

Instituto de
Computação



NeuroMat

Introdução à análise de séries temporais: 4. Séries não-lineares e teoria da informação

ARTHUR VALENCIO

Pós-doutorando IC/Unicamp

CEPID NeuroMat

Projeto TOPE Unicamp, Campinas, 21 de Outubro de 2019



>>Resumo do curso

Dia 1: Introdução à ferramentas de análise de dados

- Planilhas
- Linguagens de programação de propósito geral
- Linguagens de programação voltadas à dados

Dia 2: Técnicas convencionais de análise

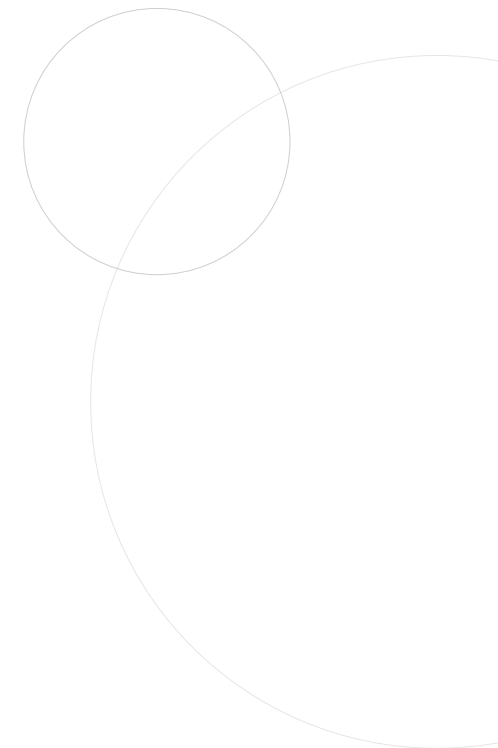
- análise espectral
- wavelet
- correlação de Pearson e Spearman
- regressão

Dia 3: Análise multivariada

- correlação parcial
- análise fatorial
- análise de componentes principais
- análise de componentes independentes

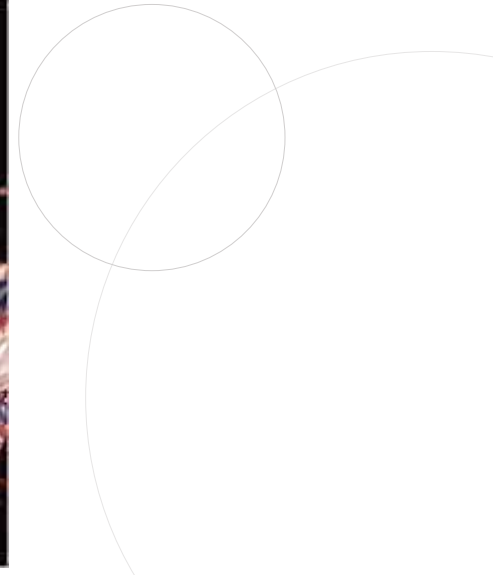
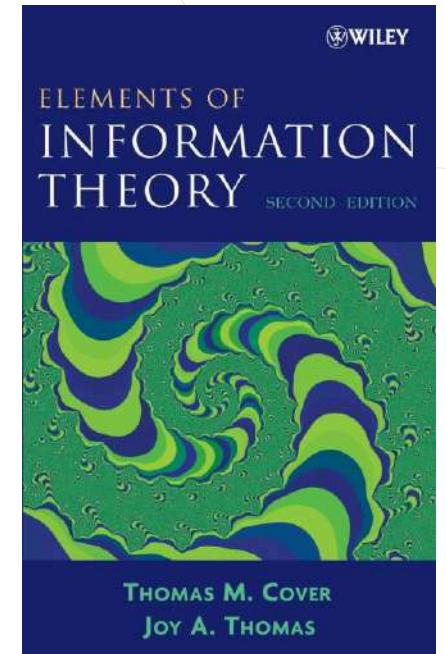
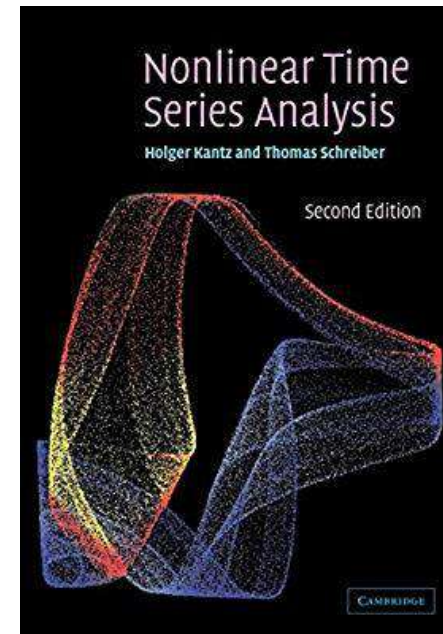
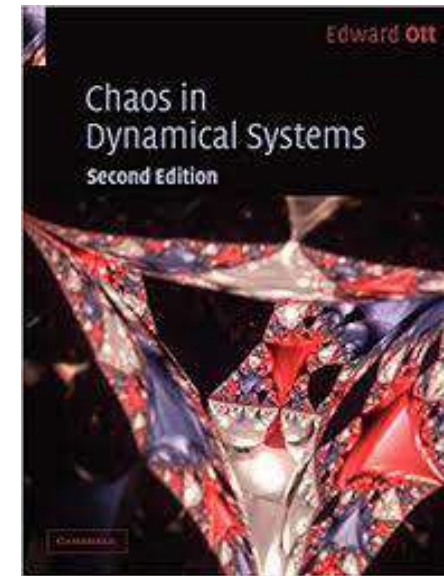
Dia 4: Séries não-lineares e teoria da informação

- reconstrução do espaço de fase
- expoentes de Lyapunov
- entropia de Shannon
- informação mútua, entropia de transferência, CaMI, direcionalidade





>>Quick refs





>>“Dinâmica”?

Digamos que temos um sistema de n variáveis \mathbf{x} que evoluem com o tempo t . Chamaremos de dinâmica f a lei que dita como essas variáveis evoluem.

O sistema dinâmico pode ser categorizado como:

- Fluxos: $\dot{\mathbf{x}} = f(\mathbf{x})$
- Mapas: $\mathbf{x}_{n+1} = f(\mathbf{x}_n)$

(Embora que, mesmo para fluxos, precisamos discretizar o tempo quando trabalhamos com um computador, mas aqui estamos falando da “lei natural” regendo o fenômeno)

Além disso a dinâmica pode ser:

- *Determinística*: não existem variáveis aleatórias, todas as variáveis são “conhecidas”
- *Estocástica*: variáveis aleatórias fazem parte da descrição do sistema



>>“Não-linear”?

