Lecture NINE

Data: $y_0, y_1, \ldots, y_{n-1}$ (Python indexing)

$$y_t = \beta_0 + \beta_1 \cos 2\pi f t + \beta_2 \sin 2\pi f t + \underbrace{\varepsilon_t}_{N(0,\sigma^2)}$$

$$\underbrace{\beta_0, \beta_1, \beta_2, \sigma}_{\text{nuisance parameters}}, \quad f$$

$$RSS(f) = \min_{\beta_0,\beta_1,\beta_2} \sum_{t=0}^{n-1} \left(y_t - \beta_0 - \beta_1 \cos 2\pi f t - \beta_2 \sin 2\pi f t\right)^2$$

$\hat{f} \rightarrow$ minimizer of $RSS(f)$

$$\text{posterior}(f) \propto \left(\frac{1}{RSS(f)}\right)^{\frac{n-3}{2}} \left|X_f^T X_f\right|^{-1/2} I(0<f<\tfrac{1}{2})$$

We need some candidate values for f to compute $RSS(f)$ (& then subsequent minimizati or posterior computation). There are two options:

① Take a dense grid of points in $[0, 0.5]$

② Take the grid of Fourier frequencies: $\left\{\frac{1}{n}, \frac{2}{n}, \ldots\right\}$ in $(0, \frac{1}{2})$.

Then $RSS(f)$, $f \in$ Fourier grid $\left\{\frac{1}{n}, \frac{2}{n}, \ldots\right\} \cap (0, \frac{1}{2})$ can be efficiently computed using the FFT algorithm.

$$\text{DFT}: \quad y_0, y_1, \ldots, y_{n-1} \;\longrightarrow\; b_0, b_1, \ldots, b_{n-1}$$

This step uses the FFT algorithm & is very fast.

$$b_j = \sum_{t=0}^{n-1} y_t \exp\left(-2\pi i \frac{j}{n} t\right)$$

Periodogram: $I\left(\frac{j}{n}\right) = \frac{|b_j|^2}{n}$

$$\boxed{RSS\left(\frac{j}{n}\right) = \sum_{t=0}^{n-1} (y_t - \bar{y})^2 - 2 I\left(\frac{j}{n}\right)}$$

Minimize $RSS\left(\frac{j}{n}\right)$ to get $\hat{f} = \frac{\hat{j}}{n}$

$$\text{posterior}\left(\frac{j}{n}\right) \propto \left(\frac{1}{RSS\left(\frac{j}{n}\right)}\right)^{\frac{n-3}{2}} \left|X_{\frac{j}{n}}^T X_{\frac{j}{n}}\right|^{-1/2} I\left(0 < \frac{j}{n} < \frac{1}{2}\right)$$

Lecture 7, we saw that if $f \in \left(0, \frac{1}{2}\right)$ is a Fourier frequency,

$$X_f^T X_f = \begin{bmatrix} n & 0 & \\ 0 & \frac{n}{2} & \\ & & \frac{n}{2} \end{bmatrix} \leftarrow \text{does not depend on } f$$

$$X_f = \left[\begin{array}{c|c|c} 1 & \cos(2\pi f t), & \sin(2\pi f t) \\ \vdots & t = 0, 1, \ldots, n-1 & t = 0, 1, \ldots, n-1 \\ 1 & & \end{array}\right]$$

$$\text{posterior}\left(\frac{j}{n}\right) \propto \left(\frac{1}{RSS\left(\frac{j}{n}\right)}\right)^{\frac{n-3}{2}} I\left(0 < \frac{j}{n} < \frac{1}{2}\right)$$

This reduction to Fourier frequencies is only used for COMPUTATIONAL PURPOSES.

More Nonlinear Regression Models

① Two or more sinusoids

$$y_t = \beta_0 + \beta_{11} \cos 2\pi f_1 t + \beta_{12} \sin 2\pi f_1 t + \beta_{21} \cos 2\pi f_2 t + \beta_{22} \sin 2\pi f_2 t + \varepsilon_t$$

$$\boxed{f_1, f_2}$$

$$RSS(f_1, f_2) = \underset{\beta}{\operatorname{argmin}} \|y - X_{f_1, f_2} \beta\|^2$$

$$X = \begin{bmatrix} 1 & \cos 2\pi f_1 t & \sin 2\pi f_1 t & \cos 2\pi f_2 t & \sin 2\pi f_2 t \\ \vdots & & & & \\ 1 & & & & \end{bmatrix}$$

$$(\hat{f}_1, \hat{f}_2) = \underset{f_1, f_2}{\operatorname{argmin}} RSS(f_1, f_2)$$

$$\text{posterior}(f_1, f_2) \propto \left(\frac{1}{RSS(f_1, f_2)}\right)^{\frac{n-5}{2}} |X_{f_1, f_2}^T X_{f_1, f_2}|^{-1/2} I(0 < f_1, f_2 < \tfrac{1}{2})$$

① Restrict both f_1 & $f_2 \in (0, \frac{1}{2})$ to Fourier frequencies

② Take another possibly denser grid for both f_1 & f_2.

