

BSADF and GSADF

⚙ Status	Not Started
🎯 Project	<u>Statistic</u>
🏷 Tags	

Idea

Note 1

The PWY Test for Bubbles

BSADF là gì ?

The Rolling Window GSADF Test for Bubbles

GSADF xác định khoảng xảy ra bubbles kiểu gì

Note 2

Idea



Ý tưởng của việc phát hiện bubbles : stationary process chỉ ra rằng giá sẽ quay trở về fundamental value của nó, trong khi nonstationary unit root có đặc điểm gợi ý rằng sự sai lệch giá vẫn tồn tại trong fundamental value, cái mà chỉ ra rằng có bubbles

→ Việc của chúng ta là dùng GSADF model để check tính stationary của dữ liệu (

$\beta < 0$ nếu là stationary)

Về cơ bản : ADF sẽ check đc tính stationary tuy nhiên nó nằm trong khoảng nào ??

Khi đó GSADF sẽ giúp chúng ta biết được khoảng đó.

Khi có được khoảng xuất hiện bubbles thực hiện việc dịch tham số ngày kết thúc trong model LPPLS để tìm bursting point.

Note 1

TABLE 1
THE ASYMPTOTIC AND FINITE SAMPLE CRITICAL VALUES OF THE SADF AND GSADF TESTS AGAINST AN EXPLOSIVE ALTERNATIVE

	SADF			GSADF		
	90%	95%	99%	90%	95%	99%
Asymptotic critical values						
$r_0 = 0.190$	1.10	1.37	1.88	1.67	1.89	2.37
$r_0 = 0.137$	1.12	1.41	2.03	1.78	2.01	2.48
$r_0 = 0.100$	1.20	1.49	2.07	1.97	2.19	2.69
$r_0 = 0.074$	1.21	1.51	2.06	1.99	2.20	2.62
$r_0 = 0.055$	1.23	1.51	2.06	2.08	2.30	2.74
Finite sample critical values						
$T = 100, r_0 = 0.190$	0.98	1.30	1.99	1.65	2.00	2.57
$T = 200, r_0 = 0.137$	1.12	1.40	1.90	1.84	2.08	2.70
$T = 400, r_0 = 0.100$	1.19	1.49	2.05	1.92	2.20	2.80
$T = 800, r_0 = 0.074$	1.25	1.53	2.03	2.10	2.34	2.79
$T = 1600, r_0 = 0.055$	1.28	1.57	2.22	2.19	2.41	2.87

NOTES: The asymptotic critical values are obtained by numerical simulations with 2000 replications. The Wiener process is approximated by partial sums of $N(0, 1)$ with 2000 steps. The finite sample critical values are obtained from Monte Carlo simulation with 2000 replications.

trong này con số 90%, 95%, .. gọi là mức ý nghĩa (khi $p\text{-value} \leq 100-95$) thì gọi là nonstatility.

Definition and interpretation [\[edit \]](#)

Definition [\[edit \]](#)

The p -value is the probability under the null hypothesis of obtaining a real-valued test statistic at least as extreme as the one obtained. Consider an observed test-statistic t from unknown distribution T . Then the p -value p is what the prior probability would be of observing a test-statistic value at least as "extreme" as t if null hypothesis H_0 were true. That is:

- $p = \Pr(T \geq t \mid H_0)$ for a one-sided right-tail test-statistic distribution.
- $p = \Pr(T \leq t \mid H_0)$ for a one-sided left-tail test-statistic distribution.
- $p = 2 \min\{\Pr(T \geq t \mid H_0), \Pr(T \leq t \mid H_0)\}$ for a two-sided test-statistic distribution. If the distribution of T is symmetric about zero, then $p = \Pr(|T| \geq |t| \mid H_0)$.

<https://en.wikipedia.org/wiki/P-value>

Vậy tức là nó lấy $p = P(T \geq 0.95)$

The PWY Test for Bubbles

PWY Test dựa vào lặp lại việc estimation ADF model trên một chuỗi mẫu mở rộng tiến về phía trước, và kết quả kiểm định được thu được dưới dạng giá trị sup(Maximum) của chuỗi thống kê ADF tương ứng.



Về cơ bản SADF có điểm bắt đầu r_1 cố định, kích thước của 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 $r_0 \rightarrow r_w = r_2$

2. Khi tăng kích thước của 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 lớn nhất (supremum) của chuỗi

ADF_{r_2} đối với các $r_2 \in [r_0, 1]$:

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

Việc lấy supremum của chuỗi ADF_{r_2} được gọi là right-sided unit root tests.

BSADF là gì ?

Cụ thể, BSADF là backward SADF test perform dựa trên sup ADF test trên một chuỗi mẫu mở rộng ngược (backward expanding sample sequence) điểm kết thúc mỗi sample cố định ở r_2 (tương ứng với điểm cuối của cửa sổ) và điểm bắt đầu biến thiên từ 0 đến $r_2 - r_0$ (điểm bắt đầu cửa sổ).

Khi đó thống kê ADF cho chuỗi tương ứng là :

$$\{ADF_{r_1}^{r_2}\}_{r_1 \in [0, r_2 - r_0]} \quad (1)$$

Thống kê BSADF sau đó sẽ được xác định là sup của thống kê ADF phương trình (1) và được viết lại như sau:

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

The Rolling Window GSADF Test for Bubbles



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 $[0, r_2 - r_0]$.

