

Multivariate Verteilungen

- a) Sei $X \sim N_p(\mu, \Sigma)$. Wie sind dann die einzelnen Komponenten X_i verteilt?
2. Tutorium S. 7: Normalverteilt.
- b) Wie hängen die Wishart-Verteilung und die multivariate Normalverteilung zusammen?
Was ist das eindimensionale Pendant?
2. Tutorium S. 8: $x_1, \dots, x_m \stackrel{\text{iid}}{\sim} N_p(0, \Sigma), \sum_{i=1}^m x_i x_i^T = M \sim W(\Sigma, m)$; χ^2 -Verteilung.

Diskriminanzanalyse

- a) Was versteht man unter Diskriminanzanalyse und was ist deren Idee?
3. Tutorium S. 4: Jedes Individuum soll in eine der k Klassen zugeordnet werden. vgl. Logit-Modell.
- b) Was versteht man unter Fehlklassifikation?
3. Tutorium S. 7: Individuum gehört eigentlich zur Klasse r , wird aber fälschlicherweise zur Klasse s zugeordnet, ϵ_{rs} ist die Fehlklassifikationswahrscheinlichkeit.
- c) Was minimiert die Bayes-Zuordnung?
3. Tutorium S. 9: Gesamt-Feherrate ϵ .
- d) Wann sind ML- und Bayes-Zuordnung äquivalent?
3. Tutorium S. 13: Bei gleichen a-priori-Wahrscheinlichkeiten.
- e) Was ist der Unterschied zwischen linearer und quadratischer Diskriminanzanalyse?
3. Tutorium S. 17 & 18: Diskriminanzfunktion, linear vs. quadratisch.

Clusteranalyse

- a) Was ist der Unterschied zur Diskriminanzanalyse?
4. Tutorium S. 4: Die Klassen sind vorab nicht bekannt und werden gesucht.
- b) Was ist der Unterschied zwischen Distanzmaß und Linkage-Methode?
4. Tutorium S. 7 & 10: Das Distanzmaß bestimmt, wie man die Distanz zwischen den Individuen misst und die Linkage-Methode, wie man die Distanz zwischen den Clustern bestimmt.
- c) Wie unterscheiden sich Single-Linkage und Complete-Linkage?
4. Tutorium S. 10: Single-Linkage ist der kürzester Abstand zwischen Individuen zweier Clustern, Complete-Linkage ist der längster Abstand zwischen Individuen zweier Clustern.

Multivariate Regression

- a) Worin unterscheidet sich das multivariate Regressionsmodell vom univariaten Regressionsmodell?
5. Tutorium S. 4 & 5: Wie der Name schon sagt, mehrere Zielvariablen vs. eine Zielvariable.
- b) Was ist der Vorteil des multivariaten Regressionsmodells gegenüber dem Univariaten?
5. Tutorium S. 8 & 11: Berücksichtigung der Kovarianzstruktur und flexible Hypothesenbildung.
- c) Unterscheiden sich die Regressionskoeffizienten, wenn man anstelle eines multivariaten Regressionsmodells einzelne univariate Regressionsmodelle berechnet?
5. Tutorium S. 9: Nein.
- d) Wie testet man lineare Hypothesen?
Welche Matrizen restringieren die Regressionskoeffizientenmatrix $\mathbf{B}$?
5. Tutorium S. 11: C und D, wobei das erstere die Kovariablen auswählt und das letztere die Zielvariablen selektiert.

Hauptkomponentenanalyse

- a) Wie wird die erste Hauptkomponente bestimmt?
Welche Nebenbedingung muss erfüllt sein?
6. Tutorium S. 7: Finde die Linear-Kombination, sodass die Varianz maximiert wird;
 $a_1^T a_1 = 1$.
- b) Welche Nebenbedingung muss für die zweite Hauptkomponente zusätzlich gelten?
6. Tutorium S. 8: $Cov(y_1, y_2) = 0 \Leftrightarrow a_1 \perp a_2$.
- c) Welchem bekannten Problem aus der linearen Algebra entspricht die Hauptkomponentenanalyse?
6. Tutorium S. 10: Eigenwertproblem.
- d) Wozu dient ein Screeplot?
6. Tutorium S. 18: Visuelle Bestimmung von Anzahl der Hauptkomponenten.