# Time Series Analysis

Eric Feigelson
Astrostatistics Summer School

# Outline

# Time series in astronomy

- Periodic phenomena: binary orbits (stars, extrasolar planets); stellar rotation (radio pulsars); pulsation (helioseismology, Cepheids)
- Stochastic phenomena: accretion (CVs, X-ray binaries, Seyfert gals, quasars); scintillation (interplanetary & interstellar media); jet variations (blazars)
- Explosive phenomena: thermonuclear (novae, X-ray bursts), magnetic reconnection (solar/stellar flares), star death (supernovae, gamma-ray bursts)

# Difficulties in astronomical time series

Gapped data streams:

Diurnal & monthly cycles; satellite orbital cycles; telescope allocations

Heteroscedastic measurement errors:

Signal-to-noise ratio differs from point to point

Poisson processes:

Individual photon/particle events in high-energy astronomy

# Important Fourier Functions

Discrete Fourier Transform

$$d(\omega_j) = n^{-1/2} \sum_{t=1}^{n} x_t exp(-2\pi i t \omega_j)$$

$$d(\omega_j) = n^{-1/2} \sum_{t=1}^{n} x_t cos(2\pi i \omega_j t) - i n^{-1/2} \sum_{t=1}^{n} x_t sin(2\pi i \omega_j t)$$

Classical (Schuster) Periodogram

$$I(\omega_j) = |d(\omega_j)|^2$$

Spectral Density

$$f(\omega) = \sum_{h=-\infty}^{h=\infty} exp(-2\pi i \omega h) \gamma(h)$$

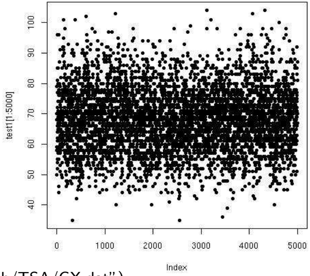
Fourier analysis reveals nothing of the evolution in time, but rather reveals the variance of the signal at different frequencies.

It can be proved that the classical periodogram is an estimator of the spectral density, the Fourier transform of the autocovariance function.

Formally, the probability of a periodic signal in Gaussian noise is $P \propto e^{d(\omega_j)/\sigma^2}$.

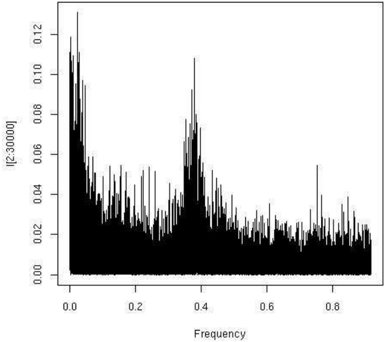
# Ginga observations of X-ray binary GX 5-1

GX 5-1 is a binary star system with gas from a normal companion accreting onto a neutron star. Highly variable X-rays are produced in the inner accretion disk. XRB time series often show 'red noise' and 'quasi-periodic oscillations', probably from inhomogeneities in the disk. We plot below the first 5000 of 65,536 count rates from Ginga satellite observations during the 1980s.




```
gx=scan("~/Desktop/CASt/SumSch/TSA/GX.dat")
t=1:5000
plot(t,gx[1:5000],pch=20)
```

Fast Fourier Transform of the GX 5-1 time series reveals the 'red noise' (high spectral amplitude at small frequencies), the QPO (broadened spectral peak around 0.35), and white noise.




f = 0:32768/65536
I = (4/65536)*abs(fft(gx)/sqrt(65536))^ 2
plot(f[2:60000],I[2:60000],type="l",xlab="Frequency")

# Limitations of the spectral density

But the classical periodogram is not a good estimator! E.g. it is formally 'inconsistent' because the number of parameters grows with the number of datapoints. The discrete Fourier transform and its probabilities also depends on several strong assumptions which are rarely achieved in real astronomical data: evenly spaced data of infinite duration with a high sampling rate (Nyquist frequency), Gaussian noise, single frequency periodicity with sinusoidal shape and stationary behavior. Formal statement of strict stationarity: $P\{x_{t_1} \leq c_1, ...s_{x_K} \leq c_k\} = P\{x_{t_{1+h}} \leq c+1, ..., x_{t_{k+h}} \leq c_k\}$.

Each of these constraints is violated in various astronomical problems. Data spacing may be affected by daily/monthly/orbital cycles. Period may be comparable to the sampling time. Noise may be Poissonian or quasi-Gaussian with heavy tails. Several periods may be present (e.g. helioseismology). Shape may be non-sinusoidal (e.g. elliptical orbits, eclipses). Periods may not be constant (e.g. QPOs in an accretion disk).

