MGTF 404 Final Review Junxiong Gao

I. Prob theory. Statistics & Heuristics

Random variable X. $\xleftrightarrow{\text{joint}}$ Random variable Y.

X, R.V. population distribution
$\updownarrow$
x observation, data

Want to know, how X, Y correlated, how change of X affect Y

$E(X) \longrightarrow$ sample estimation

$\bar{X} = \frac{1}{T}\sum X_t$

$Var(X)$

$std_x = \frac{1}{T-1}\sum(X_t - \bar{X})^2$

$E(X^n)$

$Y|X$. conditional distribution
Data Generating Process
$\updownarrow$
model e.g. linear regression

$E(Y|X) = \beta X$

$y_t = \beta x_t + \varepsilon_t$

Data: Time-series. cross-section. Panel

We focus on the time-series sample
(Not purely random sample)
$\Downarrow$ Thus regularity condition

$\{X_t\}_{t=1}^T$ distribution $\xRightarrow{\text{Based on}}$

OLS estimation method

$\hat{\beta}_{OLS} = \arg\min (y_t - \beta x_t)^2 = \frac{\sum x_t y_t}{\sum x_t^2}$

$\hat{\beta}_{OLS} - \beta = \frac{\sum x_t \varepsilon_t}{\sum x_t^2}$ Bias

a. covariance stationary doesn't change with Time, shrink

1). $E(X_t) = \mu$

2). $\cancel{\text{cor}}$ cov $(X_t, X_{t-j}) = \gamma(j)$

We need stationary & ergodic to derive distribution of Bias!

b. ergodic Sample estimate does not explode.

$\frac{1}{T}\sum X_t \to$

$\frac{1}{T}\sum(X_t - \bar{X}) \to$ Converge to the true distribution

Assume $E(X\varepsilon) = 0$. Apply CLT

$\frac{1}{T}\sum x_t \varepsilon_t \to N(0, S)$

$\frac{1}{T}\sum x_t^2 \to Q$ assume homoskedacity

$S = \sigma^2 \cdot Q$

Iteration Law $E[E(Y|X)] = E(Y)$

$\Rightarrow \hat{\beta} - \beta \to N(0, \sigma^2 \cdot Q^{-1})$

Law of total variance: $Var(Y) = Var(E(Y|X)) + E(Var(Y|X))$

spurious regression is the first thing to get rid of :

Don't work with non-stationary series. If want, take difference to make it stationary !

Test for super-consistency : Augmented Dicker-Fuller test (unit-root)

ADF distribution
simulated critical value
sample size dependent.

~~$P_t = \phi P_{t-1} +$~~ set a specific lag p

~~$P_t = \sum \phi_p P_{t-p}$~~ $y_t = \sum \phi_p y_{t-p}$

Hypothesis testing: $H_0: \phi = 1$ $H_1: \phi < 1$

get rid of other lagged value

ADF: $y_t = \phi_1 y_{t-1} + \xi_1 \Delta y_{t-1} + \xi_2 \Delta y_{t-2} + \cdots + \xi \Delta y_{t-p} + u_t$

$\Delta y_{t-1} = y_{t-1} - y_{t-2}$ like return for price

critical value from simulation

ADF : $\Delta y_t = \eta y_{t-1} + \cdots$

$\eta = 0$ v.s. $\eta < 0$

III. Some empirical results in finance

price & dividends : Campbell-Shiller model

$$R_{t+1} + 1 = \frac{D_{t+1} + P_{t+1}}{P_t} = \frac{P_{t+1}}{P_t} + \frac{D_{t+1}}{P_t} = \frac{P_{t+1} + D_{t+1}}{P_{t+1}} \cdot \frac{P_{t+1}}{P_t}$$

$$\underbrace{\log(1 + R_{t+1})}_{r_{t+1}} = \log\left(\frac{P_{t+1}}{P_t}\right) + \log\left(\frac{P_{t+1} + D_{t+1}}{P_{t+1}}\right)$$

$$= p_{t+1} - p_t + \log(1 + e^{d_{t+1} - p_{t+1}})$$

taylor expansion

$$\approx k + \rho p_{t+1} - p_t + (1-\rho) d_{t+1}$$

$p_t = k + \rho^j p_{t+j} - \sum \rho^i r_{t+1+i} + \sum_{i=0}^{i} \rho^i (1-\rho) d_{t+1+i}$

$j \to \infty \to 0$ no bubble condition

$d_t - p_t = k - \sum \rho^i (1-\rho) \Delta d_{t+1+i} + \sum \rho^i r_{t+1+i}$

motivation: DY predict returns? persistent predictor

long term regression

CCAPM and risk-return tradeoff / equity premium puzzle.

