Abgabe Blatt 9

Jean-Marco Alameddine, Johannes Kollek, Max Pernklau

Tuesday, January 17, 2017

Aufgabe 1

a)

Die Likelihoodfunktion ergibt sich zu

$$L = \prod_{n_i} \frac{\lambda^{n_i}}{n_i!} e^{-\lambda},$$

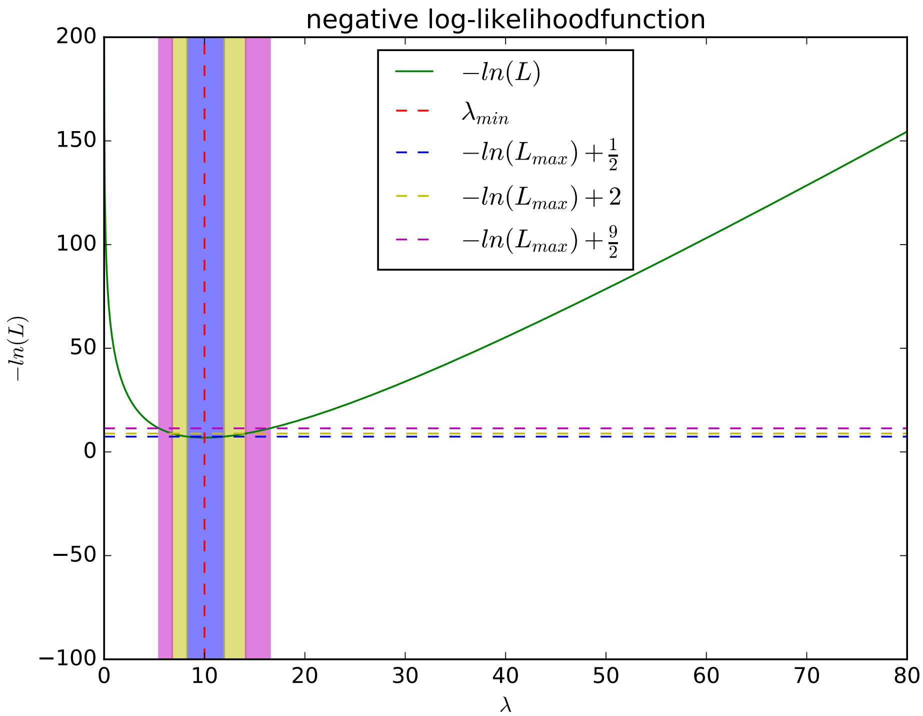
wobei

$$n_i = \{13, 8, 9\}$$

annehmen kann. Aus der Likelihoodfunktion folgt schließlich

$$\begin{aligned} -\ln(L) &= -\sum_{n_i} \ln\left(\frac{\lambda^{n_i}}{n_i!} e^{-\lambda}\right) \\ &= 3\lambda - \ln\left(\frac{\lambda^{13}}{13!}\right) - \ln\left(\frac{\lambda^{8}}{8!}\right) - \ln\left(\frac{\lambda^{9}}{9!}\right) \\ &= 3\lambda - 30\ln(\lambda) + \ln(13! \cdot 9! \cdot 8!) \end{aligned}$$

für die negative Log-Likelihoodfunktion, welche in der nächsten Abbildung geplottet wird.



b)

Im Folgenden wird das Minimum der Funktion bestimmt.

$$\frac{\partial(-\ln(L))}{\partial\lambda} = 3 - \frac{30}{\lambda} \stackrel{!}{=} 0$$
$$\Rightarrow \lambda_{min} = 10.$$

c)

