Week 3: Sep. 26.

Lecture 4:

Trend estimation:

* Part I: model estimation
* Part II: smoothing (parametric)
* Part III: Isotonic regression (non-parametric)

Part I: Model Estimation

$$X_t = \underbrace{\mu_t}_{\text{Trend}} + \underbrace{e_t}_{\text{noise}}, \quad \text{assume } e_t \sim WN(0, \sigma^2), \quad \text{ACF of } \bar{X} \text{ is } \rho(h)$$

① Constant: $\mu_t \equiv \mu$

estimator: $\hat{\mu}_t \equiv \bar{X} = \frac{1}{n}\sum_{t=1}^{n} X_t \Rightarrow E(\hat{\mu}_t) = \mu \qquad Var(\hat{\mu}_t) = \frac{\gamma(0)}{n}\left(1 + 2\sum_{t=1}^{n-1}\left(1-\frac{t}{n}\right)\rho(t)\right)$

Remark: when X_t uncorrelated, $Var(\hat{\mu}_t) = \frac{\gamma(0)}{n}$

t \ h	1	2	3	⋯	n-1
1	$\gamma(1)$	$\gamma(2)$	⋯	⋯	$\gamma(n-1)$
2	$\gamma(1)$	⋮			
3	⋮	⋮			
⋮	⋮	$\gamma(2)$			
n-1	$\gamma(1)$				

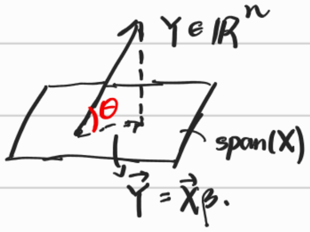
② $\mu_t = \beta_0 + \beta_1 t$ or $Y = X\beta + \varepsilon = \begin{pmatrix}\varepsilon_1\\ \vdots \\ \varepsilon_n\end{pmatrix}$, where $\beta = \underset{}{\operatorname{argmin}} \|Y - X\beta\|^2$

$Y = \begin{pmatrix}Y_1\\ \vdots \\ Y_n\end{pmatrix}$, $X = \begin{pmatrix}1 & X_{11} & \cdots & X_{1p}\\ & \vdots & & \end{pmatrix}$, $\beta = \begin{pmatrix}\beta_1\\ \vdots \\ \beta_n\end{pmatrix}$

$\Rightarrow \hat{\beta} = (X^T X)^{-1} X^T Y$

$$\hat{\beta}_1 = \frac{\sum(\mu_t - \bar{\mu})(t - \bar{t})}{\sum_{t=1}^{n}(t-\bar{t})^2}, \quad \bar{t} = \frac{n+1}{2}, \quad \hat{\beta}_0 = \bar{\mu} - \hat{\beta}_1 \bar{t}$$

Geometric Implementation of Least Square.



① $1 \in X$ (with intercept)

$\frac{1}{n}\sum_{i=1}^{n} \hat{Y}_i = \frac{1}{n}\sum_{i=1}^{n} Y_i = \bar{Y}$.

$e_i = Y_i - \hat{Y}_i$: e residual

$e \perp X \ni 1 \to 1 \perp e \Rightarrow \langle e, 1 \rangle = 0 \Rightarrow \sum e_i = 0 \Rightarrow \sum \hat{Y}_i = \sum Y_i$

Thus: $\frac{1}{n}\sum \hat{\mu}_t = \frac{1}{n}\sum \mu_t \Rightarrow$ unbiased

② evaluation

Cov(·,·) 满足内积定义.

则 Var(·) 可理解为模长的平方

* $Corr(\hat{Y}, Y)$
* $\cos\theta$
* $R^2 = \frac{Var(\hat{Y})}{Var(Y)} = 1 - \frac{Var(Y-\hat{Y})}{Var(Y)} = Corr^2(\hat{Y}, Y) = \cos^2\theta$

$= \frac{Cov^2(\hat{Y}, Y)}{Var(\hat{Y})Var(Y)}$

Exercise: Y, X_1, X_2

$Y \sim X_1 \Rightarrow R_1^2 = 0.1$ Q: $Y \sim X_1$ and X_2 $R^2 \in [?, ?]$

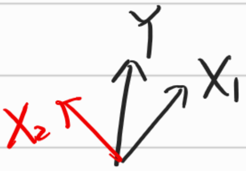
$Y \sim X_2 \Rightarrow R_2^2 = 0.2$ 0.2 1

ensemble → $R^2 \uparrow$

→ $Var(\bar{X}_n) = \frac{1}{n} Var(X_i)$, 方差↓

→ MSE↓ = var + bias.

Solution:

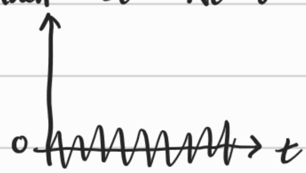


旋转X_2, X_1, X_2, Y 可共面. $\Rightarrow R^2 = 1$

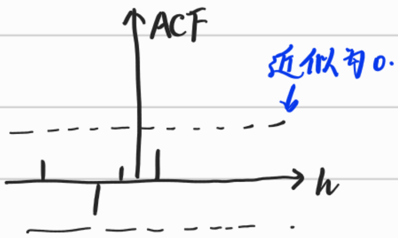
当X_1与Y的关系(R^2)确定, R^2不会变得更差 $\Rightarrow R^2 = 0.2$.

