

Dokumentation Programm korrelations_koeffizient

Armin Ulrich

26. November 2017

1 Sinn des Programms

Um Aussagen aus Datenreihen zu bekommen, stellt man die Daten meistens in Diagrammform dar. Diese werden dann oft kommentiert mit Sätzen wie: 'Ein Blick auf das Diagramm sagt uns, daß die beiden Größen voneinander abhängig sind.' Manchmal trügt er Augenschein, daher wäre es gut, wenn man ein statistisches Werkzeug hätte, das uns mit einer Zahl aussagt, ob zwei Größen zusammenhängen können oder nicht.

2 Umgang mit dem Programm

Wie schon beim Testprogramm müssen alle Dateien in das Verzeichnis geladen und kompiliert werden, zuzätzlich benötigt man das Programm Gnuplot zur graphischen Darstellung der errechneten Werte, die Schritte werden hier wieder angegeben:

2.1 Anwendung des Programms

1. Zuerst entscheiden Sie sich für eine der Programmiersprachen fortran 95 oder C.
2. Dann erstellen Sie ein Verzeichnis, in dem Sie das Programm 'Korrelationskoeffizient' laufen lassen wollen.
3. In diesem Verzeichnis benötigen Sie die alle Quelldateien, die Datei 'steuer.txt' und den 'makefile'. Es bietet sich an, das Archiv herunterzuladen und im Verzeichnis zu öffnen. Das Archiv entht den makefile, die jeweilige Quelldateien, die Steuerdatei und die Dokumentation als *.pdf. Natrlich knnen Sie die Dateien entweder einzeln anklicken und speichern, was aber mhselig ist. Der Makefile beinhaltet die Befehle zum Erstellen der ausführbaren Datei und ist spezifisch für die jeweilige Programmiersprache. Eine C-Quelldatei kann nicht mit dem fortran 95-Makefile kompiliert werden und umgekehrt. Im C-Verzeichnis liegt der richtige Makefile für den C-Quellcode bereit, ebenso im fortran 95-Verzeichnis.
4. Dann öffnen Sie ein Konsolenfenster in diesem Verzeichnis (rechter Mausklick im Fenster des Verzeichnisses, dann **Terminal hier öffnen** klicken).

5. Führen Sie einen Virensan über dieses Verzeichnis durch. Z.B. ich benutze clamscan mit dem Befehl **clamscan -r -i**. Sie können natrlich einen anderen Virens scanner benutzen.
6. Jetzt werden die Quelldateien mit dem Befehl **make** compiliert, es wird eine ausführbare Datei erzeugt.
7. Führen Sie einen erneuten Virensan über dieses Verzeichnis durch.
8. Die Eingabedaten befinden sich in der Datei 'steuer.txt', sie müssen nicht mit der Tastatur eingegeben werden.
9. Die berechneten Werte werden in der Konsole ausgegeben, gleichzeitig wird eine Ausgabedatei erstellt.

2.2 Anwenden des Graphikprogramms Gnuplot

Das Darstellungsprogramm Gnuplot läßt sich für das Betriebssystem UBUNTU wieder sehr einfach in der Konsole installieren.

Meistens verwendet man Gnuplot um zwei oder mehrere Spalten einer Datei graphisch zuzuordnen. Hier wären das die erste Spalte mit dem Zählindex, der jeweils zwei Spalten mit den beiden Zufallsreihen bzw. den beiden Sinusreihen zugeordnet wird. Falls Ihnen ein anderes Darstellungsprogramm vertraut und 'lieber' ist, können Sie natürlich auch dieses verwenden.

