TSDA Lecture 5

For more notes on the discrete FT, see the pdf titled `05_discrete-fourier-transform.pdf` in Lecture notes

For all work on Fourier series, see `vanderplas_lombscargle.pdf` in Reading.

In this lecture:

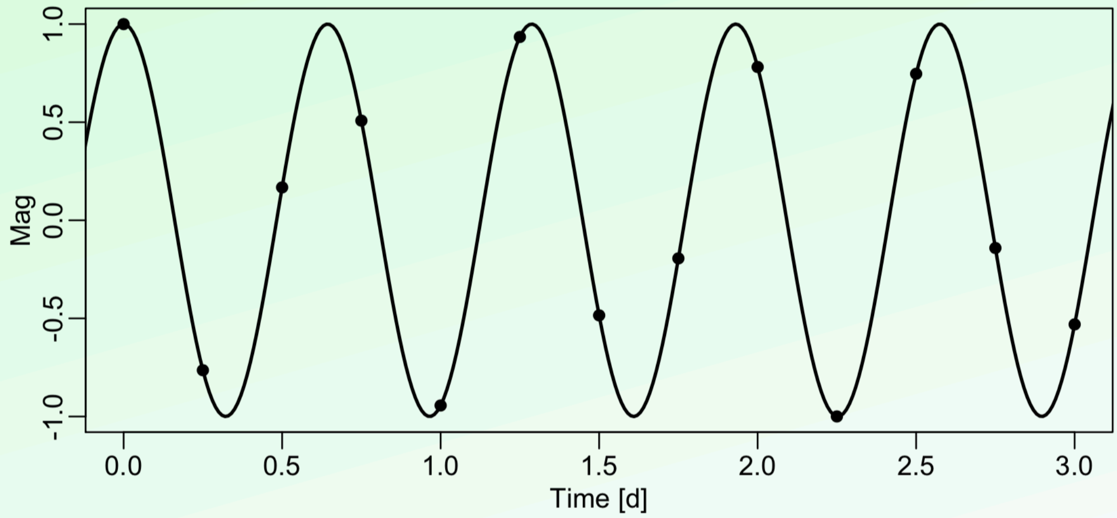
1. The discrete Fourier transform
2. The Nyquist frequency
3. What is aliasing?
4. The periodogram
5. Lomb Scargle for unevenly sampled data

Discrete Fourier transforms

Take board notes from here.

Nyquist frequency & aliasing

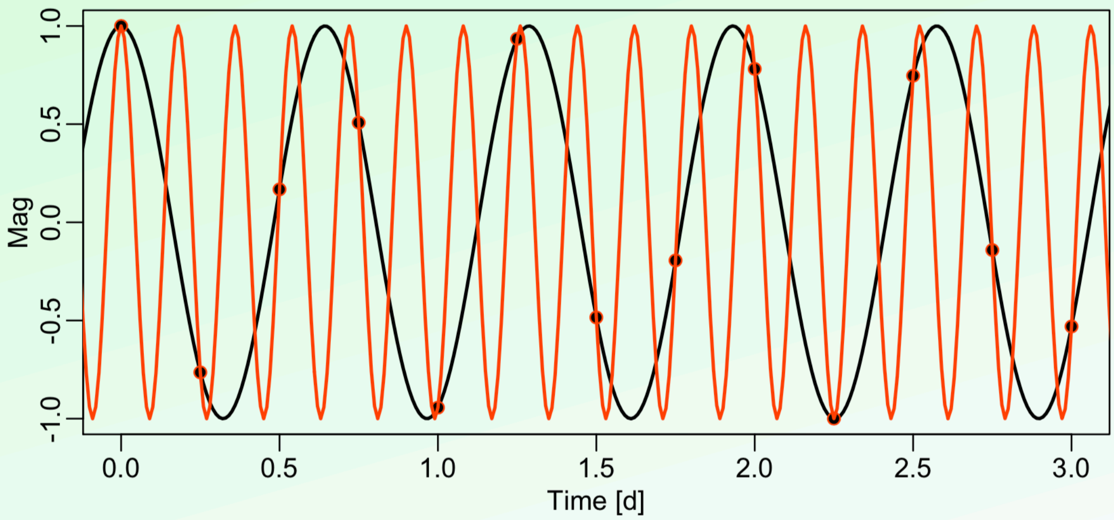
Suveges (LSST data science)



Sampling rate: 4 c/d (Gaia spin rate), Nyquist frequency: 2 c/d, $f_{\mathrm{signal}} = 1.55$ c/d

The highest detectable frequency in a time series sampled at f_s is $f_s/2 = f_c$ (Nyquist frequency).

