Lehrstuhl für Analytics and Statistics
Prof. Dr. Oliver Grothe
SS 2022

Karlsruher Institut für Technologie

Aufgaben zur Vorlesung

# Multivariate Verfahren

## Übungsblatt V

---

1. Beantworten Sie folgende Fragen bzw. bearbeiten Sie folgende Arbeitsanweisungen:
   (a) Zeigen Sie, dass bei der metrischen Multidimensionalen Skalierung $HMH = HM'H = 0$ gilt.
   (b) Kann die klassische metrische Multidimensionale Skalierung auch für nicht euklidische Abstände angewendet werden?
   (c) Wie berechnet sich das STRESS-Maß?
   (d) Wie funktioniert die Aggregation der Wahrnehmungsräume mehrerer Personen?
2. Bei Ihrer Arbeit in der Marketingabteilung eines großen Industrieunternehmens wollen Sie die Wahrnehmungsräume von Personen aus einer Umfrage mit Hilfe der nicht-metrischen Multidimensionalen Skalierung erstellen. Sie haben dazu die Rangfolge von drei Objekten erhoben:
$$u_{1,2} = 1, u_{1,3} = 2, u_{2,3} = 3.$$
   Sie wollen eine zweidimensionale Darstellung des Wahrnehmungsraumes und erzeugen zufällig die folgende Startkonfiguration:
$$x_1^0 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, x_2^0 = \begin{pmatrix} 4 \\ 5 \end{pmatrix}, x_3^0 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}.$$
   (a) Erstellen Sie ein Shepard-Diagramm. Ist die Monotoniebedingung erfüllt? (Hinweis: Benutzen Sie die euklidische Distanz.)
   (b) Berechnen Sie die Disparitäten.
   (c) Wie groß ist der Datenverdichtungskoeffizient $Q$? Reicht die Anzahl der Objekte nach diesem Kriterium für eine zweidimensionale Darstellung aus?
3. Mit Hilfe eines Computerprogramms haben Sie eine nicht-metrische Multidimensionale Skalierung für vier Objekte durchgeführt. Das STRESS-Maß zeigt eine sehr gute Anpassungsgüte. Jedoch scheint das Programm verbuggt zu sein, da die finale Konfiguration des Wahrnehmungsraumes nicht ausgegeben wird. Ihnen stehen lediglich die Konfiguration aus der vorletzten Iteration, sowie die zugehörigen Distanzen und Disparitäten zur Verfügung:
$$x_1^{i-1} = \begin{pmatrix} 3 \\ 2 \end{pmatrix}, x_2^{i-1} = \begin{pmatrix} 2 \\ 7 \end{pmatrix}, x_3^{i-1} = \begin{pmatrix} 1 \\ 3 \end{pmatrix}, x_4^{i-1} = \begin{pmatrix} 10 \\ 4 \end{pmatrix},$$

| Objektpaar $(k,l)$ | $d_{kl}$ | $\hat{d}_{kl}$ |
|---|---|---|
| (1,2) | 5.1 | 5.1 |
| (1,3) | 2.2 | 3.15 |
| (1,4) | 7.3 | 7.9 |
| (2,3) | 4.1 | 3.15 |
| (2,4) | 8.5 | 7.9 |
| (3,4) | 9.1 | 9.1 |

Berechnen Sie die Koordinaten der Objekte für die finale Konfiguration. Verwenden Sie $\alpha = 3$.

4. Gegeben sei folgende Tabelle mit Distanzen und Disparitäten:

| Objektpaar $(k,l)$ | $d_{kl}$ | $\hat{d}_{kl}$ |
|---|---|---|
| (1,2) | 4 | 4 |
| (1,3) | 2 | 3 |
| (1,4) | 7 | 8 |
| (2,3) | 3.5 | 3 |
| (2,4) | 8.5 | 8 |
| (3,4) | 10 | 10 |

(a) Berechnen Sie $STRESS_1$ und $STRESS_2$.

(b) Wie beurteilen Sie die Anpassungsgüte?

Angenommen, in einer Iteration der nicht-metrischen Multidimensionalen Skalierung wäre die Monotoniebedingung erfüllt.

(c) Skizzieren Sie ein Shepard-Diagramm zu dieser Situation.

(d) Wie groß wären die Disparitäten? Wie groß wären $STRESS_1$ und $STRESS_2$?

(e) Wie viele weitere Iterationen wären bis zum Abbruch nötig?

5. Rechnerübung: Sie arbeiten bei einem Hersteller für Navigationsgeräte und sollen eine Deutschlandkarte mit den Städten Basel, Berlin, Frankfurt, Hamburg, Hannover, Kassel, Köln, München, Nürnberg und Stuttgart erstellen. Da Sie keine Lust haben, die Entfernungen der einzelnen Städte zueinander sowie deren Lage nachzuschlagen, kommen Sie auf die Idee, aus Ihrem Bauchgefühl heraus ein Ranking der Entfernungen zu jedem Städtepaar zu erstellen. Dieses wollen Sie mit Hilfe der nicht-metrischen Multidimensionalen Skalierung auswerten.

   Laden Sie aus Ilias den Datensatz *staedte* herunter. Dieser enthält die Daten Ihres Rankings. Führen Sie eine nicht-metrische Multidimensionale Skalierung durch.