*: Evaluation — Residual: $\hat{e}_t = X_t - \hat{\mu}_t$

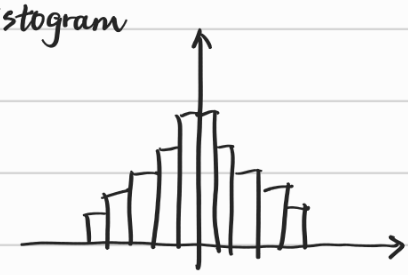
① residual plot:



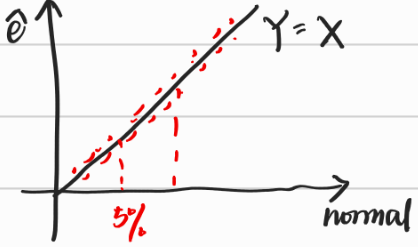
② ACF plot:



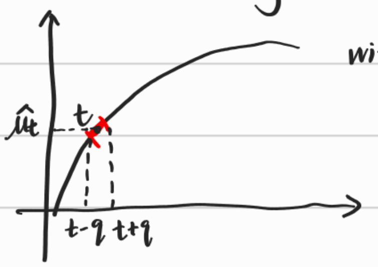
③ density plot / histogram



④ QQ-plot :



Part II : Smoothing



with non-linear Trend

if linear.

$$\bar{\mu}_t = \frac{1}{2q+1}\sum_{h=-q}^{q} X_{t+h} = \frac{1}{2q+1}\sum_{h=-q}^{q}(X_{t+h} - \mu_{t+h} + \mu_{t+h}) = \frac{1}{2q+1}\sum_{h=-q}^{q}\mu_{t+h} + \frac{1}{2q+1}\sum_{h=-q}^{q} e_t$$

t自己

Remark : ① $q\uparrow \Rightarrow e_t$ influence $\downarrow$, but might violate linearity

② end point issue.

③ smoothing $\in$ linear filter

$$\hat{\mu}_t = \sum_{h=-q}^{q} a_h X_{t+h}$$

- equal weight
- binomial
- exponential

Part III. istonic regreession

$$\min_{a_1,\dots,a_n} : \sum_{t=1}^{n} (X_t - a_t)^2$$

s.t. $a_1,\dots,a_n$ monotonic

① Pooled Adjacency : Violatim Algo V.

No tuning parameters

No end-point issue.

when $q\uparrow$ { bias $\uparrow$, variance $\downarrow$ }

Bias-Variance Trade-off

$$MSE = E(\hat{\mu}-\mu)^2 = \underbrace{(E(\hat{\mu})-\mu)^2}_{\text{bias}} + \underbrace{E(\hat{\mu}-E(\hat{\mu}))^2}_{\text{variance.}}$$

Assume noises are correlated

$$Var(\hat{\mu_t}) = Var(\bar{X}) = Var\left(\frac{1}{n}\sum_{t=1}^{n} X_t\right) = \frac{1}{n^2} Var\left(\sum_{t=1}^{n} X_t\right)$$

$$= \frac{1}{n^2}\left(\sum_{t=1}^{n} Var(X_t) + \right.$$

$$X_t = \mu_t + e_t, \quad e_t \sim WN(0, \sigma^2)$$

$$\rho(h) = \frac{\gamma(h)}{\gamma(0)} = \frac{Cov(X_t, X_{t+h})}{Var(X_t)}, \forall t$$

$$\Rightarrow \gamma(1) = \rho(1)\gamma(0)$$

$$Var(X_1 + \cdots + X_n)$$

$$= Var(X_1) + Var(X_2 + \cdots + X_n) + 2\,Cov(X_1, X_2 + \cdots + X_n)$$

$$= Var(X_1) + Var(X_2 + \cdots + X_n) + 2\left[\rho(1)\gamma(0) + \rho(2)\gamma(0) + \cdots + \rho(n-1)\gamma(0)\right]$$

$$= Var(X_1) + Var(X_2 + \cdots + X_n) + 2\gamma(0)\left[\rho(1) + \cdots + \rho(n-1)\right]$$

$$= \sum_{t=1}^{n} Var(X_t) + 2\gamma(0)\left[\left.\begin{array}{l} \rho(1) + \cdots + \rho(n-1) \\ \rho(1) + \cdots + \rho(n-2) \\ \vdots \\ \rho(1) \end{array}\right\} n-1 \right]$$

$$= n\cdot\gamma(0) + 2\gamma(0)\cdot\sum_{h=1}^{n-1}(n-h)\rho(h)$$

$$\Rightarrow \frac{1}{n^2}\left[n\gamma(0) + 2\gamma(0)\sum_{h=1}^{n-1}(n-h)\rho(h)\right]$$

$$\Rightarrow \frac{\gamma(0)}{n}\left[1 + 2\sum_{h=1}^{n-1}\left(1 - \frac{h}{n}\right)\rho(h)\right)$$