Risk-Return trade off $\operatorname{cov}(R_{t+1}, \sigma^2_t) > 0$ Important

Consumption CAPM R^*, λ sup corr

$E[R_{t+1}] - R_f = \gamma \operatorname{cov}(R_{t+1}, \lambda) = \beta_\theta [E(R^*) - R^0]$

$E[r_{t+1}] - r_f = \gamma \operatorname{cov}[r_{t+1}, \log(\lambda)] - \frac{\sigma_r^2}{2}$

$\gamma = 50$! Equity risk premium puzzle.

IV. Volatility model & MLE

Break of homoskedacity: heteroskedaerty

$Y_t = E[y_t | I_{t-1}] + u_t \quad u_t \sim N(0, h_t)$

ARCH test $\hat{u}_t^2 = h_t + V_t$ regress on explanatory factors

$\hat{u}_t^2 = \sum \phi_p \hat{u}_{t-p}^2 + V_t \qquad h_t = \sum \phi_p u_{t-p}^2$

Generalize $h_t = \xi + \alpha u_{t-1}^2 + \beta h_{t-1}$

persistence from $\alpha + \beta$

~~① factor model~~

ϕ close to 1 persistent time series

small sample bias in AR 1 $y_t = \phi y_{t-1} + \varepsilon_t$

Downward Bias

$E(\hat{\phi} - \phi) = -\frac{1+3\phi}{T} + O(1/T^2)$. correction $\phi = \hat{\phi} + \frac{1+3\hat{\phi}}{T}$

Similarly, persistent predictor in predictive regression

$y_t = \beta x_{t-1} + \varepsilon_t$

$x_t = \phi x_{t-1} + u_t$

upward Bias

$E(\hat{\beta} - \beta) = -\frac{\sigma_{\varepsilon u}}{\sigma_u^2} \frac{1+3\phi}{T} > 0$ by $\sigma_{\varepsilon u}$ empirically < 0

correction $\beta = \hat{\beta} + \frac{\sigma_{\varepsilon u}}{\sigma_u^2} \cdot \frac{1+3\hat{\phi}}{T}$

② factor model $y_t = \beta x_t + \varepsilon_t$

model specification: same as ①

want to add factors make ε_t white noise.

~~test~~ CAPM $R_{it} - R_f = \beta_i (R_{mt} - R_f) + \varepsilon_t$

test CAPM: ① scatter ② α ③ adding factors

sample 2010

④ joint test CAPM

③ Extension: VAR

$$\begin{bmatrix} y_t \\ x_t \end{bmatrix} = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} + \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix} \begin{bmatrix} y_{t-1} \\ x_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix}$$

granger-causality $\Leftrightarrow$ real causality

Fed Bank contraction policy $\Leftrightarrow$ stock returns

$\boxed{\text{spurious regression}}$

II. Regression & statistical inference

~~①~~ inference $R^2 = 1 - \frac{\sum e_t^2}{\sum (y_t - \bar{y})^2}$

~~T-stat : simple question~~

$X_t = \underset{(0.23)}{0.9}\, x_t + e_t \qquad T = \frac{0.9 - 0}{0.23} \approx 3.9$ > 1.96 Two-side 95%

> 2.576 Two-side 99%

Two regression

① AR model : capture characteristics of time series

. model specification : $y_t = \sum \phi_q y_{t-q} + \varepsilon_t$

add lags to make ε_t white noise. Back / forward testing

Test white noise : Durbin-Watson test $DW = \frac{\sum (\hat{e}_t - \hat{e}_{t-1})^2}{\sum \hat{e}_t^2} \approx 2(1 - r(1))$

Or, directly look at correlation in $\{y_t\}$

Box-Pierce $Q_{BP} = T \sum_{r=1}^{q} \hat{\rho}^2(r) \sim \chi^2(q)$ depends on the

Ljung-Box $Q_{LB} = T(T+2) \sum^{q} \frac{\hat{\rho}^2(r)}{T-r} \sim \chi^2(q)$ lag q set by model

AR 1 : $y_t = c + \phi\, y_{t-1} + \varepsilon_t$ persistency in $\{\varepsilon_t\}$