Nyquist frequency & aliasing



Sampling rate: 4 c/d (Gaia spin rate), Nyquist frequency: 2 c/d, $f_{\mathrm{signal}} = 1.55$ c/d
$f_{\mathrm{highfreq}} = 5.55$ c/d

Suveges (LSST data science)

Nyquist frequency

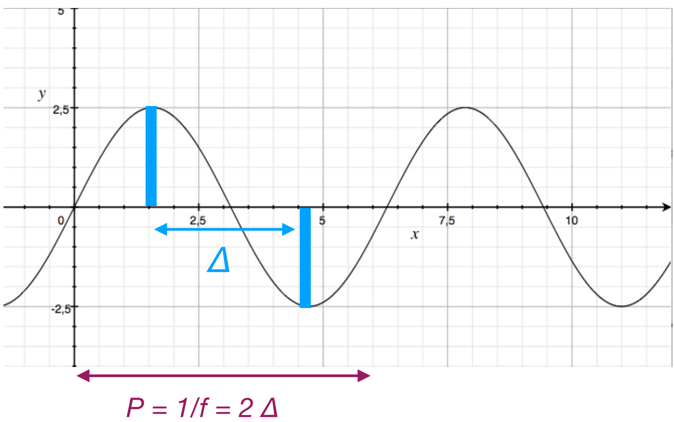
$$f_c \equiv \frac{1}{2\Delta}$$

where *Δ* is the sampling interval.

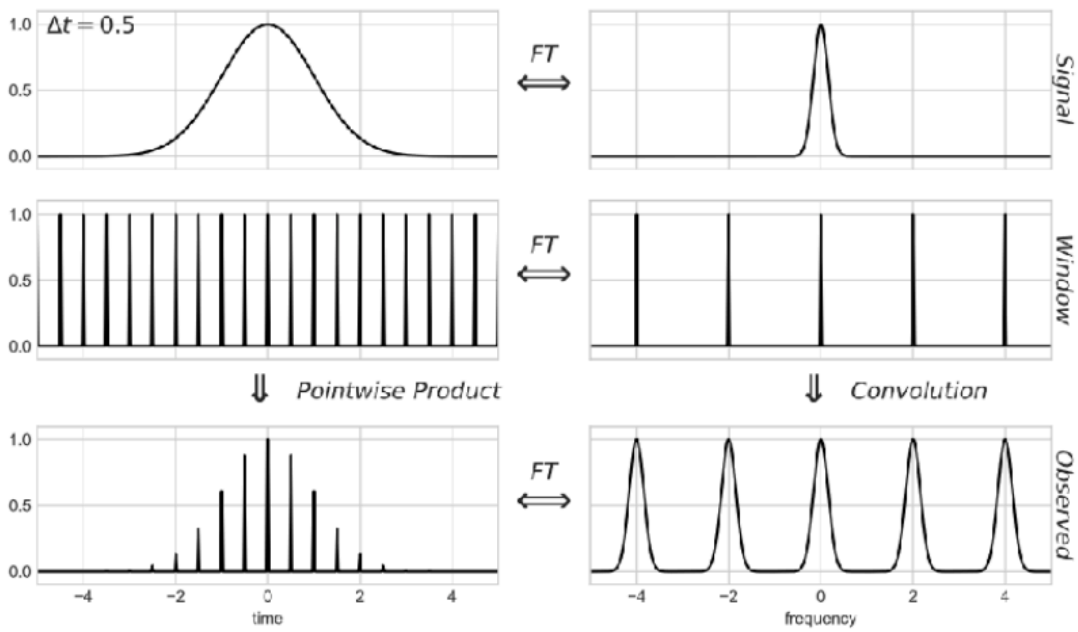
If a sine wave of frequency f_c is sampled at its positive peak value, the next sample in *Δ* will be at the trough.

Critical sampling of a sine wave is 2 points per cycle.

Undersampling of the sinusoid allows there to be a lower-frequency alias, which is a different function that produces the same set of samples.



DFT - sampling



Fast Fourier Transform

- Algorithm that does the transform of N points in a time proportion to $N \log N$, rather than the N^2 timing of a brute force implementation. Cooley & Tukey 1965
- Probably one of the most well used transforms on the planet.

