Power law estimation

T. Dudok de Wit (Univ. Orléans)
Why look for power laws?

Basic property: scale invariance

\[ y(\lambda t) \sim \lambda^k y(t) \]

In the spectral domain

\[ \tilde{y}(f) \propto f^{-\beta} \]
Why look for power laws?

- Lots of fascinating properties
  - scale-invariance (no characteristic scale)
  - large deviations (rare events)
  - turbulence (intermittency, energy cascades, ...)

- Self-similarity is ubiquitous in plasmas
  - solar flare energy distribution
  - plasma turbulence
  - bursty bulk flows in the magnetosphere
  - photospheric magnetic field
  - ...
Occurrence frequency distributions of hard X-ray peak count rates from 3 instruments.

\[ \alpha_p = 1.73 \pm 0.07 \]
Cosmic ray spectrum between 1 GeV and 1 ZeV

Fluxes of Cosmic Rays

(1 particle per m²-second)

Knee
(1 particle per m²-year)

Ankle
(1 particle per km²-year)

(Swordy - U.Chicago)

Aschwanden et al., 2014
Power spectral density of the magnetic field in solar wind turbulence

\[
\text{Power spectral density: } nT^2/\text{Hz}
\]

\[
\text{frequency (Hz)}
\]

\[
\begin{align*}
\text{Sahraoui et al., 2009}
\end{align*}
\]
Why look for power laws?

- The power law itself is totally uninteresting

- What physical process gave rise to it?
  - where does the scaling start?
  - where does the scaling end?
  - what is its slope?
Example: solar wind magnetic field recorded by Anik F1R
Solar wind observations

- Anik F1R (2005-today): Canadian geostationary satellite
  - experimental magnetic field sensor (fluxgate, 1 component)
  - high cadence: < 200 samples/second
  - some observations made in the solar wind
Magnetic field observations

Raw data (essentially $B_y$)
Scalogram

Modulus of continuous wavelet transform (Morlet wavelets)

scalogram  n=10000  n_sca=120  a_min=4  a_max=2500

amplitude

scale

log_{10} (ampl)

time [sampling periods]
To identify intermittent bursts of activity: estimate the local intermittency measure [Farge, 1992; 1996]

\[ B(t) \rightarrow \mathcal{B}(a, \tau) \rightarrow \text{LIM}(a, \tau) = \frac{|\mathcal{B}(a, \tau)|^2}{\langle |\mathcal{B}(a, \tau')|^2 \rangle_{\tau'}} \]

- Large LIM = local concentration of magnetic energy at scale a and time \( \tau \)
Scalogram

Local intermittency measure

![Scalogram](image)

- Local intermittency measure
- $n=10000$  $ns=120$
Clear evidence for a double power law, with a break

-2.574 ± 0.039

-3.895 ± 0.026

PSD [nT²/Hz]

frequency [Hz]
All of this is just a dream
Disclaimer

- My record was generated by a simple stochastic (autoregressive) process:
  no intermittency or self-similarity whatsoever!

- The true power spectral density is a log-normal one:
  no power laws here!
The true power spectral density

\[ \tilde{y}(f) = P_0 \exp\left(-\frac{(f - f_0)^2}{\Delta f^2}\right) \]
Intermediate conclusion
Intermediate conclusion

We are completely fooled by our frenzy to detect straight lines
The problem

1) Estimate the power spectral density
\[ |\tilde{y}(f)|^2 \propto f^{-\alpha} \]

need a good estimator of the slope \( \alpha \)

2) Estimate the slope (spectral index) \( \alpha \)

test against a null hypothesis (« no power law model »)

3) Prove that the power law model is indeed appropriate
2 cases occur in practice

\begin{align*}
\text{case 1 : } \textbf{data} &= \text{series of events}: \\
&\text{their } \textbf{probability density function} \text{ follows a power law} \\
&\text{(e.g. solar flare distribution)}
\end{align*}

\begin{align*}
\text{case 2 : } \textbf{data} &= \text{time series}: \\
&\text{their } \textbf{power spectral density} \text{ follows a power law} \\
&\text{(e.g. E or B field record)}
\end{align*}
Case 1
Example: number of Coronal Mass Ejections (CMEs) per apparent angular width [Robbrecht et al., ApJ 2009]
Case 1: the data consist of events

- Classical approach
  - bin events to get a histogram (what bin size?)
  - fit by least-squares a line to log(# of events) vs log(width)

