Lecture FOURTEEN

① Model 1 $\rightarrow$ HIGH-DIMENSIONAL LINEAR REGRESSION

② Model 2 $\rightarrow$ VARIANCE MODEL

③ Model 3 $\rightarrow$ VARIANCE MODEL ON DFT / SPECTRUM MODEL

$y_1 \dots y_n$

Model 1:

$$y_t \overset{\text{ind}}{\sim} N(\mu_t, \sigma^2), \quad t = 1, \dots, n$$

parameters: $\mu_1 \dots \mu_n$ & σ^2

$$\underset{\mu_1 \dots \mu_n}{\text{Minimize}} \sum_{t=1}^{n} (y_t - \mu_t)^2 \doteq (\hat{\mu}_1 \dots \hat{\mu}_n)$$

$$\hat{\mu}_t = y_t$$

If we want more structure in the estimates of $\{\mu_t\}$ e.g smoothness

$$\sum_{t=1}^{n} (y_t - \mu_t)^2 + \lambda \sum_{t=3}^{n} \left[(\mu_t - \mu_{t-1}) - (\mu_{t-1} - \mu_{t-2})\right]^2$$

or

$$+ \lambda \sum_{t=3}^{n} \left| \underbrace{(\mu_t - \mu_{t-1}) - (\mu_{t-1} - \mu_{t-2})} \right|$$

μ_t : trend estimate.

$$(\mu_t - \mu_{t-1}) - (\mu_{t-1} - \mu_{t-2})$$

$$\mu = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_n \end{pmatrix} = X\beta$$

$$X = \begin{bmatrix} 1 & 0 & & 0 \\ 1 & 1 & & 0 \\ 1 & 2 & \cdots & \vdots \\ \vdots & \vdots & & 0 \\ 1 & n-1 & & 1 \end{bmatrix}$$

$\|y - X\beta\|^2 + \lambda \sum \beta_j^2 \rightarrow$ Ridge estimate of β

$\hat{\beta}(\lambda)$

$$\boxed{\hat{\mu} = X\hat{\beta}(\lambda)}$$

Model TWO $y_1 \dots y_n$ **VARIANCE Model**

$\boxed{y_t \overset{\text{independent}}{\sim} N(0, \tau_t^2)}$, $t = 1 \dots n$

parameters: $\tau_1 \dots \tau_n > 0$

τ_t : magnitude or strength of y_t.

High-dimensional model.

Qn: How to estimate $\tau_1 \dots \tau_n$?

Likelihood: $\rightarrow \prod_{t=1}^{n} \frac{1}{\sqrt{2\pi}\,\tau_t} \exp\left(-\frac{y_t^2}{2\tau_t^2}\right)$

$\longrightarrow$ only depends on y_t^2 or $|y_t|$

$(y_1^2, \dots, y_n^2)$: sufficient statistic for the model.

Rewrite the model: $\boxed{y_t^2 \overset{\text{independent}}{\sim} \tau_t^2 \chi_1^2, \; t = 1 \ldots n}$

$\boxed{\chi_m^2 = \Gamma\left(\frac{m}{2}, \frac{1}{2}\right)}$

log-likelihood

$$= \sum_{t=1}^{n} \left(-\frac{y_t^2}{2\tau_t^2} - \log \tau_t \right)$$

Negative log-likelihood

$$= \sum_{t=1}^{n} \left(\frac{y_t^2}{2\tau_t^2} + \log \tau_t \right)$$

$\alpha_t := \log \tau_t$ or $\tau_t = e^{\alpha_t}$

$$= \sum_{t=1}^{n} \left\{ \frac{y_t^2}{2} e^{-2\alpha_t} + \alpha_t \right\}$$

$$-y_t^2 e^{-2\alpha_t} + 1 = 0$$

$\boxed{\text{OVERFITTING}}$ or $$e^{2\alpha_t} = y_t^2$$

$$e^{\alpha_t} = |y_t| \text{ or } \boxed{\tau_t = |y_t|}$$

$\boxed{\tau_t = |y_t|}$ $\quad$ $\boxed{\alpha_t = \log |y_t|}$

Regularization:

$$\sum_{t=1}^{n}\left\{\frac{y_t^2}{2}e^{-2\alpha_t}+\alpha_t\right\}+\lambda\sum_{t=2}^{n-1}\left((\alpha_{t+1}-\alpha_t)-(\alpha_t-\alpha_{t-1})\right)^2$$

or

$$+\lambda\sum_{t=2}^{n-1}\left|(\alpha_{t+1}-\alpha_t)-(\alpha_t-\alpha_{t-1})\right|$$

Data: $\{y_t\}$

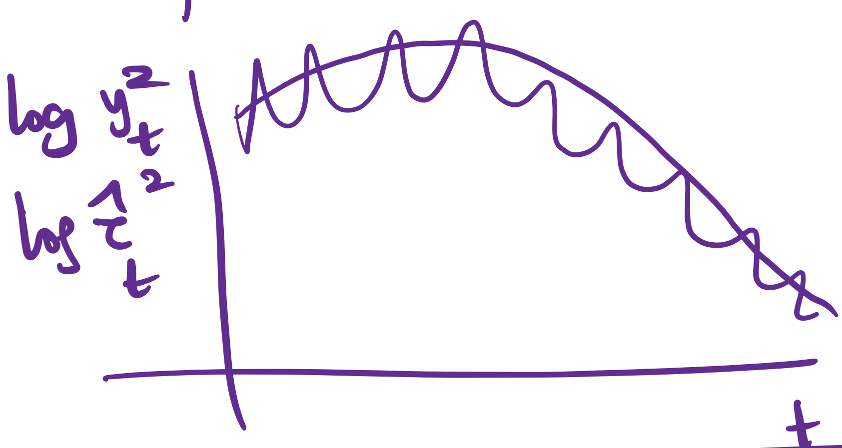
Compute estimates of $\alpha_t = \log \tau_t$:

