LECTURE SEVEN

Sinusoidal Model

$$y_t = \beta_0 + \beta_1 \cos(2\pi f t) + \beta_2 \sin(2\pi f t) + \varepsilon_t$$

f : frequency

Sinusoid

$$s_t = \beta_0 + R\cos(2\pi f t + \Phi)$$

① R (Amplitude) : Maximum deviation from the center line β_0

② f (Frequency) : Numbers of oscillations in one unit of time.

sometimes we use Hz as the unit of f.
(# oscillations in 1 sec)

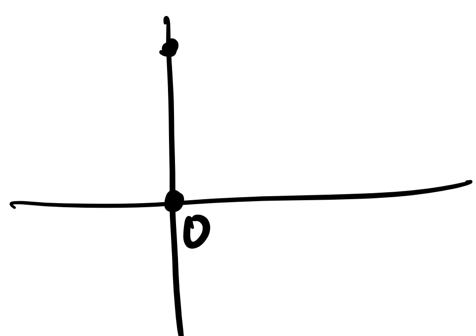
e.g: $f = \frac{1}{11}$ in the sunspots data
per year

③ $\frac{1}{f}$: Period of oscillation

④ Φ : Phase of oscillation

e.g $\Phi = 0$: $\beta_0 + R\cos(2\pi f t)$

$\sin(\pi f t)$



$\cos(\alpha+\beta)$
$= \cos\alpha\cos\beta - \sin\alpha\sin\beta$

$$\beta_0 + R\cos(2\pi f t + \Phi)$$
$$= \beta_0 + \underbrace{(R\cos\Phi)}_{\beta_1}\cos 2\pi f t + \underbrace{(-R\sin\Phi)}_{\beta_2}\sin 2\pi f t$$
$$= \beta_0 + \beta_1 \cos 2\pi f t + \beta_2 \sin 2\pi f t$$

Our time series will always be sampled at equally spaced points (with no gaps)

$$\left.\begin{array}{l} y_1, y_2, y_3, y_4, \ldots, y_n \\ t = 1, 2, \ldots, n \end{array}\right\}$$

For such data, $s_t = R\cos(2\pi f t + \Phi)$, $t = 1, \ldots, n$, then we can always restrict $f \in [0, \frac{1}{2}]$.

Fact: For every $f \in \mathbb{R}$ & Φ, there exist $f_0 \in [0, \frac{1}{2}]$ & Φ_0 such that

$$s_t = R\cos(2\pi f t + \Phi) = R\cos(2\pi f_0 t + \Phi_0)$$
for all $t = 1, \ldots, n$.

Reason: $f = 2.3$

$$R\cos(2\pi(2.3)t + \Phi)$$
$$= R\cos(4\pi t + 2\pi(0.3)t + \Phi)$$
$$= R\cos(2\pi(0.3)t + \Phi)$$

integer 0.8 $\quad R\cos(2\pi(0.8)t+\Phi)$

$\cos(2\pi t-\alpha) = \cos\alpha$ $\quad = R\cos(2\pi t - 2\pi(0.2)t+\Phi)$

$= R\cos(2\pi(0.2)t-\Phi)$

$f_0 \in [0,\frac{1}{2}]$ is called an ALIAS of f.

$$y_t = \beta_0+\beta_1\cos 2\pi f t + \beta_2 \sin 2\pi f t + \varepsilon_t$$

$\varepsilon_t \sim N(0,\sigma^2)$ $\quad f\in[0,\frac{1}{2}]$

Remark:

$f=0$: $\beta_0+\beta_1$: constant

$f=\frac{1}{2}$: $\beta_0+\beta_1(-1)^t$: max possible oscillation

Parameter Estimation

$y = X_f\beta+\varepsilon$: Linear Regression

Last lecture

$y = X_c\beta+\varepsilon$

$$X_f = \begin{bmatrix} 1 & \cos 2\pi f(1) & \sin 2\pi f(1) \\ \vdots & \vdots & \vdots \\ 1 & \cos 2\pi f(n) & \sin 2\pi f(n) \end{bmatrix}$$

$$RSS(f) = \min_\beta \|y - X_f\beta\|^2$$
$$= \|y - X_f\hat\beta_f\|^2, \quad \hat\beta_f = (X_f^T X_f)^{-1}X_f^T y$$

① Take a grid of values of f in $[0,\frac{1}{2}]$

$0, 0.0001, 0.0002, \ldots, 0.5$

② Calculate $RSS(f)$ for each f in Grid

③ $\hat f$ = minimizer of $RSS(f)$ in grid

$$\text{posterior}(f) \propto |X_f^T X_f|^{-1/2} \left(\frac{RSS(\hat{f})}{RSS(f)}\right)^{\frac{n-p}{2}}$$

$$I(0 \le f \le \tfrac{1}{2})$$

prior (f) : Unif $[0, \frac{1}{2}]$

do not compute this for f near 0 or near $\frac{1}{2}$.

