

# BÁO CÁO CHUỖI THỜI GIAN

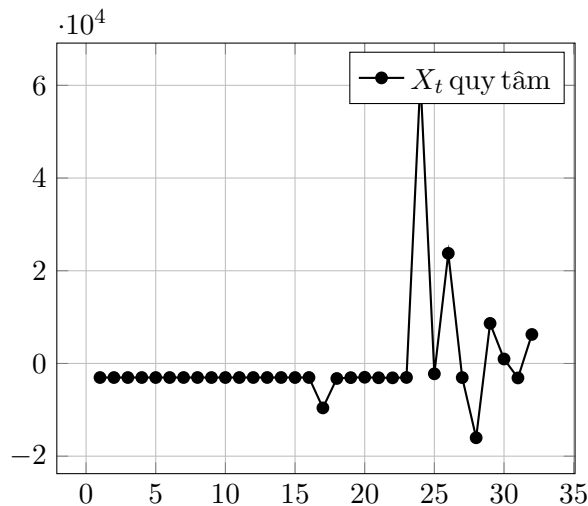
Nguyễn Công Thịnh - 20170924

Ngày 21 tháng 1 năm 2021

## 1 Dữ liệu quan sát. Xử lý thô chuỗi thời gian.

1	2	3	4	5	6	7	8
3.335316	1.778595	1.085326	6.695278	1.698594	5.037846	7.181806	9.533516
9	10	11	12	13	14	15	16
1.963198	3.847622	3.089793	7.453181	7.600316	9.742097	1.450212	9.604956
17	18	19	20	21	22	23	24
-6548.799	-201.3188	-19.45226	57.56791	-45.96305	-61.44261	17.30406	64408.31
25	26	27	28	29	30	31	32
806.715	26794.59	-10.17868	-12992.35	11681.29	3998.828	-84.71642	9295.035

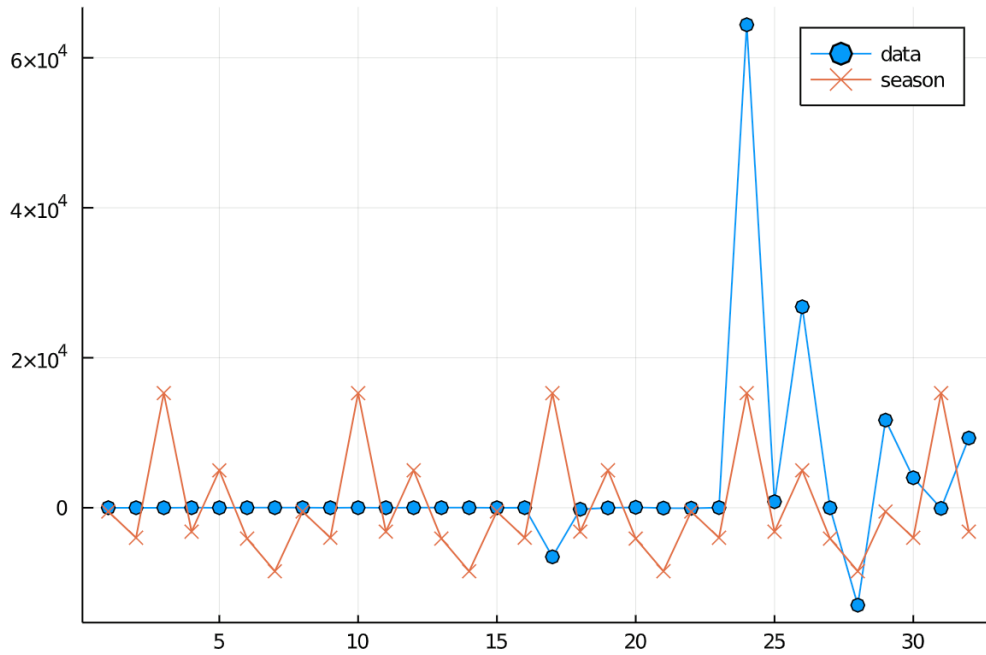
Bảng 1: Bảng dữ liệu thực tế.



Hình 1: Đồ thị biểu diễn chuỗi thời gian.

Bước đầu của việc xử lý chuỗi thời gian là xét tính có khuynh và có mùa của bộ dữ liệu.