Uma função é chamada *linear* se satisfizer:

$$a) f(x + y) = f(x) + f(y)$$

$$b) f(\alpha x) = \alpha f(x)$$



A maior parte dos problemas não são lineares, mas muitos se tornam *linearizáveis* após mudança de variáveis e aproximações

Aqueles que não podem ser linearizados são chamados **sistemas não-lineares**

- Multiestabilidade
- Transiência
- **Caos**
- Sólitons
- Histerese
- Singularidades

Exemplo simples:

$$f(x) = x_{n+1} = rx(1 - x)$$

Mapa logístico

(população de gafanhotos, R. May, 1976)



>>Reconstrução do espaço de fase

ou espaço de estado?

Floris Takens (1981)

Estudando fluidos turbulentos



Visitante regular do IMPA

Versão “correta”:

Seja $f: M \rightarrow M$ uma função suave que descreve uma dinâmica com um atrator estranho A . Existe um difeomorfismo ϕ que mapeia $A \rightarrow \mathbb{R}^k$. Mais especificamente, seja α uma função de observação, diferenciável pelo menos até segunda ordem, que associe um número real a um ponto do atrator A . Então a função:

$$\phi_T(x) = (\alpha(x), \alpha(f(x)), \dots, \alpha(f^{k-1}(x)))$$

é uma imersão do atrator estranho A .

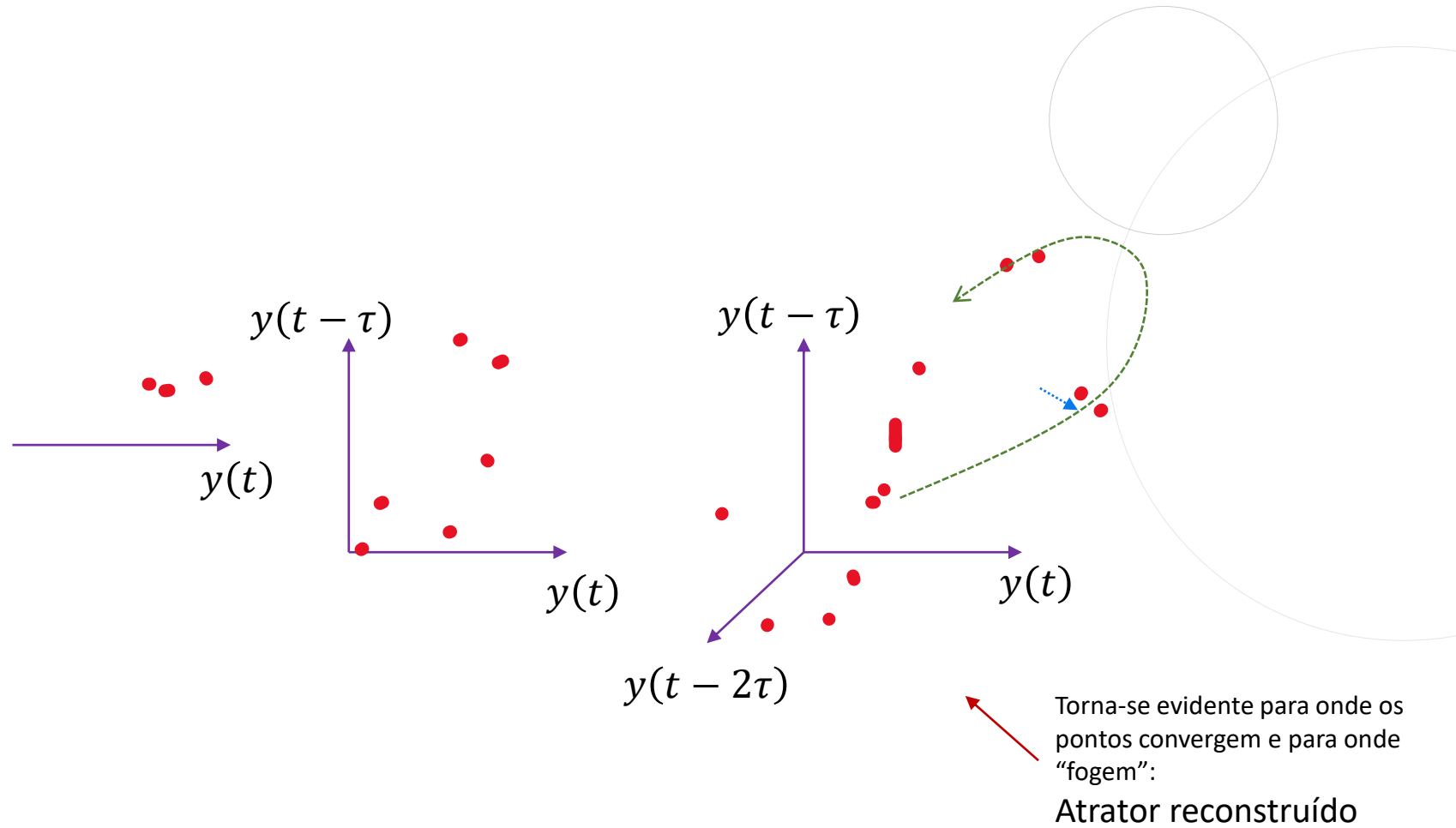
Versão prática:

Suponha que nós temos um observável unidimensional y de uma evolução possivelmente n -dimensional x de uma dinâmica *determinística* (p.ex. medidas do valor numérico da velocidade (speed) de uma partícula se movendo em 3 dimensões).

- Vamos construir um vetor k -dimensional formado pelos valores atuais $y(t)$ e seus valores passados $y(t - \tau), y(t - 2\tau), \dots, y(t - (k - 1)\tau)$
- Conforme $k \rightarrow \infty$, a dinâmica do y no espaço k -dimensional torna-se mais previsível
- A dinâmica gerada é topologicamente equivalente à do espaço de estado original (do x)