$\hat{\alpha}^{Ridge}(\lambda)$ $\quad \hat{\alpha}^{LASSO}(\lambda)$

$\hat{\tau}^{Ridge}(\lambda)$ $\quad \hat{\tau}^{LASSO}(\lambda)$

$$\boxed{y_t \sim N(0,\tau_t^2)}$$

$$\log\frac{y_t}{y_{t-1}} = \log y_t - \log y_{t-1}$$



$$y_t^2 = \tau_t^2 x_1^2$$

$$\log y_t^2 = \boxed{\log \tau_t^2} + \log x_1^2$$

MODEL 3

$y_0 \cdots y_{n-1}$

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

$b_0 = \sum y_t, \quad b_{n-j} = \overline{b_j}$

Periodogram $\rightarrow$ $\boxed{I\left(\frac{j}{n}\right) = \frac{|b_j|^2}{n}, \quad 0 < \frac{j}{n} < \frac{1}{2}}$

$Re(b_j), Im(b_j) \overset{iid}{\sim} N(0, \gamma_j^2)$

n : odd

$m = \frac{n-1}{2}$.

Model: $b_1 \dots b_m$ are independent

$\longrightarrow$ & $Re(b_j), Im(b_j) \overset{iid}{\sim} N(0, \gamma_j^2)$

$j = 1, \dots, m$

Parameters: $\gamma_1^2, \dots, \gamma_m^2$

Likelihood:

$$\prod_{j=1}^{m} \left(\frac{1}{\sqrt{2\pi}\gamma_j} \exp\left[-\frac{(Re\, b_j)^2}{2\gamma_j^2}\right] \frac{1}{\sqrt{2\pi}\gamma_j} \exp\left[-\frac{(Im\, b_j)^2}{2\gamma_j^2}\right] \right)$$

$$\propto \prod_{j=1}^{m} \frac{1}{\gamma_j^2} \exp\left[-\frac{(Re\, b_j)^2 + (Im\, b_j)^2}{2\gamma_j^2}\right]$$

$$= \prod_{j=1}^{m} \frac{1}{\gamma_j^2} \exp\left[-\frac{n\, I(j/n)}{2\gamma_j^2} \right]$$

$\longrightarrow$ Only depends on the periodogram

$\Rightarrow$ Periodogram is the Sufficient Statistic.

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

$$= \frac{(Re(b_j))^2 + (Im(b_j))^2}{n}$$

$$\sim \frac{\gamma_j^2}{n} \chi_2^2$$

$Re(b_j)$
$Im(b_j) \sim N(0, \gamma_j^2) = \gamma_j N(0,1)$

$$\left.\boxed{I\left(\frac{j}{n}\right) \overset{\text{indep}}{\sim} \frac{\gamma_j^2}{n} \chi_2^2}\right\}$$

$$\chi_2^2 = \text{Gamma}\left(1, \frac{1}{2}\right) \propto x^{1-1} e^{-\frac{1}{2}x} I\{x>0\}$$

$$= \exp\left(-\frac{1}{2}x\right) I(x>0)$$

$$\boxed{I\left(\frac{j}{n}\right) \overset{\text{ind}}{\sim} \frac{\gamma_j^2}{n} \underbrace{\text{Exp}\left(\frac{1}{2}\right)}}$$

$\text{Exp}(\lambda)$

$\lambda e^{-\lambda x} I\{x>0\}$

$$\left.\boxed{I\left(\frac{j}{n}\right) \overset{\text{ind}}{\sim} \boxed{\frac{2\gamma_j^2}{n}} \text{Exp}(1)}\right\}$$

$x \sim \text{Exp}(\lambda)$

then $\lambda X \sim \mathrm{Exp}(1)$

$$\prod_{j=1}^{m} \frac{1}{\gamma_j^2} \exp\left[-\frac{n\, I(j/n)}{2\gamma_j^2}\right] = \text{likelihood}$$

$$\text{log-lik} = \sum_{j=1}^{m} \left[-\frac{n\, I(j/n)}{2\gamma_j^2} - 2\log\gamma_j\right]$$

$$\text{Negative log lik} = \sum_{j=1}^{m} \left[\frac{n\, I(j/n)}{2\gamma_j^2} + 2\log\gamma_j\right]$$

$$\alpha_j = \log\gamma_j$$

$$\sum_{j=1}^{m} \left[\frac{n}{2} I(j/n)\, e^{-2\alpha_j} + 2\alpha_j\right]$$

Regularization:

$$+\ \lambda \sum \left((\alpha_j - \alpha_{j-1}) - (\alpha_{j-1} - \alpha_{j-2})\right)^2$$

$$\text{or } \lambda \sum \left|(\alpha_j - \alpha_{j-1}) - (\alpha_{j-1} - \alpha_{j-2})\right|$$