Aus der Bedingung

$$-\ln(\lambda_{min}) + a = -\ln(\lambda)$$

folgt mit $\lambda_{min} = 10$

$$3\lambda + 30(\ln(\lambda) - 1 - \ln(10)) - a = 0,$$

wobei

$$a = \left\{\frac{1}{2}, 2, \frac{9}{2}\right\}$$

ist. Die Nullstelle wird numerisch über das Newton-Verfahren mit Kenntnis der Ableitung der negativen Log-Likelihoodfunktion bestimmt, wobei jeweils rechts und links von λ_{min} gestartet wird. Daraus ergeben sich die Grenzen der Intervalle

$$[8.284, 11.939] \text{ und } len = 3.655,\ \text{für } a = \frac{1}{2};$$
$$[6.779, 14.109] \text{ und } len = 7.33,\ \text{für } a = 2;$$
$$[5.474, 16.52] \text{ und } len = 11.046,\ \text{für } a = \frac{9}{2}.$$

Diese Intervalle sind ebenfalls im Plot eingezeichnet. Diese Intervalle können Konfidenzintervalle bzgl. der Schätzung des Erwartungswertes λ darstellen. Liegt das tatsächliche λ außerhalb dieses gewählten Bereiches kann die Schäzung verworfen werden.

d)

Das zweite Taylorpolynom ergibt sich zu

$$T_2(10, -\ln(L)) = 10 - 30\ln(10) + \frac{3}{20}(\lambda - 10)^2 + \ln(13! \cdot 9! \cdot 8!).$$

Diese Funktion ist in der nächsten Abbildung zu sehen.

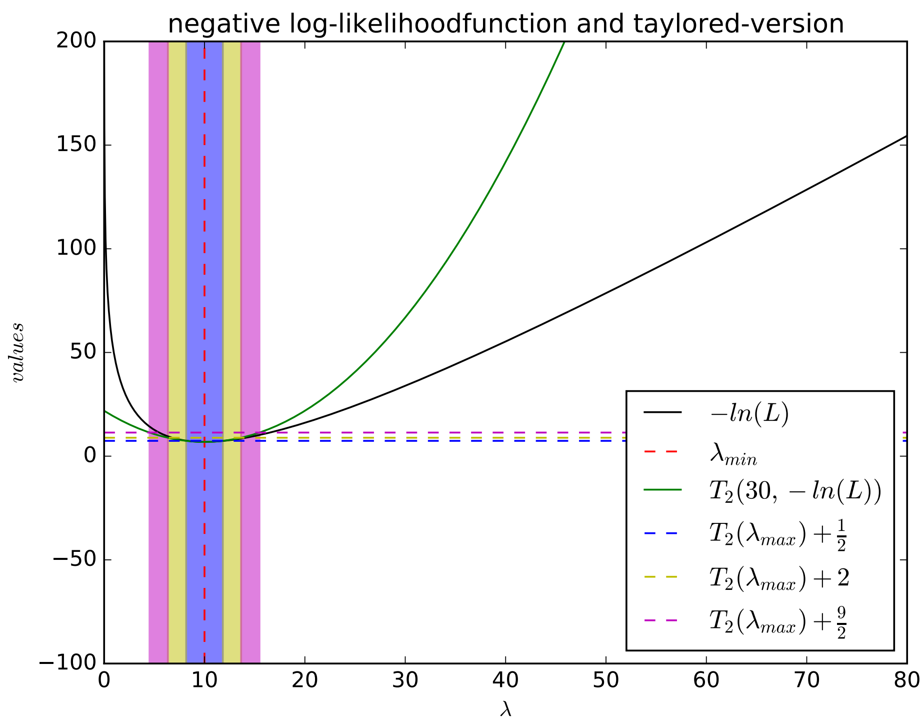


Figure 1: Negative Log-Likelihoodfunktion und Taylorpolynom.