>>Reconstrução do espaço de fase

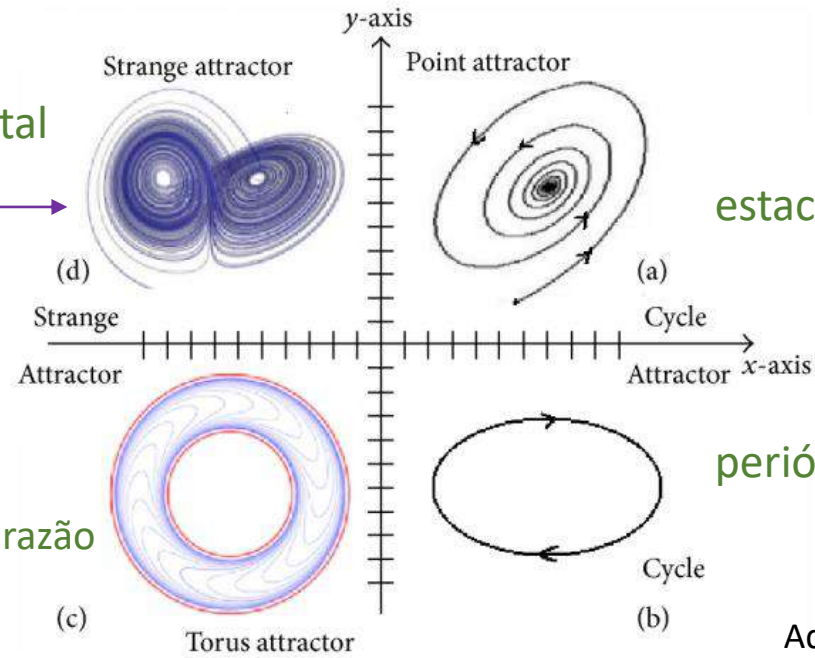




>>Reconstrução do espaço de fase

Estrutura fractal
Provavelmente →
caóticos*

quasi-periódico/
composição de
frequências com razão
irracional



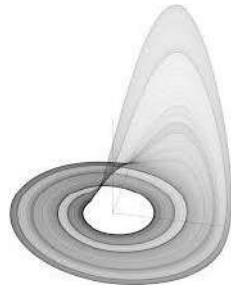
estacionaridade

periódico

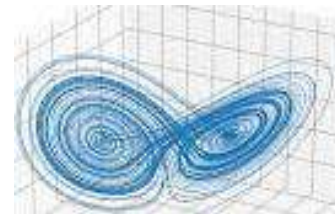
Adewumi et al (2016)

*Grebogi, Ott, Pelikan & Yorke (1984), *Physica D*, traz contra-exemplo

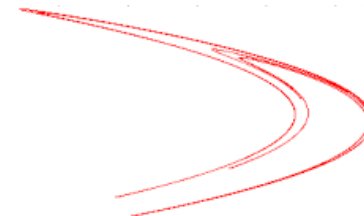
Rössler
(reações químicas)



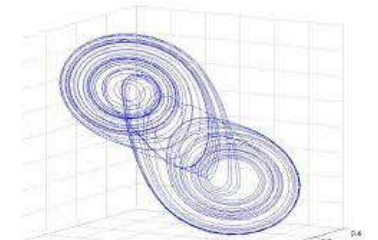
Lorenz
(convecção atmosférica)
(mas tbm lasers, motores, química,
osmose, etc)



Hénon
(simplificação de um
comportamento do Lorenz)
(hj aplicado para efeitos em MEMS)



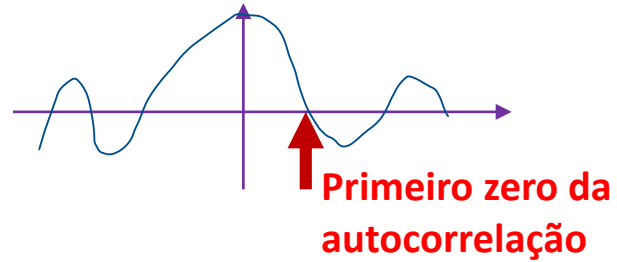
Double-scroll
(circuito de Chua)
(eng elétrica/eletrônica)





>>Reconstrução do espaço de fase

- Qual o melhor delay τ que posso utilizar?



ou

Primeiro mínimo da (auto-) informação mútua

- Qual o número de dimensões k que preciso chegar?

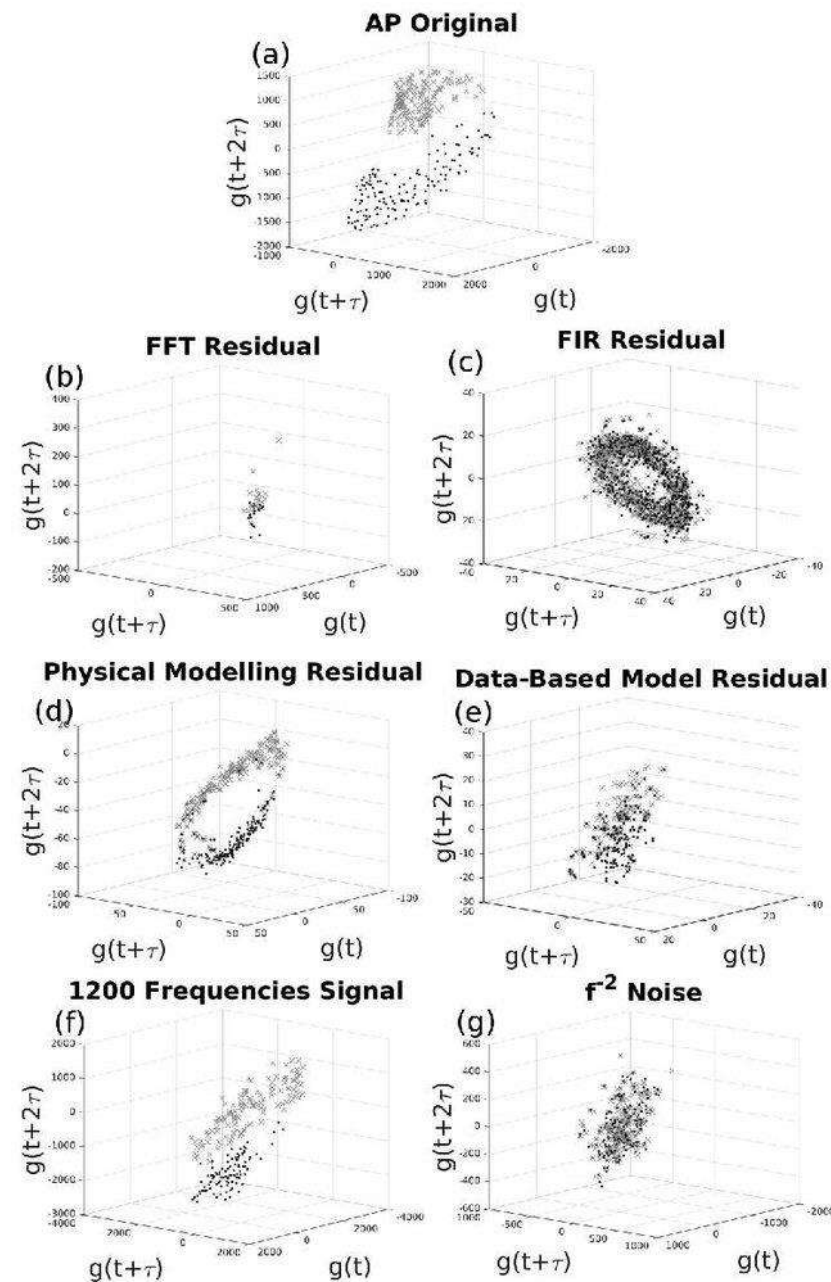
Método dos falsos vizinhos (False Nearest Neighbours, Kennel, Brown & Abarbanel (1992))

Ir reconstruindo para valores crescentes de k e vendo, para cada pequena vizinhança de pontos arbitrariamente escolhidos, quantos vizinhos são *falsos* (i.e. vem de pontos distantes na série original).