Exercise

Use direct numerical integration to do a numerical FT of a sine wave.

Compare the timing with an off-the-shelf FFT routine.

How many oscillations can you fit into your region of integration before the FFT accelerates away from the direct method?

Nyquist good news: Sampling theorem

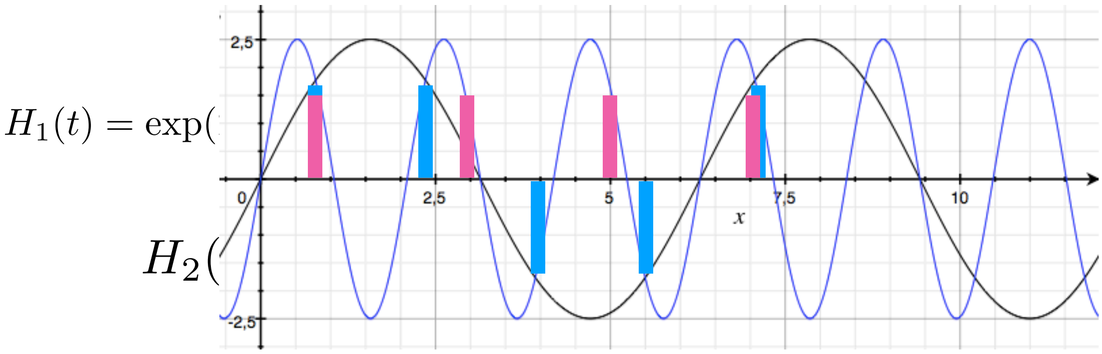
- If a continuous function $h(t)$, which is sampled at an interval Δ, happens to be bandwidth limited to frequencies $f < f_c$, then $h(t)$ is completely determined by its samples h_n.
- This is the **sampling theorem**.
- Think about a signal passed through an amplifier with a known frequency response.

Nyquist bad news: Aliasing

- If you sample a continuous function that is **not** bandwidth limited to less than the Nyquist frequency, then power at frequencies outside the range $-f_c < f < f_c$ is spuriously moved into that range.
- This is called **aliasing**

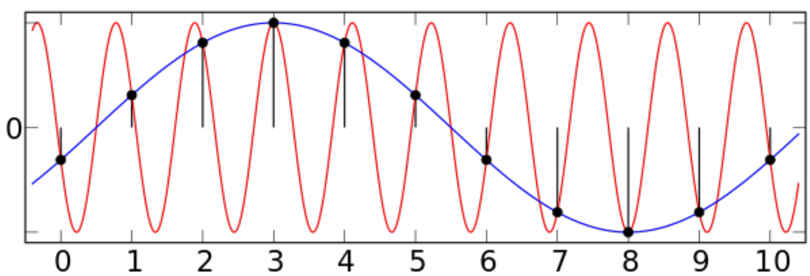
Nyquist bad news: Aliasing

- Consider two signals of frequencies f_1 and f_2
- These signals give the same samples only if f_1 and f_2 differ by multiples of $1/\Delta$.
- In our example: $f_2 = 3 f_1$



Nyquist bad news: Aliasing

- Consider a real signal (red) of frequency 9 cycles per 10 seconds, i.e. $f_{red} = 0.9$.
- You are sampling at $\Delta = 1s$, i.e. $f_s = 1$.
- The real frequency (f_{red}) is above the Nyquist frequency ($f_c = 0.5$) so you don't detect it
- But the real period can beat with the sampling period to give a frequency below the Nyquist frequency $f_{blue} = f_s - f_{red}$
- So power from a frequency outside your critical range has "leaked" or "aliased" into your range.



Power Spectral Density & Periodogram

- The *power spectral density* (or power spectrum) is given by:

$$\mathcal{P}_g \equiv |\mathcal{F}\{g\}|^2$$

where $F(g) = \hat{g}$ (see earlier in course notes) is the Fourier transform of g (a function of t)

- The *classical periodogram* is defined by Schuster (1898) as:

$$P_S(f) = \frac{1}{N}\left|\sum_{n=1}^{N} g_n e^{-2\pi i f t_n}\right|^2$$

while having a similar form as above, it's defined in the context of the discrete Fourier transform.