Die λ werden auch hier numerisch bestimmt, wobei diese hier auch analytisch gewonnen werden können. Die Intervalle lauten

$$
\begin{aligned}
[8.174, 11.826] \text{ und } len = 3.651, \ \text{für } a = \frac{1}{2};\\
[6.349, 13.651] \text{ und } len = 7.303, \ \text{für } a = 2;\\
[4.523, 15.477] \text{ und } len = 10.954, \ \text{für } a = \frac{9}{2}.
\end{aligned}
$$

Die relativen Abweichungen zum exakten Ergebnis lauten

$$
\begin{aligned}
0.093\,\%, \ \text{für } a = \frac{1}{2};\\
0.369\,\%, \ \text{für } a = 2;\\
0.828\,\%, \ \text{für } a = \frac{9}{2}.
\end{aligned}
$$

Diese Abweichungen sind sehr gering und somit ist das Taylorpolynom eine gute Näherung an die analytisch kompliziertere negative Log-Likelihoodfunktion. Der Vorteil am Polynom ist, dass damit viel besser gerechnet werden kann.

Aufgabe 2

Die Aufgabe wurde handschiftlich bearbeitet. Wo Rechenschritte fehlen, beispielsweise beim Berechnen vom Lösungsvektor, wird die Rechnung im Code in der Datei *aufg2.py* durchgeführt. Diese enthält als Ausgabe ebenfalls die in der Aufgabenstellung gefragten Werte.

Nr. 2

a) $f(\varphi) = a_1 f_1(\varphi) + a_2 f_2(\varphi)$

$= a_1 \cos(\varphi) + a_2 \sin(\varphi)$

$$A = \begin{pmatrix} f_1(\varphi_1) & f_2(\varphi_1) \\ \vdots & \vdots \\ f_n(\varphi_n) & f_n(\varphi_n) \end{pmatrix} = \begin{array}{c} \begin{pmatrix} 1 & 0 \\ \sqrt{3}/2 & 1/2 \\ 1/2 & \sqrt{3}/2 \\ 0 & 1 \\ -1/2 & \sqrt{3}/2 \\ -\sqrt{3}/2 & 1/2 \\ -1 & 0 \\ -\sqrt{3}/2 & -1/2 \\ -1/2 & -\sqrt{3}/2 \\ 0 & -1 \\ 1/2 & -\sqrt{3}/2 \\ \sqrt{3}/2 & -1/2 \end{pmatrix} \end{array} \begin{array}{c} 0^\circ \\ 30^\circ \\ 60^\circ \\ 90^\circ \\ 120^\circ \\ 150^\circ \\ 180^\circ \\ 210^\circ \\ 240^\circ \\ 270^\circ \\ 300^\circ \\ 330^\circ \end{array}$$

b) Lösungsvektor: $\hat{\vec{a}} = (A^T A)^{-1} A^T \vec{y}$

$$A^T A = \begin{pmatrix} 6 & 0 \\ 0 & 6 \end{pmatrix} \rightsquigarrow (A^T A)^{-1} = \begin{pmatrix} 1/6 & 0 \\ 0 & 1/6 \end{pmatrix}$$

mit $\vec{y} = \begin{pmatrix} -0{,}032 \\ 0{,}010 \\ 0{,}057 \\ \vdots \end{pmatrix}$

$$\Rightarrow \hat{\vec{a}} = \begin{pmatrix} -0{,}0375063 \\ 0{,}07739978 \end{pmatrix} = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$$

c) Asymmetriewerte sind (vermutlich) unkorelliert mit $\sigma = 0{,}011$

$$\Rightarrow V[\hat{\vec{a}}] = \sigma^2 (A^T A)^{-1}$$

$$= (0{,}011)^2 \cdot \begin{pmatrix} 1/6 & 0 \\ 0 & 1/6 \end{pmatrix} = \begin{pmatrix} 2{,}0167 \cdot 10^{-5} & 0 \\ 0 & 2{,}0167 \cdot 10^{-5} \end{pmatrix}$$

$$= V[\hat{\vec{a}}]$$

$\Rightarrow$ Fehler von a_1/a_2 ist $\sqrt{2{,}0167 \cdot 10^{-5}} \approx 0{,}0044907$