Como no *scree plot*, a curva de percentual de falsos vizinhos irá saturar num determinado valor baixo, e define-se a dimensão k a partir desta saturação



>>Reconstrução do espaço de fase

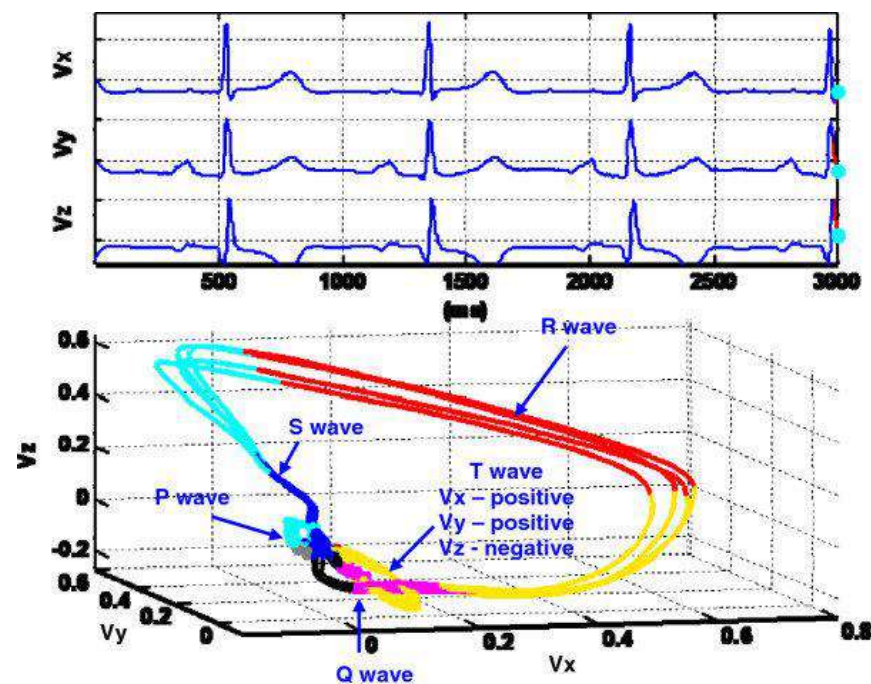


Valencio, Baptista & Grebogi (2017)

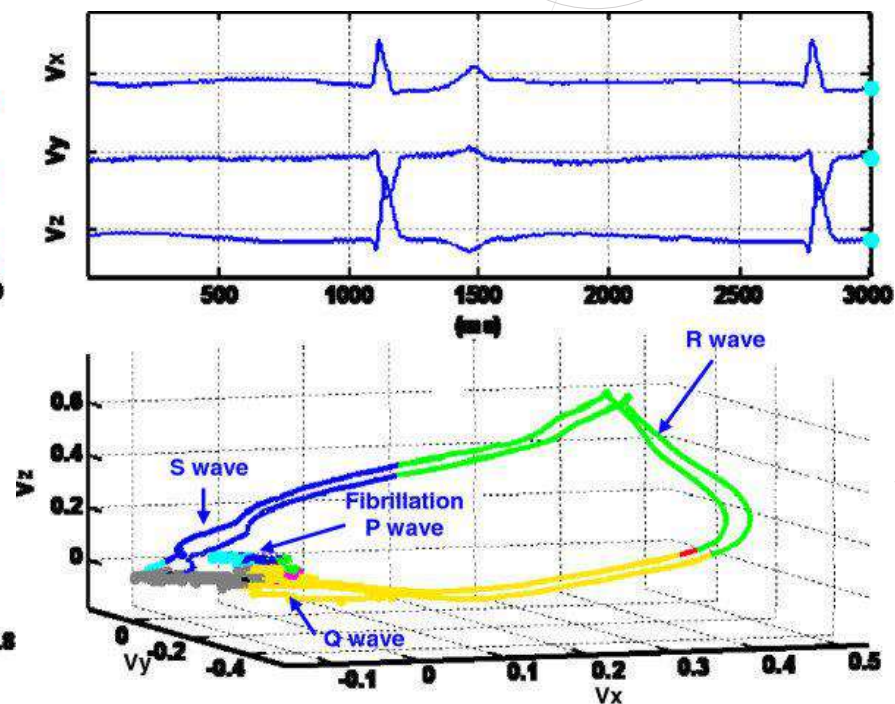


>>Reconstrução do espaço de fase

ECG sadio



ECG arritmia e fibrilação auricular



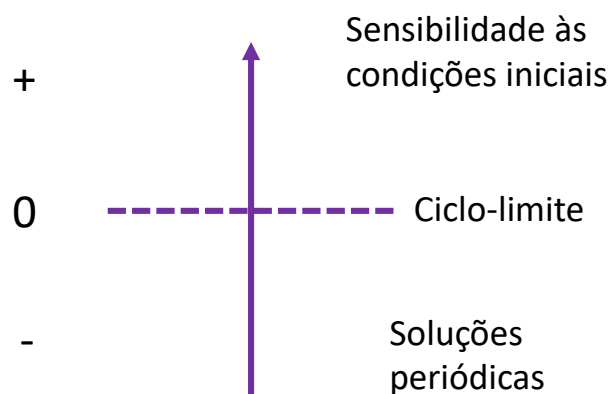
Yang, Bukkapatnam & Romanduri (2012)



>>Expoentes de Lyapunov

Objetivo: determinar o quanto pontos vizinhos divergem ao longo da trajetória no espaço de fase

Maior expoente de Lyapunov



- Seja uma dinâmica $\dot{x}_i = f_i(x)$.
- Definimos a matriz jacobiana pelos elementos $J_{ij} = \left. \frac{df_i(x)}{dx_j} \right|_{x(t)}$
- Com isso definimos a evolução dos vetores tangentes Y através de: $\dot{Y} = JY$, com $Y(t=0) = I$
- Definimos então uma matriz $\Lambda = \lim_{t \rightarrow \infty} \frac{1}{2t} \ln(Y(t)Y^T(t))$
- Os autovalores λ_i de Λ são os expoentes de Lyapunov

→ O maior valor informa sobre a sensibilidade às condições iniciais

→ Os demais informam sobre a taxa de produção de entropia e a dimensão fractal do sistema

