MATH5802M
Chapter 7: Computing the periodogram

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

January 2017

Introduction

We shall use a tool called the *discrete Fourier transform* (DFT) to draw *periodograms*. These graphs will help us explore which frequencies contribute to our time series.

The DFT is a method for decomposing a time series into periodic components. Assume we have observations X_t made at times $t = 0, 1, 2, \ldots, n-1$.

Definition 7.1 (Fourier frequencies) The frequencies $f_j = j/n$ are called *Fourier frequencies*.

7.1 The DFT

Defintion 7.2 (DFT) The discrete Fourier transform (DFT) of $X_0, \ldots, X_{n-1}$ is given by

$$\begin{aligned} \widehat{X}_j &= \frac{1}{\sqrt{n}} \sum_{t=0}^{n-1} X_t e^{-2\pi i f_j t} \qquad \text{for } j = 0, \ldots, n-1 \\ &= \frac{1}{\sqrt{n}} \sum_{t=0}^{n-1} X_t \left[\cos(2\pi f_j t) - i \sin(2\pi f_j t)\right]. \end{aligned}$$

Inverse DFT

To invert the DFT, we shall need
Lemma 7.1

$$\sum_{t=0}^{n-1} e^{2\pi i f_j t} e^{-2\pi i f_k t} = \begin{cases} \text{n} & \text{if j} = \text{k} \pmod{n} \\ 0 & \text{else.} \end{cases}$$

Definition 7.3 (Inverse DFT) Using lemma 7.1, we can recover $\{X_t\}$ from $\{\widehat{X}_j\}$:

$$X_t = \frac{1}{\sqrt{n}} \sum_{j=0}^{n-1} \widehat{X}_j e^{2\pi i f_j t}.$$

This is called the *inverse DFT*.

Example 7.1 Let $n = 4$, $X = 0, 1, 2, 3$. Then

$$\widehat{X}_0 = 3, \quad \widehat{X}_1 = -1 + i, \quad \widehat{X}_2 = -1, \quad \widehat{X}_3 = -1 - i.$$

So the DFT of $\{0, 1, 2, 3\}$ is $\{3, -1 + i, -1, -1 - i\}$.

For the inverse,

$$\begin{aligned} X_0 &= \frac{1}{2} \sum_{j=0}^{3} \widehat{X}_j e^{2\pi i f_j 0} = \frac{1}{2} \sum_{j=0}^{4} \widehat{X}_j \\ &= \frac{1}{2}(3 \quad -1 + i \quad -1 \quad -1 - i) = 0, \end{aligned}$$

as it should.

Alternative definition

Some people (including R) define the DFT as

$$\tilde{X}_j = \sum_{t=0}^{n-1} X_t e^{-2\pi i f_j t}.$$

In this case,

$$X_t = \frac{1}{n} \sum_{j=0}^{n-1} \tilde{X}_j e^{2\pi i f_j t}.$$

Properties of DFT

1. The DFT is *linear*, i.e.

$$Y_t = cX_t \Rightarrow \widehat{Y}_j = c\widehat{X}_j$$
$$Z_t = X_t + Y_t \Rightarrow \cdots \Rightarrow \widehat{Z}_j = \widehat{X}_j + \widehat{Y}_j.$$

2. We have "*conservation of energy*":

$$\sum_{j=0}^{n-1} |\widehat{X}_j|^2 = \sum_{t=0}^{n-1} |X_t|^2,$$

and hence $\|\widehat{X}\| = \|X\|$.

Properties of DFT (ctd...)

3. *Time shifts*. Let $Y = (X_{n-1}, X_0, X_1, \ldots, X_{n-2})$, i.e. $Y_t = X_{t-1 \bmod n}$. Then

$$\widehat{Y}_j = e^{-2\pi i f_j} \widehat{X}_j \qquad \Rightarrow \qquad |\widehat{Y}_j|^2 = |\widehat{X}_j|^2;$$

only the "phase" is changed, the modulus is the same. Therefore, it doesn't matter what point in the cycle our data start from.

Now let $Z = (X_{-1}, X_0, X_1, \ldots, X_{n-2})$ for some $X_{-1} \in \mathbb{R}$:

$$\|\widehat{Y} - \widehat{Z}\| = \|\widehat{Y - Z}\| = \|Y - Z\| = \sqrt{(X_{n-1} - X_{-1})^2} = |X_{n-1} - X_{-1}|.$$

Therefore the relative difference between $\widehat{Y}$ and $\widehat{Z}$ goes to zero as $n \to \infty$.