Korellationskoeffizient ist 0, da die Nebendiagonalelemente der Kovarianzmatrix verschwinden

d) $f(\varphi) = a_1 \cos(\varphi) + a_2 \sin(\varphi)$

$A_0 \cdot \cos(\varphi + \delta) = A_0 \cdot \cos(\varphi) \cos(\delta) - A_0 \sin(\varphi) \cdot \sin(\delta)$

$\Rightarrow A_0 \cdot \cos(\delta) = a_1 \; (\mathrm{I}) \wedge -A_0 \cdot \sin(\delta) = a_2 \quad (\mathrm{II})$

dh $\frac{\mathrm{II}}{\mathrm{I}}$: $\quad -\tan(\delta) = \frac{a_2}{a_1} \quad \Rightarrow \quad \underline{\delta = \arctan\left(-\frac{a_2}{a_1}\right)}$

und: aus I: $A_0 \cdot \cos\left(\arctan\left(-\frac{a_2}{a_1}\right)\right) = a_1$

$\Leftrightarrow A_0 \cdot \frac{1}{\sqrt{a_2^2/a_1^2 + 1}} = a_1$

$\Rightarrow \underline{\underline{A_0 = a_1 \sqrt{a_2^2/a_1^2 + 1}}}$

$\Rightarrow \boxed{\begin{aligned} A_0 &= -0{,}086008 \\ \delta &= 1{,}1195615 \end{aligned}}$

Transformation? $V[A_0, \delta] = J \cdot V[a_1, a_2] \cdot J^T$ mit $V[a_1, a_2] = \begin{pmatrix} 2{,}0167 \cdot 10^{-5} & 0 \\ 0 & 2{,}0167 \cdot 10^{-5} \end{pmatrix}$

$J = \begin{pmatrix} \frac{\partial A_0}{\partial a_1} & \frac{\partial A_0}{\partial a_2} \\ \frac{\partial \delta}{\partial a_1} & \frac{\partial \delta}{\partial a_2} \end{pmatrix}$

$\frac{\partial A_0}{\partial a_1} = \left(\frac{a_2^2}{a_1^2} + 1\right)^{-1/2} \approx 0{,}436077$

$\frac{\partial A_0}{\partial a_2} = \frac{a_2}{a_1} \left(\frac{a_2^2}{a_1^2} + 1\right)^{-1/2} \approx -0{,}89991$

$\frac{\partial \delta}{\partial a_1} = \frac{a_2}{a_1^2 + a_2^2} \approx 10{,}46304$

$\frac{\partial \delta}{\partial a_2} = \frac{-a_1}{a_1^2 + a_2^2} \approx 5{,}07017$

$\Rightarrow V[A_0, \delta] = \begin{pmatrix} 2{,}016 \cdot 10^{-5} & -1{,}12707 \cdot 10^{-10} \\ -1{,}12707 \cdot 10^{-5} & 2{,}72528 \cdot 10^{-3} \end{pmatrix}$

$\Rightarrow$ • $\sigma_{A_0} \approx 0{,}00449$

• $\sigma_\delta \approx 0{,}0522$

• $\rho_{A_0,\delta} = \frac{\sigma^2_{A_0,\delta}}{\sigma_{A_0} \sigma_\delta} = \frac{-1{,}12707 \cdot 10^{-10}}{\sigma_{A_0} \cdot \sigma_\delta} \approx -0{,}00205$ als Korrelationskoeffizient

Zudem führt *auf2_test.py* den zuvor berechneten Fit durch, um zu Zeigen dass die Methode erfolgreich war.

