Journal of Statistical Software

MMMMMM YYYY, Volume VV, Issue II. doi: 10.18637/jss.v000.i00

A **Python/Zig** optimized and customizable implementation for the ρ_{DCCA} and DMC_x^2 methods

Fernando Ferraz Ribeiro
Universidade Federal da Bahia

Gilney Figueira Zebende
Universidade Estadual de Feira de Santana

Abstract

This paper presents tha **Zebende**, a Python package writen in Python and Zig, that calculates the DFA, $DCCA$ ρ_{DCCA} and the DMC_x^2. The package presents an optimized algorithm that significantly improves the calculations speed. A comparison with ohter packages that calculates the .The package is also the firts to implement the DMC_x^2 coefficient for any number of series.

Keywords: ρ_{DCCA}, DMC_x^2, optimization, Python, Zig.

1. introduction

The ρ_{DCCA} (Zebende 2011) is a widely used coefficient that measures the cross-correlation between tow non-stationary time series. It´s an extension of the Detrended Fluctuation Analysis (DFA) (Peng, Buldyrev, Havlin, Simons, Stanley, and Goldberger 1994) and the Detrended Cross-correlation Analysis ($DCCA$) (Podobnik and Stanley 2008): while the DFA calculates the self-affinity and long-memory properties of a time series data, and the $DCCA$ analyses power-law cross correlations between two different non-stationarity time series, the ρ_{DCCA} coefficient quantifies this cross-correlation in simple values ranging from -1 to 1, where -1 indicates a perfect anti-correlation between the series, 1 a perfect correlation and zero (0) no correlation at all.

The Detrended Multiple Cross-Correlation Coefficient (Zebende and Silva 2018) (DMC_x^2) is a generalization of the ρ_{DCCA} coefficient that correlates one time series (dependent variable) a number of time series (independent variables). The DMC_x^2 values ranges from zero (0), indicating no correlation to 1, meaning perfect correlation or anti-correlation between the dependent and the independent variables.

This paper presents the **Zebende** Python package, an implementation of the DFA, $DCCA$, ρ_{DCCCA}, DMC_x^2 and utility functions related to the methods. In section 2 the steps for calculating the ρ_{DCCCA} and DMC_x^2 are presented and discussed. Section 3 shows how this library was implemented, the optimization technics and the recommended steps to use the library. In Section 4 the **Zebende** package is compared with other packages for Python and R that calculates the ρ_{DCCA} in terms of performance and usability, leading to the conclusions in 5.

2. Algorithms of the coefficients

$$DMC_x^2 \equiv \rho_{Y,X_i}(n)^T \times \rho^{-1}(n) \times \rho_{Y,X_i}(n) \tag{1}$$

$$\rho^{-1}(n) = \begin{pmatrix} 1 & \rho_{X_1,X_2}(n) & \rho_{X_1,X_3}(n) & \dots & \rho_{X_1,X_j}(n) \\ \rho_{X_2,X_1}(n) & 1 & \rho_{X_2,X_3}(n) & \dots & \rho_{X_2,X_j}(n) \\ \vdots & \vdots & \vdots & \dots & \vdots \\ \rho_{X_j,X_1}(n) & \rho_{X_j,X_2}(n) & \rho_{X_j,X_3}(n) & \dots & 1 \end{pmatrix}^{-1} \tag{2}$$

$$\rho_{Y,X_k}(n)^T = [\rho_{Y,X_1}(n), \rho_{Y,X_2}(n), \cdots, \rho_{Y,X_j}(n)] \tag{3}$$

$$\rho_{DCCA}(n) = \frac{F^2_{DCCA\ (x\alpha,\ x\beta)}(n)}{F_{DFA\ (x\alpha)}(n) \times F_{DFA\ (x\beta)}(n)} \tag{4}$$