Power Spectral Density & Periodogram

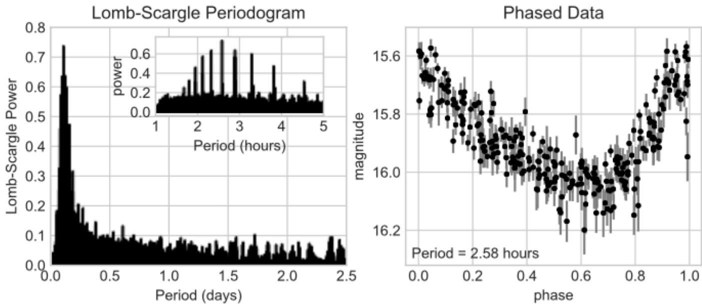
- While astronomers often use the terms *power spectrum* and *periodogram* interchangeably, we should be aware that the periodogram is an estimate of the power spectrum.

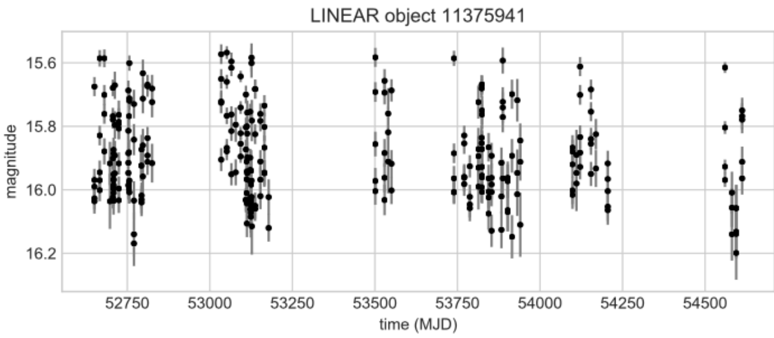
- This is because we're estimating the power spectrum (which is for a continuous underlying function) by discrete samples of this continuous function.

- Also, there are a few other quirks related to periodograms…

Lomb-Scargle periodogram

- Widely used in astronomy
- Can deal with uneven sampling





Lomb-Scargle periodogram

$$P_{LS}(f) = \frac{1}{2}\left\{ \left(\sum_n g_n \cos(2\pi f[t_n - \tau])\right)^2 \Big/ \sum_n \cos^2(2\pi f[t_n - \tau]) + \left(\sum_n g_n \sin(2\pi f[t_n - \tau])\right)^2 \Big/ \sum_n \sin^2(2\pi f[t_n - \tau]) \right\}$$

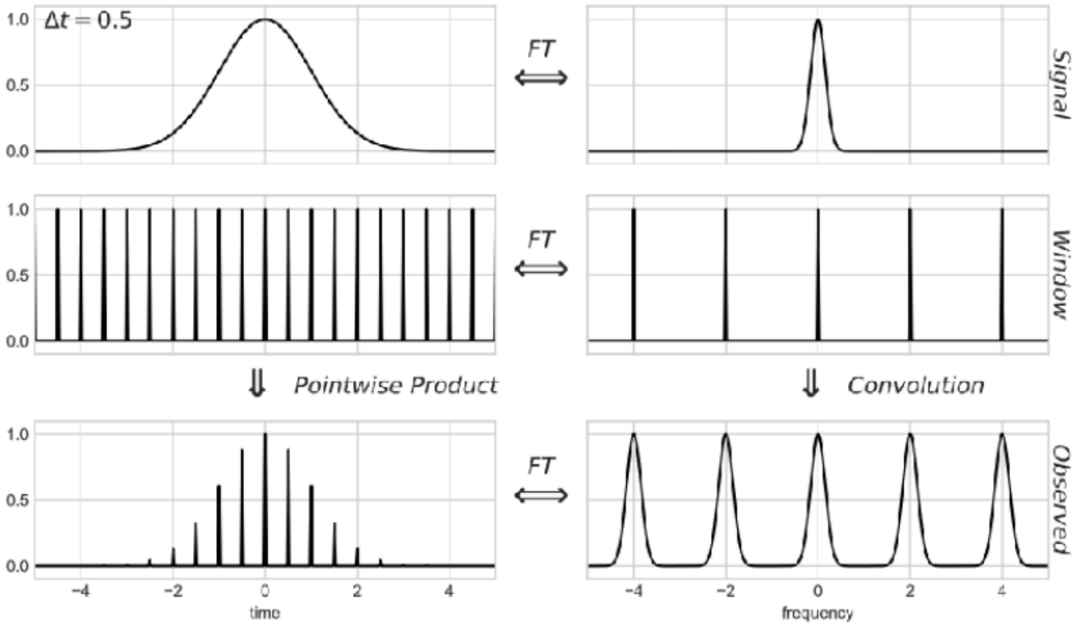
$$\tau = \frac{1}{4\pi f}\tan^{-1}\left(\frac{\sum_n \sin(4\pi f t_n)}{\sum_n \cos(4\pi f t_n)}\right)$$