# Improving the spectral density I

The estimator can be improved with **smoothing**,

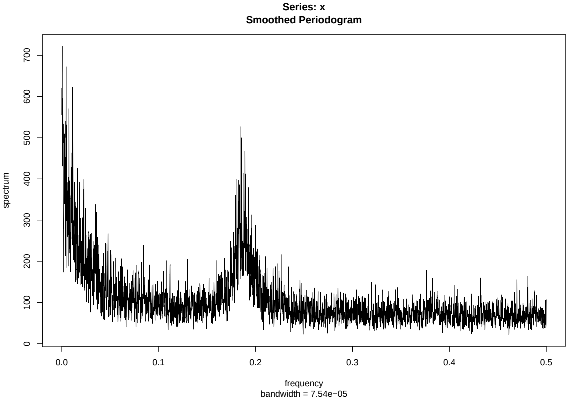
$$\hat{f}(\omega_j) = \frac{1}{2m_1} \sum_{k=-m}^{m} I(\omega_{j-k}).$$

This reduces variance but introduces bias. It is not obvious how to choose the smoothing bandwidth $m$ or the smoothing function (e.g. Daniell or boxcar kernel).

**Tapering** reduces the signal amplitude at the ends of the dataset to alleviate the bias due to leakage between frequencies in the spectral density produced by the finite length of the dataset. Consider for example the cosine taper

$$h_t = 0.5[1 + cos(2\pi(t - \bar{t})/n)]$$

applied as a weight to the initial and terminal $n$ datapoints. The Fourier transform of the taper function is known as the spectral window. Other widely used options include the Fejer and Parzen windows and multitapering. Tapering decreases bias but increases




```
postscript(file="~/Desktop/GX_sm_tap_fft.eps")
k = kernel("modified.daniell", c(7,7))
spec = spectrum(gx, k, method="pgram", taper=0.3, fast=TRUE, detrend=TRUE, log="no")
dev.off()
```

# Improving the spectral density II

**Pre-whitening** is another bias reduction technique based on removing (filtering) strong signals from the dataset. It is widely used in radio astronomy imaging where it is known as the CLEAN algorithm, and has been adapted to astronomical time series (Roberts et al. 1987).

A variety of **linear filters** can be applied to the time domain data prior to spectral analysis. When aperiodic long-term trends are present, they can be removed by spline fitting (high-pass filter). A kernel smoother, such as the moving average, will reduce the high-frequency noise (low-pass filter). Use of a parametric autoregressive model instead of a nonparametric smoother allows likelihood-based model selection (e.g. BIC).

# Improving the spectral density III

Harmonic analysis of unevenly spaced data is problematic due to the loss of information and increase in aliasing.

The **Lomb-Scargle periodogram** is widely used in astronomy to alleviate aliasing from unevenly spaced:

$$d_{LS}(\omega) = \frac{1}{2}\left(\frac{[\sum_{t=1}^{n} x_t cos\omega(x_t - \tau)]^2}{\sum_{i=1}^{n} cos^2\omega(x_t - \tau)} + \frac{[\sum_{t=1}^{n} x_t sin\omega(x_t - \tau)]^2}{\sum_{i=1}^{n} sin^2\omega(x_t - \tau)}\right)$$

where $\quad tan(2\omega\tau) = (\sum_{i=1}^{n} sin2\omega x_t)(\sum_{i=1}^{n} cos2\omega x_t)^{-1}$

$d_{LS}$ reduces to the classical periodogram $d$ for evenly-spaced data. Bretthorst (2003) demonstrates that the Lomb-Scargle periodogram is the unique sufficient statistic for a single stationary sinusoidal signal in Gaussian noise based on Bayes theorem assuming simple priors.

# Some other methods for periodicity searching

**Phase dispersion measure** (Stellingwerf 1972) Data are folded modulo many periods, grouped into phase bins, and intra-bin variance is compared to inter-bin variance using $\chi^2$. Non-parametric procedure well-adapted to unevenly spaced data and non-sinusoidal shapes (e.g. eclipses). Very widely used in variable star research, although there is difficulty in deciding which periods to search (Collura et al. 1987).

**Minimum string length** (Dworetsky 1983) Similar to PDM but simpler: plots length of string connecting datapoints for each period. Related to the Durbin-Watson roughness statistic in econometrics.

**Rayleigh** and $Z_n^2$ tests (Leahy et al. 1983) for periodicity search Poisson distributed photon arrival events. Equivalent to Fourier spectrum at high count rates.

**Bayesian periodicity search** (Gregory & Loredo 1992) Designed for non-sinusoidal periodic shapes observed with Poisson events. Calculates odds ratio for periodic over constant model and most probable shape.

