

Introduction

Distance

Example

S.L. Dist. Matrices

Cluster Plots

Dendrograms

Summary

References



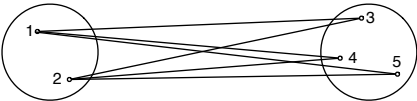
Questions

Extra Stuff

T.L. Dist. Matrices

A.L. Dist. Matrices

Distance Measures—Between Clusters

| | |
|---|--|
| <p>Single Linkage</p> <p>$d_{2,4}$</p> | $d_{(U,V),W} = \min\{d_{U,W}, d_{V,W}\}$  |
| <p>Complete Linkage</p> <p>$d_{1,5}$</p> | $d_{(U,V),W} = \max\{d_{U,W}, d_{V,W}\}$  |
| <p>Average Linkage</p> $\frac{\sum_{i=1}^2 \sum_{j=3}^5 d_{i,j}}{2 \cdot 3}$ | $d_{(U,V),W} = \frac{\sum_i \sum_j d_{i,j}}{N_U \cdot N_W}$  |

Outline

Introduction

Distance

Example

S.L. Dist. Matrices

Cluster Plots

Dendrograms

Summary

References

Questions

Extra Stuff

T.L. Dist. Matrices

A.L. Dist. Matrices

Data

Agglomerative
Hierarchical
Clustering
10/14/2010
10

Loomis &
Romanczyk

| x | y |
|---------|---------|
| -1.3508 | 0.9010 |
| -0.3674 | 1.1548 |
| -1.5895 | -0.0732 |
| -1.3615 | 0.1443 |
| -0.7088 | 0.3324 |
| 0.3155 | -0.3220 |
| 1.6638 | 0.2567 |
| 0.4751 | 0.2582 |
| 2.0778 | 0.2848 |
| 1.3015 | -1.0126 |

Outline

Introduction

Distance

Example

S.L. Dist. Matrices

Cluster Plots

Dendrograms

Summary

References

Questions

Extra Stuff

T.L. Dist. Matrices

A.L. Dist. Matrices

Example

Cluster distances using single linkage. Iteration: 1

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|----|------|------|------|------|------|------|------|------|------|------|
| 1 | 0.00 | 1.02 | 1.00 | 0.76 | 0.86 | 2.07 | 3.08 | 1.94 | 3.48 | 3.27 |
| 2 | 1.02 | 0.00 | 1.73 | 1.42 | 0.89 | 1.63 | 2.22 | 1.23 | 2.60 | 2.74 |
| 3 | 1.00 | 1.73 | 0.00 | 0.32 | 0.97 | 1.92 | 3.27 | 2.09 | 3.68 | 3.04 |
| 4 | 0.76 | 1.42 | 0.32 | 0.00 | 0.68 | 1.74 | 3.03 | 1.84 | 3.44 | 2.90 |
| 5 | 0.86 | 0.89 | 0.97 | 0.68 | 0.00 | 1.22 | 2.37 | 1.19 | 2.79 | 2.42 |
| 6 | 2.07 | 1.63 | 1.92 | 1.74 | 1.22 | 0.00 | 1.47 | 0.60 | 1.86 | 1.20 |
| 7 | 3.08 | 2.22 | 3.27 | 3.03 | 2.37 | 1.47 | 0.00 | 1.19 | 0.41 | 1.32 |
| 8 | 1.94 | 1.23 | 2.09 | 1.84 | 1.19 | 0.60 | 1.19 | 0.00 | 1.60 | 1.52 |
| 9 | 3.48 | 2.60 | 3.68 | 3.44 | 2.79 | 1.86 | 0.41 | 1.60 | 0.00 | 1.51 |
| 10 | 3.27 | 2.74 | 3.04 | 2.90 | 2.42 | 1.20 | 1.32 | 1.52 | 1.51 | 0.00 |

Outline

Introduction

Distance

Example

S.L. Dist. Matrices

Cluster Plots

Dendrograms

Summary

References

Questions

Extra Stuff

T.L. Dist. Matrices

A.L. Dist. Matrices

Example

Cluster distances using single linkage. Iteration: 2

| | 1 | 2 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
|----|------|------|------|------|------|------|------|------|------|
| 1 | 0.00 | 1.02 | 0.86 | 2.07 | 3.08 | 1.94 | 3.48 | 3.27 | 0.76 |
| 2 | 1.02 | 0.00 | 0.89 | 1.63 | 2.22 | 1.23 | 2.60 | 2.74 | 1.42 |
| 5 | 0.86 | 0.89 | 0.00 | 1.22 | 2.37 | 1.19 | 2.79 | 2.42 | 0.68 |
| 6 | 2.07 | 1.63 | 1.22 | 0.00 | 1.47 | 0.60 | 1.86 | 1.20 | 1.74 |
| 7 | 3.08 | 2.22 | 2.37 | 1.47 | 0.00 | 1.19 | 0.41 | 1.32 | 3.03 |
| 8 | 1.94 | 1.23 | 1.19 | 0.60 | 1.19 | 0.00 | 1.60 | 1.52 | 1.84 |
| 9 | 3.48 | 2.60 | 2.79 | 1.86 | 0.41 | 1.60 | 0.00 | 1.51 | 3.44 |
| 10 | 3.27 | 2.74 | 2.42 | 1.20 | 1.32 | 1.52 | 1.51 | 0.00 | 2.90 |
| 11 | 0.76 | 1.42 | 0.68 | 1.74 | 3.03 | 1.84 | 3.44 | 2.90 | 0.00 |

Outline

Introduction

Distance

Example

S.L. Dist. Matrices

Cluster Plots

Dendrograms

Summary

References

Questions

Extra Stuff

T.L. Dist. Matrices

A.L. Dist. Matrices

Example

Cluster distances using single linkage. Iteration: 3

| | 1 | 2 | 5 | 6 | 8 | 10 | 11 | 12 |
|----|------|------|------|------|------|------|------|------|
| 1 | 0.00 | 1.02 | 0.86 | 2.07 | 1.94 | 3.27 | 0.76 | 3.08 |
| 2 | 1.02 | 0.00 | 0.89 | 1.63 | 1.23 | 2.74 | 1.42 | 2.22 |
| 5 | 0.86 | 0.89 | 0.00 | 1.22 | 1.19 | 2.42 | 0.68 | 2.37 |
| 6 | 2.07 | 1.63 | 1.22 | 0.00 | 0.60 | 1.20 | 1.74 | 1.47 |
| 8 | 1.94 | 1.23 | 1.19 | 0.60 | 0.00 | 1.52 | 1.84 | 1.19 |
| 10 | 3.27 | 2.74 | 2.42 | 1.20 | 1.52 | 0.00 | 2.90 | 1.32 |
| 11 | 0.76 | 1.42 | 0.68 | 1.74 | 1.84 | 2.90 | 0.00 | 3.03 |
| 12 | 3.08 | 2.22 | 2.37 | 1.47 | 1.19 | 1.32 | 3.03 | 0.00 |

