

Determinação de Pulsção Estelar Através da Entropia de Shannon Condicional.

Gabriel Lauffer Ramos
Orientador: Fabricio Ferrari

Universidade Federal do Rio Grande - FURG
Instituto de Matemática, Estatística e Física - IMEF
Grupo de Astrofísica Teórica e Computacional - GATC

Trabalho de Conclusão de Curso

- Estrelas variáveis são estrelas cuja a magnitude varia com o tempo.
- Determinar o período de variação da magnitude é fundamental para descrever a estrela.
 - Distância (Leavitt and Pickering, 1912).
 - Densidade média (Payne, 1930).
- Porém, dependendo das condições climáticas e da disponibilidade do telescópio, nem sempre temos acesso a dados com uma variação temporal constante.

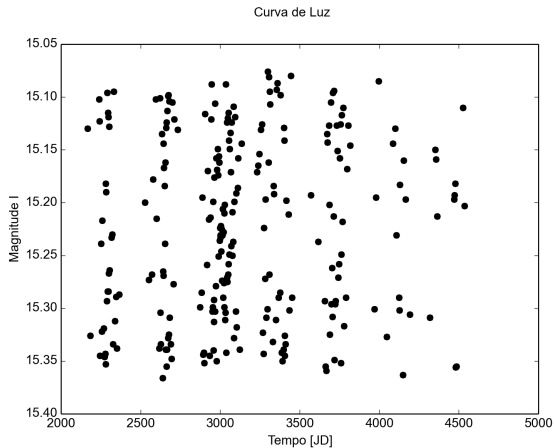


Figura 1 : Exemplo de dados obtidos do Catálogo OGLE-III para uma Cefeida Clássica.

- Existem diversos métodos que lidam com dados igualmente espaçados.
 - Análise de Fourier (Ferraz-Mello, 1981; Lomb, 1976)
- Existem alguns métodos para lidar com dados que não possuam variação temporal constante.
 - Análise do espaço de fase (Schwarzenberg-Czerny, 1989; Cincotta et al., 1995; Cincotta et al., 1999; Graham et al., 2013a)
- Nenhum método se sobressai em relação aos outros (Graham et al., 2013b).

O objetivo deste trabalho é testar um algoritmo que seja confiável para trabalhar com séries temporais astronômicas e que não seja dependente do espaçamento entre os dados observacionais.

- Este algoritmo trabalha com a entropia de Shannon condicional (Cincotta et al., 1999; Graham et al., 2013a), um método que utiliza a dispersão no espaço de fase para obter o período da série temporal através da minimização da entropia

- Estrelas cuja magnitude varia com o tempo
- Classificadas em dois grandes grupos:
 - ***Variáveis Intrínsecas***
 - ***Variáveis Extrínsecas***

- Variação da luminosidade causada por algum processo externo.
 - Variáveis eclipsantes
 - Variáveis rotacionais

O motivo da variação da magnitude está relacionada com processos internos da estrela

- Eruptivas;
- Cataclísmicas;
- Pulsantes.

- Classe de estrelas que sofrem pulsações radiais ou não-radiais ou até mesmo esses dois tipos de pulsação
- Exemplos:
 - Cefeidas;
 - RR Lyrae;
 - Miras;
 - Beta Cepheid;
 - Gamma Doradus;
 - etc.
- Objetos utilizados neste trabalho: *Cefeidas* e *RR Lyraes*.

- Estrelas que pulsam de forma radial;
- Podem possuir períodos entre 1 a 100 dias;
- Massa: $2 M_{\odot}$ a $20 M_{\odot}$
- População I (estrelas jovens, maior metalicidade)
- As cefeidas podem ser subdivididas de acordo com o seu modo de pulsação;
 - FU: modo radial fundamental
 - FO: primeiro sobre tom do modo radial

A grande importância deste tipo de estrela está na descoberta da relação entre período e luminosidade derivada por Henrietta Leavitt (Leavitt and Pickering, 1912) o que possibilitou a determinação de distâncias estelares.