# Conclusions on spectral analysis

For challenging problems, smoothing, multitapering, linear filtering, (repeated) pre-whitening and Lomb-Scargle can be used together. Beware that aperiodic but autoregressive processes produce peaks in the spectral densities. Harmonic analysis is a complicated 'art' rather than a straightforward 'procedure'.

It is extremely difficult to derive the significance of a weak periodicity from harmonic analysis. Do not believe analytical estimates (e.g. exponential probability), as they rarely apply to real data. It is essential to make simulations, typically permuting or bootstrapping the data keeping the observing times fixed. Simulations of the final model with the observation times is also advised.

# Nonstationary time series

Non-stationary periodic behaviors can be studied using **time-frequency Fourier analysis**. Here the spectral density is calculated in time bins and displayed in a 3-dimensional plot.

**Wavelets** are now well-developed for non-stationary time series, either periodic or aperiodic. Here the data are transformed using a family of non-sinusoidal orthogonal basis functions with flexibility both in amplitude and temporal scale. The resulting wavelet decomposition is a 3-dimensional plot showing the amplitude of the signal at each scale at each time. Wavelet analysis is often very useful for noise threshholding and low-pass filtering.

# Nonparametric time domain methods

**Autocorrelation function**

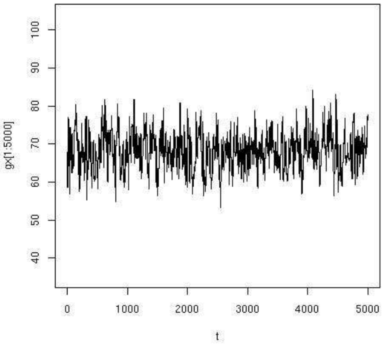
$$\hat{\rho}(h) = \frac{\hat{\gamma}(h)}{\hat{\gamma}0} \;\; \text{where}$$

$$\hat{\gamma}(h) = \frac{\sum_{i=1}^{n-h}(x_{t+h} - \bar{x})(x_t - \bar{x})}{n}$$

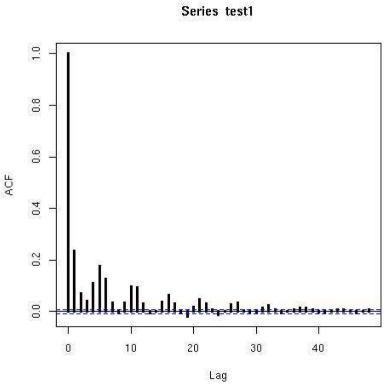
This sample ACF is an estimator of the correlation between the $x_t$ and $x_{t-h}$ in an evenly-spaced time series lags. For zero mean, the ACF variance is $Var\hat{\rho} = [n-h]/[n(n+2)]$.

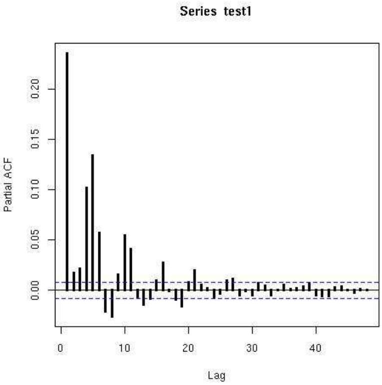
The partial autocorrelation function (PACF) estimates the correlation with the linear effect of the intermediate observations, $x_{t-1}, ..., x_{t-h+1}$, removed. Calculate with the Durbin-Levinson algorithm based on an autoregressive model.

Kernel smoothing of GX 5+1 time series
Normal kernel, bandwidth = 7 bins



# Autocorrelation functions




acf(GX, lwd=3)

pacf(GX, lwd=3)

# Parametric time domain methods: ARMA models

**Autoregressive moving average model**

Very common model in human and engineering sciences, designed for aperiodic autocorrelated time series (e.g. 1/f-type 'red noise'). Easily fit by maximum-likelihood. Disadvantage: parameter values are difficult to interpret physically.

$$\mathbf{AR(p)\ model}\quad x_t = \phi_1 x_{t-1} + \phi_2 x_{t-2} + \ldots + \phi_p x_{t-p} + w_t$$

$$\mathbf{MA(q)\ model}\quad x_t = w_t + \theta_1 w_{t-1} + \ldots + \theta_q w_{t-q}$$

The AR model is recursive with memory of past values. The MA model is the moving average across a window of size $q+1$. ARMA(p,q) combines these two characteristics.

