A notebook of financial time series models

Yuwen Long

2024

1 Introduction

This notebook focus on fundamental time series models in finance, which of course can be scaled up to other fields. The basic ideas, concepts, theorems and propositions are discussed in details to provide further insight on specific models.

Before the content, some notation is to be introduced. The price of a certain underlying at time t is denoted by P_t. It is common to use the logarithmic rate of return, $r_t = \ln\left(\frac{P_t - P_{t-1}}{P_{t-1}} + 1\right)$, in financial analysis.

2 Stationarity and White Noise Series

2.1 Stationarity

A stochastic process $\{a_t\}$ is called *stationary* if any joint distribution of a collection of its random variables $\{a_t, a_{t+1}, \ldots, a_{t+k}\}_k$ is only related to k, that is its distribution is *time invariant*. However, this *strict* stationarity is hardly satisfied in real data. So the *weak stationarity* is discussed more, which says $\mathrm{Cov}\left(a_t, a_{t+k}\right) = \gamma_k$ which is a function only of the lagged time k. And, a weak stationary series also defined to hold that $\mathbb{E}\left(a_t\right) = \mu, \mathrm{Var}\left(a_t\right) = \sigma^2 < \infty$ for all $t = 1, 2, \ldots, T$.

2.2 White Noise Series

A series of random variables $\{a_t\}$ is called *white noise series* if $\mathbb{E}\left(a_t\right) = 0, \mathrm{Var}\left(a_t\right) = \sigma^2 < \infty, \mathrm{Cov}\left(a_t, a_{t+k}\right) = 0$, which is called *Gauss-Markov Condition* in linear models. It is worthwhile to pointed that a white noise series is self-uncorrelated but not necessarily independent.

2.3 Martingale Difference Sequence

We a stochastic process $\{a_t\}$ a *Martingale Difference Sequence*(MDS) if $a_t \in \mathcal{F}_t$, where $\mathcal{F}_t$ is the information set(the information available) at time t, and $\mathbb{E}\left(a_t | \mathcal{F}_{t-1}\right) = 0$ which means we can know nothing about a_t dependent on the past information. It is obvious that an *iid* series of random variables must be MDS. And using Iterated expectation theorem, one can get a MDS must be white noises. At last, from the definition of stationarity and white noises, it is obvious s stationary series is a subset of white noise series.

3 AR Model

3.1 ACF

The correlation of a time series $\{r_t\}$ is about the components itself, so it is called *Auto Correlation Function*(ACF). First consider the covariance $\mathrm{Cov}\left(r_t, r_{t+k}\right)$. If that time series is stationary, then

$$\mathrm{Cov}\left(r_t, r_{t+k}\right) = \gamma_k. \tag{3.1}$$

And it can be seen that

$$\mathrm{Cov}\left(r_t, r_{t+k}\right) = \gamma_k = \gamma_{-k} = \mathrm{Cov}\left(r_t, r_{t-k}\right), \tag{3.2}$$

since $\mathrm{Cov}\left(r_t, r_{t-k}\right) = \mathrm{Cov}\left(r_{t-k}, r_t\right)$ and let $l = t - k$, then $\mathrm{Cov}\left(r_t, r_{t-k}\right) = \mathrm{Cov}\left(r_l, r_{l+k}\right) = \gamma_k$. Also, $\mathrm{Var}\left(a_t\right) = \gamma_0$.

Then, we can make notes about the correlation as

$$\rho_k = \frac{\mathrm{Cov}\left(r_t, r_{t+k}\right)}{\sqrt{\mathrm{Var}\left(r_t\right)\mathrm{Var}\left(r_{t+k}\right)}} = \frac{\gamma_k}{\gamma_0}, \tag{3.3}$$

Which is the auto correlation function of $\{r_t\}$. More specifically, if we have the sample of $\{r_t\}$, we can calculate the sample covariance as

$$\hat{g}amma_k = \frac{1}{T}\sum_{t=k+1}^{T}\left(r_t - \bar{r}\right)\left(r_{t-k} - \bar{r}\right), \tag{3.4}$$

where $\bar{r} = \frac{1}{T}\sum_{t=1}^{T} r_t$. And the ACF can be calculated as

$$\hat{r}ho_k = \frac{\sum_{t=k+1}^{T}\left(r_t - \bar{r}\right)\left(r_{t-k} - \bar{r}\right)}{\sum_{t=1}^{T}\left(r_t - \bar{r}\right)^2}. \tag{3.5}$$

We can calculated the ACF of a sequence of k to identify which lagged r_{t-k}s significantly influence r_t. An example is illustrated by the following figure from [1].