Outline

Introduction

Distance

Example

S.L. Dist. Matrices

Cluster Plots

Dendrograms

Summary

References

Questions

Extra Stuff

T.L. Dist. Matrices

A.L. Dist. Matrices

Example

Cluster distances using single linkage. Iteration: 4

| | 1 | 2 | 5 | 10 | 11 | 12 | 13 |
|----|------|------|------|------|------|------|------|
| 1 | 0.00 | 1.02 | 0.86 | 3.27 | 0.76 | 3.08 | 1.94 |
| 2 | 1.02 | 0.00 | 0.89 | 2.74 | 1.42 | 2.22 | 1.23 |
| 5 | 0.86 | 0.89 | 0.00 | 2.42 | 0.68 | 2.37 | 1.19 |
| 10 | 3.27 | 2.74 | 2.42 | 0.00 | 2.90 | 1.32 | 1.20 |
| 11 | 0.76 | 1.42 | 0.68 | 2.90 | 0.00 | 3.03 | 1.74 |
| 12 | 3.08 | 2.22 | 2.37 | 1.32 | 3.03 | 0.00 | 1.19 |
| 13 | 1.94 | 1.23 | 1.19 | 1.20 | 1.74 | 1.19 | 0.00 |

Outline

Introduction

Distance

Example

S.L. Dist. Matrices

Cluster Plots

Dendrograms

Summary

References

Questions

Extra Stuff

T.L. Dist. Matrices

A.L. Dist. Matrices

Example

Cluster distances using single linkage. Iteration: 5

| | 1 | 2 | 10 | 12 | 13 | 14 |
|----|------|------|------|------|------|------|
| 1 | 0.00 | 1.02 | 3.27 | 3.08 | 1.94 | 0.76 |
| 2 | 1.02 | 0.00 | 2.74 | 2.22 | 1.23 | 0.89 |
| 10 | 3.27 | 2.74 | 0.00 | 1.32 | 1.20 | 2.42 |
| 12 | 3.08 | 2.22 | 1.32 | 0.00 | 1.19 | 2.37 |
| 13 | 1.94 | 1.23 | 1.20 | 1.19 | 0.00 | 1.19 |
| 14 | 0.76 | 0.89 | 2.42 | 2.37 | 1.19 | 0.00 |

Outline

Introduction

Distance

Example

S.L. Dist. Matrices

Cluster Plots

Dendrograms

Summary

References

Questions

Extra Stuff

T.L. Dist. Matrices

A.L. Dist. Matrices

Example

Cluster distances using single linkage. Iteration: 6

| | 2 | 10 | 12 | 13 | 15 |
|----|------|------|------|------|------|
| 2 | 0.00 | 2.74 | 2.22 | 1.23 | 0.89 |
| 10 | 2.74 | 0.00 | 1.32 | 1.20 | 2.42 |
| 12 | 2.22 | 1.32 | 0.00 | 1.19 | 2.37 |
| 13 | 1.23 | 1.20 | 1.19 | 0.00 | 1.19 |
| 15 | 0.89 | 2.42 | 2.37 | 1.19 | 0.00 |

Outline

Introduction

Distance

Example

S.L. Dist. Matrices

Cluster Plots

Dendrograms

Summary

References

Questions

Extra Stuff

T.L. Dist. Matrices

A.L. Dist. Matrices

Example

Cluster distances using single linkage. Iteration: 7

| | 10 | 12 | 13 | 16 |
|----|------|------|------|------|
| 10 | 0.00 | 1.32 | 1.20 | 2.42 |
| 12 | 1.32 | 0.00 | 1.19 | 2.22 |
| 13 | 1.20 | 1.19 | 0.00 | 1.19 |
| 16 | 2.42 | 2.22 | 1.19 | 0.00 |

Outline

Introduction

Distance

Example

S.L. Dist. Matrices

Cluster Plots

Dendrograms

Summary

References

Questions

Extra Stuff

T.L. Dist. Matrices

A.L. Dist. Matrices

Example

Cluster distances using single linkage. Iteration: 8

| | 10 | 12 | 17 |
|----|------|------|------|
| 10 | 0.00 | 1.32 | 1.20 |
| 12 | 1.32 | 0.00 | 1.19 |
| 17 | 1.20 | 1.19 | 0.00 |