where *tau* is specified for each *f* to ensure time-shift invariance

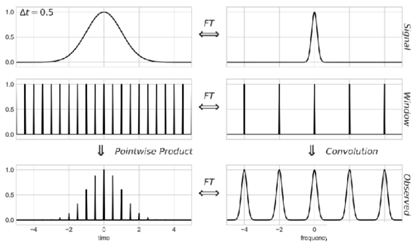
1. Periodogram reduces to classical form in the case of equally spaced observations
2. Periodogram's statistics are analytically computable
3. Periodogram is insensitive to global time-shifts in data

Lomb-Scargle periodogram

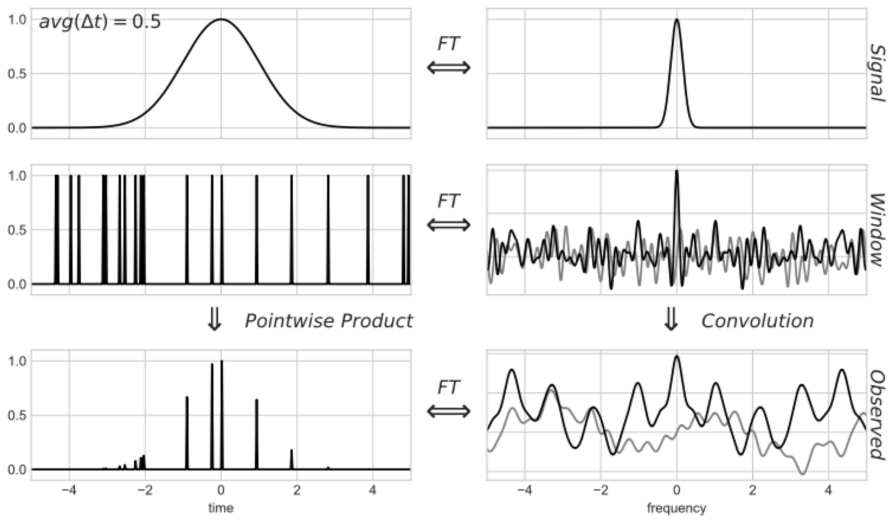
- Remember the effect of sampling? (DFT, convolution)



