AKTUELLE ANALYSIS-ANWENDUNGEN

OLIVER C. SCHNÜRER

Zusammenfassung. Durch die aktuelle Coronakrise motiviert, erkläre ich einige dafür relevante Themen aus der Analysis. Dies dient auch als Motivation für weitere Vorlesungen.

1. Methode der kleinsten Quadrate

Die Methode der kleinsten Quadrate wurde um 1800 von C. F. Gauß entwickelt, um den Zwergplaneten Ceres nach einigen Beobachtungen wiederzufinden [1]. Das Ziel war, die Parameter der elliptischen Bahn möglichst genau zu bestimmen, um damit die Position von Ceres vorherzusagen.

Wir illustrieren dies nun am Beispiel der Zahl der Infizierten einer Krankheit. Haben wir eine Tabelle mit Zeiten $(t_i)_{1\le i\le n}$ und zugehörigen Infiziertenzahlen $(x_i)_i$, so erwarten wir zunächst einen exponentiellen Zusammenhang $x(t)=\alpha\cdot\exp(\beta t)$ oder $x(t)=A\cdot 2^{Bt}$ und $x_i\approx x(t_i)$. Vielleicht ist aus der Physik das folgende Vorgehen bei exponentiellem Wachstum bekannt: Man betrachtet nicht die Punkte $(t_i,x_i)_i$, die in etwa auf einer exponentiellen Kurve liegen, sondern $(t_i,\log x_i)_i$, die in etwa auf einer Geraden liegen. Im Falle des obigen exponentiellen Zusammenhanges ergibt sich nämlich $\log x(t)=\log\alpha+\beta t$, auf der rechten Seite steht also eine Geradengleichung. Häufig sieht man nun schon recht gut am Schaubild, welche Gerade die Messpunkte gut approximiert. Wir wollen daher die Daten $(t_i,\log x_i)_i\equiv(t_i,y_i)_i$, also $y_i=\log x_i$ betrachen.

Für die Gerade machen wir den Ansatz $a+bt$. Im Idealfall können wir $a,b\in\mathbb{R}$ so wählen, dass $a+bt_i=y_i$ für alle i gilt. In der Realität klappt dies jedoch häufig nicht. Bei der Methode der kleinsten Quadrate möchte man nun die Parameter a und b so wählen, dass der Vektor $(|a+bt_i-y_i|)_i\in\mathbb{R}^n$ der Abweichungen in einer geeigneten Norm möglichst klein wird. Möglich sind beispielsweise die l^p-Normen für $1\le p\le\infty$. Bei der Methode der kleinsten Quadrate möchte man a und b so wählen, dass

$$\sum_{i=1}^n |a+bt_i-y_i|^2$$

möglichst klein wird. Wir betrachten also das Quadrat der euklidischen oder l^2-Norm.

Theorem 1.1. *Sei $n\ge 2$ und seien Zahlenpaare $(t_i,y_i)_{1\le i\le n}\subset\mathbb{R}^2$ gegeben, wobei nicht alle t_i übereinstimmen.*

Dann nimmt die Funktion

$$\begin{aligned}\varphi\colon \mathbb{R}^2 &\to \mathbb{R},\\ (a,b)&\mapsto \sum_{i=1}^n |a+bt_i-y_i|^2=\sum_{i=1}^n(a+bt_i-y_i)^2\end{aligned}$$

Date: 30. März 2020.
Vielen Dank an einen kritischen Leser für zahlreiche Korrekturen.

ihr Minimum in genau einem Punkt an. Dieser Punkt ist die eindeutige Lösung des linearen Gleichungssystems (1.1)*.*

Beweis. **Existenz:** Wir zeigen zunächst, dass die Funktion φ ihr Minimum annimmt.

(i) Ohne Einschränkung dürfen wir annehmen, dass $t_1 \neq t_2$ gilt.