Se você só estiver interessado no maior expoente de Lyapunov, num sistema discreto:

$$\lambda = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} \ln |f'(x)|$$



>>Expoentes de Lyapunov

Table 3: Largest Lyapunov exponent [bits h^{-1}] of original gravity time-series and residuals after filtering for all stations

Station	Original series	FFT res.	FIR res.	Phys. Model. res.	Data-based res.
AP	1.412 ± 0.007	0.349 ± 0.007	0.77 ± 0.03	0.93 ± 0.02	0.63 ± 0.02
BF	1.208 ± 0.005	0.29 ± 0.01	0.53 ± 0.03	0.84 ± 0.01	0.26 ± 0.01
MA	1.317 ± 0.006	0.80 ± 0.03	0.76 ± 0.02	0.73 ± 0.01	0.42 ± 0.02
NY	0.860 ± 0.009	0.35 ± 0.04	0.41 ± 0.03	0.81 ± 0.02	0.61 ± 0.02
SU	1.326 ± 0.007	0.68 ± 0.04	0.58 ± 0.03	0.90 ± 0.02	0.84 ± 0.02

Valencio, Baptista & Grebogi (2017)



>>Caos

1. Sensibilidade às condições iniciais

2. *Topologicamente transitivo*

(idéia geral: “sempre possui pontos que se movem de uma região para a outra do espaço de fase, *i.e.* o espaço de fase não pode ser decomposto em dois ou mais grupos separados)

3. *Possui órbitas periódicas densas*

(*i.e.* qualquer ponto no atrator está arbitrariamente próximo de uma órbita periódica)

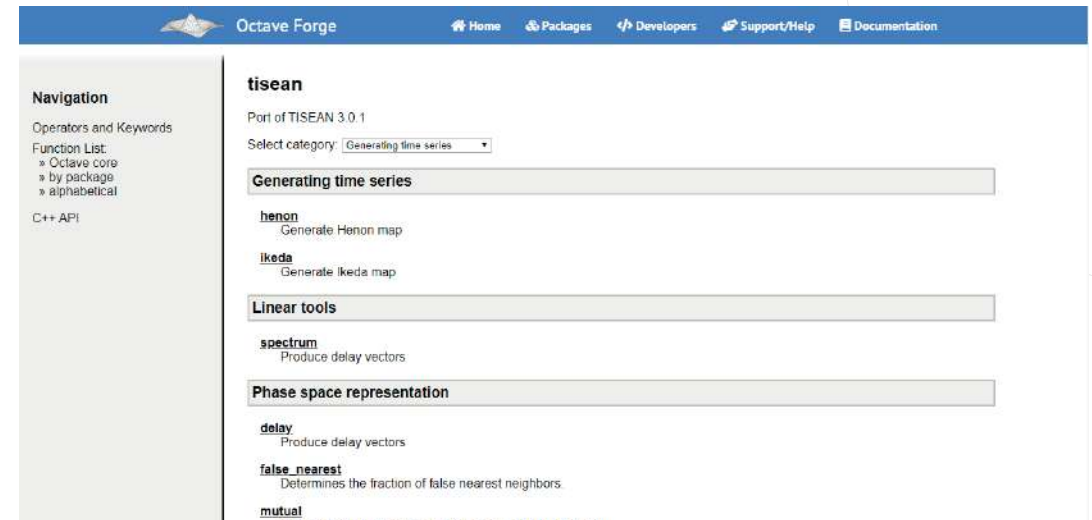
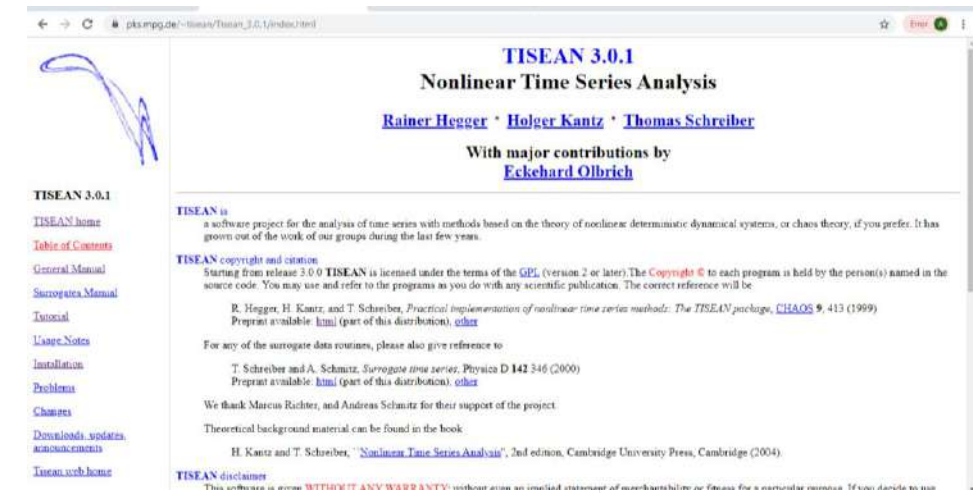
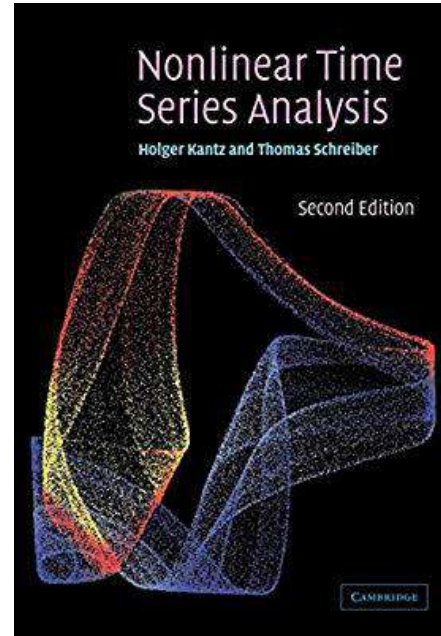
- Dois vizinhos, com trajetórias bem parecidas, a qualquer momento podem divergir para locais bem distantes
- Dois pontos, não importa o quão distantes estejam, eventualmente vão ficar próximos

Imprevisibilidade

(mesmo para um sistema determinístico)



