Anastasia August, Andreas Reiter, Andreas Prahs, Britta Nestler HsKa und KIT

Praktikum zur Vorlesung
Modellierung und Simulation im SS 2022

Praktikum 3: Ausgleichsproblem

Verwenden Sie das beigefügte Template `ausgleichsproblem.c`.

Aufgabenbeschreibung:

Anhand der in der Datei `input.dat` aufgeführten Daten des RKI (`https://www.rki.de/DE/Content/InfAZ/N/Neuartiges_Coronavirus/Daten/Fallzahlen_Kum_Tab.html`, Stand 25.11.2021) soll ein einfaches Modell für die wöchentliche 7-Tage-Indizidenz für den Spätsommer 2021 in der Altersklasse von 20-24 Jahren abgeleitet werden.

Hierfür wird zunächst als einfaches Modell eine exponentielle Wachstumskurve der Form $I(t-t_0) = a\exp\big(b(t-t_0)\big)$ mit den zwei Unbekannten Konstanten a und b postuliert. t und t_0 sind dabei die aktuelle Zeit sowie der Zeitpunkt der ersten angegebenen Messung (beide gemessen in „Kalenderwochen"). Durch passende Wahl der Konstanten a und b soll diese Kurve an die gegebenen Datenpunkte gefittet werden, um so z.B. eine Prognose für die weitere Entwicklung treffen zu können.

Im Gegensatz zu den in der Vorlesung behandelten Problemen würde es sich in dieser Form zunächst um ein nicht-lineares Ausgleichsproblem handeln. Aufgrund seiner speziellen Form kann das Modell jedoch durch „logarithmieren" wegen

$$\ln\big(a\exp\left(b(t-t_0)\right)\big) \underbrace{=}_{\ln(x\cdot y)=\ln(x)+\ln(y)} \ln(a)+\ln\big(\exp\left(b(t-t_0)\right)\big) \underbrace{=}_{\ln\left(\exp(z)\right)=z} \ln(a)+b(t-t_0)$$

in die Form

$$\ln(a) + b(t-t_0) = \ln\big(I(t-t_0)\big) \tag{1}$$

eines linearen Ausgleichsproblem für die beiden Konstanten $\ln(a)$ und b überführt werden.

Wie Sie es bereits in der Vorlesung gelernt haben, wird für dessen Lösung ein lineares Gleichungssystem der Form

$$\mathbf{A^T A}\lambda = \mathbf{A^T y} \tag{2}$$

nach dem Vektor $\lambda = (\ln(a), b)^T$ aufgestellt. Aus Gleichung (1) ist der Vektor $\mathbf{y}$ hier über $y_i = \ln\big(I(t_i - t_0)\big)$ gegeben, und die Matrix $\mathbf{A}$ ist entsprechend der Vorfaktoren in Abhängigkeit der gegebenen Kalenderwochen t_i, $i = 0, 1, ..., n-1$ wie folgt aufgebaut:

$$\mathbf{A} = \begin{pmatrix} 1 & 1 & \dots & 1 \\ (t_0 - t_0) = 0 & (t_1 - t_0) = 1 & \dots & (t_{n-1} - t_0) = n-1 \end{pmatrix}^T.$$

Bearbeiten Sie folgende Schritte:

1. Befüllen Sie den Hilfsvektor `logI` mit den Logarithmen der aus der `input.dat` Datei eingelesenen Inzidenzwerte. Hierfür kann die Bibliotheks-Funktion `log` verwendet werden, welche den natürlichen Logarithmus ihres Arguments zurückgibt.

2. Berechnen sie mit Hilfe der eingelesenen t_i-Werte die Matrix $A^T A$.

3. Lösen Sie das Normalgleichungssystem in Gleichung (1) und geben Sie die Parameter $\ln(a)$ bzw. a und b an.

4. Plotten Sie die Datenpunkte (t_i, I_i) zusammen mit der berechneten Ausgleichsfunktion, indem sie `useplot` auf 1 setzen. Um eine alternative graphische Darstellung entsprechend der Gleichung (1) zu erhalten, können Sie durch Einkommentieren der Zeile `"set logscale xy\n"` eine doppelt-Logarithmische Darstellung wählen.

   Wie bewerten Sie die Qualität des Modells? Wie sieht es aus, wenn Sie einige der späteren Messpunkte entfernen?

Hinweis: Für die Lösung der Teilaufgabe 3) muss das 2×2-Gleichungssystem (2) für $(\lambda_1, \lambda_2)^T$ gelöst werden. Nachdem es sich hier um ein sehr kleines Gleichungssystem handelt, können Sie dies z.B. tun, indem sie sich die Inverse der Matrix $\mathbf{A^T A}$ berechnen. Hierfür zur Erinnerung: ist die Matrix $\begin{pmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{pmatrix}$ invertierbar, so ist deren Inverse über $\frac{1}{m_{11}m_{22} - m_{21}m_{12}} \begin{pmatrix} m_{22} & -m_{12} \\ -m_{21} & m_{11} \end{pmatrix}$ gegeben.