Periodograms for Multiband Time Series

Jake VanderPlas
#LGAtstats; June 3rd 2015
The Bomb-Scargle Periodogram!!!
Mapping the MW with RR Lyrae

SDSS II Stripe 82:
- 483 RR Lyrae to r\~22
- 300 deg$^2$
- d \~ 100 kpc

Supports the idea of an early-forming smooth inner halo, and late-forming accreted outer halo.

Sesar et al. 2010
RR Lyrae in LSST

SDSS II
300 deg$^2$
r ~ 22 mags
d ~ 100 kpc
483 RR Lyrae

LSST
~20,000 deg$^2$
r ~ 24 mags
d ~ 300 kpc
> 30,000 RR Lyrae?

? ? ? ?
Potential for vastly increased understanding of structure of our halo & detailed constraints on models of Milky Way formation history!
Naive single-band approach for LSST:
- 1 year: 50% completeness at $g \sim 22$
- 10 years: 50% completeness at $g \sim 24.5$
Can we do better?
The issue:
For LSST-style data (1 band each visit), single band approaches fail!

We need a method that utilizes all bands at once.
Let’s think about a periodogram which can utilize multiple bands simultaneously . . .
The Lomb-Scargle Periodogram

If you’ve ever come across the Lomb-Scargle Periodogram, you’ve probably seen something like this...

\[
P_N(\omega) = \frac{1}{2 V_y} \left[ \frac{\sum_k (y_k - \bar{y}) \cos \omega(t_k - \tau)}{\sum_k \cos^2 \omega(t_k - \tau)} \right]^2 + \frac{\sum_k (y_k - \bar{y}) \sin \omega(t_k - \tau)}{\sum_k \sin^2 \omega(t_k - \tau)} \right]^2
\]

But this obfuscates the beauty of the algorithm: the classical periodogram is essentially the \( \chi^2 \) of a single sinusoidal model-fit to the data:

\[
y(t|\omega, \theta) = \theta_1 \sin(\omega t) + \theta_2 \cos(\omega t).
\]

\[
\chi_{min}^2(\omega) = \chi_0^2[1 - P_N(\omega)]
\]
Standard Lomb-Scargle

Periodogram peaks are frequencies where a sinusoid fits the data well:

cf. Lomb (1976), Scargle (1982)
Connection between Fourier periodogram and least squares allows us begin generalizing the periodogram . . .
Floating Mean Model

... in which we simultaneously fit the mean

\[ y(t \mid \omega, \theta) = \theta_0 + \theta_1 \sin \omega t + \theta_2 \cos \omega t \]

Figure: VanderPlas & Ivezic 2015

Truncated Fourier Model

... in which we fit for higher-order periodicity

\[ y(t | \omega, \theta) = \theta_0 + \sum_{n=1}^{N} [\theta_{2n-1} \sin(n\omega t) + \theta_{2n} \cos(n\omega t)] \]
Regularized Model

... in which we penalize coefficients to simplify an overly-complex model.

The “trick” is adding a strong prior which pushes coefficients to zero: higher terms are only used if actually needed!
Putting it all together: The Multiband Periodogram

\[
y_k(t|\omega, \theta) = \theta_0 + \sum_{n=1}^{N_{\text{base}}} \left[ \theta_{2n-1} \sin(n\omega t) + \theta_{2n} \cos(n\omega t) \right] + \\
\theta_0^{(k)} + \sum_{n=1}^{N_{\text{band}}} \left[ \theta_{2n-1}^{(k)} \sin(n\omega t) + \theta_{2n}^{(k)} \cos(n\omega t) \right]. \quad (18)
\]
Putting it all together: The Multiband Periodogram

\[ y_k(t|\omega, \theta) = \theta_0 + \sum_{n=1}^{N_{\text{base}}} [\theta_{2n-1} \sin(n\omega t) + \theta_{2n} \cos(n\omega t)] + \theta_0^{(k)} + \sum_{n=1}^{N_{\text{band}}} \left[ \theta_{2n-1}^{(k)} \sin(n\omega t) + \theta_{2n}^{(k)} \cos(n\omega t) \right]. \] (18)

- define a truncated Fourier base component which contributes equally to all bands.
Putting it all together:
The Multiband Periodogram

\[ y_k(t | \omega, \theta) = \theta_0 + \sum_{n=1}^{N_{\text{base}}} \left[ \theta_{2n-1} \sin(n\omega t) + \theta_{2n} \cos(n\omega t) \right] + \]

\[ \theta^{(k)}_0 + \sum_{n=1}^{N_{\text{band}}} \left[ \theta^{(k)}_{2n-1} \sin(n\omega t) + \theta^{(k)}_{2n} \cos(n\omega t) \right]. \quad (18) \]

- for each band, add a truncated Fourier band component to describe deviation from base model
Putting it all together:
The Multiband Periodogram

- **Regularize** the band component to drive common variation to the base model.
Putting it all together: The Multiband Periodogram

- **Regularize** the band component to drive common variation to the base model.

Key: Regularization reduces added model complexity & pushes common variation into the base model.
Multiband Periodogram on sparse, LSST-style data . . .

Detects period with high significance when single-band approaches fail!
The Money Plot: Prospects for LSST

Based on simulated LSST cadence & photometric errors; see VanderPlas & Ivezic (2015, in prep)
The Money Plot: Prospects for LSST

Based on simulated LSST cadence & photometric errors; see VanderPlas & Ivezic (2015, in prep)

Fraction Recovered

- 6 months
- 1 year
- 2 years
- 5 years

g-band mag

- multiband model
- Oluseyi (2012) approach

e.g. after 2 years: ~0% → ~75% completeness at survey limit!
The Money Plot: Prospects for LSST

Based on simulated LSST cadence & photometric errors; see VanderPlas & Ivezic (2015, in prep)

~2 mag improvement in effective depth of LSST!
Code to reproduce the study & figures:

http://github.com/jakevdp/multiband_LS/

Python periodogram implementation:

http://github.com/jakevdp/gatspy/

“If it’s not reproducible, it’s not science.”
~ Thank You! ~

Email: jakevdp@uw.edu

Twitter: @jakevdp

Github: jakevdp

Web: http://vanderplas.com/

Blog: http://jakevdp.github.io/
Extras ...
Oluseyi 2012 Simulated LSST Measurements:

Faintest RR-Lyrae: Pessimistic period recovery even with 5-10 years of LSST data!