Outline

Introduction

Distance

Example

S.L. Dist. Matrices

Cluster Plots

Dendrograms

Summary

References

Questions

Extra Stuff

T.L. Dist. Matrices

A.L. Dist. Matrices

Example

Cluster distances using single linkage. Iteration: 9

| | 10 | 18 |
|----|------|------|
| 10 | 0.00 | 1.20 |
| 18 | 1.20 | 0.00 |

Outline

Introduction

Distance

Example

S.L. Dist. Matrices

Cluster Plots

Dendrograms

Summary

References

Questions

Extra Stuff

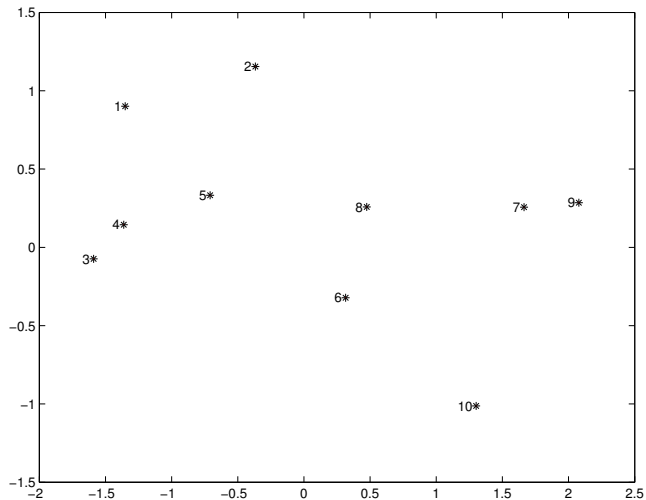
T.L. Dist. Matrices

A.L. Dist. Matrices

Example—Linkage Step 1

Agglomerative
Hierarchical
Clustering
10/14/2010
20

Loomis &
Romanczyk



Single Linkage Complete Linkage Average Linkage

Outline

Introduction

Distance

Example

S.L. Dist. Matrices

Cluster Plots

Dendrograms

Summary

References

Questions

Extra Stuff

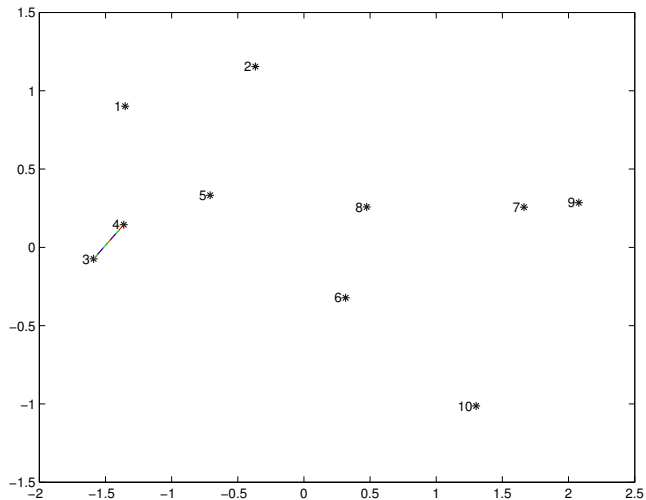
T.L. Dist. Matrices

A.L. Dist. Matrices

Example—Linkage Step 2

Agglomerative
Hierarchical
Clustering
10/14/2010
21

Loomis &
Romanczyk



Single Linkage Complete Linkage Average Linkage

Outline

Introduction

Distance

Example

S.L. Dist. Matrices

Cluster Plots

Dendrograms

Summary

References

Questions

Extra Stuff

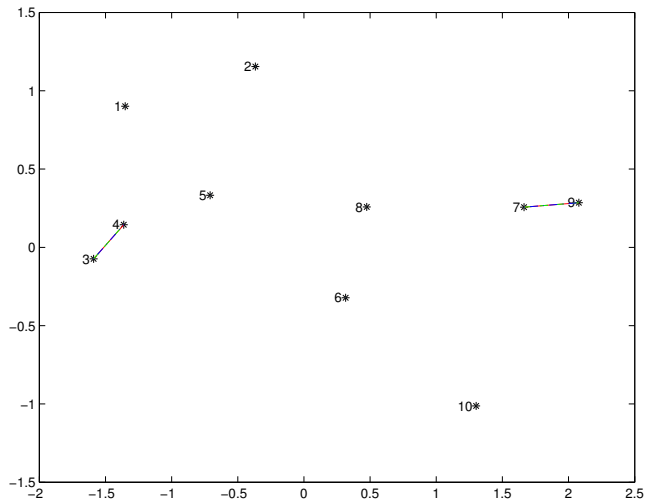
T.L. Dist. Matrices

A.L. Dist. Matrices

Example—Linkage Step 3

Agglomerative
Hierarchical
Clustering
10/14/2010
22

Loomis &
Romanczyk



Single Linkage Complete Linkage Average Linkage

Outline

Introduction

Distance

Example

S.L. Dist. Matrices

Cluster Plots

Dendrograms

Summary

References

Questions

Extra Stuff

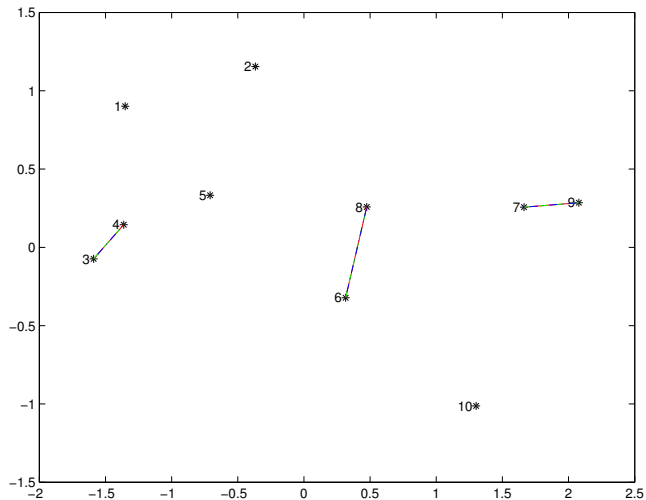
T.L. Dist. Matrices

A.L. Dist. Matrices

Example—Linkage Step 4

Agglomerative
Hierarchical
Clustering
10/14/2010
23

Loomis &
Romanczyk



Single Linkage Complete Linkage Average Linkage

Outline

Introduction

Distance

Example

S.L. Dist. Matrices

Cluster Plots

Dendrograms

Summary

References

Questions

Extra Stuff

T.L. Dist. Matrices

A.L. Dist. Matrices

Example—Linkage Step 5

Agglomerative
Hierarchical
Clustering
10/14/2010
24

Loomis &
Romanczyk

Outline

Introduction