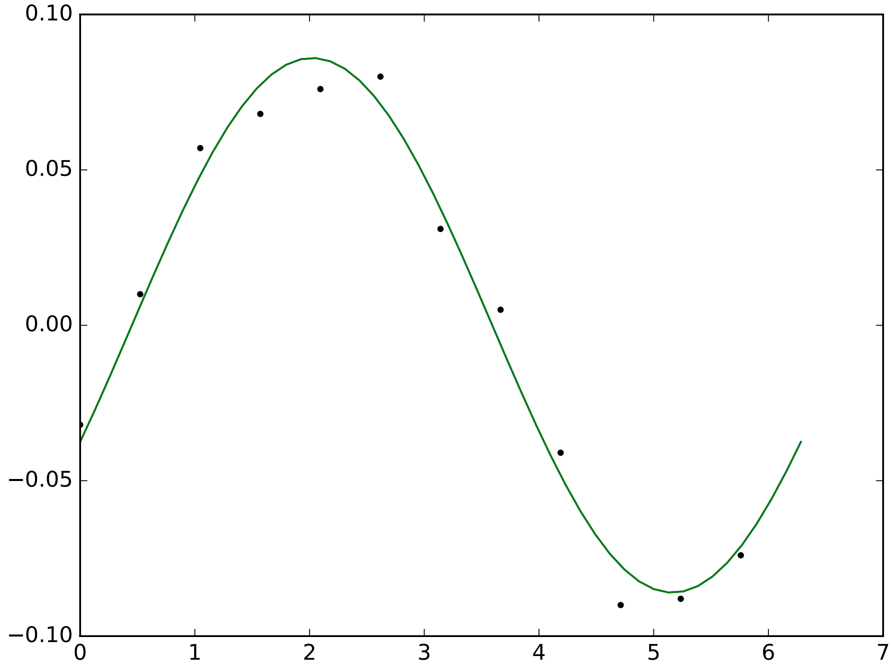


Figure 2: Aufgabe 2- Fit.

Aufgabe 3

a)

Zunächst werden die Daten mit der Methode der kleinsten Quadrate gefittet. Der Lösungsvektor ergibt sich aus

$$\hat{a} = (A^\top A)^{-1} A^\top \vec{y} \tag{1}$$

Die Koeffizienten lauten

$$a_0 = -6.744\,532\,69 \times 10^{-2} \tag{2}$$
$$a_1 = 6.096\,090\,41 \times 10^{-1} \tag{3}$$
$$a_2 = -5.137\,482\,17 \times 10^{-1} \tag{4}$$
$$a_3 = 2.105\,665\,23 \times 10^{-1} \tag{5}$$
$$a_4 = -4.520\,077\,56 \times 10^{-2} \tag{6}$$
$$a_5 = 4.785\,680\,54 \times 10^{-3} \tag{7}$$
$$a_6 = -1.962\,881\,98 \times 10^{-4} \tag{8}$$

(*Anmerkung des Autorenteams*: Die Koeffizienten und alle folgenden Werte werden auch von der Konsole beim Ausführen von *aufg3.py* ausgeben. Im Zweifel sind die dort ausgegebenen Werte die richtigen, da Kopierfehler immer passieren können...)

b)

Als nächstes wird eine Regularisierung genutzt. Dabei wird der Parameter λ varriert.

Der Schätzparameter ergibt sich zu

$$\hat{a}^{reg} = (A^\top A + \lambda(CA)^\top(CA))^{-1}A^\top\vec{y} \tag{9}$$

wobei C Teil einer Regularisierung mithilfe der zweiten numerischen Ableitung ist (siehe Skript, Seite 92, Kapitel Testen.)

(*Anmerkung des Autorenteams*: Uns ist nicht klar, wieso C so aussehen muss, um eine zweite Ableitung zu repräsentieren, bzw. ob C hier richtig angewendet wird. Eine Aufklärung in der Übungs wäre vorteilhaft!)

