



Tutorium

Allgemeines Lineares Modell

BSc Psychologie SoSe 2023

3. Termin: (1) Regression (Teil 2)

Sean Mulready

Gibt es eine Methode, das 'optimale' k für den Grad eines Ausgleichspolynoms zu bestimmen?

- es gibt theoretisch viele Möglichkeiten, ein “optimales” k zu finden
- eine Möglichkeit: bei n Datenpunkten kann ich Grad $k = (n - 1)$ wählen
 - dann geht die Funktion exakt durch alle Punkte
- “optimal” im Sinne einer Abwägung: Güte vs. Vorhersagekraft
 - allgemein kann man sagen, dass das Ausgleichspolynom ab $k > 4$ immer “überangepasster” wird und an Vorhersagekraft verliert

weitere Anmerkungen von Prof. Ostwald:

- generell ist das Finden eines optimalen Grades ein Aspekt des offenen Forschungsthemas *Modellselektion*
 - mehr dazu [hier](#)
- Literaturempfehlung: Burnham, K.P. & Anderson, D.R. (2003). *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*. Springer Science & Business Media.

8. Erläutern Sie die Motivation des einfachen linearen Regressionsmodells in Bezug auf die Ausgleichsgerade.
9. Definieren Sie das Modell der einfachen linearen Regression.
10. Geben Sie das Theorem zur Datenverteilung der einfachen linearen Regression wieder.
11. Skizzieren das Modell der einfachen linearen Regression per Hand.
12. Skizzieren Sie eine Realisierung des Modells der einfachen linearen Regression per Hand.
13. Geben Sie das Theorem zur ML-Schätzung der Parameter der einfachen linearen Regression an.
14. Skizzieren Sie den Beweis des Theorems zur ML-Schätzung der Parameter der einfachen linearen Regression.

8. Erläutern Sie die Motivation des einfachen linearen Regressionsmodells in Bezug auf die Ausgleichsgerade.

Eine Ausgleichsgerade erlaubt Aussagen über unbeobachtete y Werte für x Werte. Der Wert von $q(\hat{\beta})$ quantifiziert die Güte der Ausgleichsgeradenpassung. Eine Ausgleichsgerade erlaubt allerdings nur implizite Aussagen über die mit der Anpassung verbundene Unsicherheit.

In der einfachen linearen Regression wird die Idee einer Ausgleichsgerade um eine probabilistische Komponente (normalverteilte Fehlervariable) erweitert, um quantitative Aussagen über die mit einer Ausgleichsgeradenanpassung verbundene Unsicherheit machen zu können.

Weiterhin erlaubt die einfache lineare Regression, einen Hypothesentest-basierten Zugang zur Einschätzung der angepassten Parameterwerte $\hat{\beta}_0$ und $\hat{\beta}_1$ sowie das Bestimmen von Konfidenzintervallen, die eine quantitative Aussage über die mit dem Schätzwert assoziierte Unsicherheit ermöglichen.

9. Definieren Sie das Modell der einfachen linearen Regression.

Definition (Modell der einfachen linearen Regression)

Für $i = 1, \dots, n$ sei

$$v_i = \beta_0 + \beta_1 x_i + \varepsilon_i \quad (1)$$

wobei

- $x_i \in \mathbb{R}$ fest vorgegebene sogenannte *Prädiktorwerte* oder *Regressorwerte* sind,
- $\beta_0, \beta_1 \in \mathbb{R}$ wahre, aber unbekannte, Parameterwerte sind und
- $\varepsilon_i \sim N(0, \sigma^2)$ unabhängige und identisch normalverteilte nicht-beobachtbare Zufallsvariablen mit wahrem, aber unbekanntem, Parameter $\sigma^2 > 0$ sind.

Dann heißt (1) *Modell der einfachen linearen Regression*.

10. Geben Sie das Theorem zur Datenverteilung der einfachen linearen Regression wieder.

Theorem (Datenverteilung der einfachen linearen Regression)

Das Modell der einfachen linearen Regression

$$v_i = \beta_0 + \beta_1 x_i + \varepsilon_i \text{ mit } \varepsilon_i \sim N(0, \sigma^2) \text{ u.i.v. für } i = 1, \dots, n \quad (2)$$

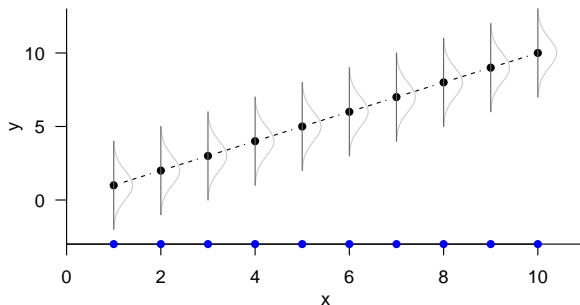
lässt sich äquivalent in der Form

$$v_i \sim N\left(\beta_0 + \beta_1 x_i, \sigma^2\right) \text{ u.v. für } i = 1, \dots, n \quad (3)$$

schreiben.