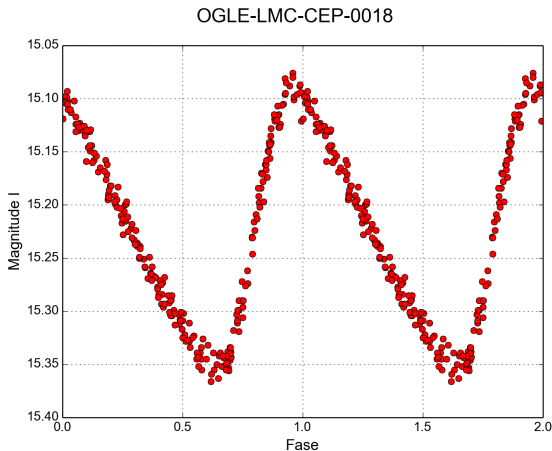


Figura 2 : Curva de luz no espaço de fase para uma Cefeida clássica ($P = 4,0478$).

- Podem possuir períodos entre 0,2 a 1 dia;
- Massa: $0,6 M_{\odot}$ a $0,8 M_{\odot}$
- População II (estrelas antigas, menor metalicidade)
- Utilizadas para determinar distância de sistemas estelares antigos
- Classificas em AB (modo radial fundamental) e C (primeiro sobre-tom)

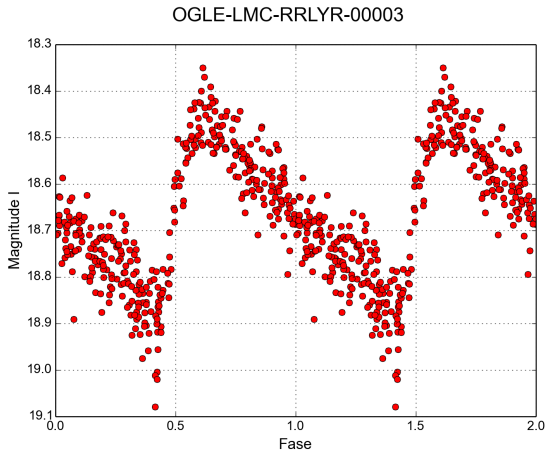


Figura 3 : Curvas de luz no espaço de fase para uma RR Lyrae tipo AB ($P = 0,6565$)

- Estrela com comportamento periódico \rightarrow ciclos iguais
- Ciclo = Fase
- Não importa qual ciclo estamos observando, apenas onde estamos no ciclo

O espaço de fase é uma representação de todos os ciclos observados em apenas uma fase, ou em apenas um ciclo.

Os pontos de sobrepõem e formam uma oscilação geral da estrela. Este espaço de fase é calculado pela seguinte expressão

$$\phi_i = \frac{t_i}{P} - \left[\frac{t_i}{P} \right] \quad (1)$$

- Na teoria de informação, a entropia ou entropia de Shannon (Shannon and Shannon, 1998) é a medida de incerteza de uma variável;
- Essa grandeza mede o grau de desordem de um sinal.

Aplicado para curvas de luz:

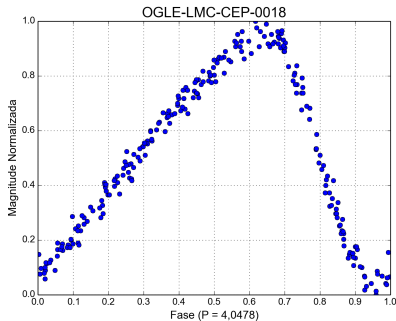
- Curva de luz \rightarrow Sinal periódico;
- Transformando para o espaço de fase
 - Período correto: Espaço de fase possui ordem
 - Período incorreto: Espaço de fase desordenado

Fazendo m repartições no espaço de fase, a entropia é definida por:

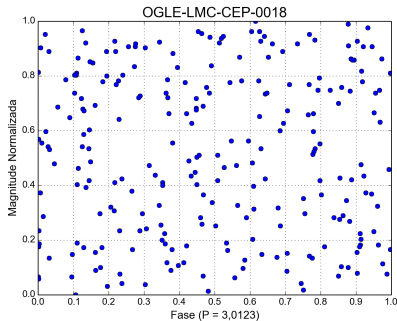
$$H = - \sum_i^m \mu_i \ln \mu_i \quad (2)$$

em que μ_i representa a probabilidade de ocupação da repartição i .

Entropia de Shannon III



(a) Período correto, $H_c = 1,0762$



(b) Período errado, $H_c = 1,5943$

Figura 4 : Exemplos da distribuição de pontos espaço de fase para a Cefeida OGLE-LMC-CEP-0018 do catálogo OGLE.

A entropia de Shannon condicional surgiu da necessidade de contornar um problema bem conhecido da análise de curvas de luz:

- Erro no período $P = 1$ dia

Devido as observações serem efetuadas durante a noite ocasionando um espaçamento de um dia entre os conjuntos de pontos.