(ii) Mit Hilfe der Substitution $\tilde{t} = t - t_1$ in der Zeit dürfen wir ohne Einschränkung annehmen, dass $t_1 = 0$ gilt. Dabei wird t_i durch $t_i - t_1$ ersetzt und die Parameter der Gerade ändern sich gemäß der Formel
$$a + bt = a + bt_1 + b(t - t_1) \equiv \tilde{a} + b\tilde{t}.$$

(iii) **Behauptung:** Für alle $M \ge 0$ gibt es ein $R > 0$ mit $\varphi(a,b) \ge M$ für alle $(a,b) \in \mathbb{R}^2 \setminus B_R(0)$.

(iv) Wir zeigen nun diese Behauptung. Sei also $(a,b) \in \mathbb{R}^2 \setminus B_R(0)$ für ein noch zu bestimmendes $R > 0$. Dann gilt $a^2 + b^2 \ge R^2$ und daher $a^2 \ge \frac{R^2}{2}$ oder $b^2 \ge \frac{R^2}{2}$. Es folgt $|a| \ge \frac{R}{\sqrt{2}}$ oder $|b| \ge \frac{R}{\sqrt{2}}$. Sei $C \ge 0$ so fixiert, dass $|y_1| \le C$ und $|y_2| \le C$ gelten.

(v) Ist $|a| \ge \frac{R}{\sqrt{2}}$, so folgt
$$|a + t_1 b - y_1| = |a + 0 \cdot b - y_1| \ge |a| - |y_1| \ge \frac{R}{\sqrt{2}} - C.$$
Für $R \ge \sqrt{2}(\sqrt{M} + C)$ erhalten wir $|a + t_1 b - y_1| \ge \sqrt{M}$ und daraus
$$\varphi(a,b) = \sum_{i=1}^{n} |a + bt_i - y_i|^2 \ge |a + t_1 b - y_1|^2 \ge M$$
wie gewünscht.

(vi) Der letzte Schritt zeigt, dass die Behauptung im Fall $|a + t_1 b - y_1| \ge \sqrt{M}$ trivial ist. Für das weitere Argument dürfen wir also ohne Einschränkung annehmen, dass $|a + t_1 b - y_1| < \sqrt{M}$ gilt.

(vii) Sonst ist $|b| \ge \frac{R}{\sqrt{2}}$ und wir erhalten
$$\begin{aligned}|a + t_2 b - y_2| &= |a + t_1 b - y_1 + (t_2 - t_1)b + (y_1 - y_2)|\\ &\ge |t_2 - t_1| \cdot |b| - |a + t_1 b - y_1| - |y_1 - y_2|\\ &\ge |t_2 - t_1|\frac{R}{\sqrt{2}} - \sqrt{M} - |y_1| - |y_2|\\ &\ge |t_2 - t_1|\frac{R}{\sqrt{2}} - \sqrt{M} - 2C.\end{aligned}$$
Für $R \ge \frac{\sqrt{2}}{|t_2 - t_1|}(2\sqrt{M} + 2C)$ folgt daher $|a + t_2 b - y_2| \ge \sqrt{M}$ und somit analog zu oben ebenfalls $\varphi(a,b) \ge |a + t_2 b - y_2|^2 \ge M$.

(viii) Für alle
$$R > \max\left\{\sqrt{2}(\sqrt{M} + C), \frac{\sqrt{2}}{|t_2 - t_1|}(2\sqrt{M} + 2C)\right\}$$
gilt also die gewünschte Abschätzung, was die Behauptung zeigt.

(ix) Setzen wir nun insbesondere $M := \varphi(0,0) + 1$, so gibt es nach der Behauptung ein $R > 0$, so dass für alle $(a,b) \in \mathbb{R}^2 \setminus B_R(0)$ die Ungleichung
$$\varphi(a,b) \ge M > \varphi(0,0) \ge \inf_{(a,b) \in \overline{B_R(0)}} \varphi(a,b)$$

gilt. Die Funktion φ kann ihr Infimum daher höchstens in Punkten $(a,b) \in \overline{B_R(0)}$ annehmen. Da die Funktion φ stetig ist, nimmt sie auf dem Kompaktum $\overline{B_R(0)}$ ihr Minimum an. Folglich nimmt die Funktion $\varphi\colon \mathbb{R}^2 \to \mathbb{R}$ ihr Minimum in einem Punkt von $\overline{B_R(0)}$ an.

