

Checkliste Lernziele:

- Mit welchen Hilfsmitteln lässt sich in R das Seriationsproblem lösen?
- Wie lassen sich in R Daten einfach ordnen?
- Wie lässt sich Lineare Programmierung umsetzen?
- Wie kann man Lineare Programmierung verwenden, um eine Reihenfolge zu erstellen?
- Wie lassen sich Score-Werte für Merkmalswerte berechnen?

Aufgaben:

1. Laden Sie den Datensatz `arch1.dat` und verwenden Sie die Hamming-Distanz, um Tabelle 5.1-2 nachzubilden.
2. Installieren und laden Sie das Paket `seriation` und schauen Sie sich die Hilfe für den Befehl `seriate` an. Achten Sie hier insbesondere auf die Angaben zur Option `method`. Warum ist das Verfahren der metrischen eindimensionalen Skalierung nicht implementiert? Wenden Sie den Befehl `seriate` auf die in der vorhergegangenen Aufgabe erzeugten Abstandsmatrix an. Verwenden Sie hierbei die Option `MDS`.
3. Laden Sie den Datensatz `dom1.dat`. Berechnen Sie die Zeilensummen der Dominanzmatrix und versuchen Sie, anhand dieser eine Ordnung zu konstruieren. Welche Probleme stellen sich?
4. Installieren und laden sie das Paket `lpSolve`.
5. Bilden Sie die Koeffizienten c'_{ij} entsprechend des Vorgehens von Michaud. Wie müssen diese abgespeichert werden, um sie mit dem Befehl `lp` benutzen zu können.
6. Bilden Sie drei Matrizen, die jeweils eine der folgenden Restriktion abbilden:

- $y_{ii} = 0$
- $y_{ij} + y_{ji} = 1; i \neq j$
- $y_{ij} + y_{jk} - y_{ik} \leq 1; i \neq j; i \neq k; j \neq k$

7. Kombinieren Sie die drei Matrizen aus der letzten Aufgabe zu einer Matrix, die Sie an den Befehl `lp` übergeben können. Erstellen Sie einen korrespondierenden Vektor. Schauen Sie sich die Hilfe für den Befehl `lp` an. Was müssen Sie noch erstellen, um den Befehl ausführen zu können?
8. Maximieren Sie mittels des Befehls `lp` das Zielkriterium unter den oben angegebenen Bedingungen. Konstruieren Sie aus dem Ergebnis eine Reihenfolge. Permutieren Sie die ursprüngliche Dominanzmatrix so, dass Sie dieses Ergebnis darstellt. Verwenden Sie den Befehl `pimage` aus dem Paket `seriate`, um eine grafische Darstellung zu erhalten. Vergleichen Sie das Resultat mit der ursprünglichen Dominanzmatrix.
9. Laden Sie wieder den Datensatz `arch1.dat`. Erstellen Sie für jede der Variablen eine Adjazenzmatrix nach dem Muster von Tabelle 5.3-1. Berechnen Sie die Abstände zwischen den Adjazenzmatrizen.
10. Geben Sie die auf S. 82 angegebene Matrix \mathbf{F} ein. Berechnen Sie die Matrix \mathbf{A} und die Singularwertzerlegung. Berechnen Sie die Score-Werte entsprechend Formel (6.4). Überlegen Sie, welchen Wert $\bar{s} = \sum_i \sum_j f_{ij} s_j / f_{..}$ annimmt. Überprüfen Sie Ihre Überlegungen, indem Sie \bar{s} berechnen.
11. Laden Sie den Datensatz `klaus1.dat` und erstellen Sie aus der erreichten Gesamtpunktzahl eine qualitative Variable. Schätzen Sie für diese eine Score-Regression, mit den Punktzahlen aus den einzelnen Aufgaben als unabhängigen Variablen. Interpretieren Sie die resultierenden Score-Werte und Koeffizienten.