

Knowledge Discovery in Databases II
WS 2013/2014

Übungsblatt 2: Feature Reduction

Aufgabe 2-1 *Hauptachsentransformation*

Rechnen Sie noch einmal das Beispiel zur Hauptachsentransformation durch.

Gegeben seien folgende Daten:

$$X = \{(-3, -2), (-2, -1), (-1, 0), (0, 1), (1, 2), (2, 3), \\ (-2, -2), (-1, -1), (0, 0), (1, 1), (2, 2), \\ (-2, -3), (-1, -2), (0, -1), (1, 0), (2, 1), (3, 2)\}$$

- Berechnen Sie die Korrelationsmatrix M .
- Berechnen Sie die Eigenwerte und Eigenvektoren von M .
- Bestimmen Sie den kleinsten Eigenwert und streichen Sie dessen Eigenvektor. Der resultierende Eigenvektor bildet die Basis für den neuen Unterraum.
- Transformieren Sie die Vektoren der Daten X in diesen neuen Unterraum, indem Sie die Vektoren aus X nach dieser neuen Unterraumbasis entwickeln.

Aufgabe 2-2 *Hauptachsentransformation*

Führen Sie die Hauptachsentransformation (PCA) auf dem folgenden Datensatz durch:

- $A(1, 0, 3), B(0, 0, 3), C(1, 0, 1), D(0, 0, 1)$

Was für ein Problem ergibt sich bei diesem Datensatz? Wie kann es behoben werden?

Aufgabe 2-3 *Singular Value Decomposition*

Ein weiteres zentrales Konzept in der Feature Reduktion ist die Singular Value Decomposition. Gegeben sei eine Matrix M und ihre SVD-Zerlegung:

$$M = T * S * D'$$

wobei

$$M = \begin{bmatrix} 1 & 2 \\ 6 & 3 \\ 0 & 2 \end{bmatrix} \quad T = \begin{bmatrix} -0.2707 & 0.5458 \\ -0.9509 & -0.2797 \\ -0.1497 & 0.7899 \end{bmatrix} \\ S = \begin{bmatrix} 7.0257 & 0 \\ 0 & 2.1539 \end{bmatrix} \quad D = \begin{bmatrix} -0.8507 & -0.5257 \\ -0.5257 & 0.8507 \end{bmatrix}$$

Führen Sie nun nach dem in der Vorlesung beschriebenen Verfahren eine Reduktion auf 1 Feature durch.