Eindeutigkeit: Nun zeigen wir, dass die Funktion φ ihr Minimum in genau einem Punkt annimmt.

(x) Sei $(a,b) \in \mathbb{R}^2$ so gewählt, dass die Funktion φ dort ihr Minimum annimmt. Dann verschwindet dort die Ableitung von φ – man sagt dann, dass (a,b) ein *kritischer Punkt* von φ ist – und somit verschwinden auch ihre *partiellen Ableitungen* $\frac{\partial}{\partial a}\varphi(a,b)$ und $\frac{\partial}{\partial b}\varphi(a,b)$. Partielle Ableitungen sind Ableitungen, bei denen alle anderen Variablen als Konstanten betrachtet werden. Wir erhalten also

$$\begin{aligned} 0 &= \frac{1}{2}\frac{\partial}{\partial a}\varphi(a,b) \\ &= \sum_{i=1}^{n}(a + bt_i - y_i)\cdot 1 \\ &= na + b\sum_{i=1}^{n} t_i \; - \sum_{i=1}^{n} y_i \end{aligned}$$

und

$$\begin{aligned} 0 &= \frac{1}{2}\frac{\partial}{\partial b}\varphi(a,b) \\ &= \sum_{i=1}^{n}(a + bt_i - y_i)\cdot t_i \\ &= a\sum_{i=1}^{n} t_i + b\sum_{i=1}^{n} t_i^2 - \sum_{i=1}^{n} y_i t_i. \end{aligned}$$

Dies schreiben wir als Matrizengleichung äquivalent um:

$$\underbrace{\begin{pmatrix} n & \sum_{i=1}^{n} t_i \\ \sum_{i=1}^{n} t_i & \sum_{i=1}^{n} t_i^2 \end{pmatrix}}_{=:A} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^{n} y_i \\ \sum_{i=1}^{n} y_i t_i \end{pmatrix}. \tag{1.1}$$

Um zu zeigen, dass die Matrix A invertierbar ist, betrachten wir die Determinante

$$\det A = n\sum_{i=1}^{n} t_i^2 - \left(\sum_{i=1}^{n} t_i\right)^2.$$

Aus

$$0 < \frac{1}{2}(t_1 - t_2)^2 = \frac{1}{2}t_1^2 + \frac{1}{2}t_2^2 - t_1 t_2$$

folgt die strikte Ungleichung in

$$\begin{aligned} \left(\sum_{i=1}^{n} t_i\right)^2 &= \sum_{i=1}^{n}\sum_{j=1}^{n} t_i t_j \\ &< \sum_{i,j=1}^{n}\left(\frac{1}{2}t_i^2 + \frac{1}{2}t_j^2\right) \end{aligned}$$

$$=\frac{n}{2}\sum_{i=1}^n t_i^2+\frac{n}{2}\sum_{j=1}^n t_j^2=n\sum_{i=1}^n t_i^2.$$

Somit ist $\det A>0$, also insbesondere $\det A\neq 0$. Damit ist A eine invertierbare Matrix und unser lineares Gleichungssystem (1.1) ist eindeutig lösbar.

(xi) Daraus folgt die Eindeutigkeit: Jedes Minimum von φ ist ein kritischer Punkt von φ und damit eine Lösung des eindeutig lösbaren Gleichungssystems (1.1). □

Abstrakter kann man Theorem 1.1 wie folgt formulieren. Man bekommt daraus Theorem 1.1, indem man $I=\{1,\ldots,n\}$ und $(\tau(i),m(i))=(t_i,y_i)$ betrachtet.