- **Khuynh:** Dựa vào bảng số liệu cũng như thông qua đồ thị, dễ thấy chuỗi thời gian ứng với bộ dữ liệu đã cho không có khuynh.
- **Mùa:** Do chênh lệch rất lớn giữa giá trị lớn nhất, giá trị nhỏ nhất của bộ số liệu và các giá trị khác nên dựa vào trực quan không thể thấy được yếu tố mùa trên đồ thị. Tuy nhiên, nếu ta phân đồ thị thành các đoạn có độ dài là 7, thành phần mùa sẽ dễ nhận biết hơn



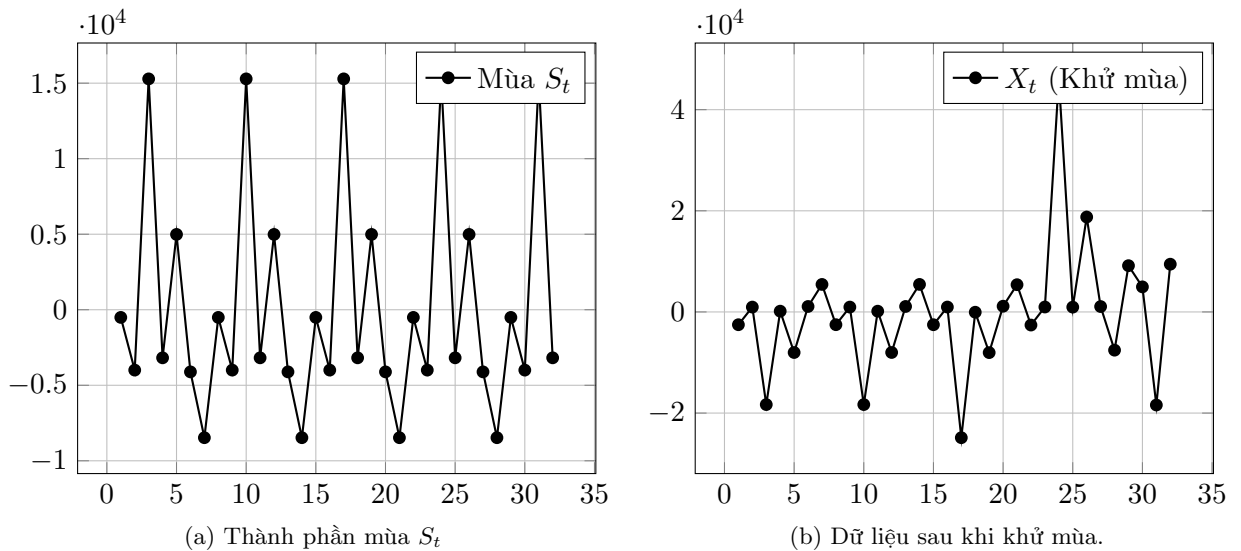
Hình 2: Biểu diễn đồ thị bao gồm thành phần mùa.

Để khử thành phần mùa trong bộ dữ liệu, đầu tiên ta xét mô hình cộng tính:

$$X_t = S_t + R_t$$

trong đó,  $S_t$  là thành phần mùa của mô hình,  $R_t$  là phần dư sau khi khử mùa.

Ta khử thành phần mùa của mô hình trên dựa theo *phương pháp S1 (ước lượng các thành phần khuynh và mùa)*, kết quả thu được gồm thành phần mùa  $S_t$  và phần dư  $R_t$  như Hình 3.



Hình 3: Kết quả sau khi khử thành phần mùa của mô hình.

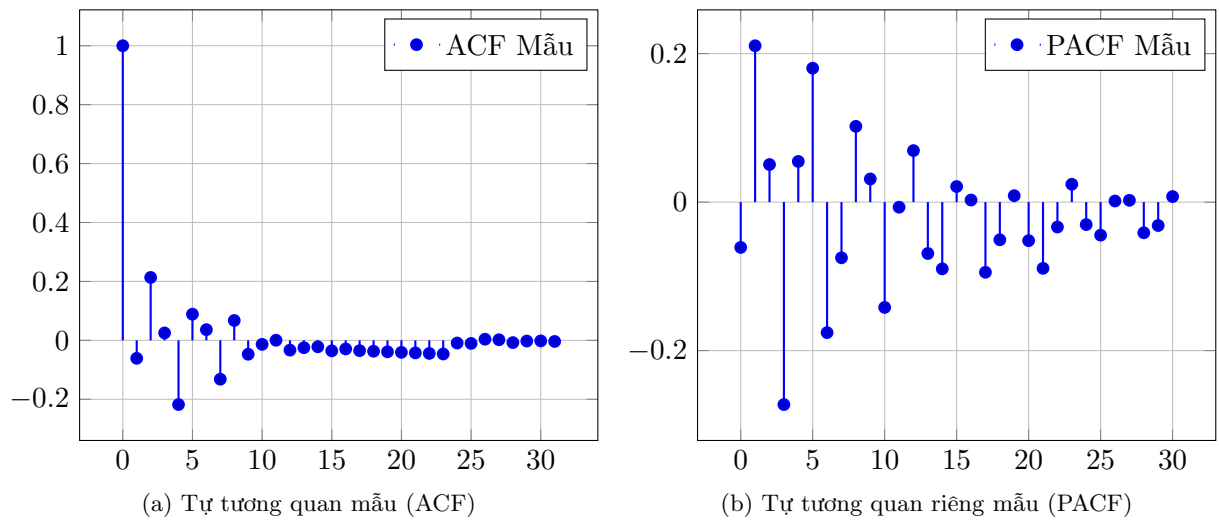
Thực hiện kiểm định Dickey-Fuller mở rộng (ADF), dựa trên việc so sánh giá trị Statistical Test và các giá trị ngưỡng (Critical Value). Do giá trị tuyệt đối của Statistical Test lớn hơn giá trị tuyệt đối của các Critical Value nên ta kết luận chuỗi đã dừng.