$\hat{f}$ f

$RSS(f) = (1.1) RSS(\hat{f})$

$$\boxed{\left(\frac{1}{1.1}\right)^{\frac{n-p}{2}}}$$

Efficient Computation of RSS(f)

n = sample size

Grid: $f \in [0, \frac{1}{2}]$

$$0, \frac{1}{n}, \frac{2}{n}, \frac{3}{n}, \frac{4}{n}, \dots, \frac{[n/2]}{n}$$

$[\frac{n}{2}]$ = integer part of $\frac{n}{2}$

Fourier Frequencies: any frequency f for which nf is an integer

325

$0, \frac{1}{325}, \frac{2}{325}, \dots$

$0, \frac{1}{300}, \frac{2}{300}, \dots$

We will focus only on Fourier Frequencies for computational reasons.

Fix f in a Fourier frequency.

How to compute $\underbrace{RSS(f)}$ efficiently?

$$RSS(f) = \min_{\beta} \|y - X_f \beta\|^2$$

sm.OLS(y, X_f)
sum(resid^2)

$$= \|y - X_f \hat{\beta}_f\|^2 \qquad \hat{\beta}_f = (X_f^T X_f)^{-1} X_f^T y$$

$$= (y - X_f \hat{\beta}_f)^T (y - X_f \hat{\beta}_f)$$

$$= y^T y - y^T X \hat{\beta} - \hat{\beta}^T X^T y + \hat{\beta}^T X^T X \hat{\beta}$$

$$= y^T y - y^T X (X^T X)^{-1} X^T y - \cancel{y^T X (X^T X)^{-1} X^T y} + \cancel{y^T X (X^T X)^{-1} X^T X (X^T X)^{-1} X^T y}$$

$$= y^T y - y^T X_f (X_f^T X_f)^{-1} X_f^T y$$

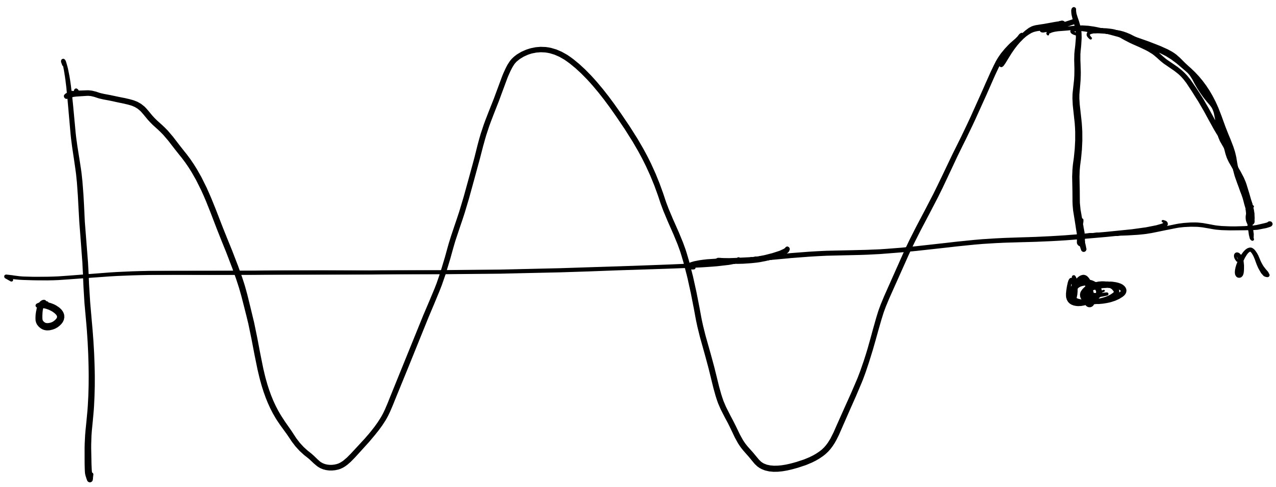
$$RSS(f) = y^T y - y^T X_f (X_f^T X_f)^{-1} X_f^T y$$

