

Maschinelles Lernen und Data Mining
Sommersemester 2014
Übungsblatt 5

Besprechung des Übungsblattes am 15.05.2014

Aufgabe 5-1 PCA

- Beschreiben Sie, welchen Zweck die PCA verfolgt bzw. unter welchen Umständen eine PCA sinnvoll ist.
- Welche möglichen negativen Auswirkungen nehmen Sie in Kauf, wenn Sie die PCA auf einem Datensatz unbekannter Struktur verwenden?

Aufgabe 5-2 PCA

Gegeben sei die untenstehende Designmatrix $\mathbf{X} \in \mathbb{R}^{M \times N}$ für 6 Datenpunkte $\mathbf{x}_i \in \mathbb{R}^2$. Entgegen der bisherigen Konvention stehen hier die Muster in den Spalten.

dim 1	1	2	3	5	6	7
dim 2	0	0	0	6	6	6

Führen Sie eine PCA auf den gegebenen Daten durch. Geben Sie dabei die Eigenvektoren, Eigenwerte und die Kovarianzmatrix an und visualisieren Sie die Daten vor und nach der PCA.

Aufgabe 5-3 Eigenfaces

Unter den *Eigenfaces* versteht man die Eigenvektoren der normalisierten Kovarianzmatrix.

- Bestimmen Sie die Eigenfaces des Zahlen-Bild-Datensatzes aus der Datei `numberMatrix.RTable`.
- Wieviele Hauptkomponenten werden benötigt um den Datensatz zu rekonstruieren? Sind dafür die Eigenfaces ausreichend? Sind alle Muster gleich gut rekonstruierbar?