23/04

Status	Not Started
Project	Statistic
Tags	

Cointegration

Right-tailed statistic

Stationary Processes

https://www.probabilitycourse.com/chapter10/10_1_4_stationary_processes.php

Serial Correlation

https://www.sfu.ca/~dsignori/buec333/lecture 17.pdf

http://home.bi.no/fgl96027/l24.pdf

Unit root test

https://www.sfu.ca/~dsignori/buec333/lecture 17.pdf

https://faculty.washington.edu/ezivot/econ584/notes/unitroot.pdf

https://www.stata.com/meeting/uk20/slides/UK20_Otero.pdf

Augmented Dickey-Fuller test

$$y_t = y_{t-1} + \epsilon_t \qquad Random\ Walk$$

$$y_t = \alpha + y_{t-1} + \epsilon_t \quad RW drift$$

$$y_t = \alpha + \gamma t + y_{t-1} + \epsilon_t \quad RW\ drift, Trend$$

Dickey-fuller test

$$y_t = \alpha + \gamma t + \beta y_{t-1} + \epsilon_t$$

$$y_t - y_{t-1} = \alpha + \gamma t + (\beta - 1) y_{t-1} + \epsilon_t$$

$$\Delta y_t = \alpha + \gamma t + \delta y_{t-1} + \epsilon_t$$

If $\delta = 0$ → $\beta = 1$ (non-stationary)

If $\delta < 0$ → $\beta < 1$ (stationary)

Augmented Dickey-Fuller

$$\Delta y_t = \alpha + \gamma t + \delta y_{t-1} + \sum_{i=1}^{k} \phi_i \Delta y_{t-i} + \epsilon_t$$

If $\delta = 0$ → $\beta = 1$ (non-stationary) → Null Hypothesis

If $\delta < 0$ → $\beta < 1$ (stationary) → Alternative Hypothesis

Cointegration-based tests

Cointegration is data testing that finds if there's a relationship between two or more time-related series.

GSADF

The general idea is to use a right-tailed ADF unit root test. Empirical tests of rational bubbles before the development of the Phillips et al. (2011) test often used the standard unit root test, which does not allow for extreme behavior. The main idea in these tests is to employ ADF-style regressions that shift the start and end dates of a rolling window.

A stationary process indicates that a price returns to its fundamental value, whereas a nonstationary unit root result suggests persistent price deviation from the fundamental value, which indicates the existence of bubbles. The largest test statistics are used as the right-tailed statistic.

ADF regression

For example, if we let y_t be a time-series variable with T total observations, $r_1$and r_2 are the starting position and ending position of the rolling windows, and the size of the window is $r_w = r_1 - r_2$. The typical ADF regression can be written as follows:

$$\Delta y_t = c_{r_1,r_2} + \beta_{r_1,r_2} y_{t-1} + \sum_{i=1}^{k} \phi^i_{r_1,r_2} \Delta y_{t-i} + \epsilon_t$$

where y_t denotes the logarithm of the real price or return on assets, Δy_t represents the first differences, c, β and ϕ are parameters to be estimated, and the error term is expected to follow a normal distribution, i.e. $\varepsilon_t \sim iidN(0, \sigma^2_{r_1,r_2})$. k lagged difference terms are included to control for serial correlation, which is determined by the information criteria. Parameters r_1 and r_2 are the fractions of the total sample size and denote the starting and ending points of a subsample period.

The null hypothesis of a standard ADF test is that y_t has a unit root against the alternative of $\beta < 0$; alternatively, it is stationary. Phillips et al. (2011) extend the basic ADF principle to allow for the right-sided alternative $\beta > 0$ or to consider an extreme alternative hypothesis. As discussed in the study by Phillips et al.2011, right-sided unit root tests are informative about mildly explosive behavior in data, and hence, they can be used as a form of market warning alert against mispricing. Therefore, we have the following test statistics:

$$ADF_{r_1}^{r_2} = \frac{\hat{\beta}_{r_1,r_2}}{s.e.(\hat{\beta}_{r_1,r_2})} = \frac{\hat{\beta}_{r_1,r_2}}{SE(\hat{\beta}_{r_1,r_2})}$$