$$X_f = \begin{bmatrix} 1 & \cos(2\pi f t) & \sin 2\pi f t \\ \vdots & t = 1 \ldots n & t = 1 \ldots n \\ 1 & & \end{bmatrix}$$

$$X_f^T X_f = \begin{bmatrix} 1 & - - - & 1 \\ & \cos 2\pi f t & \\ & \sin 2\pi f t & \end{bmatrix} \begin{bmatrix} 1 & \cos 2\pi f t & \sin 2\pi f t \\ \vdots & & \\ 1 & & \end{bmatrix}$$

$$= \begin{bmatrix} n & \sum_{t=1}^{n} \cos 2\pi f t & \sum_{t=1}^{n} \sin 2\pi f t \\ \boxed{} & \sum_{t=1}^{n} \cos^2 2\pi f t & \sum_{t=1}^{n} \cos 2\pi f t \sin 2\pi f t \\ \boxed{} & \boxed{} & \sum_{t=1}^{n} \sin^2 2\pi f t \end{bmatrix}$$

$$\sum_{t=1}^{n} \cos(2\pi f t) = 0$$ when f is a Fourier frequency



$$\sum_{t=1}^{n} \sin 2\pi f t = 0$$

$$\sum_{t=1}^{n} \cos^2 2\pi f t = \sum_{t=1}^{n} \frac{1+\cos 4\pi f t}{2}$$

$$= \frac{n}{2} + \frac{1}{2} \underbrace{\sum_{t=1}^{n} \cos(4\pi f t)}_{=0}$$

$$= \frac{n}{2}$$

$$\sum_{t=1}^{n} \sin^2 2\pi f t = \frac{n}{2}$$

$$\sum_{t=1}^{n} \sin 2\pi f t \cos 2\pi f t$$

$$= \frac{1}{2} \sum_{t=1}^{n} \sin(4\pi f t) = 0$$

$$X_f^T X_f = \begin{bmatrix} n & 0 & 0 \\ 0 & \frac{n}{2} & 0 \\ 0 & 0 & \frac{n}{2} \end{bmatrix}$$

$$(X_f^T X_f)^{-1} = \begin{bmatrix} \frac{1}{n} & 0 & 0 \\ 0 & \frac{2}{n} & 0 \\ 0 & 0 & \frac{2}{n} \end{bmatrix}$$

$$RSS(f) = y^T y - y^T X_f \underbrace{(X_f^T X_f)^{-1}} X_f^T y$$

$$= \sum y_t^2 - \left(\sum y_t \quad \sum y_t \cos 2\pi f t \quad \sum y_t \sin 2\pi f t\right) \begin{pmatrix} \frac{1}{n} & & 0 \\ & \frac{2}{n} & \\ 0 & & \frac{2}{n} \end{pmatrix} \begin{pmatrix} \sum y_t \\ \sum y_t \cos 2\pi f t \\ \sum y_t \sin 2\pi f t \end{pmatrix}$$

$$= \underbrace{\sum y_t^2 - \frac{1}{n}\left(\sum y_t\right)^2} - \frac{2}{n}\left(\sum y_t \cos 2\pi f t\right)^2 - \frac{2}{n}\left(\sum y_t \sin 2\pi f t\right)^2$$

$$\sum y_t^2 - \frac{1}{n}(n\bar{y})^2$$

$$= \sum y_t^2 - n(\bar{y})^2$$

$$= \sum (y_t - \bar{y})^2$$

$$RSS(f) = \sum (y_t - \bar{y})^2 - \frac{2}{n}\left(\sum y_t \cos 2\pi f t\right)^2 - \frac{2}{n}\left(\sum y_t \sin 2\pi f t\right)^2$$

$$I(f) = \frac{1}{n}\Big(\underbrace{\sum y_t \cos 2\pi f t}\Big)^2 + \frac{1}{n}\Big(\sum y_t \sin 2\pi f t\Big)^2$$

$$= \frac{1}{n}\left|\Big(\sum y_t \cos 2\pi f t\Big) - i\Big(\sum y_t \sin 2\pi f t\Big)\right|^2$$

$$= \frac{1}{n}\left|\sum y_t e^{-2\pi i f t}\right|^2$$

DFT of the data