# State space models

Often we cannot directly detect $x_t$, the system variable, but rather indirectly with an observed variable $y_t$. This commonly occurs in astronomy where $y$ is observed with measurement error (errors-in-variable or EIV model). For AR(1) and errors $v_t = N(\mu, \sigma)$ and $w_t = N(\nu, \tau)$,

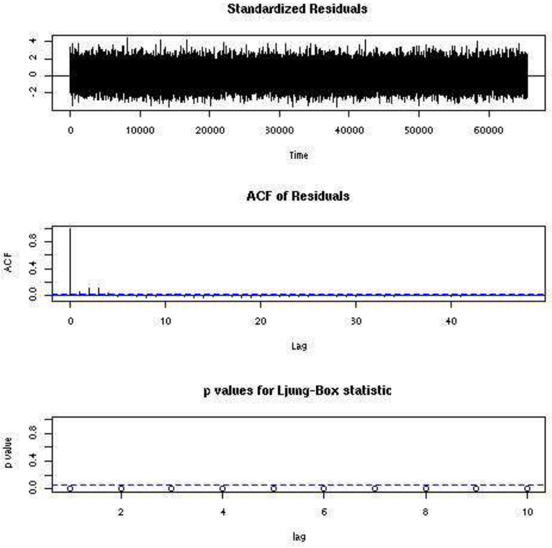
$$y_t = Ax_t + v_t \qquad x_t = \phi_1 x_{t-1} + w_t$$

This is a state space model where the goal is to estimate $x_t$ from $y_t$, $p(x_t | y_t, \ldots, y_1)$. Parameters are estimated by maximum likelihood, Bayesian estimation, Kalman filtering, or prediction.

# GX 5+1 modeling

ar(x = GX, method = "mle")
Coefficients:
1 2 3 4 5 6 7 8
0.21 0.01 0.00 0.07 0.11 0.05 -0.02 -0.03

arima(x = GX, order = c(6, 2, 2))
Coefficients:
ar1 ar2 ar3 ar4 ar5 ar6 ma1 ma2
0.12 -0.13 -0.13 0.01 0.09 0.03 -1.93 0.93
Coeff s.e. = 0.004 $\sigma^2 = 102$ log L = -244446.5
AIC = 488911.1




Although the scatter is reduced by a factor of 30, the chosen model is not adequate: Ljung-Box test shows significant correlation in the residuals. Use AIC for model selection.

# Other time domain models

- Extended ARMA models: VAR (vector autoregressive), ARFIMA (ARIMA with long-memory component), GARCH (generalized autoregressive conditional heteroscedastic for stochastic volatility)
- Extended state space models: non-stationarity, hidden Markov chains, etc. MCMC evaluation of nonlinear and non-normal (e.g. Poisson) models

# Statistical texts and monographs

D. Brillinger, Time Series: Data Analysis and Theory, 2001
C. Chatfield, The Analysis of Time Series: An Introduction, 6th ed., 2003
G. Kitagawa & W. Gersch, Smoothness Priors Analysis of Time Series, 1996
J. F. C. Kingman, Poisson Processes, 1993
J. K. Lindsey, Statistical Analysis of Stochastic Processes in Time, 2004
S. Mallat, A Wavelet Tour of Signal Processing, 2nd ed, 1999
M. B. Priestley, Spectral Analysis and Time Series, 2 vol, 1981
R. H. Shumway and D. S. Stoffer, Time Series Analysis and Its Applications (with R examples), 2nd Ed., 2006

# Astronomical references


Bretthorst 2003, "Frequency estimation and generalized Lomb-Scargle periodograms", in Statistical Challenges in Modern Astronomy

Collura et al 1987, "Variability analysis in low count rate sources", ApJ 315, 340

Dworetsky 1983, "A period-finding method for sparse randomly spaced observations ....", MNRAS 203, 917

Gregory & Loredo 1992, "A new method for the detection of a periodic signal of unknown shape and period", ApJ 398, 146

Kovacs et al. 2002, "A box-fitting algorithm in the search for periodic transits", A&A 391, 369

Leahy et al. 1983, "On searches for periodic pulsed emission: The Rayleigh test compared to epoch folding",. ApJ 272, 256

Roberts et al. 1987, "Time series analysis with CLEAN ...", AJ 93, 968

Scargle 1982, "Studies in astronomical time series, II. Statistical aspects of spectral analysis of unevenly spaced data", ApJ 263, 835

Scargle 1998, "Studies in astronomical time series, V. Bayesian Blocks, a new method to analyze structure in photon counting data, ApJ 504, 405

Stellingwerf 1972, "Period determination using phase dispersion measure",
ApJ 224, 953
Vio et al. 2005, "Time series analysis in astronomy: Limits and potentialities,
A&A 435, 773