Dr. R. Tumulka, M. Fromm
Mathematisches Institut, Universität Tübingen

Wintersemester 2005/06
2.2.2006

Mathematik I für Biologen, Geologen und Geoökologen

Computer-Übungsblatt 13

Aufgabe C26. Lineare Regression am Beispiel von Grillen

Durch ihren Flügelschlag produzieren Grillen das charakteristische Zirpen. Je schneller eine Grille ihre Flügel aufeinander schlägt, desto höher ist das Zirpen, dabei bewegen Grillen ihre Flügel bei höheren Umgebungstemperaturen schneller als bei niedrigen. Finden Sie den (hier als linear angenommen) Zusammenhang zwischen Zirpfrequenz f_Z und der Lufttemperatur T anhand der Meßwerte im Datensatz `grille.dat` (Dieser kann von der Webseite zur Vorlesung heruntergeladen werden.). Gehen Sie dabei wie folgt vor: Laden Sie den Datensatz `grille.dat` in *octave* (siehe Übungsblatt **C3**); dieser enthält in der ersten Spalte die Zirpfrequenz (1/s), in der zweiten Spalte die gemessene Temperatur T (in Fahrenheit). Stellen Sie die Daten in einem Scatterplot-Diagramm dar. Lineare Regression läßt sich in *octave* durch den Befehl `p = polyfit(x,y,1)` durchführen; dabei sind `x,y` Vektoren mit den Stützpunkten (bzw. Meßwerten), und `p = (p1,p2)` enthält die zu berechnenden Parameter p_1, p_2 der Gleichung $y = p_1 x + p_2$. Finden Sie p_1, p_2 und plotten sie die *Ausgleichsgerade* in dasselbe Diagramm. (5 Punkte)

Aufgabe C27.

a) Die lineare Regression betrachtet die x-Werte als exakt und nähert die y-Werte an. Sind die x-Werte selbst nicht verlässlich, so ist es ratsam, die beiden Variablen mal die Rollen tauschen zu lassen. Vertauschen Sie in Aufgabe **C26** die beiden Spalten, führen Sie lineare Regression durch, vertauschen nochmals die Variablen (um wieder f_Z nach rechts und T nach oben abzutragen), und plotten Sie beide Ausgleichsgeraden in ein gemeinsames Diagramm.

b) Gehen Sie nun davon aus, dass die berechneten Parameter p_1 und p_2 jeweils mit einem Fehler von 10% behaftet sind, d.h., dass für die Parameter q_1, q_2 der wahren Gerade gilt $q_i \in [p_i - \Delta p_i, p_i + \Delta p_i]$ mit $\Delta p_i = p_i/10$, $i = 1, 2$. Stellen Sie wie im Plot von **C26** weitere 100 Geraden dar mit anderen Parametern im Rahmen der Ungenauigkeit. Nutzen Sie dazu folgenden Programm-Code:

```
p1 = [    ]; errechneter p1-Wert
p2 = [    ]; errechneter p2-Wert
dp1 = p1/10; dp2 = p2/10;
r = 0:.1:1; phi = 0:2*pi/10:2*pi;
[R,Phi] = meshgrid(r,phi);
q1 = p1 + dp1*R.*cos(Phi);
q2 = p2 + dp2*R.*sin(Phi);
x= -3:0.1:3;
for i =1:1:length(r)
for j = 1:1:length(phi)
y((i-1)*length(phi)+j,:) = q1(i,j)*x+q2(i,j);
end
end
for i=1:1:100
plot(x,y(i,:),'r;;'); hold on
end
```

Die Felder $\boxed{}$ müssen Sie noch ausfüllen.

c) Teil b) hat Ihnen einen Eindruck gegeben, wie sich Ungenauigkeiten in den Parametern auswirken: die Geraden innerhalb der Ungenauigkeit füllen einen gewissen Streifen aus. So, wie man bei Punkten in einem Scatter-Plot Fehlerbalken einzeichnet, zeichnet man bei Regressionsgeraden zur Veranschaulichung der Ungenauigkeit gern die Ränder dieses Streifens ein. Plotten Sie dazu die *Einhüllenden* $f_\pm$ der in Teil b) gezeichneten Geraden, die den Gleichungen

$$f_\pm(x) = p_1 x + p_2 \pm \sqrt{\Delta p_1^2 x^2 + \Delta p_2^2}$$

genügen, andersfarbig in dasselbe Diagramm. (5 Punkte)

Abgabe: Donnerstag, 9.2.2006, zu Beginn der Vorlesung. Bitte geben Sie immer auch einen Ausdruck des von Ihnen verfassten Programmcodes ab.