$y_t = \sum_{s=0}^{t} \phi^s \varepsilon_{t-s} + \sum_{s=0}^{t} \phi^s \cdot c \qquad \frac{\partial y_t}{\partial \varepsilon_{t-s}} = \phi^s$ $\Leftrightarrow$ persistency in $\{y_t\}$

~~$E(y_t)$~~ If $0 < |\phi| < 1$, the white noise shrink $\rightarrow$ stationary

$E(y_t) = \frac{c}{1-\phi} \qquad var(y_t) = \frac{\sigma^2}{1-\phi^2}$

$E(y_t | y_{t-1}) = c + \phi\, y_{t-1}$

h_t hidden variable: cannot estimate GARCH by OLS

MLE: assume a distribution of data $N(\mu, \sigma^2)$

efficient. reach C-R lower bound

can handle latent variables

GARCH: $y_t = c + \phi y_{t-1} + \varepsilon_t$ $\quad \varepsilon_t \sim N(0, h_t)$

$$h_t = \xi + \alpha \varepsilon_{t-1}^2 + \beta h_{t-1}$$

conditional log-likelihood:

$$f_{t|t-1} = -\log\left(\sqrt{2\pi(\xi + \alpha\varepsilon_{t-1}^2 + \beta h_{t-1})}\right) - \frac{(y_t - c - \phi y_{t-1})^2}{2(\xi + \alpha\varepsilon_{t-1}^2 + \beta h_{t-1})}$$

$$LL = \sum_{t=2}^{T} f_{t|t-1}$$

$f_{1|0}$ you can assume

Variance Risk premium / Stochastic Vol / MIDAS

V. Kalman Filter: filter out latent variable. exactly like GARCH except one extra updating "nowcast" step

$y_t = A'x_t + H'z_t + w_t$ $\quad E(w_t v_t') = 0$ $\quad E(v_t v_t') = Q$ $\quad E(w_t w_t') = R$

$z_t = F'z_{t-1} + v_t$

Basically, z_t is hidden.

first forecast $z_{t|t-1}$, based on all the works before to forecast z_t

↓

Use $z_{t|t-1}$ to get $y_{t|t-1}$

↓

~~Correct~~ Adjust $z_{t|t-1} \to z_{t|t}$ from $y_t \Leftrightarrow y_{t|t-1}$

↓

Use $z_{t|t} \to z_{t+1|t}$

MGTF 404 Review II Junxiong

Kalman Filter: a single variable case Risk-Return Tradeoff

~~Rt~~ = $r_t = q + K\sigma_{t-1}^2 + u_t$ Rewrite $z_t = \sigma_{t-1}^2$

$\sigma_{t-1}^2 = \alpha + \beta\sigma_{t-2}^2 + v_t$ $z_t = \alpha + \beta z_{t-1} + v_t$

① $P_{t|t-1} = E\left[(z_t - z_{t|t-1})^2\right]$

② $r_{t|t-1} = q + K\sigma_{t-1}^2$ $r_{t|t-1} \sim N(q + Kz_{t|t-1}, K^2 P_{t|t-1} + R)$

$E\left[(r_t - r_{t|t-1})\right]^2 = K^2 P_{t|t-1} + R$

③ $z_{t|t} = z_{t|t-1} + \beta_{KF}(r_t - r_{t|t-1})$ $\beta_{KF} = \dfrac{E\left[(z_t - z_{t|t-1})(r_t - r_{t|t-1})\right]}{E\left[(r_t - r_{t|t-1})^2\right]}$

$P_{t|t} = E\left\{\left[(z_t - z_{t|t-1}) - \beta(r_t - r_{t|t-1})\right]^2\right\}$

$= P_{t|t-1} - \dfrac{K^2 P_{t|t-1}^2}{K^2 P_{t|t-1} + R}$

④ $z_{t+1|t} = \alpha + \beta z_{t|t}$

$P_{t+1|t} = \beta^2 P_{t|t} + Q$

V. GMM / Instrumental Variable / Endogenous

A model is designed to make: $E[g(x, \theta)] = 0$

$\theta_{n\times 1}$ Parameter space

$g_1 \cdots g_k$ moment conditions $k \geq n$

$\hat{g}_k = \frac{1}{T}\sum_{t=1}^{T} g_k(x_t, \theta)$ $\hat{G} = \begin{bmatrix} \hat{g}_1 \\ \vdots \\ \hat{g}_k \end{bmatrix}_{k\times 1}$ $\boxed{\hat{G}' W \hat{G} = 0}$

Write Down moment conditions for models.