Result For (long) stationary time series, we can use cyclic shifts instead of normal ones.

7.2 Spectral density and periodogam

Defintion 7.4 (Spectral density and periodogram) The function $I : [0,1] \to [0,\infty)$ with $I(f_j) = |\widehat{X}_j|^2$ is called the *spectral density* of X. A plot of I as a function of f is called a *periodogram*.

The spectral density $I(f_j)$ is a measurement of how strongly the frequency f_j is represented in the data.

Example 7.2 Let $X = \{0,1,2,3\}$ as in example 7.1. Then we have $\widehat{X} = \{3, -1+i, -1, -1-i\}$, so

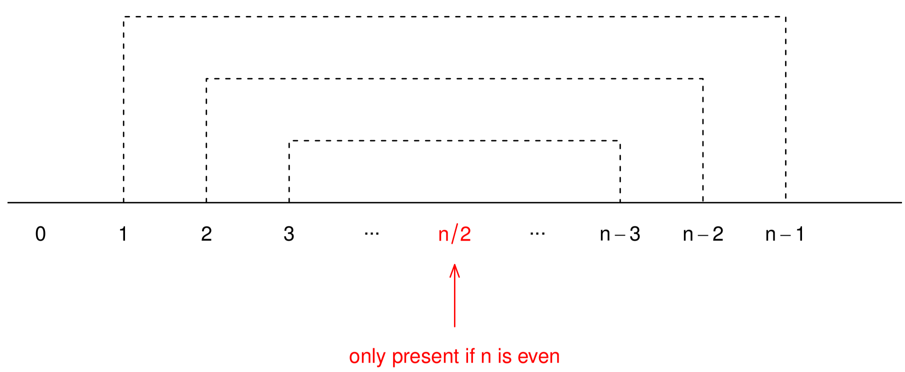
$$\begin{aligned} I(0) &= |\widehat{X}_0|^2 = 3^2 = 9, \\ I(1/4) &= |\widehat{X}_1|^2 = (-1+i)(-1-i) = 2, \\ I(2/4) &= |\widehat{X}_2|^2 = (-1)^2 = 1, \\ I(3/4) &= |\widehat{X}_3|^2 = 2. \end{aligned}$$

Nyquist again

For real data $\{X_t\}$, we have

$$\widehat{X}_{n-j} = \overline{\widehat{X}_j},$$

so only half the frequencies are needed. (This is expected, since f_j for $j > n/2$ is bigger than the Nyquist frequency $\frac{1}{2\Delta} = \frac{1}{2}$.)



Frequency $f = 0$

Note that, defining $\tilde{X}$ to be the mean of $\{X_t\}$,

$$\widehat{X}_0 = \frac{1}{\sqrt{n}} \sum_{t=0}^{n-1} X_t \underbrace{e^{2\pi i f_0 t}}_{=1} = \frac{1}{\sqrt{n}} \sum_{t=0}^{n-1} X_t = \sqrt{n}\tilde{X},$$

so $\widehat{X}_0 \in \mathbb{R}$.

Similarly, if n is even then $\widehat{X}_{n/2} = \overline{\widehat{X}}_{n-n/2} = \overline{\widehat{X}_{n/2}}$ and so $\widehat{X}_{n/2} \in \mathbb{R}$.

Alternative formulation

Let $\widehat{X}_j = a_j + ib_j \Rightarrow \widehat{X}_{n-j} = a_j - ib_j$. Then

$$X_t = \sum_{0 \leqslant j \leqslant \frac{n}{2}} [A_j \cos(2\pi f_j t) + B_j \sin(2\pi f_j t)],$$

where

$$\begin{array}{lll} A_0 = \dfrac{a_0}{\sqrt{n}} & B_0 = 0 & \\ A_j = \dfrac{2a_j}{\sqrt{n}} & B_j = -\dfrac{2b_j}{\sqrt{n}} & \text{for } 1 \leqslant j \leqslant \dfrac{n}{2} \\ A_{\frac{n}{2}} = \dfrac{a_{n/2}}{\sqrt{n}} & B_{\frac{n}{2}} = 0 & \text{if } n \text{ is even} . \end{array}$$

Example

Run the R script “fft.R”, available in the VLE.