Korollar 1.2. *Für $(a,b)\in\mathbb{R}^2$ betrachte die Funktion $g_{(a,b)}\colon\mathbb{R}\to\mathbb{R}$, $t\mapsto a+bt$, deren Graph eine affine Gerade ist. Sei*

$$\mathcal{G}:=\{g_{(a,b)}\mid(a,b)\in\mathbb{R}^2\}$$

die Menge aller solcher Geradenfunktionen. Sei I eine endliche Menge und seien $\tau\colon I\to\mathbb{R}$ und $m\colon I\to\mathbb{R}$ zwei Funktionen. Es bestehe $\tau(I)$ aus mindestens zwei Elementen. Dann besagt Theorem 1.1: Es gibt genau ein $g_0\in\mathcal{G}$ mit

$$\|g_0\circ\tau-m\|=\min_{g\in\mathcal{G}}\|g\circ\tau-m\|.$$

Hierbei ist

$$\|f\|=\left(\sum_{i\in I}(f(i))^2\right)^{1/2}$$

die l^2-Norm auf dem Raum $\mathbb{R}^I$ aller reellwertigen Funktion $f\colon I\to\mathbb{R}$.

Bemerkung 1.3.

(i) Im Beweis von Theorem 1.1 haben wir partielle Ableitungen verwendet. Diese sind Thema in Analysis II.

(ii) Das Gleichungssystem (1.1) ist leicht explizit lösbar. Dazu benutzen wir die Formel $\frac{1}{ad-bc}\left(\begin{smallmatrix}d&-b\\-c&a\end{smallmatrix}\right)$ für die Inverse einer invertierbaren 2×2-Matrix $\left(\begin{smallmatrix}a&b\\c&d\end{smallmatrix}\right)$. Damit folgt

$$\begin{pmatrix}a\\b\end{pmatrix}=\frac{1}{n\sum\limits_{i=1}^n t_i^2-\left(\sum\limits_{i=1}^n t_i\right)^2}\begin{pmatrix}\sum\limits_{i=1}^n t_i^2 & -\sum\limits_{i=1}^n t_i\\ -\sum\limits_{i=1}^n t_i & n\end{pmatrix}\begin{pmatrix}\sum\limits_{i=1}^n y_i\\ \sum\limits_{i=1}^n y_it_i\end{pmatrix}.$$

Somit sehen wir, dass der Beweis auch eine Methode liefert, a und b explizit zu bestimmen.

(iii) Wir geben in Kapitel 4 ein `sage`-Programm an, das diese Rechnungen durchführt. Das Software-Paket `sage` ist unter `www.sagemath.org` frei verfügbar.

(iv) Alternativ hätte man $\det A>0$ auch mit Hilfe der Cauchy-Schwarzschen Ungleichung zeigen können. Für

$$v=\begin{pmatrix}1\\1\\\vdots\\1\end{pmatrix}\quad\text{und}\quad w=\begin{pmatrix}t_1\\t_2\\\vdots\\t_n\end{pmatrix}$$

gilt nämlich

$$\begin{aligned}\left(\sum_{i=1}^{n} t_i\right)^2 &= \langle v, w\rangle^2 \\ &\leq |v|^2 \cdot |w|^2 \\ &= \left(\sum_{i=1}^{n} 1\right)^{2/2} \cdot \left(\sum_{i=1}^{n} t_i^2\right)^{2/2} \\ &= n \cdot \sum_{i=1}^{n} t_i^2.\end{aligned}$$

Hierbei gilt Gleichheit genau dann, wenn v und w linear abhängig sind. Daher erhalten wir im Falle, dass nicht alle t_i's gleich sind, die strikte Ungleichung und somit $\det A > 0$.

(v) Die Eindeutigkeit folgt leicht aus der strikten Konvexität von φ. Eine direkte Rechnung liefert nämlich, dass die Hessesche

$$\begin{pmatrix} \frac{\partial^2}{\partial a \partial a}\varphi & \frac{\partial^2}{\partial a \partial b}\varphi \\ \frac{\partial^2}{\partial b \partial a}\varphi & \frac{\partial^2}{\partial b \partial b}\varphi \end{pmatrix}$$

von φ durch $2A$ mit A wie in (1.1) gegeben ist. Diese ist auf Grund des Minorantenkriteriums (wegen $n > 0$ und $\det A > 0$) positiv definit, was die strikte Konvexität liefert.

