MGTF 404 Final Review

Jiahui Shui
jishui@ucsd.edu

Rady School of Management, UCSD

December 6, 2024

Agenda

- Brief Review the lecture slides
- Review sample finals
- Review Lecture Slides!!!

OLS

- Conditional Expectation: $\mathbb{E}[\mathbb{E}[X|Y]] = \mathbb{E}[X]$
- Consider following regression model

$$y_t = x_t'\beta + \varepsilon_t \tag{1}$$

- Critical Assumptions for consistency: $\text{Cov}(x_t, \varepsilon_t) = 0$
- Derivation for OLS estimator:

$$Q(\beta) := \frac{1}{T}\sum_{t=1}^{T}(y_t - x_t'\beta)^2, \quad \hat{\beta}^{OLS} = \underset{\beta}{\text{argmin}}\, Q(\beta) \tag{2}$$

- First Order Condition (FOC):

$$\frac{\partial Q}{\partial \beta_i} = 0, \quad \forall i = 1, \cdots, K \tag{3}$$

OLS - Hypothesis Test

- t-test: Suppose that we estimate the following regression:

$$y_t = \beta_0 + \beta_1 x_t + \varepsilon_t \tag{4}$$

we get $\hat{\beta}_1 = 1.841$, the standard error is $\mathrm{SE}(\hat{\beta}_1) = 0.423$. Also, we have $T = 30$ observations. We want to test

$$H_0 : \beta_1 = 1 \quad \text{v.s.} \quad H_1 : \beta_1 \neq 1 \tag{5}$$

- t-statistics: $t = \frac{\hat{\beta}_1 - 1}{\mathrm{SE}(\hat{\beta}_1)} = (1.841 - 1)/0.423 = 1.988$
- Degree of freedom: $T - K = 30 - 2 = 28$

Stationary

- Let $\{x_t\}_{t=0}^{\infty}$ be a time series.
- If $\{x_t\}$ satisfies
  - $\mathbb{E}[x_t] = \mu$, does not depend on t.
  - $\forall j$, $\mathrm{Cov}(x_t, x_{t-j}) = \gamma(j), \forall t > j$, does not depend on t.
- Then we call $\{x_t\}$ is covariance stationary, or weakly stationary.
- We will take about non-stationary processes later.
- AR(p) process:

$$Y_t = \sum_{i=1}^{p} \phi_i Y_{t-i} + \varepsilon_t \tag{6}$$

- ε_t has to be white noise.
- Tests? Box-Pierce Test; Ljung-Box; Durbin Watson Test

Bias

- When ϕ is close to 1, then AR(1) is persistent.
- Consider the following regression system

$$\begin{cases} y_t = \alpha + \beta x_{t-1} + \varepsilon_t \\ x_t = c + \phi x_{t-1} + u_t \end{cases} \tag{7}$$

- Then

$$\mathbb{E}[\hat{\phi} - \phi] = -\frac{1 + 3\phi}{T} + O(1/T^2), \quad \mathbb{E}[\hat{\beta} - \beta] = \frac{\sigma_{\varepsilon u}}{\sigma_u^2} \mathbb{E}[\hat{\phi} - \phi] \tag{8}$$

VAR

- Consider the following regression

$$Z_t = \alpha + \Phi Z_{t-1} + \varepsilon_t \tag{9}$$

- Stationary condition: $\max \quad |\lambda(\Phi)| < 1$
- Granger Casuality Test:

$$\begin{cases} r_t = \alpha_1 + \beta_{11} r_{t-1} + \beta_{12} \sigma_{t-1} + \varepsilon_{1,t} \\ \sigma_t = \alpha_2 + \beta_{21} r_{t-1} + \beta_{22} \sigma_{t-1} + \varepsilon_{2,t} \end{cases} \tag{10}$$

- Hypothesis 1: "σ_{t-1} Granger-causes r_t", we test $\beta_{12} = 0$
- Hypothesis 2: "r_{t-1} Granger-causes σ_t", we test $\beta_{21} = 0$

Non-stationary Processes/ DF Test

- Suppose that p_t and q_t are non-stationary. Consider regression: $p_t = \gamma q_t + \varepsilon_t$
- $\hat{\gamma}$ does not converge to true γ.
- t-stats is also not consistent.
- Tests for non-stationarity: consider the following regression:

$$p_t = c + \phi p_{t-1} + \varepsilon_t \tag{11}$$

- Dickey-Fuller (DF) Test: $H_0 : \phi = 1$, $H_1 : \phi < 1$
- Test: $t = \frac{\hat{\phi}-1}{\mathrm{SE}(\hat{\phi})}$
- t is **NOT** asymptotic normal.