Ex 1. $y_t = \beta x_t + \varepsilon_t$

$E(x_t \varepsilon_t) = 0$ property of OLS

Estimated by $\frac{1}{T} \sum x_t \varepsilon_t = 0$ Linear Projection Theorem

$\Rightarrow \hat{\beta}_{GMM} = \frac{\sum x_t y_t}{\sum x_t^2} = \hat{\beta}_{OLS}$

Ex 2. ARCH

$y_t = c + \phi y_{t-1} + u_t$

$u_t^2 = \zeta + \alpha u_{t-1}^2 + w_t$

moment condition for the first Regression

$E[(y_t - c - \phi y_{t-1})] = 0$

$E[(y_t - c - \phi y_{t-1}) y_{t-1}] = 0$

moment condition for ARCH 1

$E[u_t^2 - \frac{\zeta}{1-\alpha}] = 0$ z_t instrumental variable.

$E[(u_t^2 - \zeta - \alpha u_{t-1}^2) z_t] = 0$ lagged value

Instrumental variable & Endogenous

Simultaneous Bias.

$y_t = \beta' x_t + \varepsilon_t$ $y = x\beta + \varepsilon$ x Endogenous

Find IV z: ① relevant $E(z_t x_t) \neq 0$

② Valid $E(z_t \varepsilon_t) = 0$

Estimation can be done by:

① GMM $E[(y_t - x_t'\beta) z_t] = 0$

$$\beta = \sum x_t y_t \left(\sum x_t z_t'\right)^{-1}$$

② 2SLS 2 Stage Least Square

Idea. Use IV z_t. do a projection on x_t "clean up"

$$x_t = \hat{\delta}' z_t + e_t \qquad \hat{\delta} = \frac{\sum x_t z_t}{\sum z_t^2} \qquad \hat{x}_t = \hat{\delta} z_t$$

$$\hat{\beta}_{IV} = \frac{\sum y_t \hat{x}_t}{\sum \hat{x}_t^2} = \frac{\sum z_t y_t}{\sum z_t^2} \cdot \hat{\delta}^{-1} = \frac{\sum z_t y_t}{\sum x_t z_t}$$

Estimation Methods: GMM / OLS / MLE

OLS: Best linear Project. works for linear model. Cannot handle latent Variables

MLE: Can Recursively apply. can handle latent variable.
Efficient. Limited by Assumption

GMM: Very General. good for complicated Model. Trade-off

Bootstrap & Simulation

Difference ?

① Initial $z_{1|0}$ $z_{t|t-1}$

$$P_{t|t-1} = E[(z_{t|t-1} - z_t)(z_{t|t-1} - z_t)']$$

② forecasting

$$E(y_t \mid x_t, z_t) = A'x_t + H'z_t$$

$$y_{t|t-1} = A'x_t + H'z_{t|t-1}$$

$$y_t - y_{t|t-1} = H'(z_t - z_{t|t-1}) + w_t$$

$$E[(y_t - y_{t|t-1})^2] = H'P_{t|t-1}H + R$$

③ nowcasting

$$z_{t|t} = z_{t|t-1} + \beta(y_t - y_{t|t-1}) \qquad \beta = \frac{E[(z_t - z_{t|t-1})(y_t - y_{t|t-1})]}{E[(y_t - y_{t|t-1})^2]}$$

$$= P_{t|t-1} \cdot H' \, (H'P_{t|t-1}H + R)^{-1}$$

$$P_{t|t} = E[(z_t - z_{t|t})(z_t - z_{t|t})']$$

$$= E[[z_t - z_{t|t-1} - \beta(y_t - y_{t|t-1})][z_t - z_{t|t-1} - \beta(y_t - y_{t|t-1})]']$$

$$= P_{t|t-1} - P_{t|t-1}H(H'P_{t|t-1}H + R)^{-1}H'P_{t|t-1}$$

④ update

$$z_{t+1|t} = F z_{t|t}$$

$$E[(z_{t+1} - z_{t+1|t})(z_{t+1} - z_{t+1|t})'] = FP_{t|t}F' + Q$$

$$= P_{t+1|t}$$

Estimation MLE: $y_{t|t-1} \sim N(A'x_t + H'\hat{z}_{t|t-1},\ H'P_{t|t-1}H + R)$

several other questions:

Theorem. $\underset{T\times 1}{Y} = \underset{T\times k_1}{X_1}\underset{k_1\times 1}{\beta_1} + \underset{T\times k_2}{X_2}\underset{k_2\times 1}{\beta_2} + \varepsilon \qquad X = [X_1 \quad X_2]_{T\times(k_1+k_2)}$

regression I: $\begin{bmatrix}\hat{\beta}_1 \\ \hat{\beta}_2\end{bmatrix} = (X'X)^{-1}(X'Y)$

regression II: ① $Y = X_1\hat{\phi}_1 + \hat{\varepsilon}_{1\,T\times 1}$

② $X_{2,i}\ (i = 1 \cdots k_2) = X_1\hat{\phi}_{2,i} + \hat{\varepsilon}_{2,i\ T\times k_2}$ k_2 regressions

③ $\hat{\varepsilon}_1 = \hat{\varepsilon}_{2,i} \cdot \beta_2^* + u$

$\hat{\beta}_2^* = \hat{\beta}_2$

Basically. regress $(y_t - \bar{y}) = (x_t - \bar{x})'\beta + \varepsilon_t$

regress $y_t = [1 \quad x_t'] \begin{bmatrix}\alpha \\ \beta\end{bmatrix} + \varepsilon_t$

gives same $\hat{\beta}$

a simple proof: single factor case $\tilde{X} = [\underset{T\times 1}{1} \quad \underset{T\times 1}{X}]$

$\tilde{X}'\tilde{X} = \begin{bmatrix}1' \\ X'\end{bmatrix}[1 \quad X] = \begin{bmatrix}1'1 & 1'X \\ X'1 & X'X\end{bmatrix} = \begin{bmatrix}T & \Sigma x_t \\ \Sigma x_t & \Sigma x_t^2\end{bmatrix}$

$(\tilde{X}'\tilde{X})^{-1} = \frac{1}{ad-bc}\begin{bmatrix}d & -b \\ -c & a\end{bmatrix} = \frac{1}{T\cdot\Sigma x_t^2 - (\Sigma x_t)^2}\begin{bmatrix}\Sigma x_t^2 & -\Sigma x_t \\ -\Sigma x_t & T\end{bmatrix}$

$$\hat{X}'Y = \begin{bmatrix} 1' \\ x' \end{bmatrix} Y = \begin{bmatrix} \sum Y_t \\ \sum x_t Y_t \end{bmatrix}$$

$$\begin{bmatrix} \hat{\alpha} \\ \hat{\beta} \end{bmatrix} = (\hat{X}'\hat{X})^{-1}(\hat{X}'Y) = \frac{1}{T\cdot \sum X_t^2 - (\sum X_t)^2} \begin{bmatrix} \sum X_t^2 & -\sum X_t \\ -\sum X_t & T \end{bmatrix} \begin{bmatrix} \sum Y_t \\ \sum X_t Y_t \end{bmatrix}$$

$$= \frac{1}{''} \begin{bmatrix} \sum X_t^2 \sum Y_t - \sum X_t \sum X_t Y_t \\ -\sum X_t \sum Y_t + T \sum X_t Y_t \end{bmatrix}$$

Note that $\sum (X_t - \bar{X})^2 = \sum (X_t^2 - 2\bar{X} X_t + \bar{X}^2)$

$$= \sum X_t^2 - 2 \underbrace{\sum X_t}_{T\cdot \bar{X}} \cdot \bar{X} + T \bar{X}^2$$

$$= \sum X_t^2 - T \cdot \bar{X}^2$$

$$\sum X_t^2 = \sum (X_t - \bar{X})^2 + T\bar{X}^2$$

$$\sum (X_t - \bar{X})(Y_t - \bar{Y}) = \sum (X_t Y_t - \bar{X} Y_t - \bar{Y} X_t + \bar{X}\bar{Y})$$

$$= \sum X_t Y_t - \bar{X} \sum Y_t - \bar{Y} \underbrace{\sum X_t}_{T\cdot \bar{X}} + T \bar{X} \bar{Y}$$

$$= \sum X_t Y_t - T \cdot \bar{X} \cdot \bar{Y}$$

$$\sum X_t Y_t = \sum (X_t - \bar{X})(Y_t - \bar{Y}) + T \bar{X} \bar{Y}$$

$$\hat{\beta} = \frac{T(\sum X_t Y_t - \frac{1}{T} \sum X_t \sum Y_t)}{T[\sum X_t^2 - \frac{1}{T}(\sum X_t)^2]} = \frac{\sum (X_t - \bar{X})(Y_t - \bar{Y})}{\sum (X_t - \bar{X})^2}$$