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

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 $r_0 = \frac{1.8}{\sqrt{t}} + 0.01$

$$r_w = r_2 - r_1$$

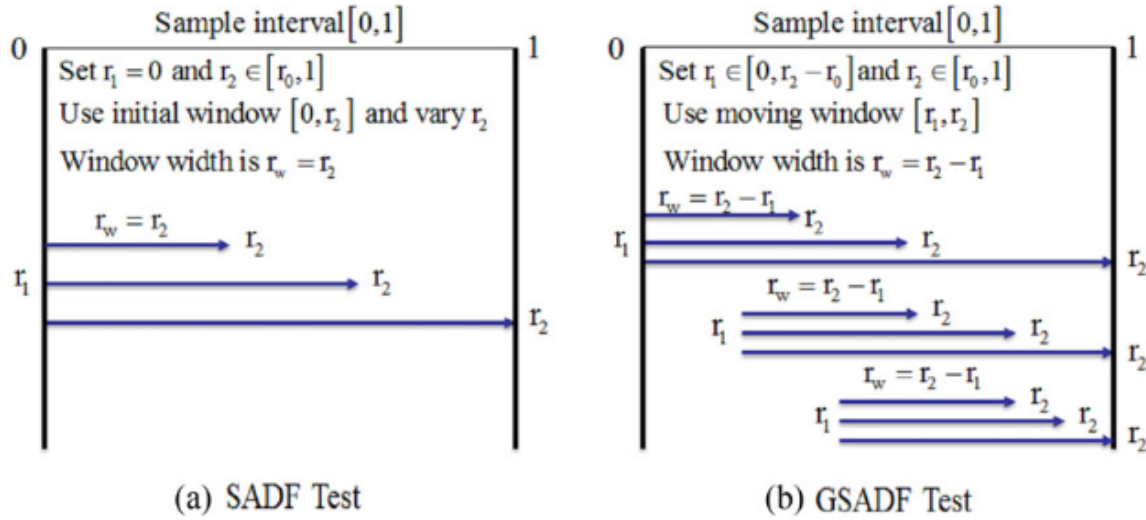


FIGURE 1

THE SAMPLE SEQUENCES AND WINDOW WIDTHS OF THE SADF TEST AND THE GSADF TEST

[1] TESTING FOR MULTIPLE BUBBLES

GSADF xác định khoảng xảy ra bubbles kiểu gì

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

Tr_2 .

Điểm bắt đầu ước tính của bong bóng là quan sát thời gian đầu tiên, ký hiệu là Tr_e , trong đó ADF_{r_2} vượt qua giá trị tới hạn tương ứng từ bên dưới, trong khi điểm kết thúc ước tính là quan sát theo thời gian đầu tiên sau Tr_e , được biểu thị bởi Tr_f , trong đó ADF_{r_2} vượt qua giá trị tới hạn (confidence interval) phía bên phải. Từ đó, các ước tính của chu kỳ xảy ra bubble được xác định bởi :

$$r_e = \inf_{r_2 \in [r_0, 1]} \{r_2 : BSADF_{r_2}(r_0) > cv_{r_2}^{\beta Tr_2}\}$$

$$r_f = \inf_{r_2 \in [r_e, 1]} \{r_2 : BSADF_{r_2}(r_0) > cv_{r_2}^{\beta_T r_2}\}$$

Trong đó : $cv_{r_2}^{\beta_T r_2}$ là giá trị tới hạn $100(1 - \beta_T)\%$ của thống kê SADF dựa trên quan sát $[Tr_2]$. $BSADF(r_0)$ với $r_2 \in [r_0, 1]$ là thống kê SADF trễ liên quan đến thống kê GSADF theo mối quan hệ sau :

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

Note 2

Đầu tiên ta phải kết luận dữ liệu bitcoin là có xu hướng và có drift (Const hay Khoảng hệ số α ban đầu) bởi nó lấy giá trị bắt đầu từ 2014 mà nên phải có 1 const > 0 khởi đầu.

→ Nó thuộc type3 của ADF : Có Drift and Trend

Hệ số lag là bao nhiêu ? theo như paper nói thì là mình quyết định một số con số như : 130 hay 260.

Vậy mô hình ADF tổng quan :

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

Và điểm quan trọng ở đây là null hypothesis khẳng định rằng nếu tồn tại unit root, hay $\delta = 1$ thì

p-value thu được phải nhỏ hơn significance level (0.05) để phủ nhận null hypothesis. Từ đó ta suy ra

$p - value \leq 0.05 \rightarrow$ stationary \rightarrow Không có Bubble.

```
quantiles <- quantile(data, probs = c(0.9, 0.95, 0.99))
```

Tính toán quantile với mức độ tin cậy 0.9, 0.95, 0.99.

Augmented Dickey-Fuller Test
alternative: stationary

```
Type 1: no drift no trend
      lag   ADF p.value
[1,]    0 -2.14  0.0346
[2,]    1 -2.24  0.0267
[3,]    2 -1.16  0.2550
Type 2: with drift no trend
      lag   ADF p.value
[1,]    0 -2.12  0.286
[2,]    1 -2.54  0.136
[3,]    2 -2.14  0.280
Type 3: with drift and trend
      lag   ADF p.value
[1,]    0 -2.39  0.409
[2,]    1 -2.61  0.329
[3,]    2 -1.54  0.742
----
```

Trong thư viện ADF_test nó sẽ thử vào model của ADF với lag = n, có drift và trend hay ko, thay số và tính ra p-value and ADF value