Die Koeffizienten für alle Lambdas werden durch die Konsole ausgegeben, als Beispiel seien die Koeffizienten für $\lambda = 0.1$ angegeben:

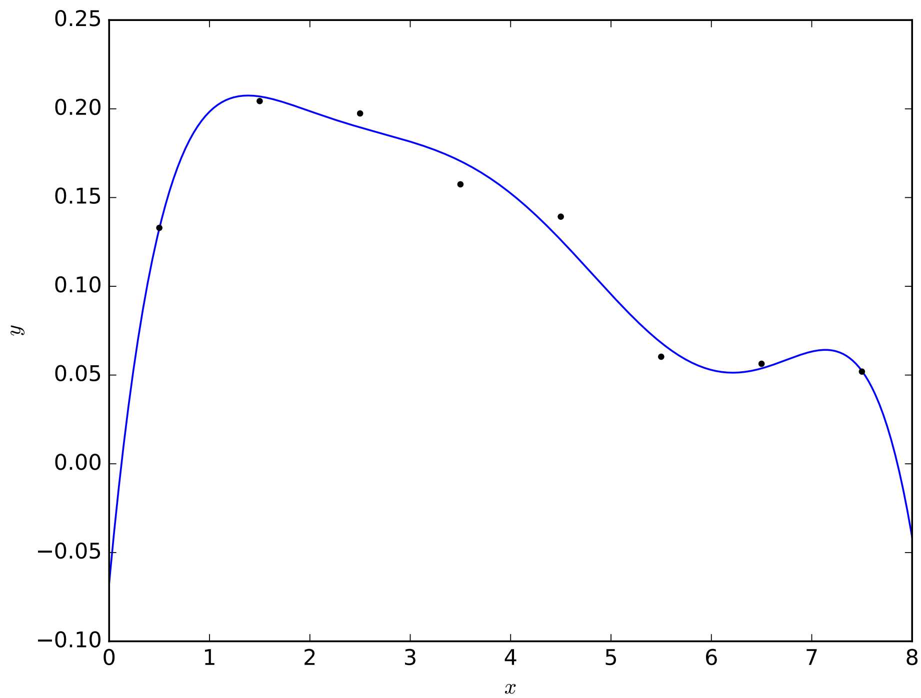


Figure 3: Aufgabe 3a - Fit.

$$a_0 = 5.279\,658\,56 \times 10^{-2} \tag{10}$$
$$a_1 = 2.595\,311\,49 \times 10^{-1} \tag{11}$$
$$a_2 = -1.932\,312\,85 \times 10^{-1} \tag{12}$$
$$a_3 = 7.696\,672\,46 \times 10^{-2} \tag{13}$$
$$a_4 = -1.716\,280\,69 \times 10^{-2} \tag{14}$$
$$a_5 = 1.903\,764\,83 \times 10^{-3} \tag{15}$$
$$a_6 = -8.103\,496\,97 \times 10^{-5} \tag{16}$$