Taking a time series $\{x_i\}$ with i ranging from 1 to N, the integrated series X_k is calculated by $X_k = \sum_{i=1}^{k} [x_i - \langle x \rangle]$ with k also ranging from 1 to N;

the X_k series is divided in $N-n$ boxes of size n(time scale),each box containing $n+1$ values, starting in i up to $i+n$;

for each box, a polynomial (usually of degree 1) best fit is calculated, getting $\widetilde{X}_{k,i}$ with $i \leq k \leq (i+n)$ (detrended values);

in each box is calculated: $f^2_{DFA}(n,i) = \frac{1}{1+n}\sum_{k=i}^{i+n}(X_k - \widetilde{X}_{k,i})^2$

for all the boxes of a time scale, the DFA is calculated as:

$F_{DFA}(n) = \sqrt{\frac{1}{N-n}\sum_{i=1}^{N-n} f^2_{DFA}(n,i)}$;

for a number of different timescales (n), with possible values $4 \leq n \leq \frac{N}{4}$ the F_{DFA} function is calculated to find a relation among $F_{DFA} \times n$

Taking two time series with the same length $\{x\alpha_i\}$ and $\{x\beta_i\}$ with i ranging from 1 to N, the integrated series $X\alpha_k$ and $X\beta_k$ are calculated by $X_k = \sum_{i=1}^{k} [x_i - \langle x \rangle]$ for each series, with k also ranging from 1 to N;

$X\alpha_k$ and $X\beta_k$ series are divided in $N-n$ boxes of size n(time scale), each box containing $n+1$ values, starting in i up to $i+n$;

for each box, a polynomial (usually of degree 1) best fit is calculated, getting $\widetilde{X\alpha}_{k,i}$ and $\widetilde{X\beta}_{k,i}$, for series $\{x\alpha_i\}$ and $\{x\beta_i\}$ respectively, with $i \leq k \leq (i+n)$ (detrended values);

in each box is calculated: $f^2_{DCCA}(n,i) = \frac{1}{1+n}\sum_{k=i}^{i+n}(X\alpha_k - \widetilde{X\alpha}_{k,i}) \times (X\beta_k - \widetilde{X\beta}_{k,i})$

for all the boxes of a time scale, the DFA is calculated as:

$F^2_{DCCA}(n) = \frac{1}{N-n}\sum_{i=1}^{N-n} f^2_{DCCA}(n,i)$;

for a number of different timescales (n), with possible values $4 \leq n \leq \frac{N}{4}$ the F^2_{DCCA} function is calculated to find a relation among $F^2_{DCCA} \times n$

3. Zebende package: implementation and optimization

4. Results

5. Summary and discussion

- As usual ...

References

Peng CK, Buldyrev SV, Havlin S, Simons M, Stanley HE, Goldberger AL (1994). "Mosaic Organization of DNA Nucleotides." **49**(2), 1685–1689.

Podobnik B, Stanley HE (2008). "Detrended cross-correlation analysis: A new method for analyzing two nonstationary time series." *Physical Review Letters*, **100**(8). ISSN 00319007. doi:10.1103/PhysRevLett.100.084102. 0709.0281.

Zebende GF (2011). "DCCA cross-correlation coefficient: Quantifying level of cross-correlation." *Physica A: Statistical Mechanics and its Applications*, **390**(4), 614–618. ISSN 03784371. doi:10.1016/j.physa.2010.10.022. URL http://dx.doi.org/10.1016/j.physa.2010.10.022.

Zebende GF, Silva AM (2018). "Detrended Multiple Cross-Correlation Coefficient." *Physica A*, **510**, 91–97. ISSN 0378-4371. doi:10.1016/j.physa.2018.06.119.

Affiliation:

Fernando Ferraz Ribeiro
Universidade Federal da Bahia
Faculty of Achitecture
Universität Innsbruck
Universitätsstr. 15
6020 Innsbruck, Austria
E-mail: fernando.ribeiro@ufba.br
and
Centro Universitário Senai-Cimatec

Journal of Statistical Software http://www.jstatsoft.org/
published by the Foundation for Open Access Statistics http://www.foastat.org/
MMMMMM YYYY, Volume VV, Issue II *Submitted:* yyyy-mm-dd
doi:10.18637/jss.v000.i00 *Accepted:* yyyy-mm-dd