This approach is full of flaws and will give biased and noisy estimates. Avoid it!
Case 1: the data consist of events

- A much better estimator (unbiased, and low variance)
  [Clauset et al., SIAM review 2006]

- select the range over which a power law will be fitted
  \[ y_{min} \leq y_i \leq y_{max} \]

- use the maximum likelihood estimator of the slope
  \[ \hat{\alpha} = 1 + n \left( \sum_{i=1}^{n} \log \frac{y_i}{y_{min}} \right)^{-1} \]

- whose standard error is
  \[ \sigma_{\hat{\alpha}} = \frac{\hat{\alpha} - 1}{\sqrt{n}} + \mathcal{O}(1/n) \]
Case 1: the data consist of events

Where does the lower bound \( (y_{\text{min}}) \) occur?

![Graph showing the apparent width distribution of CMEs with a power law fit. The text box indicates the strategy of selecting various lower bounds, fitting power laws, and determining which one matches the data best.]

Underestimating the lower bound is far worse than overestimating it.

Strategy: select various lower bounds, fit power laws, and determine which one matches the data best.
Case 2: time series

- 2 important steps for time series
  1) Estimate the spectral content of the time series
  2) Fit power law to the power spectral density
Case 2: time series

- Classical Fourier analysis is not optimal: a time series with self-similarity should be projected on self-similar basis functions (≠ sines)

- Discrete wavelet transform is more appropriate: self-similar by construction

$$\tilde{y}_{j,k} = \int_{\mathbb{R}} y(t) \psi_{j,k}(t) \, dt, \quad j, k \in \mathbb{Z}$$

where

$$\psi_{j,k}(t) = \frac{1}{2^{j/2}} \psi \left( \frac{t}{2^j} - k \right), \quad j, k \in \mathbb{Z}$$

$\psi(\ )$ is the mother wavelet
Case 2: time series

- Lots of different mother wavelets around

- The number of vanishing moments is a crucial parameter

Haar wavelet (db1)
1 vanishing moment

Daubechies db4 wavelet
4 vanishing moments
**Case 2: time series**

<table>
<thead>
<tr>
<th>Fourier</th>
<th>Discrete wavelet transform</th>
</tr>
</thead>
<tbody>
<tr>
<td>Compute transform</td>
<td><strong>Compute transform</strong></td>
</tr>
<tr>
<td>$\tilde{y}(f) = \int y(t) e^{-j2\pi ft} , dt$</td>
<td>$\tilde{y}<em>{j,k} = \int</em>{\mathbb{R}} y(t) \psi_{j,k}(t) , dt$</td>
</tr>
<tr>
<td>Power spectral density</td>
<td>Power spectral density</td>
</tr>
<tr>
<td>$S(f) =</td>
<td>\tilde{y}(f)</td>
</tr>
<tr>
<td>• can be estimated at all frequencies</td>
<td>• can be estimated per octave only (scale$_j = 1, 2, 4, 8, ...$)</td>
</tr>
<tr>
<td>• nearby values of $S(f)$ are not independent</td>
<td>• wavelet coefficients at nearby scales/positions are independent (very useful)</td>
</tr>
</tbody>
</table>
Case 2: time series

Example

- Fourier - periodogram
- Wavelet transform
- model
Why the order of the wavelet is important

If the mother wavelet does not have enough vanishing moments $n$

$$\int t^n \psi(t) \, dt = 0$$

- The wavelet coefficients $(|\tilde{y}_{j,k}|^2)$ are not independent anymore

- The estimate of the spectral index will be biased
Why the order of the wavelet is important

For estimating a power law $S(f) \sim f^{-a}$, the number of vanishing moments must be $n > (a-1)/2$
Conclusion
(not a fake one)
1. almost any gentle curve will look like a line on a log-log plot
2. plotting scalograms, local intermittency measures, power laws, etc. without proper statistical testing is nonsense.
Conclusions (2/2)

For sampled data
- Use maximum likelihood estimator
- Proper estimation of upper and lower (ymin) bounds is crucial

For time series
- Use discrete wavelet transform
- Ensure that the order of the wavelet is high enough (stay away from Haar)
- Power law should cover at least a decade to make sense.

Not addressed here:
- Use statistical tests (e.g. Kolmogorov-Smirnov) to compare power law model against other plausible models.
- Or better, use a bayesian approach (computationally expensive)
Further reading

