



Mathematisch-algorithmische Grundlagen für Data Science

Aufgabenblatt 5

Abgabe zu **zweit** vor der Vorlesung am 14. Mai 2024 oder via via E-Mail an peter.becker@h-brs.de.

Sollpunktzahl: 10 Punkte

Hinweis: Die Daten für die Aufgaben 1 bis 3 können Sie von der Homepage als Java-Arrays herunterladen.

Aufgabe 1 (Lineare Ausgleichsrechnung)

2+4+2=8 Punkte

Für die Messwerte

x_i	0.0	0.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0
y_i	5.1	7.8	7.1	7.4	9.4	8.4	8.9	6.2	4.7	3.3	0.6

vermuten wir einen Zusammenhang der Form

$$y = a \cdot x^2 + b \cdot x + c.$$

- Wie lautet das Fehlergleichungssystem?
- Stellen Sie das Normalgleichungssystem auf und lösen Sie es (hier ist ein kleines Programm vermutlich hilfreich).
- Plotten Sie die Messwerte und die Ausgleichsfunktion.

Aufgabe 2 (Polynominterpolation vs. lineare Regression)

5 Punkte

Bei einer zufällig ausgewählten Gruppe von Zuschauern an einem Basketballspiel kamen die nebenstehenden Messungen zustande.

Körpergröße	Gewicht
1.55	51
1.57	50
1.62	55
1.68	52
1.75	60
1.76	68
1.81	78
1.83	91
1.87	84
1.89	81
1.90	90
1.92	105
1.95	95
1.96	99
1.99	100
2.02	101

Schreiben Sie ein Programm das folgendes leistet:

- Es wählt von den 16 Datensätzen acht Datensätze zufällig aus,
- berechnet zu diesen acht Datensätzen die Regressionsgerade,
- berechnet weiterhin ein Polynom siebten Grades, das exakt durch die ausgewählten acht Datenpunkte verläuft (vgl. Beispiel 3.7) und
- ermittelt sowohl für die Regressionsgerade als auch für das Polynom die Summe der quadrierten Fehler der *nicht* ausgewählten acht Datensätze.

Lassen Sie Ihr Programm mehrmals laufen und schauen Sie sich die Ergebnisse an. Welche Methode ist für die Prognose des Gewichts auf Basis der Körpergröße besser geeignet?

Hinweis: Für die Berechnung des Polynoms können Sie Ihre eigenen Methoden zur Lösung linearer Gleichungssysteme verwenden oder Sie benutzen die Methoden aus der Musterlösung (siehe Aufgabenblätter 3 und 4).

Abgabe: Geben Sie an, welche acht Datensätze ausgewählt wurden, das zugehörige Interpolationspolynom, die zugehörige Regressionsgerade und die Summe der quadrierten Fehler für das Interpolationspolynom und die Gerade.

Aufgabe 3 (Logarithmisch-lineare Regression)

6 Punkte

Für die gegebenen Daten (siehe Datei `DatenB5.java` verlinkt auf der Homepage) vermuten wir folgenden funktionalen Zusammenhang:

$$f(x) = \alpha x^\beta e^{\gamma x}.$$

- Erstellen Sie einen Plot der Daten.
- Ermitteln Sie eine Ausgleichsfunktion mittels logarithmisch-linearer Regression.
- Erweitern Sie Ihren Plot aus (a) um die Ausgleichsfunktion.