ADF Test

- $p_t = c + \phi p_{t-1} + \varepsilon_t$
- Now suppose that $\varepsilon_t \sim \mathrm{ARMA}(p, q)$.
- To get rid of $\mathrm{ARMA}(p, q)$ parameters, which will influence DF test, we run the following regression:

$$p_t = \phi p_{t-1} + \zeta_1 \Delta p_{t-1} + \cdots + \zeta_k \Delta p_{t-k} + v_t \tag{12}$$

- Test statistics: $t = \frac{\hat{\phi}-1}{\mathrm{SE}(\hat{\phi})}$
- Equivalent way: Let $\Delta p_t := p_t - p_{t-1}$. Then run the following regression:

$$\Delta p_t = c + \delta p_{t-1} + \zeta_1 \Delta p_{t-1} + \cdots + \zeta_k \Delta p_{t-k} + v_t \tag{13}$$

- Test statistics: $t = \frac{\hat{\delta}}{\mathrm{SE}(\hat{\delta})}$

Spurious Regression

- Suppose that $p_t = p_{t-1} + u_t$, $q_t = q_{t-1} + v_t$ and $\text{Cov}(u_t, v_t) = 0$
- Run regression: $p_t = \gamma q_t + \varepsilon_t$
- $\hat{\gamma}$ will not converge to 0
- $t = \frac{\hat{\gamma}}{\text{SE}(\hat{\gamma})}$ will diverge
- R^2 does not converge to 0
- However, if you believe that p_t and q_t are cointegrated, i.e. the linear combination of p_t and q_t is stationary, then it is OK to run regression.
- Another way is taking first difference. $\Delta p_t = p_t - p_{t-1}$. Then run regression (if Δp_t is stationary) $\Delta p_t = \gamma \Delta q_t + \varepsilon_t$

Empirical Results in Finance

- Fama-French: Three factors:

$$R_t^i - r_f = \alpha_i + \beta_{1,i}(R_t^M - r_f) + \beta_{2,i}R_t^{\mathrm{SMB}} + \beta_{3,i}R_t^{\mathrm{HML}} + \varepsilon_t \tag{14}$$

- SMB (Small Minus Big) = Historic excess returns of small-cap companies over large-cap companies
- HML (High Minus Low) = Historic excess returns of value stocks (high book-to-price ratio) over growth stocks (low book-to-price ratio)
- We believe small firms are riskier and command a risk premium relative to large firms.
- We believe value firms are riskier and command a risk premium relative to growth firms

Empirical Results in Finance-Campbell-Shiller

- Campbell-Shiller decomposition

$$dp_t = -\frac{\kappa}{1-\rho} + \mathbb{E}_t\left[\sum_{j=0}^{\infty}\rho^j r_{t+j+1} - \sum_{j=0}^{\infty}\rho^j \Delta d_{t+j+1}\right] \qquad (15)$$

Meaning: The log dividend yield dp_t has to forecast either

  - Future returns, with a postive sign
  - Future dividend growth, with a negative sign
  - Or both

MLE

- Make sure you know the pdf of normal: $N(\mu, \sigma^2)$

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \tag{16}$$

- You should be able to derive the MLE for β in the following regression

$$y_t = \beta x_t + \varepsilon_t, \quad \varepsilon_t \sim N(0, \sigma^2), \quad \text{i.i.d} \tag{17}$$

Volatility Models

- Realized Volatility: Note that

$$\mathrm{Var}_t(r_{t+1}) = \mathbb{E}_t[(r_{t+1} - \mathbb{E}_t(r_{t+1}))^2] = \mathbb{E}_t[r_{t+1}^2] - (\mathbb{E}_t[r_{t+1}])^2$$

- One way to estimate is using residual: $e_t = r_{t+1} - \mathbb{E}_t[r_{t+1}]$
- Implied Volatility: Volatility inferred from option prices (BS-formula)
- Run AR(1) regression for volatility: R^2 is large. Volatility is much more forecastable than are simple returns!
- Relation between Bi-Power Variation (BV) and Realized Volatility (RV):

$$RV_{t+1} - BV_{t+1} \xrightarrow{\Delta \to 0} \sum_{t<s\leq t+1} \kappa^2(s) \tag{18}$$