>>TISEAN



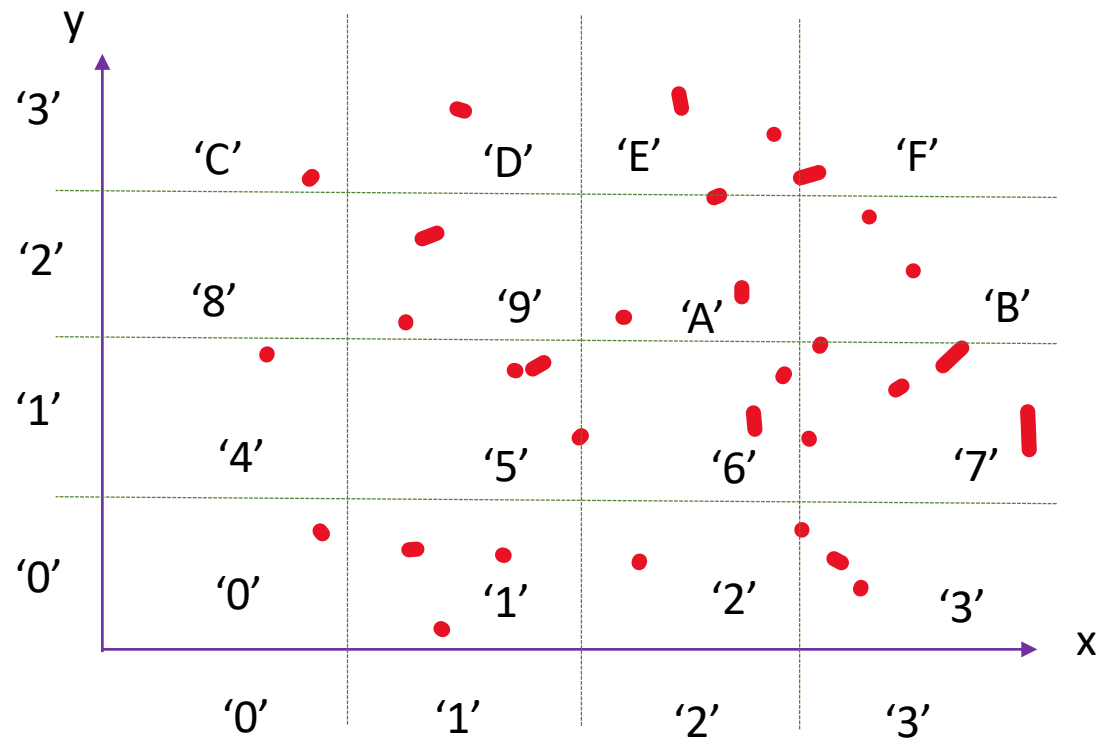
Octave: Pacote tisean

R: Pacote RTisean

Na webpage do projeto TISEAN tbm há software open source desenvolvido em C/Fortran



>>Particionamento (Symbolic encoding)



L=2

t	S(x)	ϕ_x	S(y)	ϕ_y	S(x,y)	(ϕ_x, ϕ_y)
0	0	-	3	-	C	-
1	1	'01'	3	'33'	D	'CD'
2	2	'12'	2	'32'	A	'DA'
3	2	'22'	0	'20'	2	'A2'
4	3	'23'	1	'01'	7	'27'
5	1	'31'	2	'12'	9	'79'



>>Entropia de Shannon

$$H = - \sum_i p_i \log p_i$$

Base 2: bits
Base e (i.e., \ln): nats

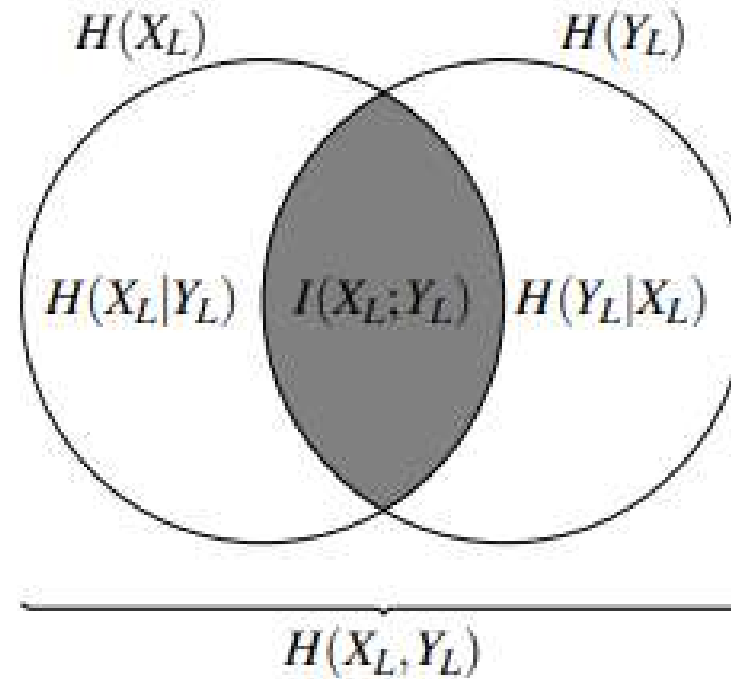
Idéia: nível de incerteza sobre uma observação da variável
(ou do sistema)



>> Informação Mútua

Idéia:

1. Redução da incerteza de uma variável quando se conhece a outra
2. Portanto equivalente à idéia genérica de “informação compartilhada entre as variáveis”



$$\begin{aligned} I(X_L; Y_L) &= H(X_L) + H(Y_L) - H(X_L, Y_L), \\ &= \sum_{\phi_x, \phi_y} p(\phi_x, \phi_y) \log \frac{p(\phi_x, \phi_y)}{p(\phi_x)p(\phi_y)}. \end{aligned}$$



>>Pausa: causalidade de Granger

Granger (1969 & 1980)
Nobel Economia de 2003

1. A causa precede o efeito
2. A causa possui alguma informação *única* sobre os futuros valores do efeito

Método:

Modelo auto-regressivo de y :

$$y_t = a_0 + a_1 y_{t-1} + a_2 y_{t-2} + \dots$$

Quantos termos?: AIC

Modelo auto-regressivo de y considerando x :

$$y_t = a_0 + a_1 y_{t-1} + a_2 y_{t-2} + \dots + b_0 x_{t-p} + b_1 x_{t-p-1} + \dots$$