Statistical Test	Critical Value 1%	Critical Value 5%	Critical Value 10%
-5.7088	-3.6614	-2.9605	-2.6193

Bảng 2: Kết quả kiểm định Dickey-Fuller

## 2 Nhận dạng mô hình

Đầu tiên, ta biểu diễn tự tương quan mẫu (ACF) và tự tương quan riêng mẫu (PACF) dưới dạng đồ thị:



Hình 4: Biểu diễn tự tương quan và tự tương quan riêng mẫu.

Dễ thấy, đồ thị tự tương quan riêng mẫu có dạng *sin* tắt dần, nên bước đầu dự đoán mô hình có dạng **ARMA(p,q)**.

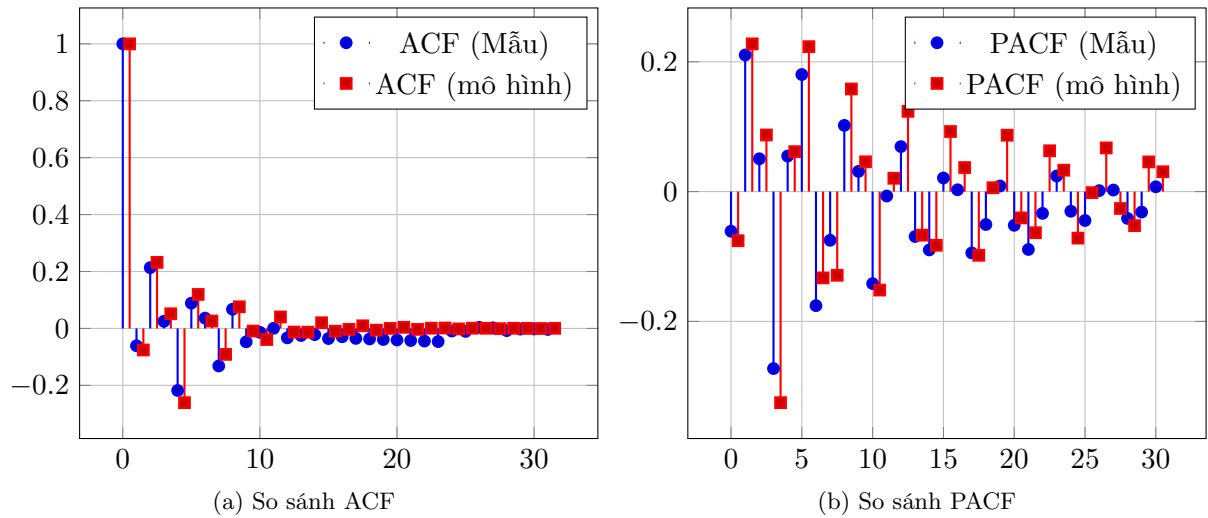
Để ước lượng bộ tham số phù hợp cho mô hình, ta cho các giá trị p và q chạy và sử dụng phương pháp hợp lý cực đại (Maximum Likelihood - ML), đánh giá AICc. Kết quả thu được cho dưới bảng sau đây:

p	q	p-value	AICc
4	2	0.925457	710.049
5	2	0.923638	713.97
5	3	0.919804	718.264
2	4	0.913197	710.152
3	3	0.908046	710.178

Bảng 3: Kết quả ước lượng bộ tham số cho mô hình.

Dựa trên Bảng 3, bộ tham số ứng với mô hình **ARMA(4,2)** có AICc bé nhất.

Tiếp theo, ta sẽ so sánh dạng tắt dần của ACF, PACF của mô hình ARMA(4,2) và ACF, PACF mẫu.