Entropia de Shannon Condicional II

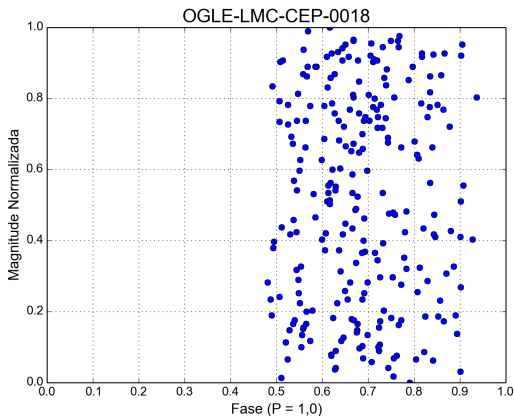


Figura 5 : Efeito de *Aliasing* devido ao período de 1 dia para a Cefeida OGLE-LMC-CEP-0018 do catálogo OGLE. $H_c = 1,5542$.

Entropia de Shannon Condicional III

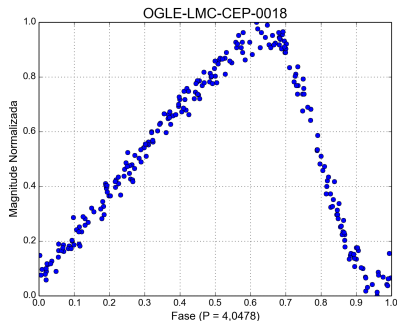
Para lidar com esse problema, Graham et al. (2013a) propuseram a entropia de Shannon condicional.

- O espaço de fase é dividido em i repartições na magnitude e j repartições na fase
- Entropia é calculada da seguinte forma:

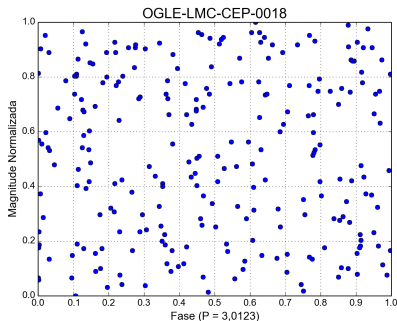
$$H_c = \sum_{i,j} p(m_i, \phi_j) \ln \left(\frac{p(\phi_j)}{p(m_i, \phi_j)} \right) \quad (3)$$

em que $p(m_i, \phi_j)$ é a probabilidade de ocupação na i -ésima repartição da magnitude e na j -ésima repartição da fase e $p(\phi_j)$ é a probabilidade de ocupação na j -ésima repartição da fase.

Entropia de Shannon Condicional IV



(a) Período correto, $H_c = 1,0762$



(b) Período errado, $H_c = 1,5943$

Figura 6 : Exemplos da distribuição de pontos espaço de fase para a Cefeida OGLE-LMC-CEP-0018 do catálogo OGLE.

Foi desenvolvido um algoritmo em Python3 para calcular a entropia condicional. O algoritmo funciona da seguinte maneira:

Para cada estrela:

- Os dados tempo e magnitude são lidos como vetores
- Um vetor com períodos é criado
- O espaço de fase é criado para cada período
- A entropia da dispersão no espaço de fase é calculado
- A menor entropia de todos os períodos é identificada
- O período associado a esta entropia é escolhido como o correto

Algoritmo aplicado para dados do Catálogo OGLE (*The Optical Gravitational Lensing Experiment*) (Udalski et al., 2008)

- 8 anos de dados observacionais;
- Área de 40 graus quadrados na direção das Nuvens de Magalhães;
- Busca por estrelas variáveis;
- Banda I e V
- Os dados são públicos

Foram analisadas um total de 25707 estrelas variáveis do Catálogo OGLE.

Tabela 1 : Quantidade de dados analisados e resultados corretos considerando uma precisão de 10^{-4} .

Estrelas	Quantidade	Acertos	Porcentagem
Cefeidas FU	1818	1817	99,94%
Cefeidas FO	1238	1231	99,43%
RRLyraes AB	17693	17540	99,14%
RRLyraes C	4958	4535	91,47%
Total	25707	25123	97,73 %

A fim de entender como os dados afetam no resultado do método, foi analisado o comportamento da entropia de Shannon para dados com diferentes níveis de ruído e pontos.