The Standard ADF Test

For the ADF test, the statistic in equation (2) is obtained by estimating regression (1) on the full sample of observations, i.e., by setting $r_1 = 0$ and $r_2 = 1$. Under the null of a unit root, the limit distribution of ADF_0^1 is given by,

$$\frac{\int_0^1 W dW}{\left(\int_0^1 W^2\right)^{1/2}},$$

where W is a Wiener process. Testing for exuberance entails comparing the ADF_0^1 statistic with the right-tailed critical value from its limit distribution. In this setting with r_1 and r_2 fixed at 0 and 1, respectively, the alternative hypothesis is that of exuberance over the entire sample. Because the standard ADF test is not consistent with changes in regime, it exhibits extremely low power in the presence of boom-bust episodes. In fact, nonlinear dynamics, such as those displayed by periodically-collapsing speculative bubbles, frequently lead to finding spurious stationarity even when the process under examination is inherently explosive (see Evans 1991).

Where SE is https://simple.wikipedia.org/wiki/Standard_error

We normalize the total observations to 1 and denote the share of data from the full sample by r_w. Therefore, the number of observations in each window is $T_w = [r_w T]$, and $[\bullet]$ represents the integer part of the number in the brackets. If the smallest window size is given as r_0, then a sequence of test statistics based on forward-expanding samples can be derived. A sup of these statistics is defined as $sup[ADF(SADF)]$ test statistics as follows:

$$SADF(r_0) = sup_{r_2 \in [0,1]} ADF_0^{r_2}$$

💡 Về cơ bản SADF có điểm bắt đầu r_1 cố định, kích thước cửa sổ trượt do người dùng quy định sau đó được mở rộng dần.
Quy trình :
1. Quan sát đầu tiên của mẫu Bắt đầu từ
$r_1 = 0$, kết thúc ở r_2 đặt theo r_0 → $r_w = r_2$
2. Khi tăng kích thước cửa sổ thì
$r_2 \in [r_0, 1]$ trong mỗi lần quan sát. Mỗi ước tính mang lại thống kê ADF được ký hiệu là ADF_{r_2}.
3. Cuối cùng
$r_2 = 1$ và thống kê ADF sẽ là ADF_1
4. SADF là giá trị nhỏ nhất (supremum) của chuỗi
ADF_{r_2} đối với các $r_2 \in [r_0, 1]$:

$$SADF(r_0) = sup_{r_2 \in [r_0,1]}\{ADF_{r_2}\}.$$

GSADF : các bước thực hiện tương tự như kiểm định SADF nhưng tổng quát hơn với điểm bắt đầu
r_1 cũng có thể thay đổi trong khoảng $[0, r_2 - r_0]$.

$$GSADF(r_0) = sup_{r_2 \in [r_0,1], r_1 \in [0, r_2 - r_0]}\{ADF_{r_2}\}.$$

PWY đề xuất so sánh từng yếu tố của chuỗi ADF_{r_2} ước tính được với các giá trị tới
hạn tương ứng với kiểm định phía phải của thống kê ADF chuẩn để xác định bong bóng bắt đầu tại thời điểm
r_2

Note :

r_0 được sử dụng để định rộng cửa sổ ban đầu ước tính $r_0 = \frac{1.8}{\sqrt{t}} + 0.01$

$r_w = r_2 - r_1$

In reality, there are often multiple bubbles. Although the SADF test performs well in identifying a single boom and bust in a series, it may not consistently identify the origination and termination when there are multiple episodes of exuberance. To cope with this problem and to detect more than one bubble in the series, Phillips et al. (2015) propose generalized SADF (GSADF) to deal with multiple events of boom and bust in a single series. Unlike the SADF method, GSADF changes not only the initial observation of the subsample but also the end point. The GSADF statistic is given as follows:

$$GSADF(r_0) = sup_{r_1 \in [0, r_2 - r_0]} SADF(r_0) = sup_{r_2 \in [0,1],\ r_1 \in [0, r_2 - r_0]} ADF_{r_1}^{r_2} \quad (1)$$

In equation (1), if $GSADF(r_0)$ is greater than the right tail critical value, we reject the null hypothesis in favor of the explosive alternative hypothesis. Phillips et al. (2015) developed the limited distribution of these test statistics and provided asymptotic critical values for both the SADF and GSADF statistics.