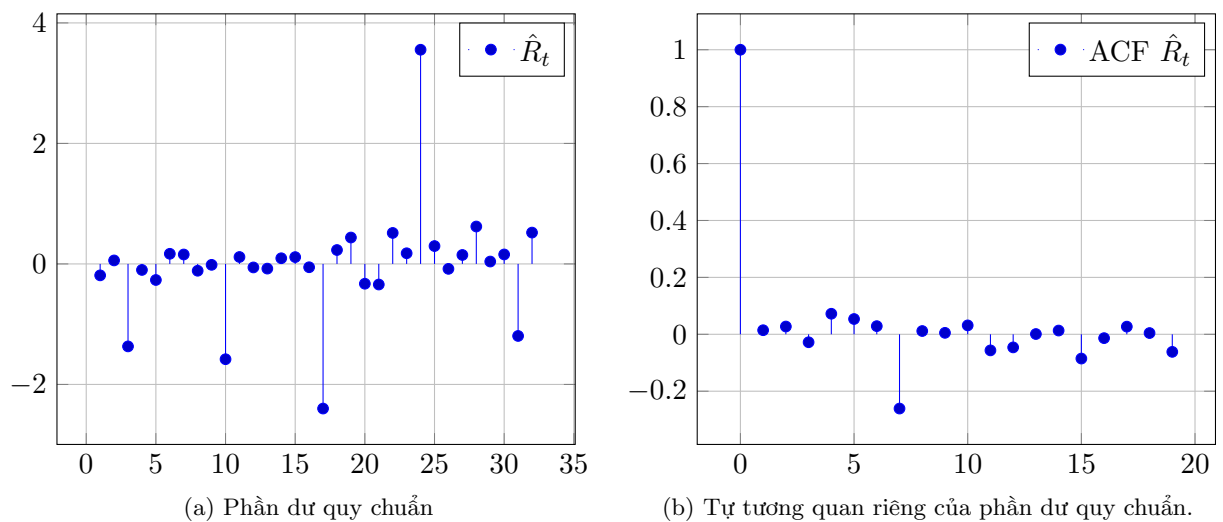
Hình 5: So sánh ACF và PACF của mô hình và mẫu.

Nhìn vào đồ thị, dễ thấy dạng đồ thị ACF, PACF của mô hình ARMA(4,2) và mẫu của  $X_t$  khá tương đồng.

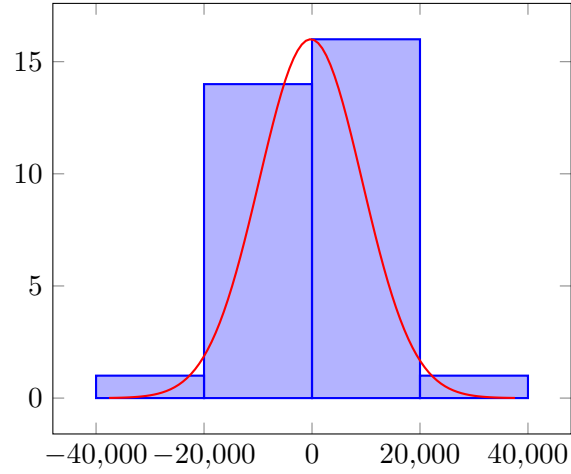
### 3 Kiểm tra tính phù hợp của mô hình

Phần dư quy chuẩn và tự tương quan của phần dư quy chuẩn tương ứng với mô hình ARMA(4,2) được cho bởi Hình 6.a, 6.b.

Dựa vào biểu đồ tần suất của phần dư được cho như Hình 7, ta thấy một đường cong phân phối chuẩn được đặt chồng lên biểu đồ tần số. Đường cong này có dạng hình chuông, phù hợp với dạng đồ thị của phân phối chuẩn. Giá trị trung bình gần bằng 0, độ lệch chuẩn gần bằng 1, như vậy có thể nói, phân phối phần dư xấp xỉ chuẩn. Do đó, có thể kết luận rằng: Giả thiết phân phối chuẩn của phần dư không bị vi phạm.



Hình 6: Biểu diễn phần dư quy chuẩn, tự tương quan của phần dư quy chuẩn.



Hình 7: Biểu đồ tần suất của phần dư.

Đồng thời, dựa trên đồ thị tự tương quan của phần dư quy chuẩn (Hình 6.b), các giá trị tự tương quan luôn nằm trong khoảng tin cậy 95%, tức là nhỏ hơn giá trị  $1.96 \times \frac{\sigma^2}{\sqrt{n}} = 1.96 \times \frac{1}{\sqrt{32}} = 0.3465$ . Vì vậy, mô hình ARMA(4,2) chấp nhận được với bộ số liệu đã cho.

Áp dụng thuật toán Hannan-Rissanen, ta ước lượng được bộ hệ số của mô hình như sau:

$$X_t = -0.3516X_{t-1} - 0.4478X_{t-2} + 0.03634X_{t-3} - 0.1387X_{t-4} + Z_t + 0.3003Z_{t-1} + 0.8187Z_{t-2}$$

với  $Z_t \sim WN(0, \sigma^2)$ ,  $\sigma^2 = 0.121970E + 09$ .