- Utilizando os dados das RRLyraes como referência, foram obtidos as grandezas:
 - $t_i = 2152.5019$ HJD,
 - $t_f = 4539.4593$ HJD,
 - quantidade de pontos $n \approx 351$
 - $dt = \frac{t_f - t_i}{n}$
 - Amostragem: $f_s = \frac{1}{dt} = 0.1473$ Hz

Possuindo a amostragem média dos dados observacionais, podemos construir dados sintéticos variando a amostragem e o ruído.

De acordo com Graham et al. (2013a) e Cincotta et al. (1995), para construir dados sintéticos semelhantes com os dados observacionais da maioria dos Surveys de estrelas variáveis, podemos utilizar a seguinte expressão,

$$m(t) = A_0 + \sum_i^3 A_n \sin\left(\frac{2n\pi t}{P}\right) + \varepsilon\eta \quad (4)$$

em que ε é um fator de escala para o ruído entre 0.0 e 1.0, η é uma distribuição gaussiana com média zero e desvio unitário. P é o período médio das RRlyraes que, segundo Soszyński et al. (2009) é de 0.576 dias.

- Foram gerados dados sintéticos variando:
 - Amostragem: variando de 0.25 a 4 com intervalo de 0.25
 - Ruído: ε de 0.0 até 1.0 com intervalo de 0.05
- Total de 300 curvas de luz.

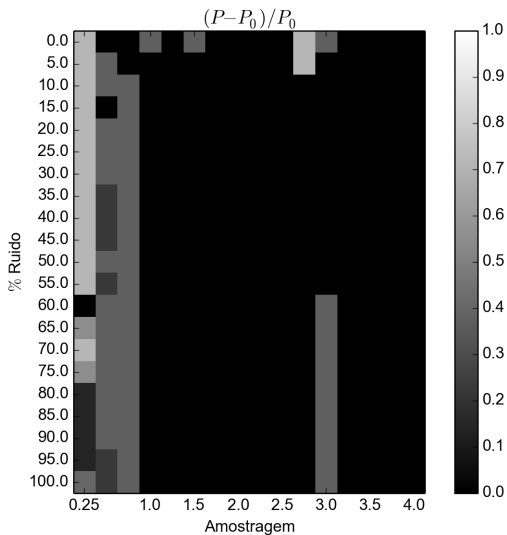


Figura 7 : Resultados obtidos em escala de cinza

A grande importância da determinação de períodos de estrelas variáveis está na possibilidade de calcular distâncias a partir da Lei de Leavitt. Utilizando essa relação,

$$\bar{M}_i = a \log P_i + b \quad (5)$$

e o modulo de distância $\mu = m - M - A_\lambda$ obtemos

$$\bar{m}_i = a \log P_i + b + \mu_i + A_{\lambda_i} \quad (6)$$

é possível calcular as constantes a e b utilizando o método dos mínimos quadrados. A extinção interestelar A_λ é dada pelo mapa de Pejcha and Stanek (2009)

Tabela 2 : Constantes da Relação PL.

Objeto	a	b'
Cefeida FU	-1,292	16,878
Cefeida FO	-1,573	16,558
RR Lyrae AB	-0,721	18,363
RR Lyrae C	-0,022	18,841

Obtendo essas relações, é possível calcular a variação na distância $\Delta\mu_i$ de cada uma das estrelas pelo residual entre a Lei de Leavitt e a magnitude média obtida pelos dados do catálogo, ou seja,

$$\Delta\mu_i = \bar{m}_{OGLE} - \bar{m}_i \quad (7)$$

considerando uma distância média de 50kpc ou $\bar{\mu} = 18,495$ (Pejcha and Stanek, 2009).

Para calcular a distribuição das estrelas é necessário fazer uma transformação de coordenadas, passando de (α, δ, r) para (x, y, z) . As transformações de coordenadas são dadas por Deb and Singh (2014) e mostradas a seguir:

$$\begin{aligned}x &= -r \sin(\alpha - \alpha_0) \cos(\delta) \\y &= r \sin(\delta) \cos(\delta_0) - r \sin(\delta_0) \cos(\alpha - \alpha_0) \cos(\delta) \\z &= r_0 - r \sin(\delta) \sin(\delta_0) - r \cos(\delta_0) \cos(\alpha - \alpha_0) \cos(\delta)\end{aligned} \tag{8}$$

O centro da Grande Nuvem de Magalhães é obtido a partir dos valores médios das coordenadas equatoriais α e δ e possui valores $\alpha_0 = 80,240^\circ$ e $\delta_0 = -69.608^\circ$.