Distance

Example

S.L. Dist. Matrices

Cluster Plots

Dendrograms

Summary

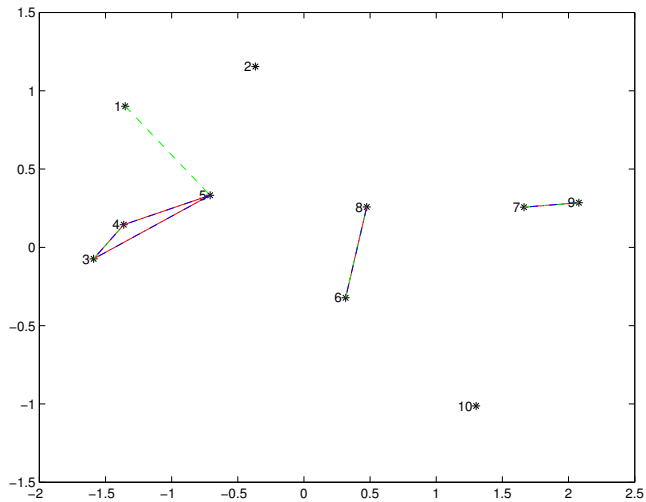
References

Questions

Extra Stuff

T.L. Dist. Matrices

A.L. Dist. Matrices

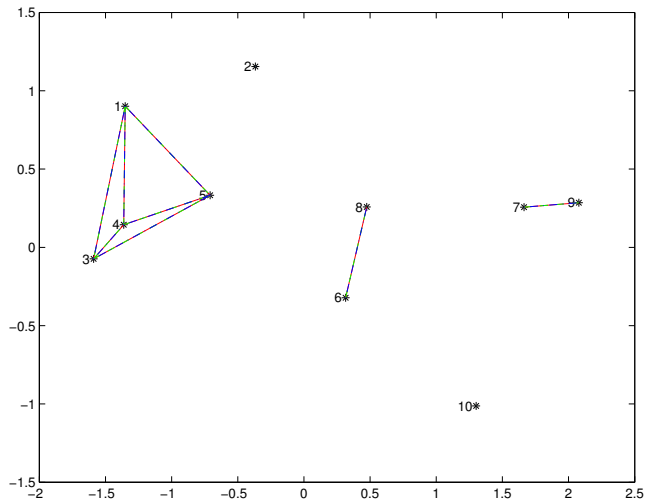


Single Linkage Complete Linkage Average Linkage

Example—Linkage Step 6

Agglomerative
Hierarchical
Clustering
10/14/2010
25

Loomis &
Romanczyk



Single Linkage Complete Linkage Average Linkage

Outline

Introduction

Distance

Example

S.L. Dist. Matrices

Cluster Plots

Dendrograms

Summary

References

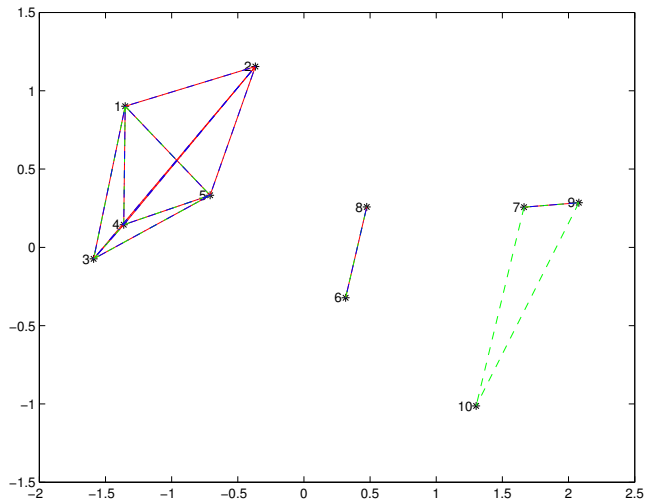
Questions

Extra Stuff

T.L. Dist. Matrices

A.L. Dist. Matrices

Example—Linkage Step 7



Single Linkage Complete Linkage Average Linkage

Outline

Introduction

Distance

Example

S.L. Dist. Matrices

Cluster Plots

Dendrograms

Summary

References

Questions

Extra Stuff

T.L. Dist. Matrices

A.L. Dist. Matrices

Example—Linkage Step 8

Agglomerative
Hierarchical
Clustering
10/14/2010
27

Loomis &
Romanczyk

Outline

Introduction

Distance

Example

S.L. Dist. Matrices

Cluster Plots

Dendrograms

Summary

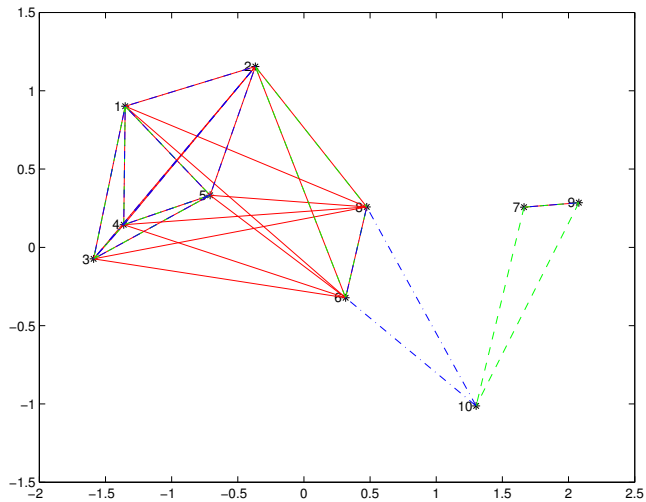
References

Questions

Extra Stuff

T.L. Dist. Matrices

A.L. Dist. Matrices

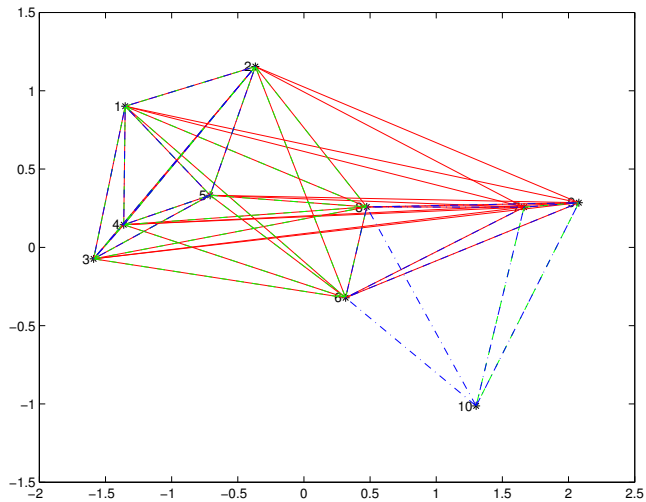


Single Linkage Complete Linkage Average Linkage

Example—Linkage Step 9

Agglomerative
Hierarchical
Clustering
10/14/2010
28

Loomis &
Romanczyk



Single Linkage Complete Linkage Average Linkage

Outline

Introduction

Distance

Example

S.L. Dist. Matrices

Cluster Plots

Dendrograms

Summary

References

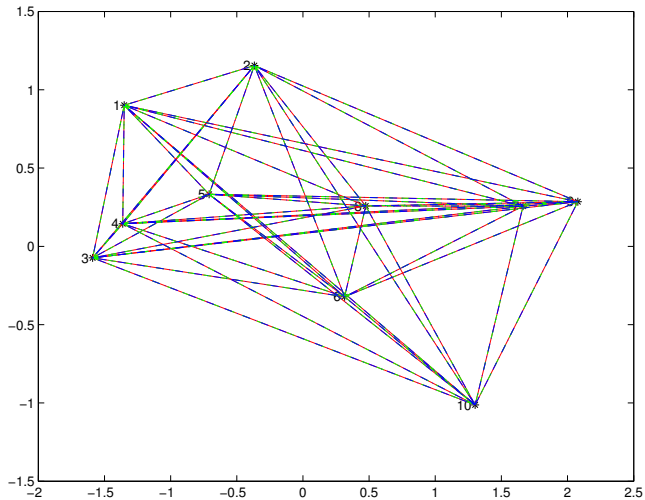
Questions

Extra Stuff

T.L. Dist. Matrices

A.L. Dist. Matrices

Example—Linkage Step 10



Single Linkage Complete Linkage Average Linkage

Outline

Introduction

Distance

Example

S.L. Dist. Matrices

Cluster Plots

Dendrograms

Summary

References

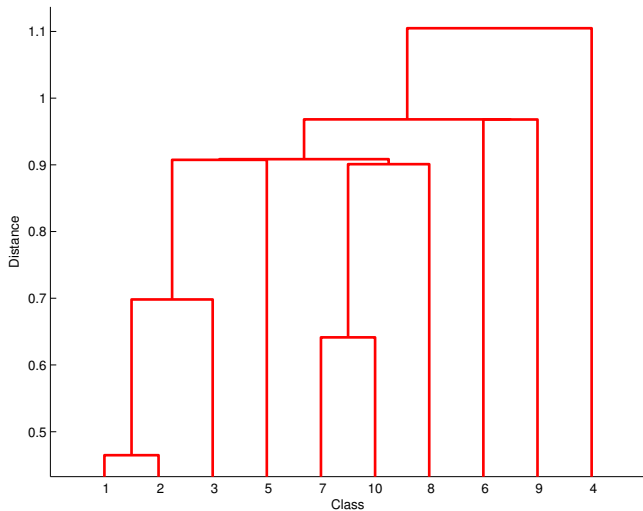
Questions

Extra Stuff

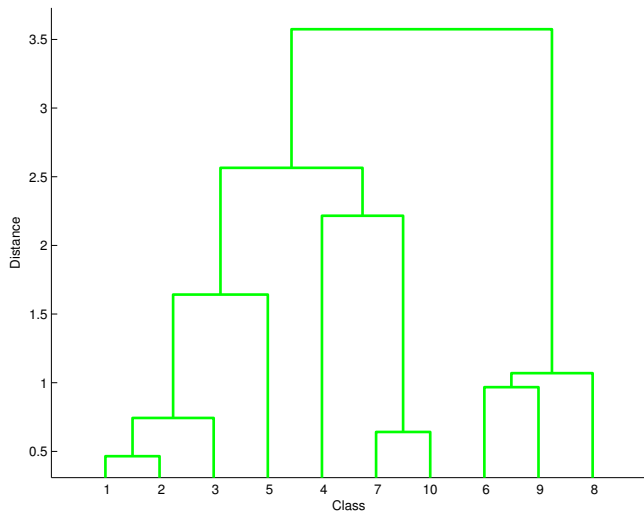
T.L. Dist. Matrices

A.L. Dist. Matrices

Example–Dendrogram



Example–Dendrogram



Outline

Introduction

Distance