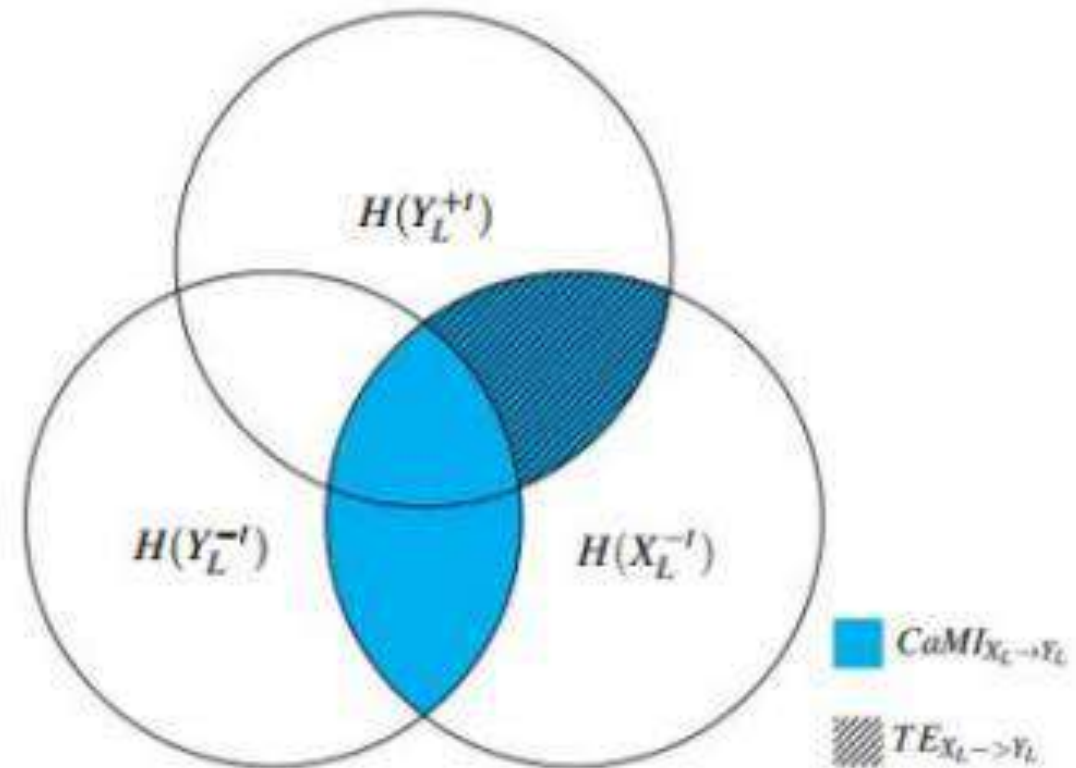
Se o modelo com o x possui maior poder explicativo dos dados (teste t),
então X Granger-causa Y



>> Entropia de Transferência, CaMI e direcionalidade

Idéia:

1. Redução da incerteza do futuro de uma variável quando se conhece o passado dela e é dado o passado da outra.
2. Equivalente à dizer que esta outra variável possui informação única sobre o futuro da primeira variável
3. Portanto, “causalidade de Granger generalizada”



Bianco-Martinez & Baptista (2018)



>> Entropia de Transferência, CaMI e direcionalidade

$$\begin{aligned} TE_{X_L \rightarrow Y_L} &= \sum_{\phi_x^{-t}, \phi_y^{-t}, \phi_y^{+t}} p(\phi_x^{-t}, \phi_y^{-t}, \phi_y^{+t}) \log \frac{p(\phi_y^{+t} | (\phi_x^{-t}, \phi_y^{-t}))}{p(\phi_y^{+t} | \phi_y^{-t})}, \\ &= I(Y_L^{+t}; (X_L^{-t} | Y_L^{-t})), \end{aligned}$$

$$\begin{aligned} CaMI_{X_L \rightarrow Y_L} &= \sum_{\phi_x^{-t}, \phi_y^{-t}, \phi_y^{+t}} p(\phi_x^{-t}, \phi_y^{-t}, \phi_y^{+t}) \log \frac{p(\phi_x^{-t}, \phi_y^{-t}, \phi_y^{+t})}{p(\phi_x^{-t}) p(\phi_y^{-t}, \phi_y^{+t})}, \\ &= I(X_L^{-t}; (Y_L^{+t}, Y_L^{-t})). \end{aligned}$$

Idéia:

1. Fluxo líquido de informação causal
2. Portanto, direção principal entre variáveis causantes e seus efeitos

$$\begin{aligned} DI &= TE_{X_L \rightarrow Y_L} - TE_{Y_L \rightarrow X_L} \\ &= CaMI_{X_L \rightarrow Y_L} - CaMI_{Y_L \rightarrow X_L} \end{aligned}$$

Schreiber, T. (2000)
Bianco-Martinez & Baptista (2018)



>>Entropia de Transferência, CaMI e direcionalidade

PHYSICAL REVIEW E 97, 022303 (2018)

Inference of topology and the nature of synapses, and the flow of information in neuronal networks

F. S. Borges,^{1,2} E. L. Lameu,³ K. C. Iarosz,^{1,4} P. R. Protachevicz,⁵ I. L. Caldas,¹ R. L. Viana,⁶ E. E. N. Macau,^{3,7}
A. M. Batista,^{1,4,5,8} and M. S. Baptista⁴

¹Physics Institute, University of São Paulo, São Paulo, SP 05508-090, Brazil

²Center of Mathematics, Computation and Cognition, Federal University of ABC, São Bernardo do Campo, SP 09606-045, Brazil

³National Institute for Space Research, São José dos Campos, SP 12227-010, Brazil

⁴Institute for Complex Systems and Mathematical Biology, University of Aberdeen, SUPA, AB24 3FX, United Kingdom

⁵Post-Graduation in Science, State University of Ponta Grossa, Ponta Grossa, PR 84030-900, Brazil

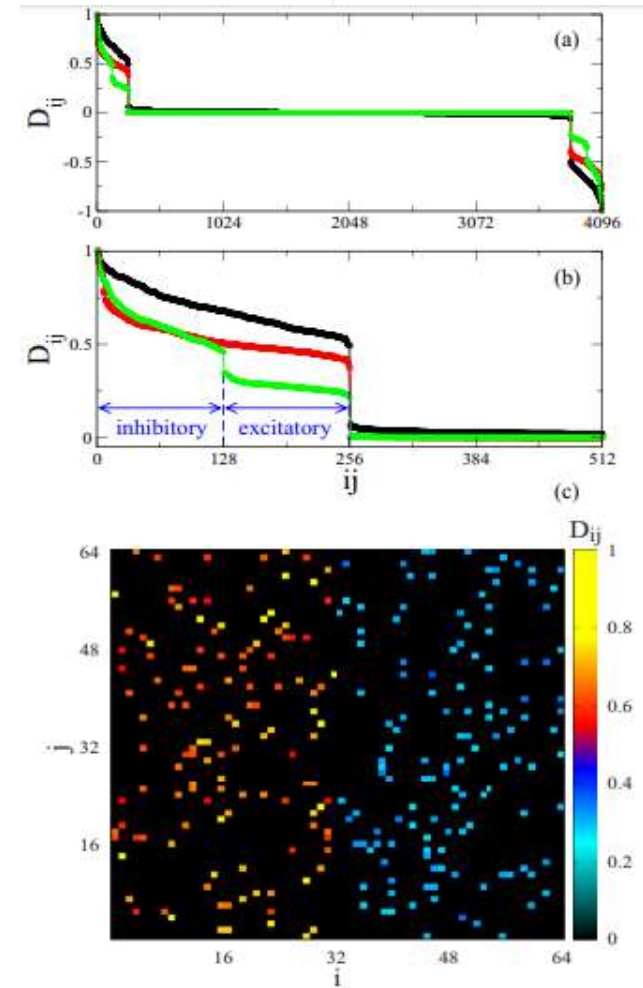
⁶Physics Department, Federal University of Paraná, Curitiba, PR 81531-980, Brazil