Auch ohne Differentialrechnung sieht man, dass φ konvex ist, dass also

$$\varphi(\tau a + (1-\tau)\alpha, \tau b + (1-\tau)\beta) \leq \tau\varphi(a, b) + (1-\tau)\varphi(\alpha, \beta)$$

für alle $\tau \in [0, 1]$ gilt.

(vi) Sprachlich genauer wäre es, statt von der *Methode der kleinsten Quadrate* von der *Methode der kleinsten Quadratsumme* zu sprechen.

(vii) Varianten dieser Methode benutzen andere Ansätze als $a + tb$, eine andere Norm, die minimiert werden soll, oder Gewichte in der Summe.

2. Wachstumsgleichungen

In dieser Anwendung benutzen wir gewöhnliche Differentialgleichungen, die wir in Analysis III behandeln werden.

Ein Problem bei dem obigen Ansatz ist, dass wir unsere Daten durch eine Funktion $x(t) = a\exp(bt)$ approximieren, die für $a, b > 0$ und $t \to \infty$ unendlich wird. Spätestens für $x(t) \geq 10^{10}$ prognostiziert dieses Modell also mehr Infizierte als die aktuelle Weltbevölkerung.

Auf solche exponentiellen Lösungen kommt man durch ein Modell

$$\dot{x}(t) = bx(t),$$

was $x(t) = a\exp(bt)$ als allgemeine Lösung hat. Es besagt, dass die zeitliche Änderung proportional zur aktuellen Zahl der Infizierten ist. Ein allererstes Modell, was den zeitlichen Zuwachs begrenzt, sobald man sich einer gewissen Schwelle M annähert, ist

$$\dot{x}(t) = b \cdot x(t) \cdot (M - x(t)).$$

Dabei ist der Zuwachs sowohl proportional zur Anzahl der Infizierten als auch zur Anzahl der noch nicht Infizierten.

Erstaunlicherweise kann man diese Gleichung sogar explizit lösen. Das Verfahren nennt sich Trennung der Variablen und führt in den folgenden Schritten zu einer

Lösung. Auch wenn wir die Notation zwischendurch nicht mathematisch rechtfertigen, sieht man am Ende mit direktem Nachrechnen doch, dass eine Lösung herauskommt.

$$\begin{aligned}
\frac{dx}{x(M-x)} &= b\,dt,\\
\left(\frac{1}{x}+\frac{1}{M-x}\right)dx &= Mb\,dt,\\
\log x - \log(M-x) &= Mb(t-t_0),\\
\frac{x}{M-x} &= e^{Mb(t-t_0)},\\
x(t)\cdot\left(1+e^{Mb(t-t_0)}\right) &= Me^{Mb(t-t_0)},\\
x(t) &= \frac{Me^{Mb(t-t_0)}}{1+e^{Mb(t-t_0)}}
\end{aligned}$$

für ein $t_0 \in \mathbb{R}$.

Wir fassen nun $x(t)$ auch als Funktion von t_0 auf und schreiben dafür in diesem Absatz $x(t, t_0)$. Es gilt $x(0, t_0) \to 0$ für $t_0 \to +\infty$ und $x(0, t_0) \to M$ für $t_0 \to -\infty$. Somit gibt es aufgrund der Stetigkeit und Monotonie, die man leicht aus $\frac{\partial x(t,t_0)}{\partial t_0} = -\frac{\partial x(t,t_0)}{\partial t}$ und der Differentialgleichung bekommt, für jedes $x_0 \in (0, M)$ genau ein $t_0 \in \mathbb{R}$, so dass $x(0, t_0) = x_0$ gilt.