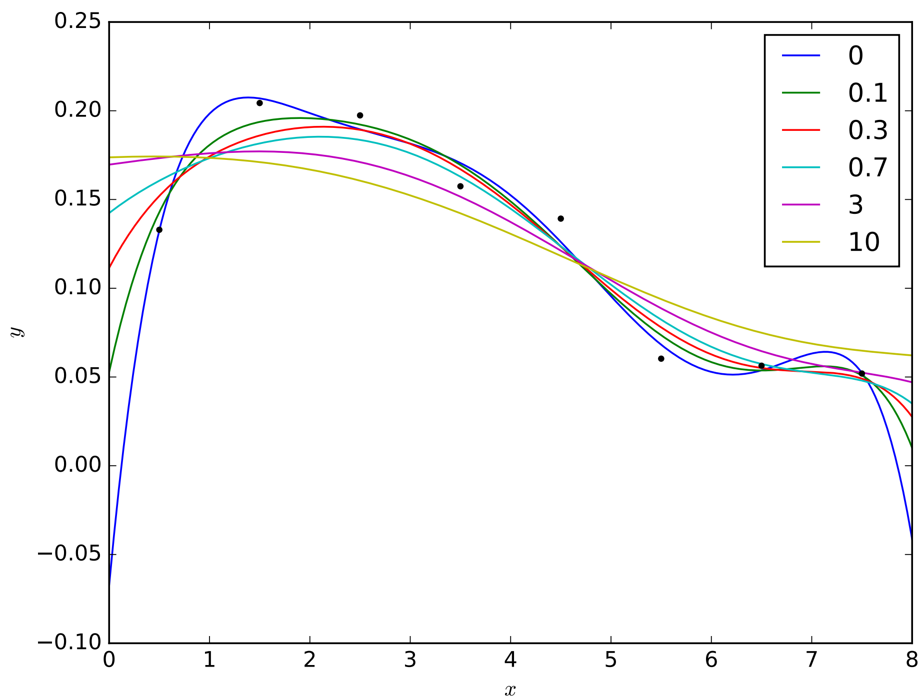


Figure 4: Aufgabe 3b - Fit.

c)

Zuletzt soll an einer größeren Datenmenge die gewichtete Methode zum Fitten genutzt weden.

Gefittet wird dabei jeweils an den Mittelwert der Messdaten zu jedem x, das Quadrat der Inverse vom Fehler des Mittelwertes wird dabei jeweils als Wichtungsparameter verwendet. (Ist das Qudarat hier richtig/nötig?)

Als Schätzer ergibt sich nun

$$\hat{a} = (A^\top W A)^{-1} A^\top W \vec{y} \tag{17}$$

wobei die Gewichtsmatrix W jeweils als Diagonalelemente die bereits angesprochenen Gewichte enthält.

Als Gewichte ergeben sich dabei:

$$a_0 = -1.114\,069\,47 \times 10^{-1} \tag{18}$$
$$a_1 = 7.632\,110\,82 \times 10^{-1} \tag{19}$$
$$a_2 = -6.800\,067\,87 \times 10^{-1} \tag{20}$$
$$a_3 = 2.893\,991\,63 \times 10^{-1} \tag{21}$$
$$a_4 = -6.339\,377\,99 \times 10^{-2} \tag{22}$$
$$a_5 = 6.792\,775\,85 \times 10^{-3} \tag{23}$$
$$a_6 = -2.810\,139\,23 \times 10^{-4} \tag{24}$$

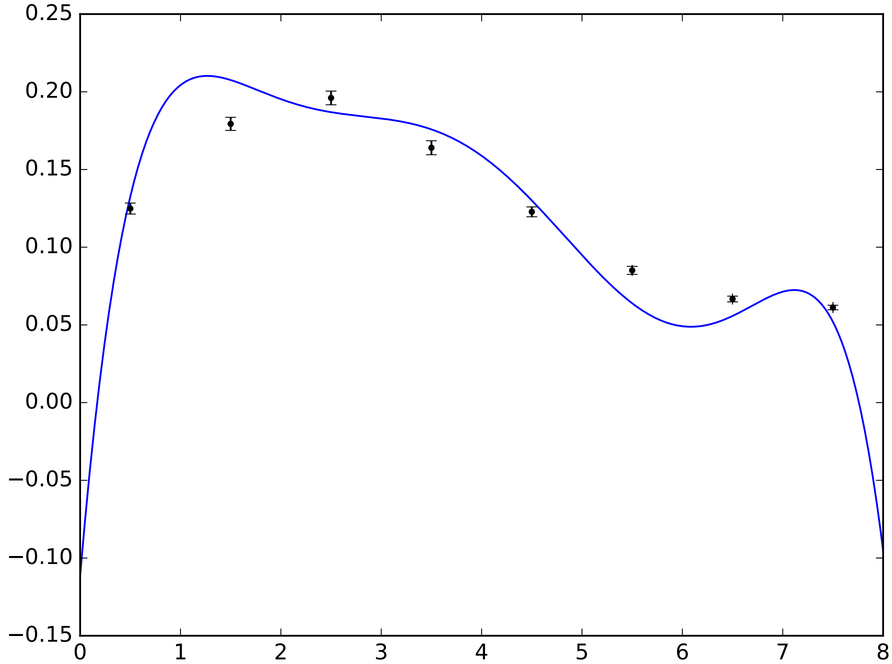


Figure 5: Aufgabe 3c - Fit.