Example

S.L. Dist. Matrices

Cluster Plots

Dendrograms

Summary

References

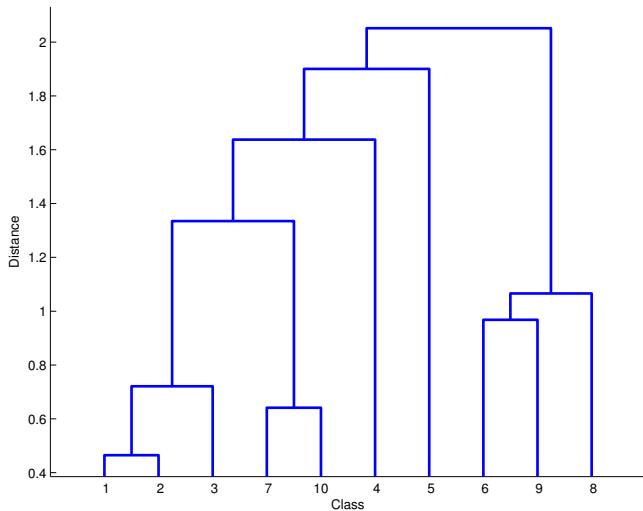
Questions

Extra Stuff

T.L. Dist. Matrices

A.L. Dist. Matrices

Example–Dendrogram



Outline

Introduction

Distance

Example

S.L. Dist. Matrices

Cluster Plots

Dendrograms

Summary

References

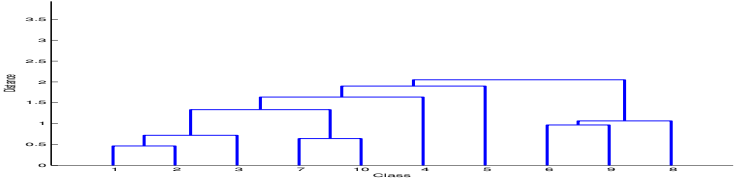
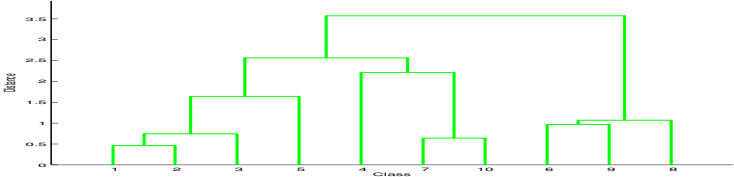
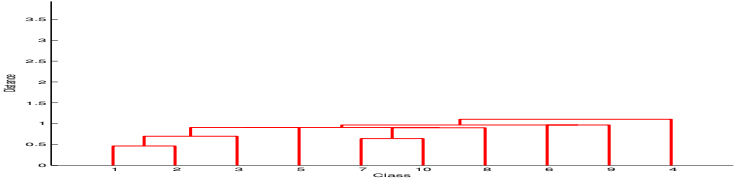
Questions

Extra Stuff

T.L. Dist. Matrices

A.L. Dist. Matrices

Example–Dendrogram



Single Linkage Complete Linkage Average Linkage

- Outline
- Introduction
- Distance
- Example
- S.L. Dist. Matrices
- Cluster Plots
- Dendrograms
- Summary
- References
- Questions

Extra Stuff
T.L. Dist. Matrices
A.L. Dist. Matrices

Example

| Single Linkage | | Complete Linkage | | Average Linkage | |
|----------------|---------------|------------------|---------------|-----------------|---------------|
| dist. | action | dist. | action | dist. | action |
| 0.3151 | {4, 3} → 11 | 0.3151 | {4, 3} → 11 | 0.3151 | {4, 3} → 11 |
| 0.4149 | {9, 7} → 12 | 0.4149 | {9, 7} → 12 | 0.4149 | {9, 7} → 12 |
| 0.6018 | {8, 6} → 13 | 0.6018 | {8, 6} → 13 | 0.6018 | {8, 6} → 13 |
| 0.6792 | {11, 5} → 14 | 0.8576 | {5, 1} → 14 | 0.8244 | {11, 5} → 14 |
| 0.7568 | {14, 1} → 15 | 1.0030 | {14, 11} → 15 | 0.8724 | {14, 1} → 15 |
| 0.8904 | {15, 2} → 16 | 1.5119 | {12, 10} → 16 | 1.2640 | {15, 2} → 16 |
| 1.1862 | {16, 13} → 17 | 1.6271 | {13, 2} → 17 | 1.3598 | {13, 10} → 17 |
| 1.1887 | {17, 12} → 18 | 2.0910 | {17, 15} → 18 | 1.4924 | {17, 12} → 18 |
| 1.2038 | {10, 18} → 19 | 3.2706 | {16, 18} → 19 | 2.4476 | {16, 18} → 19 |

Outline

Introduction

Distance

Example

S.L. Dist. Matrices

Cluster Plots

Dendrograms

Summary



References

Questions

Extra Stuff

T.L. Dist. Matrices

A.L. Dist. Matrices

-  Duda, R., Hart, P., and Stork, D. (2001).
Pattern Classification.
John Wiley and Sons, 2nd edition.
-  Johnson, R. and Wichern, D. (2007).
Applied multivariate statistical data analysis.
Prentice Hall: Upper Saddle River, NJ, 6th edition.

