Universität Greifswald
Institut für Mathematik und Informatik
https://math-inf.uni-greifswald.de

Angebot für eine Bachelor-Arbeit

Thema: Koordinatenrekonstruktion aus Euklidischen Distanzmatrizen

Gegeben seien N paarweise verschiedene Punkte in der Euklidischen Ebene mit Koordinaten (x_i, y_i) für $i = 1, \ldots, N$. Die Abstände zwischen den Punkten lauten

$$d_{ij} = \left\| \begin{pmatrix} x_i \\ y_i \end{pmatrix} - \begin{pmatrix} x_j \\ y_j \end{pmatrix} \right\| = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}$$

für $i, j = 1, \ldots, N$ mit der Euklidischen Norm $\| \cdot \|$. Diese Werte bilden eine Euklidische Distanzmatrix $D = (d_{ij}) \in \mathbb{R}^{N \times N}$. Eine solche Matrix ist offensichtlich symmetrisch und enthält nichtnegative Einträge.

Ein Beispiel ist die Lage von Städten. Die folgende Tabelle zeigt die Distanzen zwischen sechs Städten im Rheinland, gerundet auf Kilometer.

	A.	B.	D.	F.	K.	W.
Aachen	0	91	80	259	70	121
Bonn	91	0	77	175	27	84
Düsseldorf	80	77	0	232	47	29
Frankfurt	259	175	232	0	189	236
Köln	70	27	47	189	0	55
Wuppertal	121	84	29	236	55	0

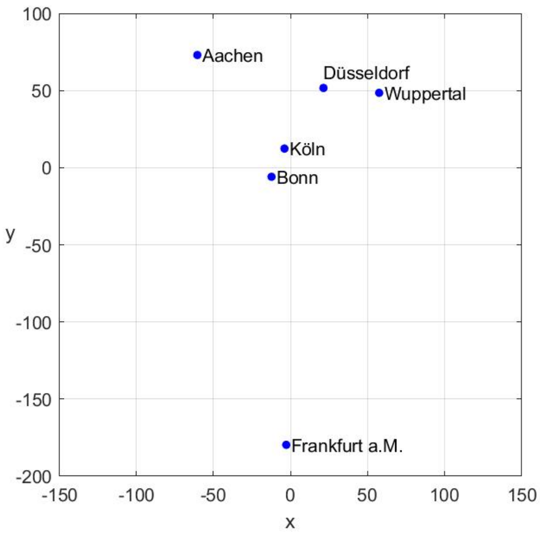
Betrachtet wird nun das inverse Problem hierzu. Gegeben sei eine Distanzmatrix D. Bestimmt werden sollen die Koordinaten (x_i, y_i) der zugehörigen Punkte. Diese Problemstellung wird auch als Multidimensionale Skalierung (MDS) bezeichnet. Die Rekonstruktion der Positionen ist nur eindeutig bis auf längenerhaltende Koordinatentransformationen (Verschiebungen, Rotationen, Spiegelungen). In der Praxis sind zudem die Distanzwerte typischerweise mit Fehlern behaftet (z.B. Messfehler, Modellfehler). Dies macht eine exakte Rekonstruktion der Positionen unmöglich. Gesucht ist dann eine möglichst gute Näherung. Dieses Aufgabenstellung kann auch im dreidimensionalen Euklidischen Raum betrachtet werden.

Ein Verfahren zur numerischen Lösung dieses inversen Problems verwendet eine Eigenwertzerlegung. Aus der Euklidischen Distanzmatrix wird die Matrix $\widehat{D} = (d_{ij}^2) \in \mathbb{R}^{N \times N}$ gebildet, welche ebenfalls symmetrisch ist. Damit folgt die Gramsche Matrix

$$G = -\tfrac{1}{2} \left(\widehat{D} - \mathbb{1} \widehat{d_1}^\top - \widehat{d_1} \mathbb{1}^\top \right),$$

wobei $\widehat{d_1}$ die erste Spalte von $\widehat{D}$ und $\mathbb{1} = (1, 1, \ldots, 1, 1)^\top$ bezeichnet. Eine Eigenwertzerlegung $G = U \Lambda U^\top$ mit Diagonalmatrix Λ und orthogonaler Matrix U liefert dann die gesuchten Positionen. Jedoch entstehen bei fehlerbehafteten Werten in D und damit in $\widehat{D}$ oft hohe relative Fehler in der Rekonstruktion.

Zum Beispiel mit den Distanzen der sechs Städte im Rheinland ergibt das Verfahren basierend auf der Eigenwertzerlegung die Positionen in der folgenden Abbildung.



In dieser Arbeit soll alternativ das inverse Problem als ein nichtlineares Ausgleichsproblem numerisch gelöst werden. Dabei bestimmen sich die Koordinaten (x_i, y_i) für $i = 1, \ldots, N$ aus einem Minimierungsproblem mit dem Funktional

$$\min_{(x_1,y_1),\ldots,(x_N,y_N)} \sum_{i=1}^{N} \sum_{j=i+1}^{N} \left(\sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} - d_{ij} \right)^2.$$

Ein Vorteil ist hier, dass die Distanzwerte d_{ij} selbst auftreten und nicht deren Quadrate. Dadurch ist ein robusteres Verhalten bei fehlerbehafteten Distanzdaten zu erwarten. Ein Nachteil besteht darin, dass das Funktional nicht überall stetig differenzierbar ist wegen der Wurzelfunktion. Die Lösung des nichtlinearen Augleichsproblems muss mit iterativen Verfahren erfolgen. Daher sind auch geeignete Startwerte für Iteration erforderlich. Da die Lösung des inversen Problems nicht eindeutig ist, erscheint der Einbezug von Termen zur Regularisierung in die Minimierung sinnvoll.

Das Verhalten des Minimierungsproblems soll mit dem Ansatz über die Eigenwertzerlegung bei fehlerbehafteten Distanzdaten verglichen werden. Dies erfolgt sowohl durch eine Sensitivitätsanalyse als auch aus den numerischen Ergebnissen für Testbeispiele.

Vorkenntnisse: Analysis I+II, Lineare Algebra I+II, Numerik I

Literatur:

I. Dokmanić, R. Parhizkar, J. Ranieri, M. Vetterli: Euclidean Distance Matrices – Essential theory, algorithms and applications. IEEE Signal Processing Magazine, Vol. 32, No. 6 (2015), pp. 12-30.
`http://dokmanic.ece.illinois.edu/assets/pdf/Dokmanic2015eg.pdf`

R. Freund, R. Hoppe: Stoer/Bulirsch: Numerische Mathematik 1. (10. Aufl.) Springer Verlag, 2007. (Abschnitt 5.5)

Ansprechpartner: Prof. Dr. Roland Pulch
Email: `pulchr@uni-greifswald.de`