Aufgabe 7-1  \textit{PAM}

Show that the algorithm PAM (Partitioning Around Medoids, Kaufman and Rousseeuw, 1987) converges.

Aufgabe 7-2  \textit{Hierarchical Clustering}

Given the following data set:

As distance function, use Manhattan Distance:

\[ L_1(x, y) = |x_1 - y_1| + |x_2 - y_2| \]

Compute two dendrograms for this data set. To compute the distance of sets of objects, use

- the single-link method
- the average-link method

Hint: with discrete distance values, nodes may have more than two children.
Aufgabe 7-3  DBSCAN

Given the following data set:

As distance function, use Manhattan Distance:

\[ L_1(x, y) = |x_1 - y_1| + |x_2 - y_2| \]

Compute DBSCAN and indicate which points are core points, border points and noise points.

Use the following parameter settings:

- Radius \( \varepsilon = 1.1 \) and \( \text{minPts} = 2 \)
- Radius \( \varepsilon = 1.1 \) and \( \text{minPts} = 3 \)
- Radius \( \varepsilon = 1.1 \) and \( \text{minPts} = 4 \)
- Radius \( \varepsilon = 2.1 \) and \( \text{minPts} = 4 \)
- Radius \( \varepsilon = 4.1 \) and \( \text{minPts} = 5 \)
- Radius \( \varepsilon = 4.1 \) and \( \text{minPts} = 4 \)

When \( \text{minPts} = 2 \), what happens to border points?

What is the relationship of DBSCAN with \( \text{minPts} = 2 \) to single-linkage clustering? Why does DBSCAN run in \( O(n^2) \) time while hierarchical clustering is usually denoted as \( O(n^3) \)? Why is this not a contradiction?