⁷Federal University of São Paulo, São José dos Campos, SP 12231-280, Brazil

⁸Mathematics and Statistics Department, State University of Ponta Grossa, Ponta Grossa, PR 84030-900, Brazil

(Received 28 September 2017; revised manuscript received 21 December 2017; published 7 February 2018)

The characterization of neuronal connectivity is one of the most important matters in neuroscience. In this work, we show that a recently proposed informational quantity, the causal mutual information, employed with an appropriate methodology, can be used not only to correctly infer the direction of the underlying physical synapses, but also to identify their excitatory or inhibitory nature, considering easy to handle and measure bivariate time series. The success of our approach relies on a surprising property found in neuronal networks by which nonadjacent neurons do “understand” each other (positive mutual information), however, this exchange of information is not capable of causing effect (zero transfer entropy). Remarkably, inhibitory connections, responsible for enhancing synchronization, transfer more information than excitatory connections, known to enhance entropy in the network. We also demonstrate that our methodology can be used to correctly infer directionality of synapses even in the presence of dynamic and observational Gaussian noise, and is also successful in providing the effective directionality of intermodular connectivity, when only mean fields can be measured.





>>Causality toolbox

- Valencio, A. (2018)

github.com/artvalencio/causality-toolbox

Branch: master | New pull request | Find file | Clone or download

artvalencio Update: simplified output + procedure. Update help | Latest commit 3a0195 on May 17, 2018

README.md	Update README.md	2 years ago
caml.m	updated call and help	last year
camir.m	Update: simplified output + procedure. Update help	last year
mir.m	Update mir.m	2 years ago
totals.m	Add files via upload	2 years ago

causality-toolbox

Functions for calculating information theory measures from time-series, with the purpose of distinguishing the cause and the effect variables on a physical process.

This version is written in Matlab/Octave.

- caml.m: calculates, for bivariate systems, the Causal Mutual Information (CaMI) in both directions, the Mutual Information, the Transfer Entropy in both directions and the Directionality Index
- mir.m: calculates the Mutual Information Rate
- camir.m: calculates, for bivariate systems, the Causal Mutual Information Rate (CaMIR) in both directions, the Mutual Information Rate and the Transfer Entropy Rate in both directions. (Note: If only interested in Mutual Information Rate,

Causality Toolbox

Calcula MI, TE, CaMI, DI
(e também as relativas
taxas conforme aumenta a
seq simbólica)

Multithread Causality

Idem, porém aplicável para quando os dados originais vêm de muitos experimentos curtos ao invés de um único experimento longo

github.com/artvalencio/multithread-causality

README.md	Update README.md	4 months ago
multithreadcaml.m	Update: fully paraleled	4 months ago
serialmultithreadcaml.m	Add files via upload	4 months ago

multithread-causality

Causality toolbox for systems observed through many parallel experiments

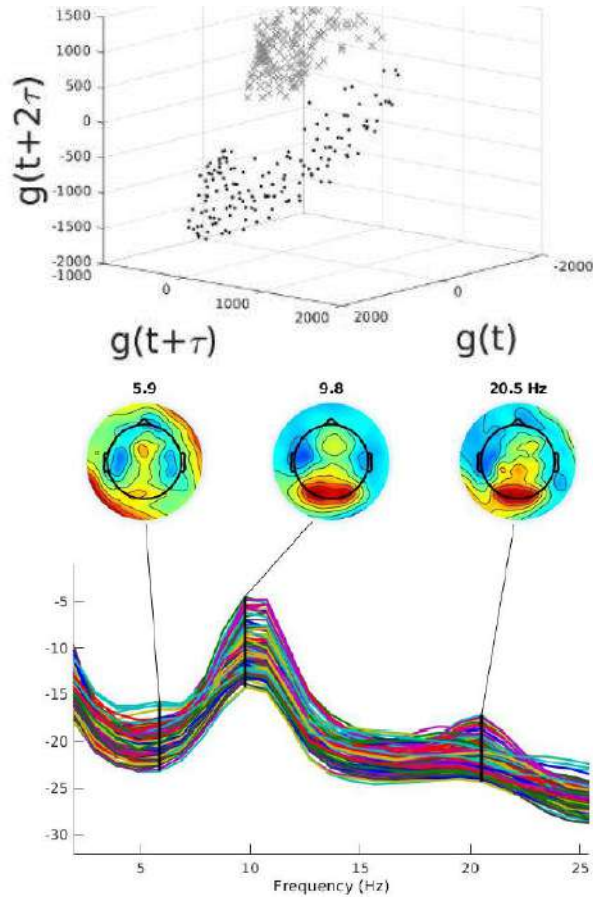
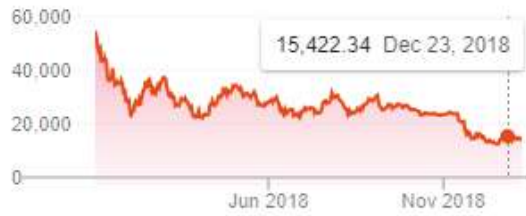
Details of each program and usage example is available by typing in the command window: doc [function-name]

*When to use Multithread Causality?

When you are investigating the dynamics of a phenomena you may: (i) have a very long time-series of the observations of this phenomena, or (ii) have lots of small observations taken independently. The original Causality Toolbox was developed for the first case. The Multithread Causality tool addresses the second. Furthermore, the Multithread Causality is backward compatible, i.e., it can also be used for one single (long) observation of the cause/effect variables.

*What "Serial" Multithread CaMI is supposed to mean?

Oftentimes it is simply faster to perform the computation using serial scheme (i.e. running on one CPU core) instead of parallel (running on all available cores). In our case it is just a matter of using "for" instead of "parfor". Also, it is the way-to-go to Matlab users without the parallel computing toolbox. However, we remain with the principle of computing each small



Instituto de
Computação



NeuroMat



Obrigado!



arthur_valencio@physics.org



<http://www.arthurvalencio.com>
<http://neuromat.numec.prp.usp.br>

A.V. agradece à FAPESP por uma bolsa de pós-doutoramento (#2018/09900-8).
 CEPID NeuroMat agradece apoio FAPESP (#2013/07699-0).

“As opiniões, hipóteses e conclusões ou recomendações expressas neste material são responsabilidade do(s) autor(es) e não necessariamente refletem a visão da Fapesp”