11. Skizzieren das Modell der einfachen linearen Regression per Hand.

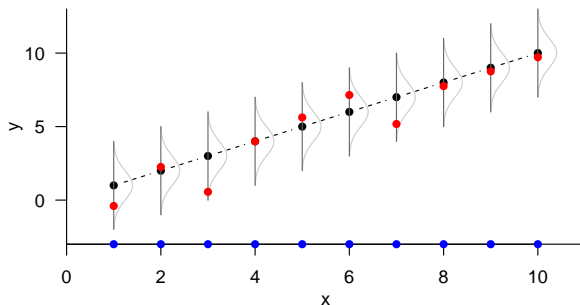
Modell der einfachen linearen Regression



• x_i • $\beta_0 + \beta_1 x_i$ für $\beta_0 := 0, \beta_1 := 1$ — $N(y_i; \beta_0 + \beta_1 x_i, \sigma^2)$ für $\sigma^2 := 1$.

12. Skizzieren Sie eine Realisierung des Modells der einfachen linearen Regression per Hand.

Realisierung des Modells der einfachen linearen Regression



13. Geben Sie das Theorem zur ML-Schätzung der Parameter der einfachen linearen Regression an.

Theorem (Maximum Likelihood Schätzung)

Es sei

$$v_i = \beta_0 + \beta_1 x_i + \varepsilon_i \text{ mit } \varepsilon_i \sim N(0, \sigma^2) \text{ u.i.v. für } i = 1, \dots, n \quad (4)$$

das Modell der einfachen linearen Regression. Dann sind Maximum Likelihood Schätzer der Modellparameter β_0, β_1 und σ^2 gegeben durch

$$\hat{\beta}_1 := \frac{c_{xy}}{s_x^2}, \quad \hat{\beta}_0 := \bar{y} - \hat{\beta}_1 \bar{x} \quad \text{und} \quad \hat{\sigma}^2 := \frac{1}{n} \sum_{i=1}^n \left(y_i - \left(\hat{\beta}_0 + \hat{\beta}_1 x_i \right) \right)^2. \quad (5)$$

14. Skizzieren Sie den Beweis des Theorems zur ML-Schätzung der Parameter der einfachen linearen Regression.

Teil 1/2: $\hat{\beta}_0$ und $\hat{\beta}_1$

Wir wollen zunächst zeigen, dass die Ausgleichsgeradenparameter $\hat{\beta}_0$ und $\hat{\beta}_1$ den entsprechenden ML Schätzern gleichen.

Um die ML Schätzer zu bestimmen, formulieren wir zunächst die Likelihood-Funktion des Modells der einfachen linearen Regression in Abhängigkeit von β_0 und β_1 .

Die Likelihood-Funktion ist definiert als der Wert der gemeinsamen Verteilung der v_1, \dots, v_n in Abhängigkeit von den Parametern β_0 und β_1 .

Aufgrund der Unabhängigkeit der v_1, \dots, v_n können wir die gemeinsame Verteilung als Produkt der einzelnen Wahrscheinlichkeitsdichtefunktionen, also als Produkt von Dichtefunktionen der univariaten Normalverteilung aufschreiben.

Die funktionale Form der Dichtefunktionen der univariaten Normalverteilung enthält eine Exponentialfunktion. Mit den Eigenschaften einer Exponentialfunktion können wir dieses Produkt umschreiben zu einer Exponentialfunktion von einem Term, der im Wesentlichen aus der negativen Summe der quadrierten Abweichungen (i.e. der Funktion q) besteht.

Weil für eine Exponentialfunktion gilt, dass für $a < b \leq 0$ gilt, dass $\exp(a) < \exp(b)$, wird der Exponentialterm der Likelihood-Funktion maximal, wenn q minimal und entsprechend $-q$ maximal wird.

Wie im Beweis der Ausgleichsgeradenform gezeigt, wissen wir, dass q für $\hat{\beta}_0$ und $\hat{\beta}_1$, wie sie auch im Theorem zur ML-Schätzung der Parameter der einfachen linearen Regression angegeben sind, minimal wird, und damit $\hat{\beta}_1$ und $\hat{\beta}_0$ die Likelihood-Funktion maximieren.

14. Skizzieren Sie den Beweis des Theorems zur ML-Schätzung der Parameter der einfachen linearen Regression.

Teil 2/2: $\hat{\sigma}^2$

Als nächstes wollen wir zeigen, dass $\hat{\sigma}^2$ dem ML-Schätzer entspricht.

Dazu betrachten wir analog zu oben die Likelihood-Funktion des Modells der einfachen linearen Regression, jedoch als Funktion von σ^2 und formulieren die entsprechende log-Likelihood-Funktion.

Wir wollen das $\hat{\sigma}^2$ bestimmen, für das die (log-)Likelihood-Funktion maximal wird.

Um die log-Likelihood-Funktion zu maximieren, bilden wir die 1. Ableitung, setzen diese gleich 0 und lösen nach σ^2 auf. Durch umstellen erhalten wir dann die Formel zur Schätzung von σ^2 , also $\hat{\sigma}^2$, wie sie im Theorem angegeben ist.