In the case of multiple bubbles, the ADF test, like earlier unit root and cointegration-based tests for explosive behavior, may find pseudo-stationary behavior and is typically less powerful in identifying more bubbles after the first. Therefore, a date-stamping strategy with a double recursive test procedure and backward supremum ADF(BSADF) test method was proposed:

$$BSADF(r_0) = sup_{r_1 \in [0, r_2 - r_0]} ADF_{r_1}^{r_2}$$

The GSADF test can then be written as follows:

$$GSADF(r_0) = sup_{r_2 \in [0,1]} [BSADF_{r_2}(r0)]$$

The sequence of the BSADF test statistics is compared to the critical value of the GSADF test. Whenever the statistic exceeds the critical value, there is a bubble, which is then considered to burst when the BSADF statistic falls below the critical value. The critical values for the GSADF test are taken from 1000 Monte Carlo simulations with the actual sample size. The simulated critical values of the 95% confidence interval in empirical studies are available in the attachment.

BSADF

LPPLS

The LPPLS model is used to measure the degree of growth in price beyond exponential growth under positive feedback, and it assumes that the asset price will rupture at critical time Tc. The logarithmic oscillation form is used because the oscillation frequency increases as the critical time approaches the linear scale, but the oscillation frequency is constant on a logarithmic scale. The model expression is as follows:

$$ln[P(t)] \approx A + B_0(T_c - T)^{\beta}\{1 + Ccos[\omega ln(T_c - T) + \phi]\}$$

$ln[P(t)]$ is the logarithm of the asset price at time T and T_c represents the most likely time for a change in the regime at which the growth rate changes, signifying a change in regime. The regime change may be the time of the crash, the burst of the bubble, or the fading of the bubble. β is the power exponent, which measures the acceleration of the price increase $(0 < \beta < 1)$.

The component $A + B_0(T_c - T)^{\beta}$ in equation captures super-exponential growth. This component is the power law singular component, and it embodies the positive feedback mechanism of bubble development. A represents the asset price when the bubble reaches critical point $T_c(ln[P(Tc) > 0]).B_0$ indicates the direction of price change. For super-exponential growth, it is required that $0 < \beta < 1$, while positive and negative bubbles are characterized by $B_0 < 0$ and $B_0 > 0$, respectively. $B_0 < 0$ indicates upward acceleration, while $B_0 > 0$ indicates downward acceleration. ωω is the angular frequency of bubble fluctuation, and ϕ is the phase parameter, which is positive and less than 2π. C represents the magnitude of the fluctuation used to quantify the degree of logarithmic periodic oscillation $(|C| < 1)$.

The positive feedback effect between investors promotes exponential growth in prices, and the herding effect eventually leads to a bubble. Moreover, the ratio of rational investors to irrational investors determines power index β, and the convergence of market views contributes to ω.

The LPPLS equation has three linear parameters (A, B_0, C) and four nonlinear parameters $(\beta, \omega, T_c, \phi)$. To simplify the estimation, the three linear parameters, i.e.,A, B_0, C, are usually estimated from the given values of the nonlinear parameters, i.e., β, ω, T_c and ϕ. A common method of estimation for the LPPLS equation is nonlinear least squares. However, the minimization of nonlinear multivariate least squares functions is nontrivial due to the presence of multiple local minima, where well studied algorithms, such as the steepest descent or Newton’s method, usually become trapped. To overcome the complexity of the minimization, a genetic algorithm is used to estimate the parameters of the LPPLS model.

Empirical findings

GSADF results

The GSADF method can be employed to accurately identify bubble periods; however, since GSADF is based on hypothesis testing and uses sample data after the event, it cannot provide us with an accurate timing of a bubble burst. Therefore, this method cannot provide decision-makers with a warning function to identify the timing of a bubble burst. As investors, if we can understand the information expected by the public, we can formulate corresponding hedging strategies to mitigate the adverse effects of a bubble burst. The LPPLS method can not only help us infer the specific possible timing of a bubble burst but also grasp the degree of the herding effect in the process of bubble formation.

LPPLS results

Although BSADF can be used to identify the starting point, the ending point and the number of bubbles, the method is still based on events that have occurred and is a statistic of extreme behavior. In fact, what investors care about is not the bubble itself but the bursting point of the bubble.

In addition, since BSADF cannot reveal the mechanism of bubble generation, the method can hardly provide more information for macro decision-makers and managers. The appearance of different bubbles may have different microscopic mechanisms, such as the degree of the herding effect and the level of oscillation frequency. These characteristics all reflect different external factors that cause bubbles. It is actually possible for market regulators to use this information to issue signals to investors to avoid worse damage caused by bubbles.

Since LPPLS is based on out-of-sample prediction and can reveal the bubble generation mechanism, we use it as the main bubble analysis method.