In Analysis III werden wir lernen, dass diese gewöhnliche Differentialgleichung genau eine Lösung besitzt und wie man die Existenz einer Lösung für allgemeinere gewöhnliche Differentialgleichungen der Form $\dot{x}(t) = f(t, x(t))$ zeigen kann.

3. Weiterführende Gedanken dazu

- Möchte man noch die räumliche Verteilung berücksichtigen, so erhält man Gleichungen wie
$$\dot{u}(x,t) = \Delta u(x,t) + f(u(x,t)).$$
Dies ist eine partielle Differentialgleichung. Den Fall $f \equiv 0$ werden wir in der Vorlesung Theorie partieller Differentialgleichungen betrachten.
- Die Richtigkeit des Theorems konnten wir beweisen. Um es anwenden zu können, muss man jedoch die Voraussetzungen kritisch hinterfragen:
  - Das Modell liefert eine reelle Zahl von Infizierten. Diese Zahl wird im ersten Ansatz sogar unbeschränkt.
  - Ist die Annahme, dass die Entwickung sich für die logarithmischen Daten linear verhält, gerechtfertigt?
  - Wie gut sind die Daten? Gibt es systematische Fehler, dass beispielsweise nur schwere Fälle erfasst werden oder die Daten zu anderen als den angegebenen Zeitpunkten gehören, weil Tests lange dauern?
- Beschreibt die gewöhnliche oder partielle Differentialgleichung das Geschehen gut genug?
- Wir haben oben eine sogenannte Heterokline gefunden; dies ist eine Lösung $x(t)$, die für $t \to \pm\infty$ gegen stationäre Lösungen konvergiert. Hier sind $x(t) \equiv 0$ und $x(t) \equiv M$ die stationären Lösungen. Das Verhalten von Lösungen von gewöhnlichen Differentialgleichungen werden wir genauer in Analysis III oder in der Vorlesung Gewöhnliche Differentialgleichungen mit geometrischen Anwendungen betrachten. Auch die Vorlesung Dynamische Systeme beschäftigt sich damit.
- Minimierungsaufgaben tauchen in der Veranstaltung Variationsrechnung häufig auf, ebenso in Optimierung. Minimalflächen sind Minimierer des

Oberflächenfunktionals und ein wichtiges Beispiel in der geometrischen Analysis.

4. Das sage-Programm

```
# 2020, Olaf Schnuerer und Oliver Schnuerer
# Daten aus Wikipedia
# https://de.wikipedia.org/wiki/COVID-19-Pandemie_in_Deutschland
# 01. April 2020 ist Zeitpunkt t=1
var('t')
tabelle=[
# [-36,16],
# [-35,18],
# [-34,21],
# [-33,26],
# [-32,53],
# [-31,66],
# [-30,117],
# [-29,150],
# [-28,188],
# [-27,240],
# [-26,400],
# [-25,639],
# [-24,795],
# [-23,902],
[-22,1139],
[-21,1296],
[-20,1567],
[-19,2369],
[-18,3062],
[-17,3795],
[-16,4838],
[-15,6012],
[-14,7156],
[-13,8198],
[-12,10999],
[-11,13957],
[-10,16662],
[-9,18610],
[-8,22672],
[-7,27436],
[-6,31554],
[-5,36508],
[-4,42288],
[-3,48582],
[-2,52547],
[-1,57298]]

prognoseende=3
maxY=120000

from matplotlib import ticker
```