Distribuição das Estrelas na LMC

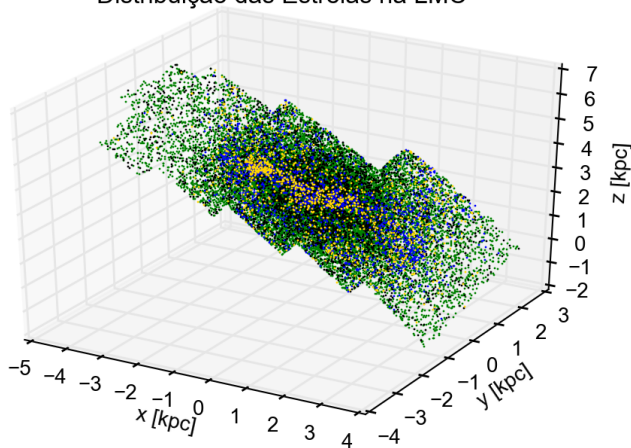


Figura 8 : Distribuição 3D das variáveis pulsantes na LMC.

Entropia de Shannon Condicional

- Técnica simples de ser aplicada;
- Embasada na Teoria da Informação
- Taxa de acerto maior do que 97%

Análise dos dados sintéticos

- O método é confiável para qualquer nível de ruído desde que a frequência de pontos dos dados seja maior do que $f_s = 0,1473$.
- Ferramenta que nos indica como os dados influenciam no resultado do método
- É Possível determinar como a observação nos telescópios devem ser conduzidas

- Cincotta, P. M., Helmi, A., Mendez, M., Nunez, J. a., and Vucetich, H. (1999). Astronomical time-series analysis – II. A search for periodicity using the Shannon entropy. *Monthly Notices of the Royal Astronomical Society*, 302(3):582–586.
- Cincotta, P. M., Mendez, M., and Nunez, J. A. (1995). Astronomical Time Series Analysis. I. A Search for Periodicity Using Information Entropy. *The Astrophysical Journal*, 449:231.
- Deb, S. and Singh, H. P. (2014). Chemical and structural analysis of the Large Magellanic Cloud using the fundamental mode RR Lyrae stars. *Monthly Notices of the Royal Astronomical Society*, 438(3):2440–2455.
- Ferraz-Mello, S. (1981). Estimation of Periods from Unequally Spaced Observations. *The Astronomical Journal*, 86:619.
- Graham, M. J., Drake, A. J., Djorgovski, S. G., Mahabal, A. A., and Donalek, C. (2013a). Using conditional entropy to identify periodicity. *Monthly Notices of the Royal Astronomical Society*, 434:2629–2635.
- Graham, M. J., Drake, A. J., Djorgovski, S. G., Mahabal, A. A., Donalek, C., Duan, V., and Maker, A. (2013b). A comparison of period finding algorithms. *Monthly Notices of the Royal Astronomical Society*, 434:3423–3444.

- Leavitt, H. S. and Pickering, E. C. (1912). Periods of 25 Variable Stars in the Small Magellanic Cloud. *Harvard College Observatory Circular*, 173.
- Lomb, N. R. (1976). Least-squares frequency analysis of unequally spaced data. *Astrophysics & Space Science*, 39:447–462.
- Payne, C. H. (1930). On the Relation of Period to Mean Density for Cepheid Variables. *Harvard College Observatory Bulletin No. 876*.
- Pejcha, O. and Stanek, K. Z. (2009). THE STRUCTURE OF THE LARGE MAGELLANIC CLOUD STELLAR HALO DERIVED USING OGLE-III RR Lyr STARS. *The Astrophysical Journal*, 704(2):1730–1734.
- Schwarzenberg-Czerny, A. (1989). On the advantage of using analysis of variance for period search. *Monthly Notices of the Royal Astronomical Society*, 241:153–165.
- Shannon, C. and Shannon, W. (1998). *Mathematical theory of communication*. University of Illinois Press, first edition edition.
- Soszyński, I., Udalski, A., Szymański, M. K., Kubiak, M., Pietrzyński, G., Wyrzykowski, Ł., Szewczyk, O., Ulaczyk, K., and Poleski, R. (2009). The Optical Gravitational Lensing Experiment. The OGLE-III Catalog of Variable Stars. III. RR Lyrae Stars in the Large Magellanic Cloud. *Acta Astronomica*, 59:1–18.
- Udalski, A., Szymanski, M. K., Soszynski, I., and Poleski, R. (2008). The Optical Gravitational Lensing Experiment. Final Reductions of the OGLE-III Data. page 13.

Obrigado!

Perguntas?