

Knowledge Discovery in Databases
SS 2011

Übungsblatt 7: Partitionierendes Clustering

Aufgabe 7-1 PAM

Zeigen Sie, dass der Algorithmus PAM konvergiert.

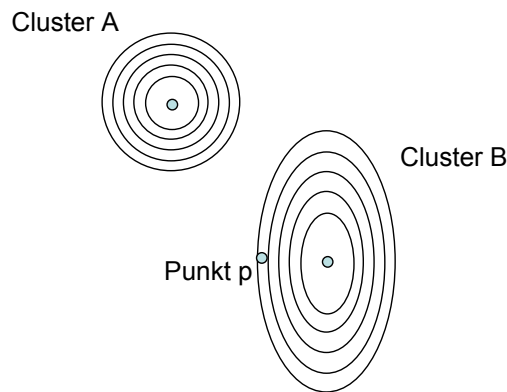
Aufgabe 7-2 EM-Algorithmus

Gegeben sei eine Datenmenge D mit 100 Punkten, die zwei Gausscluster A und B und den Punkt p enthält.

Der Cluster A ist repräsentiert durch den Mittelwert aller seiner Punkte (1,1) und die Kovarianzmatrix $\begin{pmatrix} 4 & 0 \\ 0 & 4 \end{pmatrix}$.
60 Prozent aller Punkte gehören zu diesem Cluster.

Der Cluster B ist repräsentiert durch den Mittelwert aller seiner Punkte (3,3) und die Kovarianzmatrix $\begin{pmatrix} 4 & 0 \\ 0 & 16 \end{pmatrix}$.
40 Prozent aller Punkte gehören zu diesem Cluster.

Der Punkt p ist durch die Koordinaten (2.5,2.5) gegeben. Geben Sie die beiden Wahrscheinlichkeiten an, mit der p zum Cluster A bzw. B gehört.



Achtung: Die Skizze ist nur zu Veranschaulichungszwecken gedacht und nicht maßstabsgetreu!