ARCH & GARCH

- Return:

$$r_t = \mu + \phi r_{t-1} + u_t \tag{19}$$

- ARCH: u_t^2 follows AR(p)

$$u_t^2 = \zeta + \alpha_1 u_{t-1}^2 + \alpha_2 u_{t-2}^2 + \cdots + \alpha_p u_{t-p}^2 + w_t \tag{20}$$

- Alternative representation: $u_t = \sqrt{h_t} v_t$
- GARCH(1,1):

$$h_t = \zeta + \delta h_{t-1} + \alpha u_{t-1}^2 \tag{21}$$

- GARCH(p,q):

$$h_t = \zeta + \sum_{i=1}^{q} \alpha_i u_{t-i}^2 + \sum_{j=1}^{p} \beta_j h_{t-j}$$

- IGARCH(1,1) is GARCH(1,1) when $\delta + \alpha = 1$
- AsyGARCH: Negative surprises increase volatility more than positive surprises

ARCH & GARCH

- Can you estimate ARCH with OLS? Yes
- Can you estimate GARCH with OLS? No
- Persistence? ARCH: No; GARCH: Yes
- For most practical purposes a GARCH(1,1) is GREAT
- Test of Homoskedasticity vs ARCH(p):
- Consider

$$u_t^2 = \zeta + \alpha_1 u_{t-1}^2 + \alpha_2 u_{t-2}^2 + \cdots + \alpha_p u_{t-p}^2 + w_t \tag{22}$$

- Engle (1982): Lagrange Multiplier-type test
  - Regress Y_t on Y_{t-1} to get $\hat{u}_t$
  - Regress $\hat{u}_t^2$ on $\hat{u}_{t-1}^2, \cdots, \hat{u}_{t-p}^2$.
  - Get R^2 from this regression
  - $T \cdot R^2 \sim \chi^2(m)$

ARCH & GARCH

- Testing Nested GARCH(p,q) Models:
- Estimating GARCH(1,1), get log-likelihood function $\Lambda(\theta_0)$
- Estimating GARCH(2,2), get log-likelihood function $\Lambda(\theta_1)$
- Note that GARCH(1,1) is a special case of GARCH(2,2) (take $\alpha_2 = \delta_2 = 0$)
- Hence $\Lambda(\theta_0) < \Lambda(\theta_1)$
- Likelihood ratio:

$$LR = 2(\Lambda(\theta_1) - \Lambda(\theta_0)) \sim \chi^2(2) \tag{23}$$

Sample Final (2015) Question 1

- (1) The GARCH(1,1) is a simple and elegant model of volatility and its in-sample and out-of-sample performance are difficult to beat by other, more complicated models
- True. See Lecture Slides 8 page 25.
- What are the advantages and disadvantages of GARCH? Why GARCH is successful?
- See Lecture Slides 8 page 35.

Sample Final (2015) Question 1

- (2) Stock returns at 5-minute, daily, weekly, and monthly frequency are equally serially uncorrelated.
- False. This question is related to market microstructure. At higher data frequencies, noise is less likely to be smoothed out. Consequently, high-frequency data exhibits lower serial correlation.
- (4) The CAPM is a useful benchmark because it explains a great deal of the cross-sectional variation in returns.
- Actually this question can be answered both sides. CAPM is simple but useful. However it can not explain a great deal of cross-sectional returns.

Sample Final (2015) Question 1

- (5) 3-factor Fama-French model can be estimated using only time-series regressions. In fact, any 3-factor model can be estimated using only one time-series regression.
- False. Note that CAPM and Fama-French are cross-sectional regressions. Not time series regression.
- (9) A non-linear model with more parameters will always do a better job at forecasting out-of-sample than a simpler linear model with fewer parameters.
- False. Model with more parameters is subject to large estimation errors that may lead to poor out-of-sample performance.

Sample Final (2015) Question 3

Simplified question: Suppose that you have the information of option markets. What information from the options market do you think will be useful to estimate conditional mean and conditional variance of stock return?
(1) Suppose your variable is X. Then how will you estimate $\mu_{t|t-1}$ using $X_{i,t-1}$?
(2) Why it is not a good idea to run regression:

$$R_{i,t} = \eta_i + \delta_i X_{i,t} + u_{i,t} \tag{24}$$

- Implied volatility will help us to predict returns in stock market.
- For question (2), think about the $\mathrm{Cov}(X_{i,t}, u_{i,t})$. Stock price $P_{i,t}$ is used to calculate $R_{i,t}$. Also, $P_{i,t}$ is in the derivation of $X_{i,t}$. Then there is endogeneous problem.