If f_1 & f_2 are distinct Fourier frequencies in $(0, \frac{1}{2})$ (e.g $f_1 = \frac{j_1}{n}$, $f_2 = \frac{j_2}{n}$, $j_1 \neq j_2$)

$$RSS(f_1, f_2) = \sum_{t=0}^{n-1} (y_t - \bar{y})^2 - 2I(f_1) - 2I(f_2)$$

$\rightarrow$ happens because of orthogonality of Sinusoids at Fourier frequencies.

$$X_{f_1,f_2}^T X_{f_1,f_2} = \begin{bmatrix} n & & & & \\ & \frac{n}{2} & & & 0 \\ & & \frac{n}{2} & & \\ & 0 & & \frac{n}{2} & \\ & & & & \frac{n}{2} \end{bmatrix}$$

$RSS(f_1, f_2)$ is minimized when f_1 & f_2 are the top two maximizers of the periodogram.

$$\underbrace{I(\tfrac{1}{n}), I(\tfrac{2}{n}), \dots \, I(\;)}$$

$$I(\tfrac{10}{n}) : \hat{f}_1 = \frac{10}{n}$$

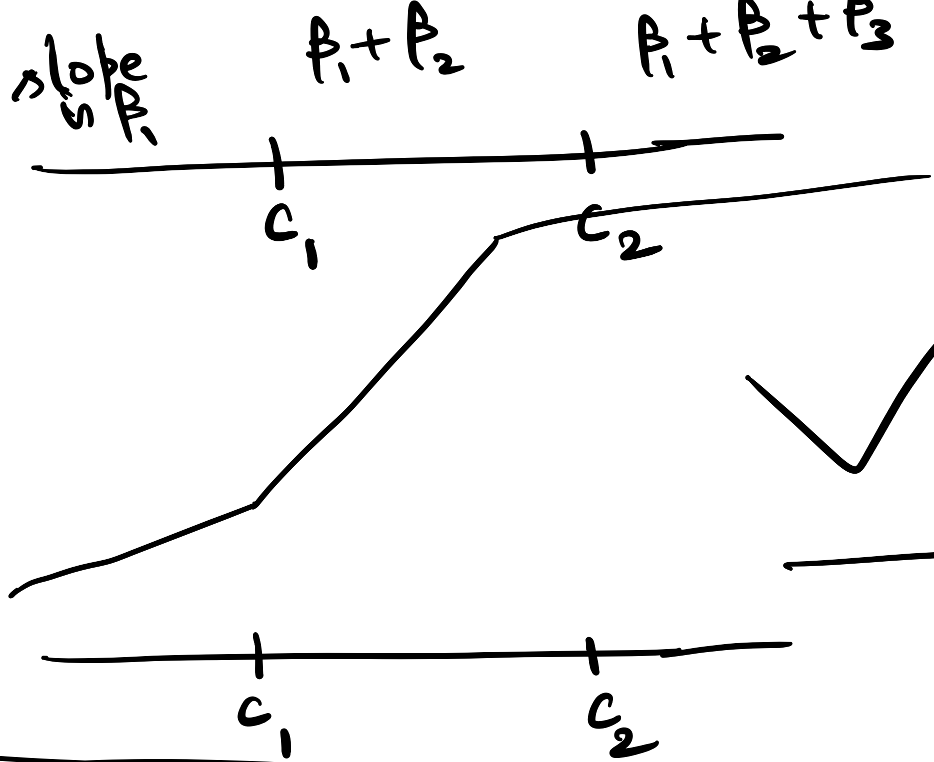
$$I(\tfrac{8}{n}) : \hat{f}_2 = \frac{8}{n}$$

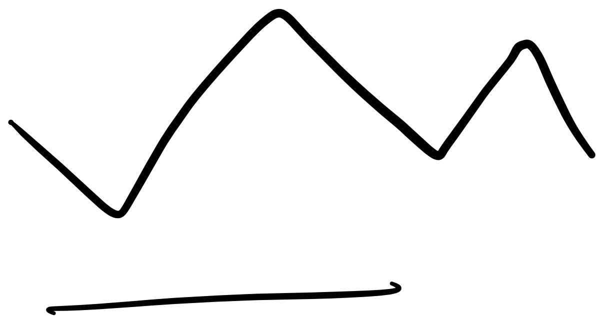
Change of Slope Model

① $y_t = \beta_0 + \beta_1 t + \beta_2 (t - c)_+ + \varepsilon_t$

$RSS(c)$

② $y_t = \beta_0 + \beta_1 t + \beta_2 (t-c_1)_+ + \beta_3 (t-c_2)_+ + \varepsilon_t$





$\boxed{RSS(c_1, c_2)}$

③ c_1, c_2, c_3

$RSS(c_1, c_2, c_3)$

$$y_t = \beta_0 + \beta_1 t + \beta_2 (t-2)_+ + \beta_3 (t-3)_+ + \cdots + \beta_{n-1} (t-(n-1))_+ + \varepsilon_t$$

$\boxed{RSS = 0}$ ⟶ High-dimensional Linear Regression Model

$\boxed{\text{Regularized}}$ Estimation

① minimize [Least squares + penalty]

② $\beta_0, \beta_1, \beta_2, \ldots, \beta_{n-1} \overset{iid}{\sim} \underbrace{\cancel{\mathrm{Unif}(-C, C)}}_{\text{more informative}}$

$\mathrm{Unif}(-\tau, \tau)$
for some $\boxed{\tau}$