Multifractal Detrended Fluctuation Analysis

Leonardo Rydin Gorjão[1, 2, *]

[1] *Forschungszentrum Jülich, Institute for Energy and Climate Research - Systems Analysis and Technology Evaluation (IEK-STE), 52428 Jülich, Germany*
[2] *Institute for Theoretical Physics, University of Cologne, 50937 Köln, Germany*

Multifractal Detrended Fluctuation Analysis in Python

SUMMARY

A common tool to unveil the nature of the fractionality of a process, natural or computer-generated, is the Detrended Fluctuation Analysis (DFA), initally developed by Peng *et al* and later extended to study multifractal processes by Kandelhardt *et al*, giving rise to Multifractal Detrended Fluctuation Analysis (MFDFA) [1, 2]. It addresses the question of the presence of long-range correlations and can be employ to study discrete processes, as auto-regressive models, as well as time-continuous stochastic processes. An extensive study of DFA and the interplay between trends in data and correlated noise can be found in [3].

Add references on Earthquake studies, climate, temperature, power-grids(?)

In order to determine the self-affinity of a stochastic process, one can study the relation between the variance of the process and time. Auto-regressive and stochastic processes diffuse with different rates, and uncovering the rates of diffusion is of importance in natural processes with power-law correlations, like temperature variability [4] earthquake frequency [5], heartbeat dynamics [6]. Fluctuation Analysis provides a method to uncover these correlations, but fails in the presence of trends in the data, which is, for example, particularly present in weather and climate data. Detrending the data via polynominal fittings allows one to uncover solely the relation between the inherent fluctuations and the time scaling of a process. Moreover, several processes might be driven by more than one time scale. They might by of a mono- or multi-fractal nature. By studying a continuum of power variations of the Detrended Fluctuation Analysis one extends into Multifractal Detrended Fluctuation Analysis, which permits the study of the fractality of the data by comparing power variations of Detrended Fluctuation Analysis, see below.

There are currently no viable flexible implementations of Multi-Fractal Detrended Fluctuation Analysis in Python, but there are several `Matlab` versions available. There is a particularly thorought introductory guide to Multifractal Detrended Fluctuation Analysis, and subsequently a source-code by Espen Ihlen from 2012 [7], which is flexible but slow. With this implementation efficiency was sought, by makng the most out of Python, reshaping the code to allow for multithreading, especially relying on `numpy`'s `polynomial`, which scales easily with modern computers having more CPU cores.

THEORY

Multifractal Detrended Fluctuation Analysis studies the fluctuation of a given process by considering increasing segments of the timeseries. Take a timeseries $X(t)$ with N elements X_i, $i = 1, 2, \dots, N$. Obtained the "detrended" profile of the process by defining

$$Y_i = \sum_{k=1}^{i} \left(X_k - \langle X \rangle \right), \text{ for } i = 1, 2, \dots, N, \tag{1}$$

i.e., the cumulative sum of X subtracting the mean $\langle X \rangle$ of the data. Section the data into smaller non-overlapping segments of length s, obtaining therefore $N_s = \mathrm{int}(N/s)$

---

* `l.rydin.gorjao@fz-juelich.de`

segments. Given the total length of the data isn't always a multiple of the segment's length s, discard the last points of the data. Consider the same data, apply the same procedure, but discard now instead the first points of the data. One has now $2N_s$ segments.

To each of this segments fit a polynomial y_v of order m and calculate the variance of the difference of the data to the polynomial fit

$$F^2(v,s) = \frac{1}{s}\sum_{i=1}^{s}[Y_{(v-1)s+i} - y_{v,i}]^2, \text{ for } v = 1,2,\dots,N_s, \tag{2}$$

where $y_{v,i}$ is the polynomial fitting for the segment i of length v. One also has the freedom to choose the order of the polynomial fitting. This gives rise to the denotes `DFA1`, `DFA2`, ..., for the orders chosen.

Notice now $F^2(v,s)$ is a function of each variance of each v-segment of data and of the different s-lenght segments chosen. On can now define, will all due generally, the q-th order fluctuation function by averaging each row of segments of size s

$$F_q^2(s) = \left\{\frac{1}{N_s}\sum_{v=1}^{N_s}[F^2(v,s)]^{q/2}\right\}^{1/q} \tag{3}$$

where Detrended Fluctuation Analysis with $q = 2$ is a subset of Multi-Fractal Detrended Fluctuation Analysis (where $q \in \mathbb{R}$). The q-th order fluctuation function $F_q^2(s)$ is our object of interest.

The inherent scaling properties of the data, is the data displays power-law correlations, can now be studying in a log-log plot of $F_q^2(s)$ versus s, where the scaling of the data obeys a power-law with exponent $h(q)$ as

$$F^2(s) \sim s^{h(q)}, \tag{4}$$

where $h(q)$ is the *self-affinity* exponent, which may dependent on q, and relates to the previously introduced Hurst exponent. The self-affinity exponent $h(q)$ is calculated by finding the slope of this curve in the log-log plots, as seen henceforth in the figures.

If the data is monofractal, the *self-affinity* exponent $h(q) = h$ is independent of q. On the other hand, if the data is multifractal, the dependence on q can be understood by studying the multifractal scaling exponent $\tau(q)$ [8]

$$\tau(q) = qh(q) - 1. \tag{5}$$

For a clearer discussion of these properties, see [2, 8].