Outline

Introduction

Distance

Example

S.L. Dist. Matrices

Cluster Plots

Dendrograms

Summary

References

Questions

Extra Stuff

T.L. Dist. Matrices

A.L. Dist. Matrices

Questions

Agglomerative
Hierarchical
Clustering
10/14/2010
36

Loomis &
Romanczyk

Outline

Introduction

Distance

Example

S.L. Dist. Matrices

Cluster Plots

Dendrograms

Summary

References

Questions

Extra Stuff

T.L. Dist. Matrices

A.L. Dist. Matrices

Example

Cluster distances using total linkage. Iteration: 1

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|----|------|------|------|------|------|------|------|------|------|------|
| 1 | 0.00 | 1.02 | 1.00 | 0.76 | 0.86 | 2.07 | 3.08 | 1.94 | 3.48 | 3.27 |
| 2 | 1.02 | 0.00 | 1.73 | 1.42 | 0.89 | 1.63 | 2.22 | 1.23 | 2.60 | 2.74 |
| 3 | 1.00 | 1.73 | 0.00 | 0.32 | 0.97 | 1.92 | 3.27 | 2.09 | 3.68 | 3.04 |
| 4 | 0.76 | 1.42 | 0.32 | 0.00 | 0.68 | 1.74 | 3.03 | 1.84 | 3.44 | 2.90 |
| 5 | 0.86 | 0.89 | 0.97 | 0.68 | 0.00 | 1.22 | 2.37 | 1.19 | 2.79 | 2.42 |
| 6 | 2.07 | 1.63 | 1.92 | 1.74 | 1.22 | 0.00 | 1.47 | 0.60 | 1.86 | 1.20 |
| 7 | 3.08 | 2.22 | 3.27 | 3.03 | 2.37 | 1.47 | 0.00 | 1.19 | 0.41 | 1.32 |
| 8 | 1.94 | 1.23 | 2.09 | 1.84 | 1.19 | 0.60 | 1.19 | 0.00 | 1.60 | 1.52 |
| 9 | 3.48 | 2.60 | 3.68 | 3.44 | 2.79 | 1.86 | 0.41 | 1.60 | 0.00 | 1.51 |
| 10 | 3.27 | 2.74 | 3.04 | 2.90 | 2.42 | 1.20 | 1.32 | 1.52 | 1.51 | 0.00 |

Outline

Introduction

Distance

Example

S.L. Dist. Matrices

Cluster Plots

Dendrograms

Summary

References

Questions

Extra Stuff

T.L. Dist. Matrices

A.L. Dist. Matrices

Example

Cluster distances using total linkage. Iteration: 2

| | 1 | 2 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
|----|------|------|------|------|------|------|------|------|------|
| 1 | 0.00 | 1.02 | 0.86 | 2.07 | 3.08 | 1.94 | 3.48 | 3.27 | 1.00 |
| 2 | 1.02 | 0.00 | 0.89 | 1.63 | 2.22 | 1.23 | 2.60 | 2.74 | 1.73 |
| 5 | 0.86 | 0.89 | 0.00 | 1.22 | 2.37 | 1.19 | 2.79 | 2.42 | 0.97 |
| 6 | 2.07 | 1.63 | 1.22 | 0.00 | 1.47 | 0.60 | 1.86 | 1.20 | 1.92 |
| 7 | 3.08 | 2.22 | 2.37 | 1.47 | 0.00 | 1.19 | 0.41 | 1.32 | 3.27 |
| 8 | 1.94 | 1.23 | 1.19 | 0.60 | 1.19 | 0.00 | 1.60 | 1.52 | 2.09 |
| 9 | 3.48 | 2.60 | 2.79 | 1.86 | 0.41 | 1.60 | 0.00 | 1.51 | 3.68 |
| 10 | 3.27 | 2.74 | 2.42 | 1.20 | 1.32 | 1.52 | 1.51 | 0.00 | 3.04 |
| 11 | 1.00 | 1.73 | 0.97 | 1.92 | 3.27 | 2.09 | 3.68 | 3.04 | 0.32 |

Outline

Introduction

Distance

Example

S.L. Dist. Matrices

Cluster Plots

Dendrograms

Summary

References

Questions

Extra Stuff

T.L. Dist. Matrices

A.L. Dist. Matrices

Example

Cluster distances using total linkage. Iteration: 3

| | 1 | 2 | 5 | 6 | 8 | 10 | 11 | 12 |
|----|------|------|------|------|------|------|------|------|
| 1 | 0.00 | 1.02 | 0.86 | 2.07 | 1.94 | 3.27 | 1.00 | 3.48 |
| 2 | 1.02 | 0.00 | 0.89 | 1.63 | 1.23 | 2.74 | 1.73 | 2.60 |
| 5 | 0.86 | 0.89 | 0.00 | 1.22 | 1.19 | 2.42 | 0.97 | 2.79 |
| 6 | 2.07 | 1.63 | 1.22 | 0.00 | 0.60 | 1.20 | 1.92 | 1.86 |
| 8 | 1.94 | 1.23 | 1.19 | 0.60 | 0.00 | 1.52 | 2.09 | 1.60 |
| 10 | 3.27 | 2.74 | 2.42 | 1.20 | 1.52 | 0.00 | 3.04 | 1.51 |
| 11 | 1.00 | 1.73 | 0.97 | 1.92 | 2.09 | 3.04 | 0.32 | 3.68 |
| 12 | 3.48 | 2.60 | 2.79 | 1.86 | 1.60 | 1.51 | 3.68 | 0.41 |

Outline

Introduction

Distance

Example

S.L. Dist. Matrices

Cluster Plots

Dendrograms

Summary

References

Questions

Extra Stuff

T.L. Dist. Matrices

A.L. Dist. Matrices

Example

Cluster distances using total linkage. Iteration: 4

| | 1 | 2 | 5 | 10 | 11 | 12 | 13 |
|----|------|------|------|------|------|------|------|
| 1 | 0.00 | 1.02 | 0.86 | 3.27 | 1.00 | 3.48 | 2.07 |
| 2 | 1.02 | 0.00 | 0.89 | 2.74 | 1.73 | 2.60 | 1.63 |
| 5 | 0.86 | 0.89 | 0.00 | 2.42 | 0.97 | 2.79 | 1.22 |
| 10 | 3.27 | 2.74 | 2.42 | 0.00 | 3.04 | 1.51 | 1.52 |
| 11 | 1.00 | 1.73 | 0.97 | 3.04 | 0.32 | 3.68 | 2.09 |
| 12 | 3.48 | 2.60 | 2.79 | 1.51 | 3.68 | 0.41 | 1.86 |
| 13 | 2.07 | 1.63 | 1.22 | 1.52 | 2.09 | 1.86 | 0.60 |