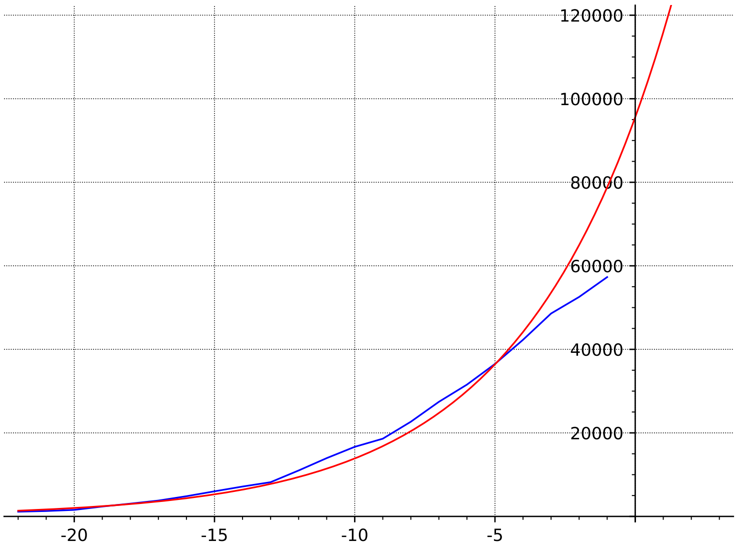
```
start=tabelle[0][0]
ende=tabelle[len(tabelle)-1][0]
n=ende-start+1

logtabelle=list([tabelle[i][0], log(1.0*tabelle[i][1])] \
  for i in range(len(tabelle)))

gesamt=list_plot(tabelle, plotjoined=True, \
  ymax=maxY, scale='semilogy')
gesamt2=list_plot(tabelle, plotjoined=True, \
  ymax=maxY, tick_formatter=(None, ticker.FormatStrFormatter('%d')))

M=Matrix([[n,sum(logtabelle[i][0] \
for i in range(len(logtabelle)))],\
[sum(logtabelle[i][0] for i in range(len(logtabelle))),\
sum(logtabelle[i][0]^2 for i in range(len(logtabelle)))]])

v=vector([sum(logtabelle[i][1] for i in\
range(len(logtabelle))),\
sum(logtabelle[i][0]*logtabelle[i][1] for i in\
range(len(logtabelle)))])

w=M.solve_right(v)

gesamt=gesamt+plot(exp(w[0]+t*w[1]),\
(t,start,prognoseende),scale='semilogy', \
rgbcolor=(1,0,0), gridlines='minor')

gesamt2=gesamt2+plot(exp(w[0]+t*w[1]),\
(t,start,prognoseende), rgbcolor=(1,0,0),\
gridlines='major')

gesamt.save('corona-germany-logarithmisch.pdf')
gesamt2.save('corona-germany.pdf')
```

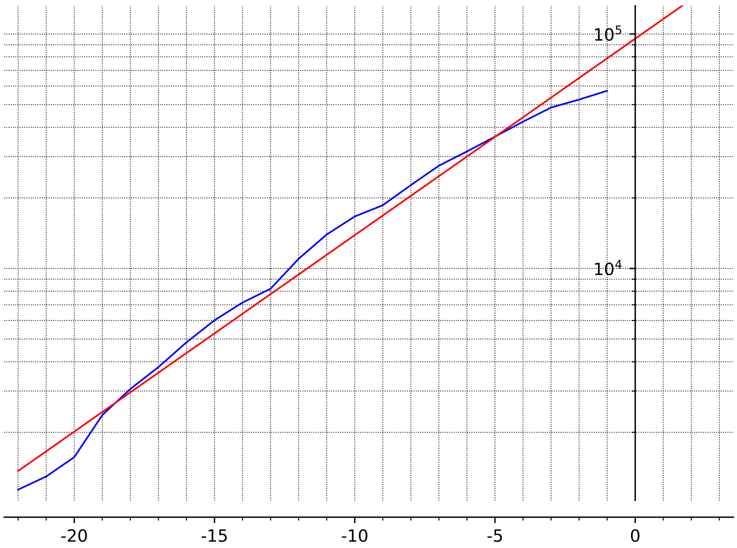
Es liefert die folgenden Diagramme:

Literatur

1. Wikipedia, http://www.wikipedia.org.

Oliver C. Schnürer, Universität Konstanz, Germany
Email address: Oliver.Schnuerer@uni-konstanz.de


120000
100000
80000
60000
40000
20000
-20
-15
-10
-5


10^5
10^4
-20
-15
-10
-5
0