



Mathematisch-algorithmische Grundlagen für Data Science

Aufgabenblatt 6

Abgabe zu **zweit** vor der Vorlesung am 21. Mai 2024.

Sollpunktzahl: 10 Punkte

Aufgabe 1 (Multiple Regression)

6 Punkte

Gegeben seien die Daten

$x_{i,1}$	$x_{i,2}$	y_i
0.35	5.81	98.9
4.84	4.58	1134.5
9.51	7.68	20916.8
7.54	7.52	7489.1
2.05	0.88	37.6
4.27	3.87	663.7
6.49	3.30	4666.0
2.36	1.54	56.6
8.55	6.45	13605.3
6.44	3.98	4339.3

Wir vermuten einen funktionalen Zusammenhang der Form

$$f(x_1, x_2) = \lambda_1(x_2 - x_1^2)^2 + \lambda_2(1 - x_1)^2$$

Ermitteln Sie eine Ausgleichsfunktion mittels multipler Regression.

Aufgabe 2 (Frequenzanalyse mittels linearer Ausgleichsrechnung) 7+3=10 Punkte

Wir wollen eine Funktion f lernen, von der wir annehmen, dass sie eine Überlagerung von Cosinus-Schwingungen zwischen 0 Hz und 10 Hz ist, also

$$f(x) = \lambda_0 + \sum_{k=1}^{10} \lambda_k \cos(kx).$$

Wir wissen nicht, welche Frequenzen k tatsächlich vorhanden sind und wie groß deren Amplitude λ_k ist.

Wir haben für $n + 1$ Stellen $-\pi = x_0 < x_1 < x_2 < \dots < x_n = \pi$ jeweils einen Messwert y_i von f , der aber durch leichtes gleichförmiges Rauschen überlagert ist, also

$$y_i = f(x_i) + r(x_i)$$

mit einer Rauschfunktion $r(x)$.

- (a) Auf der Homepage der Vorlesung finden Sie einen Link zu einem Java-Quelltextfragment, das die Werte für y_i in einem Java-Array definiert ($n = 1024$). Die zugehörigen Stützstellen sind

$$x_i = -\pi + \frac{2\pi i}{n}, \quad i = 0, \dots, n.$$

Schätzen Sie mithilfe der linearen Ausgleichsrechnung die Parameter $\lambda_0, \dots, \lambda_{10}$ und geben Sie diese Werte an.

- (b) Wenn Sie sich die Schätzung für die λ_k anschauen, werden Sie feststellen, dass einige recht nahe bei Null liegen. Wir können daher davon ausgehen, dass diese Frequenzen in der Funktion f tatsächlich nicht auftreten.

Entscheiden Sie auf Basis der Parameterschätzung von (a), welche Frequenzen in f tatsächlich vorhanden sind und berechnen Sie auf dieser Basis die Parameter λ_k neu.