wdv-notes 176
Stand: 16.SEP.1993 (3.)
Wiss.Datenverarbeitung
FREIE UNIVERSITÄT BERLIN

Ermittlung Fraktaler Dimensionen mit *FDim* (P 511).

© 1992–1993 *Karl-Heinz Dittberner*

Pro

Ein Werkzeug zum Erkennen fraktaler Natur-Prozesse ist die Ermittlung der fraktalen Dimensionen von meßbaren Größen solcher Prozesse [2, 3]. Sind diese Dimensionen *nicht ganzzahlig*, dann ist das ein Indikator, daß das betrachtete System „chaotische“ Eigenschaften haben kann.

Für derartige Analysen steht seit dem 23. April 1993 auf dem HAIRcluster das C89-Programm ***FDim*** V0.3.2 unter der POSIX-Oberfläche vom OpenVMS zur Verfügung. Mit ***FDim*** lassen sich aus geeigneten, experimentell gewonnenen Daten die zugrundeliegenden fraktalen Dimensionen berechnen bzw. abschätzen.

Das Programm wurde von John Sarraille und Peter DiFalco von der California State University, Turlock (USA) als ***FD3*** entwickelt. Eine ausführliche Beschreibung der Algorithmen ist in den wdv-notes Nr. 148 zu finden [1].

1. Aufruf von *FDim*:

Das Programm ***FDim*** kann *nur* unter der POSIX-Oberfläche – dem genormten Unix [5] – verwendet werden. Diese Benutzerumgebung wird erreicht mit:

```
VAXUS>  POSIX  <Return>
```

Nun kann ***FDim*** gestartet werden. In der Befehlszeile wird der Dateiname der Daten sowie ggf. der zu dieser Datei führende Pfad angegeben.

```
VAXUS-psx> FDim  /pfad/datei  <Return>
```

Die Ausgabe des Dimensionsreports erfolgt auf dem Terminal. Wünscht man die Ausgabe der Ergebnisse als Datei, dann startet man ***FDim*** mit einer Umleitung der Ausgabe auf eine Datei (piping). Im folgenden Beispiel stehen die Eingabe-Daten in der Datei „eeg.dat“ im aktuellen Verzeichnis.

```
VAXUS-psx> FDim eeg.dat  > eeg.rep  <Return>
```

Die Ausgabe des Dimensionsreports erfolgt in der Datei „eeg.rep“ ebenfalls im aktuellen Verzeichnis, da kein Pfad angegeben wurde. Der Report kann dann später gedruckt werden.

2. Struktur der Eingabe-Datei:

Die Daten in der Eingabe-Datei müssen in folgender Form angeordnet sein, wobei die Koordinaten der Datenpunkte auf einer Zeile durch wenigstens einen Zwischenraum oder Tabulator getrennt werden müssen.

Zeile 1:	Anzahl *N* der Datenpunkte in dieser Datei.
Zeile 2:	Koordinaten des 1. Punktes.
Zeile 3:	Koordinaten des 2. Punktes.
...	
Zeile n:	Koordinaten des (*N*–1). Punktes.
Zeile n+1:	Koordinaten des *N*. Punktes.

Beispiel einer 2-dimensionalen Eingabe-Datei:

```
10
3.34            5.56
6664.9          4.561
111145          .776
234             34.55
2.2245          0
2556.3          324.55555
432             333.0001
334.987         689.9756
8765.567        546.99874
345.45          76.098
```

Beispiel einer 3-dimensionalen Eingabe-Datei:

```
10
2.47            3.34            5.56
-3.123          6664.9          4.561
0.334231        111145          0.776
497650          234             34.55
217             2.2245          0
1.23            2556.3          324.55555
10548           432             333.0001
2815            334.987         689.9756
219.85          8765.567        546.99874
0.0042987       345.45          76.098
```

Natürlich muß die Anzahl *N* der Datenpunkte größer sein als in diesen kurzen Beispielen, um eine ausreichende Genauigkeit der berechneten fraktalen Dimensionen zu erzielen.