Lomb-Scargle periodogram

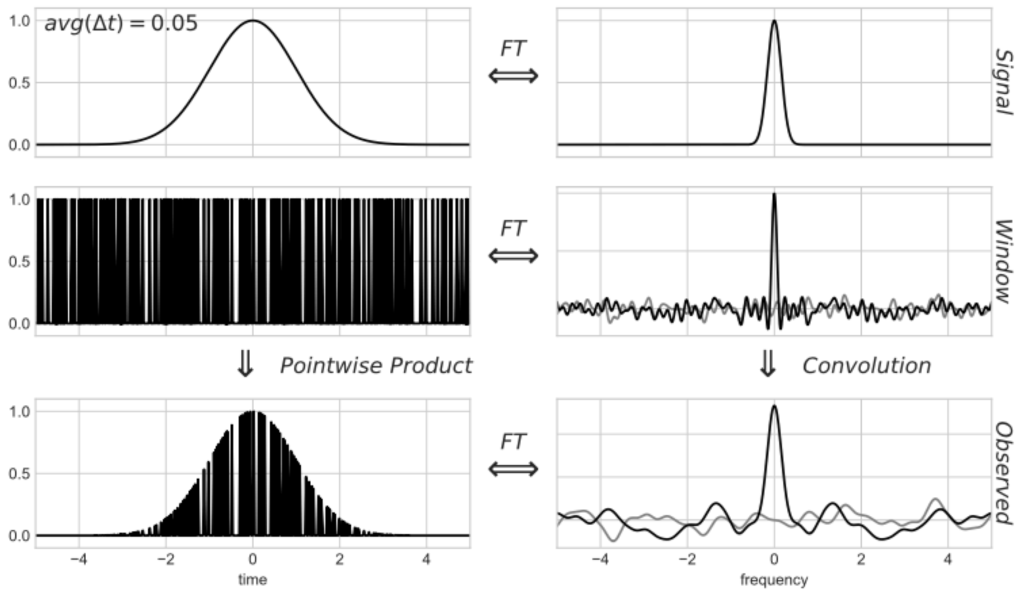


- FT of sampling looks like random noise
- No exact aliasing of the true signal



Lomb-Scargle periodogram

- Denser sampling helps, but doesn’t get rid of the noise in the FT.

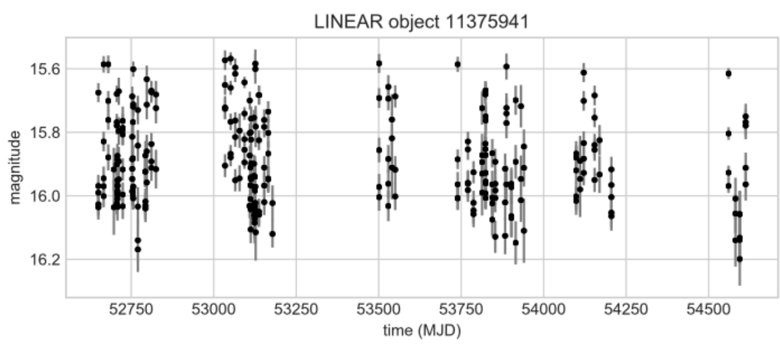


Lomb-Scargle periodogram

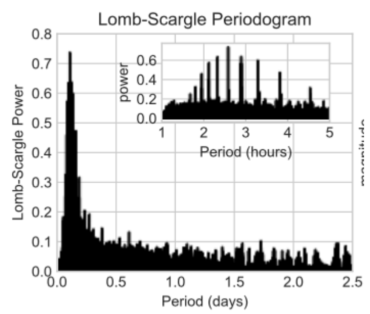
- Does the Nyquist limit exist in unevenly sampled data?

- Not in the same sense as for the DFT - it is a direct result of the even sampling of a Dirac comb.

- Some kinds of aliasing can still occur though, so best to be careful and use simulations.



Mean sampling 7 days



Period recovered 2.58 hours!

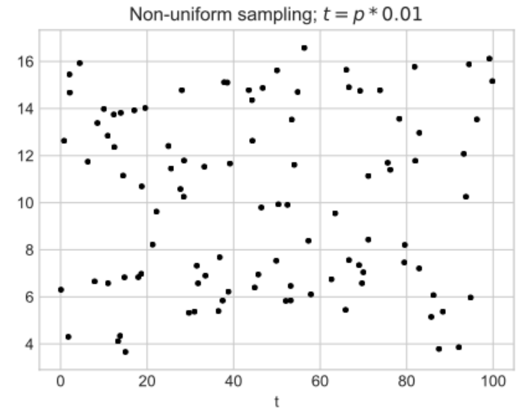
Lomb Scargle Periodogram

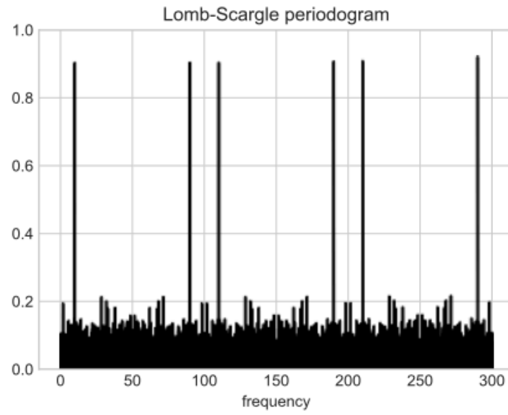
The non-uniform Nyquist limit

Eyer & Bartholdi (1999), Koen (2006)

$$\Delta t_i = t_i - t_0 = n_i p$$

Choose *p* to the largest value that allows each time interval to be written as an exact integer multiple of this factor *p*.





An exercise

1. Use an FFT algorithm to make a periodogram of the AirPassengers.csv dataset.

    1.1. What is the Nyquist frequency?

    1.2. Do you detect any frequencies in the data?

2. Run the F86_ogle234.dat file through a Lomb-Scargle routine and plot the periodogram.