Test of Math/Stats/Finance/Programming

Jingtao Zheng

April 20, 2018

Note

- Please explain your reasoning concisely and precisely
- Note that the your coding habit will also be evaluated
- Please type in LaTeXif you can
- You have 2 days to finish this test
- Submit your type-up to: zjt9101@gmail.com, along with your code in Python (preferred) or R

1 ARMA

Consider the AR(p) process

$$X_t = \mu + \sum_{k=1}^{p} a_k X_{t-k} + \epsilon_t \tag{1}$$

where $\epsilon_t \sim N(0, \sigma^2)$ iid. We assume the process is stationary and the characteristic roots are not the same.

1. What is the conditional distribtion of $P(X_t < c | X_0 = x_0)$?
2. Calculate the impulse response function (IRF) and half-life.
3. Use Monte-Carlo to verify your analytical result in Part 1, for $X_t = 0.7X_{t-1} - 0.01X_{t-2} + \epsilon_t$ and $\sigma = 1, x_0 = 1, c = 2$
4. Write the process in MA(∞)
5. Given a time series, explain how you will:
   - determine p
   - estimate parameters
   - diagnose whether $\epsilon_t \sim N(0, \sigma^2)$ iid is a good assumption
   - evaluate whether this is a good model for the given time series

2 Hidden Markov Chain

Consider the following regime-switching model. The underlying unobservable binary regime $H_t = 0, 1$ is a Markov chain with transition matrix $\begin{bmatrix} 1-p & p \\ q & 1-q \end{bmatrix}$. We can observe time series $X_t = (X_{1t}, X_{2t})$. Conditioning on regime $H_t = h$ $(h = 0, 1)$, the process X_t follows VAR(1):

$$X_t = \mu_h + A_h X_{t-1} + \epsilon_t \tag{2}$$

where $\epsilon_t \sim N(0, \Omega_h)$. Note that the parameters μ_h, A_h, Ω_h can differ from regime to regime.

1. Write the conditional distribution $P(H_t = h_t | H_0 = h_0, X_0 = x_0, X_1 = x_1, \ldots, X_t = x_t)$.
2. Suppose h_0 is drawn from the stationary distribution of the Markov chain (what is it?), write the conditional distribution $P(H_t = h_t | X_0 = x_0, X_1 = x_1, \ldots, X_t = x_t)$.
3. Write the likelihood function $L(X_0 = x_0, X_1 = x_1, \ldots, X_t = x_t | p, q, \mu_h, A_h, \Omega_h)$
4. Given time series $X_t = (x_{1t}, x_{2t}), t = 0, ..., T$, how you will estimate the parameters $p, q, \mu_h, A_h, \Omega_h$. How do you calculate the standard errors?
5. Explain how this model can be useful.

3 Principal Component Analysis

$X_t = (X_{1t}, \ldots, X_{pt})$ is stationary process with mean 0, variance-covariance matrix $V = Cov(X_t)$. We define principal component recursively:

- the first principal component is a column vector a_1 defined as the solution to

$$\max_{a_1' a_1 = 1} Var(a_1' X_t) = a_1' V a_1 \tag{3}$$

- the k-th $(2 \leq k \leq p)$ principal component is a column vector a_k, defined as the solution to

$$\max_{a_k' a_k = 1; \forall j < k, a_k' a_j = 0} Var(a_k' X_t) = a_k' V a_k \tag{4}$$

1. Explain intuitively what the principal component means.
2. Prove that the first principal component is eigen-vector corresponding to the largest eigenvalue.
3.
   - Download the data from FRED (Python) API for US Constant Maturity Rate with tenor 1,2,3,5,7,10 years at daily frequency, from 2000 to today.
   - Use rolling window of 252 days to calculate the first 3 principal components (hint: you should keep the sign convention consistent).
   - Show how the first 3 principal components evolve over time in the same plot, with multi-axes.
   - Explain why the first 3 principal components are often labeled as level, slope and curvature. Explain their movement.

4 Two Dimension Vasicek Model

Let the short-rate (instantaneously risk-free rate) be

$$r_t = X_t + Y_t \tag{5}$$

Under risk neutral measure Q

$$dX_t = (a - bX_t)dt + cdZ_1 \tag{6}$$
$$dY_t = (\mu - \gamma Y_t)dt + \eta dZ_2 \tag{7}$$

where dZ_1, dZ_2 are **independent** Brownian motion. Therefore the processes X_t, Y_t are also independent of each other. Define the time t price of the zero-coupon bond maturing at time T as:

$$P_t = E^Q[e^{-\int_t^T r_u du}] \tag{8}$$

Define the (continuously compounded) yield is

$$y_t = -\frac{1}{T-t}\log P_t \tag{9}$$

1. Solve y_t analytically.

   Hint: make sure you understand how to solve the classical Vasicek model (1 dimension case) before you try to tackle this problem. You may need to solve the PDE in this procedure

   $$xF = F_t + (a - bX_t)F_x + \frac{c^2}{2}F_{xx} \tag{10}$$

   where $F = F(t, x; T)$

5 Machine Learning

Write here if you have any finance/econ related projects where you apply sophisticated machine learning techniques. Show your result quantitatively.

6 User Interface

Write here if you have any projects involving creating interactive user interface, preferrably written in Python. Illustrate in relevant pictures.