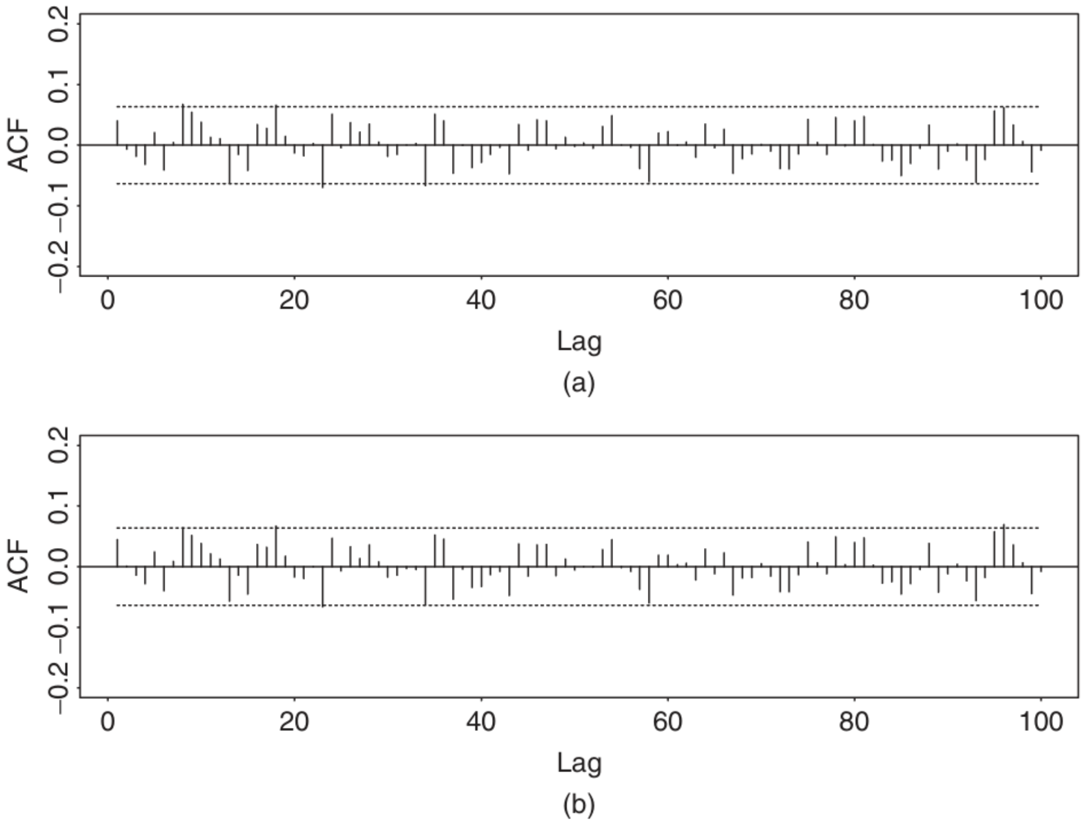

Figure 1: Sample autocorrelation functions of monthly (a) simple returns and (b) log returns of IBM stock from January 1926 to December 2008. In each plot, two horizontal dashed lines denote two standard error limits of sample ACF.

That is, those ACF which exceeds the two standard error limits is considered to be significant.

What's more, there are some test statistics for ACF.

***t*-Test**

For a given positive integer k, we test $H_0 : \rho_k = 0$ vs. $H_1 : \rho_k \neq 0$. The t-statistic is

$$t \text{ ratio} = \frac{\hat{\rho}_k}{\sqrt{\left(1 + 2\sum_{i=1}^{k-1} \hat{\rho}_k^2\right)/T}}. \tag{3.6}$$

If $\{r_t\}$ is a Gaussian stationary series satisfying $\rho_j = 0$ for $j > k$, then the t ration defined above is asymptotically distributed as a standard normal random variable. And the hypothesis test is a two-side test.

Portmanteau Test

The statistic define below test several correlations of r_t are zero jointly. The Portmanteau statistic is

$$\mathcal{Q}^*(m) = T\sum_{i=1}^{m} \hat{\rho}_i \tag{3.7}$$

which test $H_0 : \rho_1 = \cdots = \rho_m = 0$ vs. $H_1 : \rho_i \neq 0$ for some $i \in \{1, \ldots, m\}$. If $\{r_t\}$ is an iid sequence with certain moment conditions, $\mathcal{Q}^*$ is asymptotically a chi-squared random variable with m degrees of freedom.

The Portmanteau statistics is modified by Ljung and Box to increase the power of the test as

$$\mathcal{Q}(m) = T(T+2)\sum_{i=1}^{m}\frac{\hat{\rho}_i^2}{T-i}. \tag{3.8}$$

Both statistics above reject H_0 if its value is larger than $\chi^2_{1-\alpha}(m)$, the $100(1-\alpha)$th percentile of a chi-squared distribution with m degrees of freedom.

3.2 AR(1) Model

References

[1] RUEY S. TSAY. *Analysis of Financial Time Series.* A JOHN WILEY & SONS, INC., PUBLICATION.