Outline

Introduction

Distance

Example

S.L. Dist. Matrices

Cluster Plots

Dendrograms

Summary

References

Questions

Extra Stuff

T.L. Dist. Matrices

A.L. Dist. Matrices

Example

Cluster distances using total linkage. Iteration: 5

| | 2 | 10 | 11 | 12 | 13 | 14 |
|----|------|------|------|------|------|------|
| 2 | 0.00 | 2.74 | 1.73 | 2.60 | 1.63 | 1.02 |
| 10 | 2.74 | 0.00 | 3.04 | 1.51 | 1.52 | 3.27 |
| 11 | 1.73 | 3.04 | 0.32 | 3.68 | 2.09 | 1.00 |
| 12 | 2.60 | 1.51 | 3.68 | 0.41 | 1.86 | 3.48 |
| 13 | 1.63 | 1.52 | 2.09 | 1.86 | 0.60 | 2.07 |
| 14 | 1.02 | 3.27 | 1.00 | 3.48 | 2.07 | 0.86 |

Outline

Introduction

Distance

Example

S.L. Dist. Matrices

Cluster Plots

Dendrograms

Summary

References

Questions

Extra Stuff

T.L. Dist. Matrices

A.L. Dist. Matrices

Example

Cluster distances using total linkage. Iteration: 6

| | 2 | 10 | 12 | 13 | 15 |
|----|------|------|------|------|------|
| 2 | 0.00 | 2.74 | 2.60 | 1.63 | 1.73 |
| 10 | 2.74 | 0.00 | 1.51 | 1.52 | 3.27 |
| 12 | 2.60 | 1.51 | 0.41 | 1.86 | 3.68 |
| 13 | 1.63 | 1.52 | 1.86 | 0.60 | 2.09 |
| 15 | 1.73 | 3.27 | 3.68 | 2.09 | 1.00 |

Outline

Introduction

Distance

Example

S.L. Dist. Matrices

Cluster Plots

Dendrograms

Summary

References

Questions

Extra Stuff

T.L. Dist. Matrices

A.L. Dist. Matrices

Example

Cluster distances using total linkage. Iteration: 7

| | 2 | 13 | 15 | 16 |
|----|------|------|------|------|
| 2 | 0.00 | 1.63 | 1.73 | 2.74 |
| 13 | 1.63 | 0.60 | 2.09 | 1.86 |
| 15 | 1.73 | 2.09 | 1.00 | 3.68 |
| 16 | 2.74 | 1.86 | 3.68 | 1.51 |

Outline

Introduction

Distance

Example

S.L. Dist. Matrices

Cluster Plots

Dendrograms

Summary

References

Questions

Extra Stuff

T.L. Dist. Matrices

A.L. Dist. Matrices

Example

Cluster distances using total linkage. Iteration: 8

| | 15 | 16 | 17 |
|----|------|------|------|
| 15 | 1.00 | 3.68 | 2.09 |
| 16 | 3.68 | 1.51 | 2.74 |
| 17 | 2.09 | 2.74 | 1.63 |

Outline

Introduction

Distance

Example

S.L. Dist. Matrices

Cluster Plots

Dendrograms

Summary

References

Questions

Extra Stuff

T.L. Dist. Matrices

A.L. Dist. Matrices

Example

Cluster distances using total linkage. Iteration: 9

| | 16 | 18 |
|----|------|------|
| 16 | 0.00 | 3.27 |
| 18 | 3.27 | 0.00 |

Outline

Introduction

Distance

Example

S.L. Dist. Matrices

Cluster Plots

Dendrograms

Summary

References

Questions

Extra Stuff

T.L. Dist. Matrices

A.L. Dist. Matrices

Example

Cluster distances using average linkage. Iteration: 1

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|----|------|------|------|------|------|------|------|------|------|------|
| 1 | 0.00 | 1.02 | 1.00 | 0.76 | 0.86 | 2.07 | 3.08 | 1.94 | 3.48 | 3.27 |
| 2 | 1.02 | 0.00 | 1.73 | 1.42 | 0.89 | 1.63 | 2.22 | 1.23 | 2.60 | 2.74 |
| 3 | 1.00 | 1.73 | 0.00 | 0.32 | 0.97 | 1.92 | 3.27 | 2.09 | 3.68 | 3.04 |
| 4 | 0.76 | 1.42 | 0.32 | 0.00 | 0.68 | 1.74 | 3.03 | 1.84 | 3.44 | 2.90 |
| 5 | 0.86 | 0.89 | 0.97 | 0.68 | 0.00 | 1.22 | 2.37 | 1.19 | 2.79 | 2.42 |
| 6 | 2.07 | 1.63 | 1.92 | 1.74 | 1.22 | 0.00 | 1.47 | 0.60 | 1.86 | 1.20 |
| 7 | 3.08 | 2.22 | 3.27 | 3.03 | 2.37 | 1.47 | 0.00 | 1.19 | 0.41 | 1.32 |
| 8 | 1.94 | 1.23 | 2.09 | 1.84 | 1.19 | 0.60 | 1.19 | 0.00 | 1.60 | 1.52 |
| 9 | 3.48 | 2.60 | 3.68 | 3.44 | 2.79 | 1.86 | 0.41 | 1.60 | 0.00 | 1.51 |
| 10 | 3.27 | 2.74 | 3.04 | 2.90 | 2.42 | 1.20 | 1.32 | 1.52 | 1.51 | 0.00 |

Outline

Introduction

Distance

Example

S.L. Dist. Matrices

Cluster Plots

Dendrograms

Summary

References

Questions

Extra Stuff

T.L. Dist. Matrices

A.L. Dist. Matrices

Example

Cluster distances using average linkage. Iteration: 2

| | 1 | 2 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
|----|------|------|------|------|------|------|------|------|------|
| 1 | 0.00 | 1.02 | 0.86 | 2.07 | 3.08 | 1.94 | 3.48 | 3.27 | 0.88 |
| 2 | 1.02 | 0.00 | 0.89 | 1.63 | 2.22 | 1.23 | 2.60 | 2.74 | 1.58 |
| 5 | 0.86 | 0.89 | 0.00 | 1.22 | 2.37 | 1.19 | 2.79 | 2.42 | 0.82 |
| 6 | 2.07 | 1.63 | 1.22 | 0.00 | 1.47 | 0.60 | 1.86 | 1.20 | 1.83 |
| 7 | 3.08 | 2.22 | 2.37 | 1.47 | 0.00 | 1.19 | 0.41 | 1.32 | 3.15 |
| 8 | 1.94 | 1.23 | 1.19 | 0.60 | 1.19 | 0.00 | 1.60 | 1.52 | 1.97 |
| 9 | 3.48 | 2.60 | 2.79 | 1.86 | 0.41 | 1.60 | 0.00 | 1.51 | 3.56 |
| 10 | 3.27 | 2.74 | 2.42 | 1.20 | 1.32 | 1.52 | 1.51 | 0.00 | 2.97 |
| 11 | 0.88 | 1.58 | 0.82 | 1.83 | 3.15 | 1.97 | 3.56 | 2.97 | 0.16 |

