Financial Statistics coursework.

Please address all questions.
You are invited (compulsory!) to develop your own R code.

Question 1 (CAPM)

Let r_{it} be the rate of return from holding the ith stock, let r_{Mt} be the (equity) market index return, let r_{ft} be the risk-free rate.

$$r_{it} - r_{ft} = \alpha + \beta(r_{Mt} - r_{ft}) + u_t, \tag{1}$$

and

$$r_{it} - r_{ft} = \alpha + \beta_1[D_t(r_{Mt} - r_{ft})] + \beta_2[(1 - D_t)(r_{Mt} - r_{ft})] + \beta_3(r_{Mt} - r_{ft})^2 + u_t, \tag{2}$$

where $D_t = 1$ if $(r_{Mt} - r_{ft}) > 0$ and $D_t = 0$ otherwise.

- Use the dataset data coursework Q1.
- Write a programme which estimates the two above linear regression model using OLS over the whole sample period.
- With respect to the null hypothesis

  $$H_0: \ \ \beta_1 = \beta_2$$

  write down a code to perform the F test and apply the tests to the data.
- With respect to the null hypothesis

  $$H_0: \ \ \alpha = 0$$

  write down a code to perform the t test and apply the tests to the data.

Question 2 (probability of a positive asset return)

Let R_t be an equity market index return. Construct an indicator variable, that is a variable made by zeros and ones, as follows:

$$Y_t = \begin{cases} 1 & \text{if } R_t > 0 \\ 0 & \text{if } R_t \leq 0 \end{cases}$$

We know that we can re-write a conditional probability as a conditional expectation:

$$Pr(R_{t+1} > 0 \mid X_t) = E(Y(R_{t+1}) \mid X_t).$$

where now X_t denotes a set of regressors (predictors) observed at time t. We can then use the linear regression model to estimate this conditional probability

Given a sample $(Y_1, X_1, ..., Y_n, X_n)$ consider then

$$Y_{t+1} = \beta' X_t + \varepsilon_t, \quad t = 1, ..., n-1.$$

The estimated probability will then be

$$\hat{P}(R_{t+1} > 0 \mid X_t) = \hat{\beta}' X_t, \quad t = 1, 2, ..., n-1.$$

where $\hat{\beta}$ is the OLS estimator. To evaluate the forecasting performance typically we construct the hit ratio indicator Z in the following way:

$$Z_t(\alpha) = \begin{cases} 1 & \text{if } \hat{P}(R_t > 0 \mid X_{t-1}) > \alpha \text{ and } R_t > 0, \\ 1 & \text{if } \hat{P}(R_t > 0 \mid X_{t-1}) \leq \alpha \text{ and } R_t \leq 0, \\ 0 & \qquad\qquad \text{otherwise.} \end{cases}$$

for a fixed $0 < \alpha < 1$. Then set

$$Z(\alpha) = \frac{\sum_{t=2}^{n} Z_t(\alpha)}{n-1}.$$

- Use the dataset data coursework Q2.
- Write a programme which estimates the linear regression model using OLS over the whole sample period.

- For $\alpha = 1/2$, construct $Z_t(\alpha)$ and $Z(\alpha)$ trying to find the combination of predictors that gives the best performance in term of $Z(\alpha)$ (the maximum value of $Z(\alpha)$).
- Finally, for this combination of regressors found in the previous point, can you find the value of α for which $Z(\alpha)$ is maximized?

Question 3 (CIR model for the term structure of interest rate)

The discrete time version of the CIR model for the term structure postulates that the short-term interest rate r_t satisfies the following dynamic equation:

$$r_t = \mu(1-\phi) + \phi r_{t-1} + r_{t-1}^{\frac{1}{2}} u_t,$$

with $u_t \sim NID(0, \sigma^2)$.

Write the code to estimate this model using MLE, deriving also the asymptotic covariance matrix using the Gaussian loglikelihood:

$$l(\theta) = \sum_{t=2}^{T} log f(r_t \mid r_{t-1}, \theta)$$

where

$$f(r_t \mid r_{t-1}, \theta) = \frac{1}{\sqrt{2\pi r_{t-1}\sigma^2}} e^{-\frac{0.5}{r_{t-1}\sigma^2}(r_t - \mu(1-\phi) - \phi r_{t-1})^2},$$

and

$$\theta = (\mu, \phi, \sigma^2)'.$$

How does its fit go as compared with the Vasicek model

$$r_t = \mu(1-\phi) + \phi r_{t-1} + u_t,$$

with $u_t \sim NID(0, \sigma^2)$?

Summarizing:

- Evaluate the MLE, and its asymptotic covariance matrix, for the for the CIR model using both the US 1-month interest rate r_t^{US} and the UK 1-month interest rate r_t^{UK} in the dataset data coursework Q3.

- Evaluate the MLE, and its asymptotic covariance matrix, for the for the Vasicek model using both the US 1-month interest rate r_t^{US} and the UK 1-month interest rate r_t^{UK} in the dataset data coursework Q3. (You can simply adapt the previous code to estimate this latter model.)
- Comment on the results.