1. Wir wollen zuerst zwei Diagramme erzeugen, eines für die Zufallszahlen, eines für die Sinusfunktion, mit den Zufallszahlen sei angefangen.
2. Gnuplot wird in der Konsole mit dem Befehl **gnuplot** gestartet.
3. Bevor wir zum Plotten der Ausgabewerte kommen, müssen wir das Diagramm vorbereiten. Wir brauchen einen Diagrammtitel, den setzen wir mit **set title 'Zufallszahlen'**. Beachten Sie die Hochkommata.
4. Die x-Achse beschriften wir mit **set xlabel 'Zaehindex i'**, die y-Achse mit **set ylabel 'Zufallszahlen'**. Hier sind die Hochkommata wieder wichtig.
5. Der Befehl um die Zufallszahlen zu plotten ist: **plot "feld_ausgabe.dat" using 1:2, "feld_ausgabe.dat" using 1:3** . Die Zahlen sind Spaltenindizes.
6. Die erstellte Graphik läßt sich mit Gnuplot speichern, der Befehl dazu ist: **save "feld_ausgabe_zufall.gnu"**.
7. Mit **load "feld_ausgabe_zufall.gnu"** kann man diese Graphik wieder aufrufen und anzeigen.
8. Eine Änderung des Bildformates läßt sich auch bewerkstelligen, z.B. werden .jpg-Bilder erzeugt mit den vier Befehlen:
 - (a) **set term jpeg**
 - (b) **set output "feld_ausgabe_zufall.jpg"**
 - (c) **replot**
 - (d) **set term x11**

3 Physikalischer Hintergrund

Entfällt.

4 Mathematischer Hintergrund

Die Funktionen `random_seed()` , `random_number()` in fortran 95 bzw. `rand()` in C erzeugen 2 eindimensionale Felder, also Vektoren von Pseudozufallszahlen, die für jeden Programmdurchlauf gleich sind. Weiterhin erzeugt das Programm zwei Sinusfunktionen, die mit über die Steuerdatei eingelesenen Parametern phasenverschoben werden kann. Der Korrelationskoeffizient ist nach Martin Hengst, Einführung in die mathematische Statistik und ihre Anwendung, S. 236 oben, definiert mit:

$$r = r_{x,y} = \frac{\sum_{i=1}^n (x_i - \bar{x}) \cdot (y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \cdot \sum_{i=1}^n (y_i - \bar{y})^2}} \quad (1)$$

Man erwartet von den beiden Zufallszahlenreihen einen Korrelationskoeffizienten, der auf keinen Zusammenhang hindeutet, also

$$r_{x,y} \approx 0 \quad (2)$$

Der Korrelationskoeffizient zweier identischer (aufeinanderliegender) Sinusfunktionen soll natürlich von dem der Zufallszahlen abweichen.

5 Ergebnis des Programms

6 Struktogramme

6.1 Hauptprogramm

Steuerdatei richtig geöffnet?	
ja	nein
LeseSteuerdatei	A(B): Fehler beim Öffnen der Steuerdatei, Existenz ueberpruefen!
A(B): Steuerdatei wurde korrekt gelesen.	
A(Log-Datei): Steuerdatei wurde korrekt gelesen.	A(Log-Datei): Fehler beim Öffnen der Steuerdatei, Existenz ueberpruefen!
Zufallszahlen	
Sinus	∅
Schirmausgabe	
Dateiausgabe	
Programmende	

6.2 LeseSteuerdatei

Öffnen der Steuerdatei	
Steuerdatei richtig geöffnet?	
ja	nein
E(Steuerdatei): m	A(B):'Beim Oef-fenen der Datei ist ein Fehler Nr.', io_error,'aufgetreten' lese_steuerdat = .false. ∅
A(B):'Feldgroesse m=',m	
E(Steuerdatei): phi	
A(B):'Phasenverschiebung: phi [°] =',phi	
E(Steuerdatei): sig	
A(B):'Vorzeichen vor dem Sinus: sig [-] =',sig	
E(Steuerdatei): feld_ausgabe_datei	
A(B):'feld_ausgabe_datei=', feld_ausgabe_datei	
E(Steuerdatei):korrelations_koeffizienten_datei	
A(B):'korrelations_koeffizienten_datei=', korrelations_koeffizienten_datei	
lese_steuerdat = .true.	

6.3 Zufallszahlen

Speicherreservierung Felder: zufall_1, zufall_2		
A(B): Reihe 1		
n=0		
	random_seed(n)	
	random_number(zufall_1)	
A(B): Reihe 2		
n=1		
	random_seed(n)	
	random_number(zufall_2)	
	korr_koeff(zufall_1, zufall_2)	
	A(B): 'Korrelationskoeffizient Zufall=', korr_koeff_zufall	