



LISTA DE EXERCÍCIOS (10/10) – Entregar em 25/5/2020

1-Simulação de Sinais Estocásticos com GRNG1.py com N valores de medidas.

1.1. Utilize o algoritmo e gere 10 sinais para cada família com N elementos:

N1: 64; N2:128; N3:256; N4:512; N5:1024; N6: 2048; N7:4096; N8: 8192. **Total do Grupo Noise: 80.**

1.2. Escreva um algoritmo em Python que permita, tendo como entrada

cada um dos sinais acima, obter sua forma normalizada entre 0 e 1, obter o respectivo Histograma e calcular os 4 momentos estatísticos respectivos.

1.3. Organize todos os dados num dataset (instancias x atributos) e tente agrupá-los com a técnica *K-means* para caracterizar, se houver, classes nos espaço de parâmetros composto por variância, skewness e kurtosis.

2- Repita o exercício anterior considerando, entretanto, o algoritmo colorednoise.py .

Neste caso, considere N= 8192 e diversifique os dados em 3 famílias: *white noise*, *pink noise* e *red noise*.

Produza pelo menos 20 sinais para cada família. **Total do Grupo ColorNoise: 60.**

3- Repita o exercício anterior considerando, entretanto, o algoritmo pmodel.py .

Neste caso, considere N= 8192 e diversifique os dados em 2 famílias: endógeno e exógeno.

Para endógeno considere 3 ou mais valores de p na faixa (0.32-0.42) e exógeno considere

3 ou mais valores de p na faixa (0.18-0.28). Produza pelo menos 30 sinais para cada família

(10 para cada valor de p). **Total do Grupo pmnoise: 60.**

4- Espaço de Cullen-Frey e Distribuições de Probabilidades

4.1. Considere duas das séries das famílias N8 dos exercícios anteriores. Classifique as mesmas no espaço de Cullen-Frey.

4.2. Escreva (ou utilize) um algoritmo em Python que permita ajustar uma PDF (Gaussiana ou GEV) aos respectivos histogramas selecionadas em 2.1. Escreva (ou adapte um programa) que permita ler arquivo de uma série genérica com diferentes tamanho

5- Processos Estocásticos Canônicos: Caos Determinístico e Turbulência

5.1. A partir do Mapeamento Logístico e do Mapeamento de Henon gere 2 famílias de series

Temporais com 30 séries em cada uma. Para a família logística varie o parâmetro p na faixa

(3.85 a 4.05). Paragerar a família Henon varie os parâmetros a e b nas respectivas faixas:

(de 1.350 a 1.420) e (0.210-0.310). Por exemplo: pode fixar o $a=1.40$ e variar o b .

Pode fixar o $b=0.300$ e variar o a . Ou variar ambos dentro de um critério de passo ou aleatoriamente.

Aplique as respectivas análises estatísticas dos exercícios 4.1 e 4.2. **Total do Grupo chaosnoise: 60**

5.2. A partir da 2ª lei de Newton construa a equação de Navier-Stokes na sua forma mais simples.

Exercício Bonus (+1 na nota final)

Considere o caso particular da Equação de Burgers Estocástica e gere uma família de Sinais estocásticos a partir da mesma (dica: artigo Flores et al.

<http://www.aims sciences.org/article/doi/10.3934/dcdss.2020224>

6- PSD & DFA:

- 6.1. Considere as séries temporais listadas na tabela `dataset_signal` e obtenha, para cada série, os valores respectivos dos seguintes atributos: S^2 , K , β (via PSD) e α (via DFA). Confira para todas as séries se β (via PSD) está bem ajustado a partir da fórmula WKP: $\beta = 2\alpha - 1$. Construa dois espaços de parâmetros EPSB-K-means: $S^2 \times K \times \beta$ e EDF-K-means: $S^2 \times K \times \alpha$.
- 6.2. Classifique, nos espaços de parâmetros do exercício anterior, as séries temporais: (a) ST-Sol3GHz, (b) ST-surftemp504 e (c) ST-OWS_NDC_Covid1 para um país de sua escolha. Discuta comparativamente as classificações nos dois espaços.
- 6.3. Aplique k-means para todas as séries ST-OWS_NDC_Covid1 considerando os seguintes espaços de atributos: $S^2 \times \alpha$ e $K \times \alpha$. Obtenha os melhores agrupamentos, identifique os grupos e discuta os resultados.

7- Singularity Multifractal Spectra (SMS), também conhecido como MDFA.

- 7.1. Considere o programa `mfdfa.py`. Aprimore o programa para que o mesmo calcule o índice $\Psi = \Delta\alpha/\alpha_{\max}$.
- 7.2. Obtenha o espectro de singularidade para todos os sinais do `dataset_signal` e inclua os valores de Ψ como mais um atributo.
- 7.3. Com base nos valores obtidos em 7.2., repita o exercício 6.3 no espaço $S^2 \times \Psi$.

8- Global Wavelet Spectrum

- 8.1. Utilize o `Waipy` para obter o GWS (Morlet) de todas as ST do exercício 6.2.
- 8.2. Repita 9.1. utilizando uma Db8.

9- Self-Organized Criticality (SOC).

- 10.1. Gere 100 ST (50 exo e 50 endo) com p-model para $N=8192$ e aplique `SOC.py` para as 50 de cada grupo.
- 10.2. Aplique o `SOC.py` para todas as ST do exercício 6.3.

11- EXERCÍCIO BONUS: +1.0 NA NOTA FINAL

Crie um programa em Python para *Preparação Generalizada de Dados (PGD)* e incorpore uma técnica de análise (estatística ou espectral) de sua escolha dentro do paradigma de Data Lake. Teste com alguns dados de seu interesse. Considere apenas os formatos `.txt` e `.csv`. Considere `Pchip` no módulo de normalização e uma RNN LSTM no módulo de previsão. Discuta a emulação deste programa como software embarcado em minicomputador (ex. Arduino ou RaspberryPi-4).

Dataset_signal

Grupo	Tipo	N	Algoritmo Gerador	Tamanho da Familia
noise	ST	2^n n: 6,..., 13	grng1.py	80
colornoise	ST	8192	colorednoise. py	60
pmnoise	ST	8192	pmodel.py	60
chaosnoise	ST	8192	logis.py , henon.py	60