Definitionen

Gegeben sei ein fraktales Objekt (fractal measure) M im Euklidischen Raum $\mathbf{R}^E$ mit E Dimensionen [3]. M kann ein Menge, eine Attraktor oder nach [4] irgend ein anderer „fraktaler Prozeß“ sein. Auf dieses fraktale Objekt M bezogen seien $(x_N) = \{x_k\}_{k=1}^{N}$ die Datenpunkte einer langen Zeitreihe (sample), wobei die Anzahl der Datenpunkte N sehr groß, aber notwendigerweise endlich sei.

Bedeckt man den Raum $\mathbf{R}^E$ mit einem dichten Gitter aus E-dimensionalen Würfeln der Kantenlänge ε und dem Volumen ε^E, dann sei $N(\varepsilon)$ die Anzahl von Würfeln, die Punkte der Zeitreihe (x_N) enthalten. $N(\varepsilon, i)$ sei dann die Anzahl von Punkten im i-ten Würfel und $P(\varepsilon, i) = N(\varepsilon, i)/N$ die Wahrscheinlichkeit, daß sich ein Datenpunkt im i-ten Würfel befindet. Der Abstand zwischen zwei Punkten sei $s(i, j) = \| x(i) - x(j) \|$.

Die drei *fraktalen Dimensionen* sind dann durch die folgenden Gleichungen definiert:

Kapazitäts-Dimension:

$$D = \lim_{\varepsilon \to 0} \frac{\log[N(\varepsilon)]}{\log[1/\varepsilon]} \qquad (1)$$

Informations-Dimension:

$$I(\varepsilon) = \sum_{i=1}^{N(\varepsilon)} -P(\varepsilon, i) \cdot \log[P(\varepsilon, i)] \qquad (2)$$

$$D_I = \lim_{\varepsilon \to 0} \frac{I(\varepsilon)}{\log[1/\varepsilon]} \qquad (3)$$

Korrelations-Dimension:

$$C(\varepsilon) = \frac{1}{N^2} \cdot \{\#\ \text{of pairs } (i, j) \text{ with } s(i, j) < \varepsilon\} \qquad (4)$$

$$D_C = \lim_{\varepsilon \to 0} \frac{\log[C(\varepsilon)]}{\log[\varepsilon]} \qquad (5)$$

Die *Kapazitäts-Dimension* (capacity dimension) entspricht der *fraktalen Dimension D* (fractal dimension) nach Mandelbrot [2] und in den meisten Fällen auch der *Ähnlichkeits-Dimension* (similarity dimension).

Beziehung zwischen den drei Dimensionen:

$$D_C \le D_I \le D \qquad (6)$$

Generelle Dimensionen:

Nach Hentschel und Procaccia [4] existiert eine ganze Hierarchie genereller Dimensionen D_q für beliebige $q \ge 0$. Im Fall, daß q ganzzahlig ist, haben die D_q eine physikalische Bedeutung. Für $q \to 0, 1, 2$ ergeben sich die fraktalen Dimensionen D, D_I und D_C :

$$\lim_{q \to 0} D_q = D \qquad (7)$$

$$\lim_{q \to 1} D_q = D_I \qquad (8)$$

$$D_{q=2} = D_C \qquad (9)$$

3. Voraussetzungen für *FDim*:

Sarraille und DiFalco haben festgestellt, daß der Fehler von berechneten fraktalen Dimensionen kleiner als 5 % ist, wenn ein 2-dimensionaler Datensatz einen Umfang von $N \geq 200$ *unterschiedlichen* Datenpunkten hat.

Allgemein werden für eine gute Abschätzung von fraktalen Dimensionen etwa $N = 2^{4D}$ unterschiedliche Datenpunkte benötigt. Dabei ist D die wahre fraktale Dimension der zugrundeliegenden Datenmenge.

D	*Mindest N*
1	200 Punkte
2	256 Punkte
3	4096 Punkte
4	65536 Punkte

Da dieses D unbekannt ist (das Programm ***FDim*** soll ja gerade eine Abschätzung von D liefern), muß das D zur Bestimmung des mindestens erforderlichen N geschätzt werden. Dabei kann berücksichtigt werden, daß immer $D \leq E$, der Anzahl der Koordinaten eines Datenpunkts ist.