I. EXAMPLES

To exemplify the usage of Multifractal Detrended Fluctuation Analysis, let's take two common examples of stochastic processes, a continuous-time stochastic process: fractional Brownian motion and a discrete auto-regressive process:

For an example of multifractal behaviour in real-world data of sun spots timeseries, alongside a detailed explanation of MFDFA, or an application to European temperature variability, see respectively [9] and [4].

II. THE `MFDFA` LIBRARY

The Multifractal Detrended Fluctuation Analysis library `MFDFA` presented is a standalone package based integrally on python's `numpy`, thus it can avail also of `numpy`'s masked arrays `ma`, which is particularly convenient when dealing with large datasets with missing data.

ACKNOWLEDGEMENTS

L. R. G. thanks Klaus Lehnertz, M. Reza Rahimi Tabar, and Pedro G. Lind for all the help in understanding stochastic processes, Dirk Witthaut for the support, and

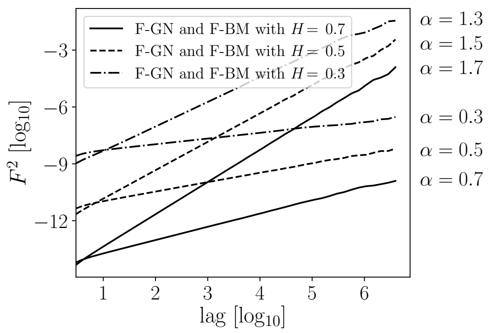

FIG. 1. Detrended Fluctuation Analysis of three examplary paths of $dB_H(t)$ with $H = 0.3$, $H = 0.5$, and $H = 0.7$ are displayed, alongside with their integrated counterparts $B_H(t)$ (with $H = 0.3$, $H = 0.5$, and $H = 0.7$). Gaussian noise $dB.$ has a generalised Hurst exponent $\alpha = H$ and Brownian motion $B.$ has $\alpha = H + 1$, given integration increases regularity by 1. The generalised Hurst exponents α are obtained my extracting the slopes of the curves. Numerically integrated with an integration step $\Delta t = 0.0001$ over 500 time units (5×10^6 data points)

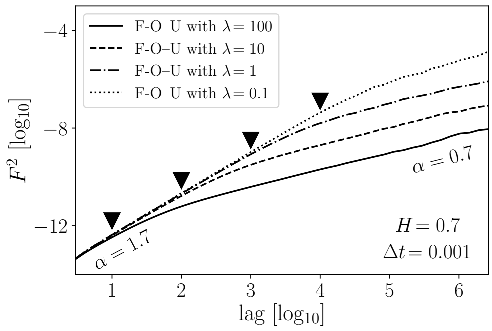

FIG. 2. Detrended Fluctuation Analysis of four examplary paths of a fractional Ornstein–Uhlenbeck process with $H = 0.7$, fixed $\sigma = 1$, fixed integration timestep $\Delta t = 0.001$, but varying $\lambda = 100, 10, 1$, and 0.1 (bottom to top). Regardless of the integration timestep, the generalised Hurst exponent α matches the expected value: $\alpha = 0.7$ for $dB_{H=0.7}(t)$, $\alpha = 1.7$ for $B_{H=0.7}(t)$. Numerically integrated with an integration step $\Delta t = 0.001$ over 5000 time units (5×10^6 data points)

Marieke Helmich for the text reviews. L. R. G. gratefully acknowledges support by the Helmholtz Association, via the joint initiative *Energy System 2050 - A Contribution of the Research Field Energy*, the grant No. VH-NG-1025, the scholarship funding from *E.ON Stipendienfonds*, and the *STORM - Stochastics for Time-Space Risk Models* project of the Research Council of Norway (RCN) No. 274410, under the supervision of Giulia di Nunno.

---

[1] C.-K. Peng, S. V. Buldyrev, S. Havlin, M. Simons, H. E. Stanley, and A. L. Goldberger, Physical Review E **49**, 1685 (1994).
[2] J. W. Kantelhardt, S. A. Zschiegner, E. Koscielny-Bunde, S. Havlin, A. Bunde, and H. Stanley, Physica A **316**, 87 (2002).
[3] K. Hu, P. C. Ivanov, Z. Chen, P. Carpena, and H. Eugene Stanley, Physical Review E **64**, 011114 (2001).
[4] P. G. Meyer and H. Kantz, Climate Dynamics **53**, 6909 (2019).
[5] S. Shadkhoo, F. Ghanbarnejad, G. Jafari, and M. Tabar, Open Physics **7**, 620 (2009).
[6] P. C. Ivanov, L. A. N. Amaral, A. L. Goldberger, S. Havlin, M. G. Rosenblum, Z. R. Struzik, and H. E. Stanley, Nature **399**, 461 (1999).
[7] E. Ihlen, Frontiers in Physiology **3**, 141 (2012).
[8] A.-L. Barabási and T. Vicsek, Physical Review A **44**, 2730 (1991).
[9] M. S. Movahed, G. R. Jafari, F. Ghasemi, S. Rahvar, and M. R. R. Tabar, Journal of Statistical Mechanics: Theory and Experiment **2006**, P02003 (2006).