MATH5802M
Chapter 6: Fitting sine waves

Peter Thwaites

with thanks to Stuart Barber for previous years' notes and slides

January 2017

Admin details

Lectures As 3802 plus Friday 9am, RSLT14.
Assessment 2.5 hour end of semester examination (80%) and an assessed practical (20%).
Trigonometry and complex numbers will be used a lot in spectral analysis.

6.1 Frequency

- In *spectral analysis*, a time series is analysed by decomposing it into sine waves of different amplitude and frequency.

  low frequency $\approx$ slowly changing trends
  high frequency $\approx$ rapid changes (noise).

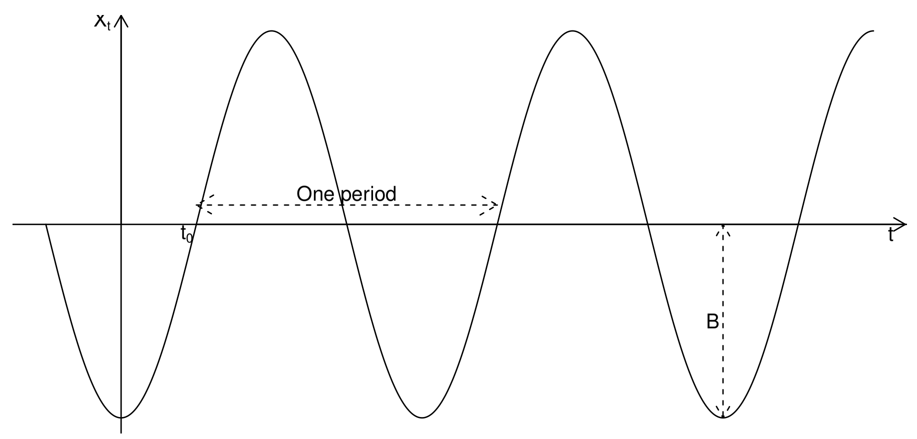
- If the data have a periodic component of known frequency, we can fit the following model:

$$X_t = B\sin(2\pi f(t - t_0)) + \mu + \varepsilon_t,$$

where

μ is a baseline or mean;
f is *frequency* (in cycles per unit time);
t_0 is the start time of one period;
B is the *amplitude*

Graphical interpretation



Example 6.1 If t is given in months since the start of measurements, then a period length of one year implies $f = 1/12$.

Parameterisation

- How many distinct parameters are there in

$$X_t = B\sin(2\pi f(t - t_0)) + \mu + \varepsilon_t?$$

- Use

$$\sin(x \pm y) = \sin(x)\cos(y) \pm \cos(x)\sin(y)$$
$$\cos(x \pm y) = \cos(x)\cos(y) \mp \sin(x)\sin(y)$$

to get

$$X_t = A'\cos(2\pi ft) + B'\sin(2\pi ft) + \mu + \varepsilon_t.$$

- Since any periodic component with frequency f can be written in this form, we can use this model for the sum of arbitrary many shifted sines/cosines with frequency f.
- Therefore we need just *two* parameters (A, B) per frequency.

6.2 Aliasing

- We have seen that only two parameters per frequency are required; frequency f is described by

$$A\cos(2\pi ft) + B\sin(2\pi ft). \tag{1}$$

A model with more parameters for frequency f cannot be distinguished from (1) by using observations.

- Since

$$A\cos(2\pi(-f)t) + B\sin(2\pi(-f)t) = A\cos(2\pi ft) - B\sin(2\pi ft),$$

the frequencies f and $-f$ are indistinguishable; they are *aliases* of each other. Hence we can assume $f \geqslant 0$.

- Assume that observations are given at times $0, \Delta, 2\Delta, 3\Delta, \ldots$, i.e. we have sampling frequency $1/\Delta$. A component with frequency f will be observed as

$$X_k = A\cos(2\pi f k\Delta) + B\sin(2\pi f k\Delta).$$

- The bigger f, the faster X_k oscillates, until $f = \frac{1}{2\Delta}$. Here, we have

$$X_k = A(-1)^k.$$

  This is the fastest possible oscillation.
- Now, $\frac{1}{2\Delta} + \varepsilon$ and $\frac{1}{2\Delta} - \varepsilon$ are aliases, so we can assume $0 \leqslant f \leqslant \frac{1}{2\Delta}$.

Definition 6.1(Nyquist frequency) The frequency $f = \frac{1}{2\Delta}$ is called the *Nyquist frequency* for the sampling frequency $1/\Delta$.