Für $D \leq 1$ werden jedoch mehr Punkte notwendig sein, so daß empfohlen wird, ein Minimum von $N = 200$ Punkten zu verwenden.

4. Zur Interpretation der Ergebnisse:

Ein Berechnungsbeispiel ist auf Seite 4 der wdv-notes Nr. 148 [1] zu finden. Dort ist ein vollständiger „Fractal Dimension Report“ von ***FDim*** wiedergegeben.

Außerdem haben die Autoren des Programms einen ausführlich kommentierten „Fractal Dimension Report“ angefertigt. Dieser Report gibt Hilfestellungen für die Interpretation der mit ***FDim*** (Originalname: ***FD3***) angefertigten Dimensionsanalysen.

5. Erste Erfahrungen mit biolog. Signalen:

Für einen ersten Versuch der Analyse von biologischen Signalen wurden einige – aus anderen Projekten stammende – EVPs (evoked potentials) willkürlich ausgewählt. Die EVPs lagen mit $N = 200$ abgetasteten Datenpunkten vor. Mit ***FDim*** wurden dabei folgende Abschätzungen der fraktalen Dimensionen gefunden:

Kapazitäts-Dimension:	$D \approx 1{,}18$
Informations-Dimension:	$D_I \approx 1{,}11$
Korrelations-Dimension:	$D_C \approx 1{,}07$

Diese Ergebnisse bestätigen zunächst, daß auch für diese EVPs die Ungleichung $D \geq D_I \geq D_C$ gilt. Die berechneten fraktalen Dimensionen liegen deutlich oberhalb des (klassischen) Erwartungswerts von 1. Es darf also sehr wohl vermutet werden, daß ein fraktaler Prozeß bei der Erzeugung der EVP's eine Rolle spielt. Es wird sich also lohnen, die fraktalen Dimensionen von EVPs systematisch zu untersuchen.

6. Sonstige Hinweise:

■ Das Programm ***FDim*** verwendet für die Ermittlung der fraktalen Dimensionen die *Anzahl der Zahlenwerte* (Koordinaten des 1. Datenpunktes), die auf der 2. Zeile der Eingabe-Datei stehen, als die „einhüllende“ (embedding) Dimension.

■ Sollten bei der Benutzung von ***FDim*** (Originalname: ***FD3***) Fehler (bugs) erkannt werden, so möchten die Autoren per E-Mail informiert werden. E-Mail: john@ishi.csustan.edu (Internet).

■ Im internationalen UseNET (wdv-notes Nr. 79) werden Fragen zu *nichtlinearen dynamischen Systemen* in der News-Gruppe »comp.theory.dynamic-sys« diskutiert und beantwortet.

7. Literatur:

[1] Sarraille, J. J.: Developing Algorithms for Measuring Fractal Dimension. FU Berlin (IfP): wdv-notes Nr. 148, 1992–1993.

[2] Mandelbrot, B.: Die fraktale Geometrie der Natur. Basel: Birkhäuser 1991 (Sonderauflage). ISBN: 3-7643-2646-8. Preis: 49,80 DM. Die Lektüre dieses Buches ist zur Einführung unerläßlich. Außerdem sind hier sehr viele weiterführende Literaturhinweise zu finden. – [BU 753]

[3] Dittberner, K.-H.: Die Fraktalen Dimensionen der Natur. FU Berlin (IfP): wdv-notes Nr. 88, 1984–1992.

[4] Hentschel, H. G. E. and Procaccia, I.: The infinite number of generalized dimensions of fractals and strange attractors. Physica 8D, 435–444 (1983).

[5] Dittberner, K.-H.: Benutzung des HAIRclusters: Die genormte POSIX-Benutzerumgebung – Teil 1. FU Berlin (IfP): wdv-notes Nr. 162, 1992–1993.

Notizen:

© khd-research • 2004 • Eine Publikation von http://www.khd-research.net/ • File: Notes/wdv-notes_176.pdf