Francisco Paz-Chinchón

NCSA- University of Illinois

A DEMO ON

TIME SERIES

LSST WORKSHOP - JUNE 20 MMXVIII

IN BRIEF...

- What we're going to do? Variability period estimation
- How? Through parametric and non-parametric methods (not autoregressive)
- Additionally, we're going to explore Gaussian Process Regression, and Cross Validation
- Afterwards, play with more data

AUTOCORRRELATION FUNCTION

- Non-parametric estimator
- Works well with sparse data
- Fast and easy to implement
- Stable

$$\rho_k = \frac{\sum_{t=1}^{N-k} (x_t - \bar{x})(x_{t+k} - \bar{x})}{\sum_{t=1}^{N} (x_t - \bar{x})}$$

LOMB SCARGLE

- Generalization of FFT to sparse data (or Least Squares)
- Multiband option
- Fast and stable

$$P_{LS}(\nu) = \frac{1}{2\sigma^2}\left[\frac{(\sum_{i=1}^{N} x_i \cos(2\pi\nu t_i))^2}{\sum_{i=1}^{N} \cos^2(2\pi\nu(t_i - \tau(\nu)))} + \frac{(\sum_{i=1}^{N} x_i \sin(2\pi\nu t_i))^2}{\sum_{i=1}^{N} \sin^2(2\pi\nu(t_i - \tau(\nu)))}\right]$$

$$tan(4\pi\nu\tau) = \frac{\sum_{i=1}^{N} \sin(4\pi\nu t_i)}{\sum_{i=1}^{N} \cos(4\pi\nu t_i)}$$

WAVELET

- Multi-scale analysis
- Good for non-stationary phenomena
- Dependent on mother wavelet
- Convolution with orthonormal basis function

$$W[x](\tau, s) = \frac{1}{\sqrt{s}} \int_{-\infty}^{\infty} x_t \, \Psi^* \left(\frac{t - \tau}{s} \right) dt$$

GAUSSIAN PROCESS REGRESSION

- Useful for construct flexible models
- Used as prior in Bayes
- Maximizes likelihood, maximizes covariance
- Distance-dependent process, easy to setup

CROSS VALIDATION

- How to train a model? (when we don’t have data to give away)
- Useful to choose the right hyperparameters
- Helps detect stable solutions

TEXT

ACCESS JUPYTER-LABS

- Go to: des.ncsa.illinois.edu:31000/easyweb/
- Open “DES JupyterLabs”

- Delete Lab + Deploy Lab (to get the most recent version)
- Go to Lab
- Open a terminal (“+” button) and type

git clone **https://github.com/paztronomer/time_series_demo.git**

TEXT

REQUIREMENTS

- If running locally, first get the Github repo. On a terminal type:
  git clone **https://github.com/paztronomer/time_series_demo.git**

- Then, you should also have the following (5) packages
  numpy, pandas, matplotlib, scipy
  - statsmodels: pip install -U statsmodels ||
    conda install -c conda-forge statsmodels
  - sklearn: pip install -U scikit-learn || conda install scikit-learn
  - gatspy: pip install gatspy
  - pywt: pip install PyWavelets || conda install pywavelets
  - vaex: pip install --pre vaex || conda install -c conda-forge vaex