Outline

Introduction

Distance

Example

S.L. Dist. Matrices

Cluster Plots

Dendrograms

Summary

References

Questions

Extra Stuff

T.L. Dist. Matrices

A.L. Dist. Matrices

Example

Cluster distances using average linkage. Iteration: 3

| | 1 | 2 | 5 | 6 | 8 | 10 | 11 | 12 |
|----|------|------|------|------|------|------|------|------|
| 1 | 0.00 | 1.02 | 0.86 | 2.07 | 1.94 | 3.27 | 0.88 | 3.28 |
| 2 | 1.02 | 0.00 | 0.89 | 1.63 | 1.23 | 2.74 | 1.58 | 2.41 |
| 5 | 0.86 | 0.89 | 0.00 | 1.22 | 1.19 | 2.42 | 0.82 | 2.58 |
| 6 | 2.07 | 1.63 | 1.22 | 0.00 | 0.60 | 1.20 | 1.83 | 1.67 |
| 8 | 1.94 | 1.23 | 1.19 | 0.60 | 0.00 | 1.52 | 1.97 | 1.40 |
| 10 | 3.27 | 2.74 | 2.42 | 1.20 | 1.52 | 0.00 | 2.97 | 1.42 |
| 11 | 0.88 | 1.58 | 0.82 | 1.83 | 1.97 | 2.97 | 0.16 | 3.36 |
| 12 | 3.28 | 2.41 | 2.58 | 1.67 | 1.40 | 1.42 | 3.36 | 0.21 |

Outline

Introduction

Distance

Example

S.L. Dist. Matrices

Cluster Plots

Dendrograms

Summary

References

Questions

Extra Stuff

T.L. Dist. Matrices

A.L. Dist. Matrices

Example

Cluster distances using average linkage. Iteration: 4

| | 1 | 2 | 5 | 10 | 11 | 12 | 13 |
|----|------|------|------|------|------|------|------|
| 1 | 0.00 | 1.02 | 0.86 | 3.27 | 0.88 | 3.28 | 2.00 |
| 2 | 1.02 | 0.00 | 0.89 | 2.74 | 1.58 | 2.41 | 1.43 |
| 5 | 0.86 | 0.89 | 0.00 | 2.42 | 0.82 | 2.58 | 1.20 |
| 10 | 3.27 | 2.74 | 2.42 | 0.00 | 2.97 | 1.42 | 1.36 |
| 11 | 0.88 | 1.58 | 0.82 | 2.97 | 0.16 | 3.36 | 1.90 |
| 12 | 3.28 | 2.41 | 2.58 | 1.42 | 3.36 | 0.21 | 1.53 |
| 13 | 2.00 | 1.43 | 1.20 | 1.36 | 1.90 | 1.53 | 0.30 |

Outline

Introduction

Distance

Example

S.L. Dist. Matrices

Cluster Plots

Dendrograms

Summary

References

Questions

Extra Stuff

T.L. Dist. Matrices

A.L. Dist. Matrices

Example

Cluster distances using average linkage. Iteration: 5

| | 1 | 2 | 10 | 12 | 13 | 14 |
|----|------|------|------|------|------|------|
| 1 | 0.00 | 1.02 | 3.27 | 3.28 | 2.00 | 0.87 |
| 2 | 1.02 | 0.00 | 2.74 | 2.41 | 1.43 | 1.35 |
| 10 | 3.27 | 2.74 | 0.00 | 1.42 | 1.36 | 2.79 |
| 12 | 3.28 | 2.41 | 1.42 | 0.21 | 1.53 | 3.10 |
| 13 | 2.00 | 1.43 | 1.36 | 1.53 | 0.30 | 1.67 |
| 14 | 0.87 | 1.35 | 2.79 | 3.10 | 1.67 | 0.44 |

Outline

Introduction

Distance

Example

S.L. Dist. Matrices

Cluster Plots

Dendrograms

Summary

References

Questions

Extra Stuff

T.L. Dist. Matrices

A.L. Dist. Matrices

Example

Cluster distances using average linkage. Iteration: 6

| | 2 | 10 | 12 | 13 | 15 |
|----|------|------|------|------|------|
| 2 | 0.00 | 2.74 | 2.41 | 1.43 | 1.26 |
| 10 | 2.74 | 0.00 | 1.42 | 1.36 | 2.91 |
| 12 | 2.41 | 1.42 | 0.21 | 1.53 | 3.14 |
| 13 | 1.43 | 1.36 | 1.53 | 0.30 | 1.75 |
| 15 | 1.26 | 2.91 | 3.14 | 1.75 | 0.57 |

Outline

Introduction

Distance

Example

S.L. Dist. Matrices

Cluster Plots

Dendrograms

Summary

References

Questions

Extra Stuff

T.L. Dist. Matrices

A.L. Dist. Matrices

Example

Cluster distances using average linkage. Iteration: 7

| | 10 | 12 | 13 | 16 |
|----|------|------|------|------|
| 10 | 0.00 | 1.42 | 1.36 | 2.87 |
| 12 | 1.42 | 0.21 | 1.53 | 3.00 |
| 13 | 1.36 | 1.53 | 0.30 | 1.69 |
| 16 | 2.87 | 3.00 | 1.69 | 0.77 |

Outline

Introduction

Distance

Example

S.L. Dist. Matrices

Cluster Plots

Dendrograms

Summary

References

Questions

Extra Stuff

T.L. Dist. Matrices

A.L. Dist. Matrices

Example

Cluster distances using average linkage. Iteration: 8

| | 12 | 16 | 17 |
|----|------|------|------|
| 12 | 0.21 | 3.00 | 1.49 |
| 16 | 3.00 | 0.77 | 2.08 |
| 17 | 1.49 | 2.08 | 0.74 |

Outline

Introduction

Distance

Example

S.L. Dist. Matrices

Cluster Plots

Dendrograms

Summary

References

Questions

Extra Stuff

T.L. Dist. Matrices

A.L. Dist. Matrices

Example

Cluster distances using average linkage. Iteration: 9

| | 16 | 18 |
|----|------|------|
| 16 | 0.00 | 2.45 |
| 18 | 2.45 | 0.00 |

Outline

Introduction

Distance

Example

S.L. Dist. Matrices

Cluster Plots

Dendrograms

Summary

References

Questions

Extra Stuff

T.L. Dist. Matrices

A.L. Dist. Matrices